

GAO

Report to the Chairman, Subcommittee on
Seapower and Strategic and Critical
Materials, Committee on Armed Services,
House of Representatives

May 1989

NAVY WEAPONS TESTING

Defense Policy on Early Operational Testing



**National Security and
International Affairs Division**

B-222886

May 8, 1989

The Honorable Charles E. Bennett
Chairman, Subcommittee on Seapower and
Strategic and Critical Materials
Committee on Armed Services
House of Representatives

Dear Mr. Chairman:

This report responds to your Committee's request that we evaluate the Navy's operational testing and evaluation (OT&E)¹ conducted before decisions are made on the full-scale development or low-rate initial production of weapon systems. Our objective was to determine whether the Navy conducted OT&E to support these decisions. Department of Defense (DOD) acquisition policy provides for flexible approaches but generally calls for early operational testing. Appendix I contains more detail on our objectives, scope, and methodology.

Results in Brief

DOD acquisition directives generally call for OT&E input at all major acquisition decision points, including full-scale development and low-rate initial production. There is some question whether OT&E can be accomplished for many systems before the full-scale development decisions are made, and we found that the Navy typically made these decisions before any OT&E was accomplished. We also found, however, that in many cases, decisions to proceed with low-rate initial production of weapon systems were made before any operational testing was completed. (See app. IV.)

Since operational testing often was not done before these early acquisition decisions were made, DOD is now encouraging the services to conduct "operational assessments" before these decision points. Operational assessments are based on computer modeling, simulation, analyses of program documents such as system requirements, engineering proposals, and design specifications, or any kind of testing that may be available but not necessarily operational testing.

¹OT&E is defined as field testing, under realistic combat conditions, of any item (or key component) of weapons, equipment, or munitions for the purpose of determining the effectiveness and suitability of the item for use in combat by typical military users and the evaluation of the results of such tests.

One reason operational assessments are being encouraged is that early operational testing requires the availability of hardware to test, generally a prototype. Prototypes of new weapons are expensive and time-consuming to build and often are not available prior to the full-scale development decision.

Independent agencies established within each service to conduct OT&E are being asked to conduct operational assessments. Both DOD and the Congress understand that these agencies conduct realistic field testing of systems or system components, and decisionmakers rely on these agencies' reports in deciding whether to proceed with a weapon system. These agencies have expressed concern that performing operational assessments may compromise their independence and cause misunderstandings over the nature and extent of testing actually performed on a weapon system.

Operational Testing Not Completed Before Early Acquisition Decisions

During fiscal years 1985 to 1988, the Navy typically approved weapon systems for full-scale development and, in many cases, for low-rate initial production before any OT&E was completed. The Navy did conduct a limited number of operational assessments to support early milestone decisions.

We reviewed 19 Navy systems to determine whether OT&E was conducted in support of early acquisition decisions. This included 10 full-scale development decisions and 10 low-rate production decisions. (For one system, both decisions were made during the period covered by our review.) OT&E was not conducted before any of the 10 full-scale development decisions and was conducted before only 3 of the 10 low-rate initial production decisions. For three of the systems that had not been operationally tested, the Navy's Operational Test and Evaluation Force (OPTEVFOR) prepared operational assessments to support initial production decisions. Two of these assessments were so limited that OPTEVFOR could not project the systems' potential effectiveness or suitability. (See app. IV.)

Operational Testing and the Acquisition Process

The concept of early OT&E has been a part of the weapon system acquisition process for many years. With strong congressional support, DOD adopted the "fly-before-buy" concept in the early 1970s and established separate OT&E phases to support corresponding phases of the acquisition process. Over the years this phased approach to OT&E has remained a

tenet of DOD's published acquisition policy, but this policy has been difficult to apply to so-called concurrent acquisition programs. (In concurrent programs, development, production, and test and evaluation overlap, to some extent, rather than occur sequentially.)

Weapon system programs are subject to review at various "milestone" decision points. (See app. II.) The milestone process is predicated on the principle that systems advance to higher acquisition phases by demonstrating that they have met prescribed technical specifications and performance thresholds, not by meeting predetermined schedules. Testing and evaluation is DOD's method of verifying that these standards have been met.

The most visible OT&E phase is the final testing phase that is legislatively required to be completed for major programs before full-rate production decisions are made. This phase is known as operational evaluation (OPEVAL) and is sometimes referred to as the final examination.

Operational testing conducted before earlier acquisition decisions are made, while not legislatively required, is also important. This testing is recognized in DOD's acquisition policy directives that state, among other things, "OT&E shall be structured to provide inputs at each decision point, including major milestones." These early decisions firmly commit a service to a weapon system design selected from competing alternatives. A Navy instruction on acquisition notes that milestone II is the "single most critical decision point, for it is here that the Navy makes a firm commitment to the program; once started, it is difficult to turn back." (Milestone II authorizes full-scale development and may authorize low-rate initial production.) Decisions to begin full production and deployment of a system are important, but they are often made after most program funds have been committed and, except in extreme cases, at a time when production can only be slowed, not stopped.

The purpose of early OT&E is to provide early projections of potential operational effectiveness and suitability for systems being considered for full-scale development or low-rate initial production. In contrast, OT&E conducted in support of a full-rate production decision is to verify operational effectiveness and suitability and, thus, to ensure that systems meet operational thresholds. Both types of testing can be conducted using prototypes. However, during early OT&E, before the full-scale development decision, prototypes may not be available or when available may be less advanced and may not fully represent the systems

or subsystems to be produced. Prototypes tested in support of full-rate production decisions are supposed to be production-representative.

In a 1987 report, the Secretary of Defense proposed to the House and Senate Committees on Armed Services that a change be made in the approach to weapon system testing and evaluation. Traditional OT&E typically was not being performed during the early phases of the acquisition cycle and consequently an early operational perspective based on actual testing was not being provided to decisionmakers. The Secretary's report proposed an approach whereby operational assessments would be made to fill this informational void. According to this proposal, OT&E conducted during the phases before OPEVAL would be called operational assessments and could be based on computer modeling, simulation, paper analyses,² or any kind of testing.

In commenting on our report, DOD said that the purpose of redefining OT&E was to enhance systems' evaluations by providing operationally oriented evaluations of valid information from any available source as early as possible in the acquisition process. We agree that OT&E agencies need to use valid information from any available source. We also agree that the emphasis on operational assessments is intended as a step forward in filling an informational void by assuring that some operationally oriented evaluations not previously available would be provided to decisionmakers at the earliest possible time. We are concerned, however, that over time operational assessments may be seen as a substitute for operational testing rather than as a supplement to fill an informational gap and that, as a result, the policy calling for early OT&E may be deemphasized.

Prototyping New Systems

Deciding whether to perform traditional OT&E or operational assessments during the early phase of the acquisition process is to some extent a question of how much program concurrency is necessary and if prototyping is the only reliable method of obtaining performance data to support program decisions that commit large amounts of money. Unlike operational assessments, traditional OT&E requires an actual working system (either a production model or a prototype) that can be tested under realistic conditions.

²OPTEVFOR officials told us that "paper analyses" refer to analyzing a variety of documents such as system requirement documents, engineering proposals, and design specifications.

The debate on whether to build prototypes is not new, and in 1986 the President's Blue Ribbon Commission on Defense Management (the "Packard Commission") again raised this issue. The Commission recommended that high priority be given to building and testing prototypes of major weapon systems before decisions are made to begin full-scale development. The Commission's underlying premise was that prototype testing is essential if the services are to exploit their technological advantages fully. Although warning that state-of-the-art technology should be used only when the benefits outweigh the risks, the Commission urged DOD to place a greater emphasis on using technology to reduce cost.

The problem, as the Commission noted, is getting the information needed to make the risks-versus-benefits analysis. According to the Commission, the only consistently reliable means of getting such information is by building prototypes that embody the new technology. The Commission concluded that such a prototype at either the system or the subsystem level should undergo operational and developmental testing to detect deficiencies before a decision is made to proceed with full-scale development.

The Commission's June 30, 1986, report, A Quest for Excellence: Final Report to the President by the President's Blue Ribbon Commission on Defense Management, stated that:

"A high priority should be given to building and testing prototype systems and subsystems before proceeding with full-scale development."

"We recommend that operational testing begin early in advanced development and continue through full-scale development, using prototype hardware."

Reaction to the Commission's recommendation was mixed. In a September 1986 report to the Secretary of Defense, the Director of OT&E stated that a renewed emphasis on cutting the cost and the length of the acquisition process had been encouraged by the Commission's recommendations, including the recommendation to use extensive prototyping.

In his September 1987 report to the House and Senate Committees on Armed Services, the Secretary reaffirmed that conclusion, also citing the recommendation to use extensive prototyping in testing as a positive contribution. However, in describing the revised approach to early OT&E that would rely heavily on operational assessments, the Secretary discussed the use of computer modeling and simulation in preproduction

OT&E but did not emphasize the testing of prototypes. The Secretary's report is not clear whether any actual testing of a preproduction item would occur until after the low-rate initial production decision is made.

The Director of OT&E told us that, although he agrees with the intent of the Commission's recommendation on preproduction OT&E, he does not believe prototypes are feasible for most acquisition programs. He said that prototypes are extremely costly and time-consuming to build and are usually not representative of the systems eventually produced. The Director said that he supports traditional early OT&E whenever it is reasonable, but, in most cases, it is better to have early performance assessments based on data obtained from realistic, high-quality simulators.

Operational Testing or Operational Assessments

There are two features of early OT&E that make it a unique source of information for DOD and the Congress: the independence of the agencies that conduct OT&E and the emphasis on realism that distinguishes operational testing from other kinds of weapon system testing. Both of these features may be affected if operational assessments become a substitute for early OT&E rather than a supplement to such testing and evaluation.

When OT&E was incorporated into the acquisition process, one agency in each service was chartered to plan, test, evaluate, and report on realistic tests of weapon system performance. That agency is OPTEVFOR in the Navy, and its Commander reports directly to the Chief of Naval Operations. By remaining separate and distinct from developing, procuring, and using commands, OPTEVFOR and the other services' OT&E agencies have no vested interest in systems under development. Thus, the unique role of the OT&E agencies is to serve as impartial judges. However, if OPTEVFOR is given responsibility for operational assessments, it will be more dependent on the results of modeling, simulation, studies, and tests conducted by agencies or contractors with a strong interest in getting systems into production.

In July 1988 testimony³ before congressional committees investigating alleged DOD procurement scandals, the Comptroller General emphasized that early weapon system testing must be performed in a manner that ensures independence and that is clearly separated from contractors that have played a key role in a system's development.

³Defense Management and Procurement Issues (GAO/T-NSIAD-88-39, July 13, 1988).

DOD's Director of OT&E told us that OPTEVFOR had objected to preparing operational assessments based on data from other sources without having actually tested the systems. The Director said he told OPTEVFOR that an OT&E agency is the logical organization for preparing operational assessments because of its independence. While this independence may make OT&E agencies desirable organizations to perform operational assessments, their independence may be compromised over time by relying on data from nonindependent sources and simulation models owned by other organizations that may not have been validated.⁴

Because operational assessments rely on information from nonindependent sources, the important distinctions between OT&E and developmental test and evaluation (DT&E) can also be blurred. These two principal methods of assessing weapon systems are important in the acquisition process, but they have different purposes, use different criteria, and are conducted under different conditions. (See app. III.)

Under the concept for operational assessments outlined in the Secretary's September 1987 report, the only test results available may be those derived from DT&E conducted by contractors or agencies responsible for developing the system. Furthermore, assessments may be based largely on computer modeling, simulation, paper analyses, or other attempts to replicate real world conditions without the actual testing of any hardware. Yet assessment reports will be issued by an OT&E agency known for independent operational testing and used by persons who may not be aware of the distinction between OT&E and operational assessments. Although OPTEVFOR's reports describe limitations to the scope of its work, summaries provided to decisionmakers sometimes omit those limitations.

In commenting on the report, DOD said that it did not believe the independence of OT&E agencies would be threatened by relying on the results of modeling, simulation, studies, or tests conducted by nonindependent sources. DOD said that the use of modeling and simulation in test and evaluation is not new and that the DOD directive on testing encourages the use of modeling and simulation to ensure the availability of operational effectiveness and suitability projections at milestone II. DOD also noted that an OT&E agency would not blindly use unvalidated data from outside sources.

⁴Validation refers to the process of determining agreement between the real world system being modeled and the model itself.

We agree that using computer modeling and simulation in test and evaluation is not new and that the independent testing agencies are permitted to use these tools in pre-milestone II OT&E. However, the directive on testing policy does not imply that the use of modeling and simulation should be a substitute for actual OT&E. While we agree that pre-milestone II OT&E cannot always be accomplished, the directive states that adequate testing and evaluation shall be accomplished and documented before a full-scale development decision and that this shall produce test data. The directive provides that, in all cases, appropriate and adequate testing and evaluation must precede the milestone II decision and that decisions to commit funds for long lead items or low-rate initial production (decisions that can be made at milestone II) must be supported by OT&E.

The need for early operational testing was recognized and emphasized by the Director of OT&E in his 1986 report to the Secretary of Defense. It stated that DOD policies should “encourage the use of simulation and modeling during all phases of the acquisition process to augment and supplement—not replace—the realistic field testing that always must be accomplished.” (Underscoring in the original.)

In commenting on our report, DOD stated that it is appropriate for an OT&E agency to provide early operational assessments before OT&E information from actual field testing is available. We agree that this may be an appropriate function for an independent testing agency such as OPTEVFOR, but its mission does not include this function. OPTEVFOR’s current mission is to (1) test and evaluate systems in the anticipated environment and against the anticipated threat, (2) develop and validate procedures and tactics for employing systems, and (3) assist developing agencies in accomplishing DT&E, when directed by the Chief of Naval Operations. Thus, if OPTEVFOR’s mission is to be expanded to include the performance of operational assessments, its mission statement should be clarified to reflect this new role.

OPTEVFOR had performed some operational assessments (see app. IV) and recognized the shortcomings in its initial approach. In 1988 it adopted an approach that requires other agencies to assist OPTEVFOR. This joint arrangement was dictated in part by OPTEVFOR’s lack of capability to develop computer modeling and simulation data or to verify such data provided by other agencies. Under this approach, OPTEVFOR established a working relationship with the Center for Naval Analyses and recommended to the Chief of Naval Operations that the Center be granted a charter for independence similar to OPTEVFOR’s. According to an OPTEVFOR

official, as of January 1989, such a charter had not been granted. However, based on the Center's first involvement with OPTEVFOR under this approach, it appears that the Center's role may go beyond developing or verifying modeling and simulation data.

OPTEVFOR's first operational assessment under this approach began in 1988. According to OPTEVFOR, the assessment would "be a pioneering effort to project operational performance from limited test data by conducting a comprehensive analysis within a well defined architecture." It was to consist of a combination of modeling, simulation, hybrids, and operational tests of individual components and extensive analysis "to extrapolate to the desired projection of operational performance." The Director of Navy Laboratories was to have prime responsibility for developing the test architecture and identifying test resources and assist in developing the operational test plan. The Center was to take the lead in evaluating test data and extrapolating results to the projection of operational performance. This operational assessment was still underway as of January 1989.

A second operational assessment using this approach was completed in January 1989. Although the goals of this operational assessment were to assess the operational concept and to project potential operational performance, OPTEVFOR's report of this assessment stated that its conclusions on the weapon system's capabilities "should be viewed as statements of opinion vice findings of fact."

Conclusions

The Navy usually did not conduct OT&E before decisions were made to begin full-scale development or low-rate initial production of weapon systems as generally called for by DOD policy. To fill the informational gap, the Director, OT&E, is encouraging the emerging practice of conducting operational assessments.

Ideally, operational testing of representative hardware should be conducted as early as possible in the acquisition cycle. This testing would include the operational testing and evaluation of prototypes to support both full-scale development and low-rate initial production decisions. We recognize, however, that there are times when this cannot be done. The cost of building prototypes may be excessive or the urgent need for a system may outweigh the time required to build and test prototypes. Nonetheless, we support DOD's current policy that calls for early OT&E at all decision points in the acquisition process. We recognize, however, that this typically has not been done in the Navy.

The concept of conducting operational assessments proposed by the Director, OT&E, is a way to fill the current informational void, particularly when a high degree of program concurrency is involved. Nonetheless, we believe that at a minimum, DOD should clarify the distinction between OT&E and operational assessments and the degree to which assessments may be relied upon in the weapon system acquisition process.

Since the services' OT&E agencies are independent, they are attractive candidates for preparing operational assessments. But because of their traditional role of conducting operational testing and decisionmakers' reliance on this independent testing, we are concerned that their independence could be compromised by a role that requires them to predict weapon system performance on the basis of data provided by contractors and agencies responsible for developing and using the systems. The likelihood of this happening is increased when key roles, such as analyzing data and projecting performance, are delegated to agencies other than the independent OT&E agencies.

Recommendations

We recommend that the Secretary of Defense reemphasize the desirability of performing OT&E as early as possible in the acquisition cycle as called for by DOD's acquisition directives. We also recommend that in so doing, the Secretary clarify when it is appropriate for decisionmakers to rely on operational assessments that may not include the operational testing of any hardware and when actual operational testing and evaluation must occur. In addressing this issue, the Secretary should assure that the independence of the services' OT&E agencies is not compromised and that the basis for weapon system assessments is fully disclosed when the assessments are reported to congressional and DOD decisionmakers.

Agency Comments and Our Evaluation

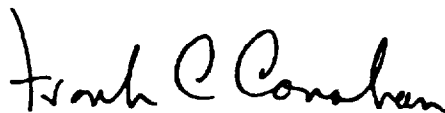
DOD partially agreed with our findings and with our recommendation that the independence of the services' OT&E agencies not be compromised. However, DOD interpreted our report as stating that conducting OT&E before full-scale development is statutorily required and that operational assessments should not be relied upon in the acquisition process. DOD disagreed with this point of view. (A copy of DOD's comments is included in app. V.)

We believe DOD may have misinterpreted our report, and we have modified the report to clarify our position. First, we never asserted a statutory requirement for OT&E before full-scale development. We also recognize that, although generally called for by DOD directives, the Navy typically had not completed operational testing and evaluation before the full-scale development decisions were made and had seldom completed any segment of OT&E before low-rate initial production decisions were made. Emphasis by DOD's Director, OT&E, on performing operational assessments when hardware is unavailable is therefore a step forward in filling a void in the availability of operationally oriented information. We continue to be concerned, however, that operational assessments not be seen as a substitute for OT&E. Such assessments should be seen instead as a means of providing supplemental information.

Copies of this report are being sent to the Secretaries of Defense, the Navy, the Army, and the Air Force; responsible congressional committees and interested parties; and to others upon request.

This report was prepared under the direction of Harold J. Johnson, Associate Director, Navy Issues. Other major contributors are listed in appendix VI.

Sincerely yours,



Frank C. Conahan
Assistant Comptroller General

Contents

Letter		1
Appendix I Objectives, Scope, and Methodology		14
Appendix II Acquisition Milestones		16
Appendix III Principal Types of Testing and Evaluation		17
Appendix IV Navy Weapon Systems Reviewed		18
Appendix V Comments From the Department of Defense	GAO Comments	20 33
Appendix VI Major Contributors to This Report	National Security and International Affairs Division Washington, D.C. Norfolk Regional Office	36 36 36
Tables	Table IV.1: Programs Authorized for Full-Scale Development From October 1984 to March 1988 Table IV.2: Programs Authorized for Low-Rate Initial Production From October 1984 to March 1988	18 19

Abbreviations

DOD Department of Defense
DT&E Development Testing and Evaluation
GAO General Accounting Office
OT&E Operational Test and Evaluation
OPEVAL Operational Evaluation
OPTEVFOR Operational Test and Evaluation Force

Objectives, Scope, and Methodology

Our objectives were to (1) determine whether Navy weapon systems were being operationally tested before being approved for full-scale development and low-rate initial production, as generally called for by published Department of Defense (DOD) and Navy policy, and (2) identify reasons for any disparities between policy and practices. Our work focused on Navy systems, although we did discuss Army and Air Force practices with officials at those services' operational test and evaluation (OT&E) agencies.

We reviewed 19 Navy systems (see app. IV) that were approved for full-scale development or low-rate initial production in fiscal years 1985 to 1988. We considered only systems in the higher acquisition categories, based on the Navy's determination of cost and importance. Systems that are normally not subject to early OT&E, such as embedded computers, were excluded.

To determine OT&E policies and practices, we reviewed applicable DOD, Navy, Army, and Air Force directives and instructions; DOD reports; Navy program documents, test plans, and test reports; legislation and congressional reports; and related documents. We interviewed officials regarding their plans for making operational assessments and potential policy changes at the following offices.

Department of Defense
Office of the Deputy Under Secretary of Defense (Test and Evaluation)
Director, Operational Test and Evaluation
Defense Advanced Research Projects Agency

Department of the Navy
Office of the Chief of Naval Operations, Office of Research, Development, and Acquisition
Commander, Operational Test and Evaluation Force
Naval Air Systems Command
Naval Sea Systems Command
Space and Naval Warfare Systems Command

Department of the Air Force
Air Force Operational Test and Evaluation Center

Department of the Army
Operational Test and Evaluation Agency

Other
Center for Naval Analyses

At OPTEVFOR we reviewed information in the database and compared the dates of the agency's reports with the dates of program decisions to determine whether OT&E was conducted before the decisions and we interviewed OPTEVFOR officials to determine why early testing and evaluation was not conducted. We also interviewed officials at four Navy project offices to obtain their rationale for omitted testing. In addition, we interviewed Army and Air Force officials regarding the status of preproduction OT&E in their services. Our review was performed from October 1987 to January 1989 in accordance with generally accepted government auditing standards.

Acquisition Milestones

Decision point	Purpose
Milestone 0	Determines mission need, approves program initiation, and provides authority to budget for a new major program.
Milestone I	Authorizes concept demonstration and validation and establishes broad cost, schedule, and performance goals and thresholds.
Milestone II	Authorizes full-scale development and may authorize low-rate initial production. It also establishes more specific cost, schedule, and performance goals and thresholds.
Milestone III	Authorizes full-rate production and deployment. (There may be multiple milestone III production decisions; e.g., milestone IIIA for low-rate initial production and milestone IIIB for full production.)
Milestone IV	Normally occurs 1 to 2 years after initial deployment and identifies actions and resources to ensure that operational readiness and support objectives are met.
Milestone V	Normally occurs 5 to 10 years after initial deployment and determines whether major upgrade or replacement is necessary.

Principal Types of Testing and Evaluation

Weapon systems undergo two principal types of testing and evaluation: OT&E and developmental test and evaluation (DT&E). OT&E differs from DT&E in at least three important respects.

- OT&E is the field test to be conducted in a realistic combat environment against a representative threat, with typical military personnel operating and maintaining production-representative systems. (Contractor employees may be involved in operational testing only to the extent that they will operate, maintain, or support the systems when deployed in combat.) DT&E may be conducted by civilian engineers and technicians in laboratories or under other highly-controlled conditions and may involve using models, simulations, or system configurations that are not production-representative.
- OT&E is conducted to assess whether weapons can actually perform their intended missions. DT&E is more narrowly focused on engineering design and on verifying that technical specifications have been attained.
- OT&E is planned, conducted, and reported by an organization within a service that is separate and independent from the organizations responsible for developing and using new systems. DT&E¹ is controlled by program managers and may be conducted by contractors, user commands, or other parties with vested interests in the systems under development.

¹The Deputy Director, Defense Research and Engineering (Test and Evaluation) has oversight responsibility for all developmental testing.

Navy Weapon Systems Reviewed

Tables IV.1 and IV.2 show the Navy systems reviewed; whether OT&E was conducted in support of the designated milestone authorization; and, if not, whether OPTEVFOR made any other kind of operational assessment.

Table IV.1: Programs Authorized for Full-Scale Development From October 1984 to March 1988

System	Performance projections based on:	
	OT&E	Operational assessment
High Frequency Anti-Jam	None	None
Sea Lance missile	None	None
SH-60F helicopter	None	None
V-22 Osprey aircraft	None	None
Penguin missile	None	None
Afloat correlation system	None	None
Standoff Land Attack missile	None	None
ALR-67 Advanced Special Receiver	None	None
NULKA decoy	None	None
AN/BSY-2 submarine combat system	None	(a)

^aOPTEVFOR's "projection" report was based on observation of developmental testing and review of test results for related systems. According to OPTEVFOR, unavailability of a system to test under operationally realistic scenarios prevented the projection of potential operational effectiveness or suitability against a projected threat.

**Appendix IV
Navy Weapon Systems Reviewed**

Table IV.2: Programs Authorized for Low-Rate Initial Production From October 1984 to March 1988

System	Performance projections based on:	
	OT&E	Operational assessment
ALQ-165 Airborne Self Protection Jammer	None	None
SQS-53C sonar	None	(a)
E-6A aircraft	None	(b)
MK-45 capsule launching system	None	None
Ocean Surveillance Information System Baseline Upgrade	None	None
TB-23 Accelerated Thinline Towed Array	None	None
AN/BSY-2 submarine combat system	None	(c)
NAVSTAR Global Positioning System	Yes	Not applicable; OT&E conducted
MK-48 ADCAP torpedo	Yes	Not applicable; OT&E conducted
Bigeye bomb	Yes	Not applicable; OT&E conducted

^aOPTEVFOR's "observation" report was based on observation of testing in the contractor's plant, review of technical testing summaries, and attendance at program meetings. On this basis, OPTEVFOR stated that it could not project the system's potential effectiveness and suitability.

^bOPTEVFOR's projection report on the system's capabilities was based on observation of contractor testing and review of "functional checkout procedures" and thus was characterized as very limited in scope.

^cSee footnote a, table IV.1

Comments From the Department of Defense

Note: GAO comments supplementing those in the report text appear at the end of this appendix.



OPERATIONAL TEST
AND EVALUATION

OFFICE OF THE SECRETARY OF DEFENSE

WASHINGTON, D.C. 20301-1700

12 December 1988

Mr. Frank C. Conahan
Assistant Comptroller General
National Security and International
Affairs Division
U.S. General Accounting Office
Washington, DC 20548

Dear Mr. Conahan:

This is the Department of Defense (DoD) response to the General Accounting Office (GAO) draft report "NAVY WEAPONS TESTING: Defense Policy on Early Operational Testing Is In Transition" dated October 13, 1988 (GAO Code 394227), OSD Case 7800.

The stated GAO objectives appear to assume that published DoD policy requires operational testing prior to Milestone II (approval to enter Full-Scale Development). This is not DoD policy. Neither 10 USC 138 nor any DoD directive calls for operational testing of a system before that system enters full-scale development. Moreover, despite the GAO assertion to the contrary, neither does the Packard Commission. It is not possible to operationally field test a system that has not yet commenced full-scale development. The DoD policy does require early and progressive assessments of operational capability, including realistic operational field testing before full-scale production may commence.

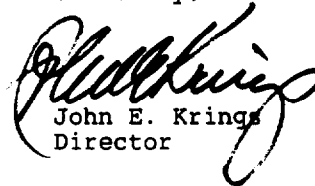
The Department agrees with the GAO statement that all early assessments of system performance are, to some degree, predictive. However, the DoD, in general, and the DOT&E, in particular, have not advocated such assessments as a replacement for operational testing. Such assessments fill an information void by providing an independent operationally oriented perspective not previously available to DoD and congressional decision makers. It is the Department's position that the operational test community should utilize all available information while awaiting system readiness for actual field tests--especially in instances where the majority of production funds will have been committed before the first production- representative system is available. The approach

See Comment 1

outlined in the September 1987 DoD report cited by the GAO provides for an increasing emphasis on operational considerations from the beginning of, and throughout, the acquisition process by utilizing all available information to make independent operationally oriented judgments concerning our new systems.

Additional DoD comments on the findings and recommendations are enclosed. The Department appreciates the opportunity to comment on the draft report.

Sincerely,



John E. Krings
Director

Attachments:
As stated

L

GAO DRAFT REPORT - DATED OCTOBER 12, 1988
(GAO CODE 394227) OSD CASE 7800

"NAVY WEAPONS TESTING: DEFENSE POLICY OF EARLY OPERATIONAL
TESTING IS IN TRANSITION

DEPARTMENT OF DEFENSE COMMENTS

FINDINGS

FINDING A: Operational Testing and the Acquisition Process. The GAO observed that operational test and evaluation (OT&E) has been an integral part of weapon system acquisitions since the early 1970s. The GAO noted the congressional concurrence with the "fly-before-buy" concept and the support for the establishment and operation of separate OT&E phases to support the corresponding acquisition process. The GAO emphasized that this approach has remained the tenet of the DoD acquisition policy and the Congress has continued to provide strong support to this approach. The GAO discusses the two principal types of test and evaluation: OT&E and developmental test and evaluation (DT&E). The GAO explained that the OT&E is conducted in a realistic combat environment against a representative threat, with typical military personnel operating and maintaining production-representative systems. The GAO characterized the DT&E by stating it may be conducted by civilian engineers and technicians in laboratories or under other highly controlled conditions, using models, simulations, or system configurations that are not production-representative. The GAO further described the distinctions between these approaches by stating that OT&E is conducted to assess whether new weapons can actually perform their intended missions, while the DT&E is more narrowly focused on engineering design and on verifying that technical specifications have been attained. The GAO observed OT&E is planned, conducted, and reported by an organization within the Service that is separate and independent from the systems, whereas DT&E is controlled by program managers and may be conducted by contractors, user commands or other parties with vested interests in the systems under development. (p. 1-7/GAO Draft Report)

Now on pp. 2-6

See comment 2

DoD RESPONSE: Partially Concur. The GAO characterization of DT&E and OT&E is not completely valid. The definition of Operational Test and Evaluation (OT&E), as found in Title 10

ENCLOSURE

1 of 11

U.S.C., is: "Operational test and evaluation means the field test, under realistic combat conditions, of any item of (or key component of) weapons, equipment, or munition for the purpose of determining the effectiveness and suitability of the weapons, equipment, or munition for use in combat by typical users; and the evaluation of the results of such tests." While OT&E is to be accomplished in an environment as operationally realistic as possible, including threat representative hostile forces, the GAO characterization that OT&E can only be conducted in a combat environment is incorrect. (It is not usually possible to insert operational tests into an actual "combat environment".) Development test and evaluation (DT&E) is explained in DoDD 5000.3 as: "that T&E conducted throughout various phases of the acquisition process to ensure the acquisition and fielding of an effective and supportable system by assisting in the engineering design and development and verifying attainment of technical performance specifications, objectives, and supportability." The Office of the Secretary of Defense (OSD) oversight of developmental testing is vested in the Deputy Director, Defense Research and Engineering (Test and Evaluation), who is charged with "responsibility and authority for all DT&E conducted within the Department of Defense, including designating Research Development, Test & Evaluation (RDT&E) programs as major for the purposes of DT&E oversight." The GAO has incorrectly implied that DT&E is conducted without OSD oversight, saying: "DT&E is controlled by Program Managers."

See comment 3

FINDING B: Importance of Operational Test and Evaluation Proposal to Use Operational Assessment. The GAO observed that the most visible OT&E phase is the final testing phase, completed before the full-rate production and deployment decision is made (Milestone III). The GAO noted that this is sometimes referred to as the final examination. The GAO concluded, however, that no less important is the testing conducted before the decisions are made to begin full-scale development and low-rate initial production (LRIP), emphasizing it is at this point that commitments are made to the system design selected from among competing alternatives. The GAO pointed out the Navy Instruction on Acquisition notes that Milestone II is the "single most critical decision point, for it is here that the Navy makes a firm commitment to the program; once started, it is difficult to turn back." The GAO further concluded that, while the subsequent Milestone III decisions to begin full production and deployment of the system are important, they are made after most program funds are committed and, except in extreme cases, at a time when production can only be slowed, not stopped. The GAO summarized that the purpose of the OT&E conducted in support of Milestone II is to provide early projections of potential operational effectiveness and suitability for the system being considered for full scale development. The GAO notes that, in contrast, the OT&E conducted in support of the Milestone III production

decision is to verify operational effectiveness and to ensure that the system meets required operational thresholds. The GAO concluded that both type of testing can be conducted with hand-crafted prototypes (noting prototypes testing in support of Milestone III decisions are supposed to be production-representative).

The GAO observed that, in 1987 a report to the House and Senate Armed Services Committees, the Secretary of Defense proposed a change in testing and evaluation functions, substituting "operational assessment" for operational testing during the early phases preceding the "final exam." The GAO explained that the operational assessments would be based on modeling, simulation, paper analyses, and any available testing (such as developmental or operational). (pp. 6-7/GAO Draft Report)

Now on pp. 6-8

See comment 4

DoD RESPONSE: Partially Concur. The GAO has inaccurately characterized the concepts outlined in the 1987 report to the House and Senate Armed Services Committees. That proposal was intended to enhance the operational evaluations of weapon systems by providing more complete operationally oriented evaluations earlier in the acquisition process. The report proposed more complete operationally oriented evaluations by defining two of the test phases as:

"INITIAL OT&E (IOT&E) is that component of T&E, initiated at program inception, conducted to forecast operational effectiveness and suitability. IOT&E is a tool to provide insights about the potential worth of a system throughout its acquisition life. IOT&E may take advantage of any test results and may use simulation, modeling, and paper analyses to develop assessments. IOT&E is reportable to the Director, Operational Test and Evaluation (DOT&E) as that an independent, objective, and candid assessment of a systems progress from an operationally oriented point of view will be available to decision makers throughout the acquisition process.

Production OT&E is that T&E that traditionally is thought of as "true" OT&E. It is the "final exam" prior to a full-production decision. Production OT&E is required to meet all the criteria established by the Congress in 10 USC 138 for OT&E, including the prohibitions on contractor-personnel participation. Production OT&E is conducted only after the program manager has certified (as required by DoDD 5000.3) that the system is ready for its final exam."

It is clear that the purpose of the redefinition quoted above is to enhance the evaluations of new systems by providing for operationally oriented evaluations of all valid information

3 of 11

on a system's expected capability as early as possible in the acquisition process. Such an approach would still provide for DOT&E reporting "at the conclusion of such operational test and evaluation," as stipulated in 10 USC 138, but would enhance DOT&E evaluations by providing for consideration of all relevant information as it became available throughout the development process.

The approach does not dilute evaluations of a system by "substituting operational assessments for operational testing" as reported by the GAO. Rather, the approach provides for preliminary operational assessments to be made when a system was not yet sufficiently developed for actual operational field tests to be conducted.

FINDING C: The Value of Prototyping. The GAO noted that whether to employ traditional OT&E or to rely on operational assessments is essentially a question of whether prototyping is the only reliable method for obtaining information needed to make program decision that commit large amounts of money. The GAO recognized that prototypes can be expensive and time consuming to build, and production models are unavailable during the early phases of development. The GAO pointed, however, to the 1986 President's Blue Ribbon Commission on Defense Management (the "Packard Commission") recommendation that high priority be given to building and testing prototypes of major weapon systems before Milestone II decision to begin full-scale development. The GAO explained that the Commission's underlying premise was that prototype testing is essential if the Services are to fully exploit their technological advantages. (The GAO notes that, according to the Commission, the problem is getting the information needed to make the risks-versus-benefits analysis. The GAO quoted that commission as stating, "The only consistently reliable means of getting such information is by building prototypes that embody the new technology." The GAO further noted that, according to the Commission, such a prototype at either the system or subsystem level should undergo operational as well as technical deficiencies before a decision is made to proceed with full-scale development. (Emphasis Added by the GAO). The GAO reported that there have been mixed reactions to the Packard Commission recommendation that prototypes be tested before full-scale development decisions are made. The GAO cited a 1986 report to the Secretary of Defense by the Director, Operational Test and Evaluation (DOT&E) and a 1987 report to the House and Senate Armed Services Committees by the Secretary of Defense that confirmed the Commission's conclusion and specifically cited the recommendation to employ extensive prototyping in testing as a positive contribution. The GAO found, however, that the September 1987 report by the Secretary of Defense also addressed the use of modeling and

4 of 11

Now on pp. 2 and 4-6

See comment 5.

Now on p. 5-6

simulation--but not testing of prototypes--in preproduction OT&E. Furthermore, the GAO referred to statements by the Director, Operational Test and Evaluation, that he does not believe prototypes are feasible for most acquisition programs. According to the GAO, the Director supports traditional OT&E whenever it is reasonable, but that in most cases it is better to have early performance assessments based on data obtained from realistic, high-quality simulators. (p. 2, pp. 8-10/GAO Draft Report)

DoD RESPONSE: Partially Concur. The GAO has inaccurately reported on the findings of the Packard Commission. In discussing prototyping (page 9 of the draft report) the GAO states: "Such a prototype at either the system or subsystem level should undergo operational and developmental testing, the Commission concluded, 'to uncover operational as well as technical deficiencies before a decision is made to proceed with full-scale development.'" (Emphasis added.) The Packard Commission does not recommend operational testing before a decision is made to proceed with full-scale development. It is quite difficult to conduct operational testing as defined by 10 USC 138 before a system even enters full-scale development. Clarification is provided by excerpts from the Packard Commission report, "A Formula for Action: A Report to the President on Defense Acquisition by the President's Blue Ribbon Commission of Defense Management," dated April 1986. On page 22 of that report is the following statement:

"A program manager should agree to a baseline for all phases of his program. For the Acquisition Executives, however, the agreement should extend only to the first two phases of the program, full scale development and low-rate production. Before a program could enter its third phase, high rate production, it must be subjected to developmental and operational testing. Operational tests are particularly critical, and should continue through full-scale development. [Emphasis added by the DoD.] The first units that come off a low-rate production line should be subjected to intensive operational testing. Low-rate production should continue during testing, but a program should not be approved for high rate production until the results of these tests are evaluated."

The Packard Commission did not recommend operational testing prior to a decision to proceed with full-scale development, as the GAO asserts. The Packard Commission report further elaborates on this in the suggestions on page 32 concerning responsibility for authorizing full-scale

5 of 11

development and high-rate production, urging a restructured Joint Requirements and Management Board (JRMB), for all joint and major Service programs, to:

"1) require the testing of prototype systems and subsystems before the authorization of full-scale development; 2) require the use of baselining for all major new programs; 3) require that operational test data be available before authorization of high-rate production [Emphasis added by the DoD]; and 4) significantly increase the use of nondevelopment items as an alternative to new development programs."

See comment 6

It can be seen from the above that the GAO characterization of the Packard Commission recommendation is inaccurate. The GAO is correct in stating that the Commission recommended the testing of prototypes; and the DoD supports such an approach when feasible and cost effective. The Commission did not conclude that a system should undergo operational testing before full-scale development. Contrary to the GAO draft report, the Commission suggested that, before a program could enter its third phase, high-rate production, it must be subjected to developmental and operational testing and that operational tests should continue through full-scale development. The Commission suggested that operational test data be available before authorization of high-rate production (not before entry into full-scale development). The cited 1986 and 1987 reports have also been misconstrued by the GAO. For example, the September 1987 report (page 5) stated: "DT&E includes T&E of components, computer software, subsystems, and hardware/software. It encompasses the use of modeling, simulations, testbeds, as well as advanced development, prototype, and full-scale engineering development models of the system." On the same page is this statement: "That portion of OT&E conducted throughout the acquisition process prior to the decision to proceed to full-rate production is called initial operational test and evaluation (IOT&E). It can be accomplished using a prototype, preproduction articles, or an LRIP article as the test item." Therefore, the GAO conclusion that those reports did not address testing of prototypes is inaccurate. Neither the reports to the Congress and the Secretary of Defense, nor the statements of the Director, Operational Test and Evaluation, advocate the substitution of other means for actual operational testing when such testing can be accomplished. The September 1987 report clearly stated (page 8) that: "IOT&E may take advantage of any test results and may use simulation, modeling, and paper analyses to develop assessments." It is clear that the Director has been advocating the use of all valid information for operationally oriented judgments as early as possible during the life cycle

6 of 11

of any system. If the Congress authorizes a concurrent approach to development and production of a system and significant funds will be expended prior to the availability of a system (or prototype) for actual operational field testing, then it is entirely appropriate to use whatever information is available to make early assessments of expected system capability. Such information could include the output from high-quality, validated simulators as well as any other reliable information source. In those instances where prototypes are representative of the actual configuration to be produced, the use in operational testing is valid. However, it is the Department's position that the operational test community should utilize all available information while awaiting system readiness for actual field tests--especially in instances where the majority of production funds will have been committed before the first production-representative system is available. The approach outlined in the September 1987 DoD report cited by the GAO provides for an increasing emphasis on operational considerations from the beginning of and throughout the acquisition process by utilizing all available information to make independent operationally oriented judgments concerning our new systems.

FINDING D: Potential Effects of Deemphasizing Operational Testing. The GAO reported that two features have long made OT&E a unique source of information for the DoD and that Congress: the independence of the agencies that conduct OT&E, and the emphasis on realism that distinguishes operational testing from other kinds of weapon system testing. The GAO concluded that both of these qualities may be affected by the new emphasis on operational assessments. The GAO observed that, when the OT&E process was incorporated into the acquisition process, one agency was chartered to independently plan, test, evaluate and report on realistic tests of weapon system performance. The GAO reported that in the Navy, the agency is the Operational Test and Evaluation Force (OPTEVFOR). The GAO concluded, however, that the added responsibility for operational assessments makes OPTEVFOR dependent on the results of modeling, simulation, studies, and tests conducted by agencies or contractors with a strong interest in getting the systems into production. (The GAO learned that the OPTEVFOR initially objected to preparing performance assessments based on data from other sources, without having actually tested the system in question.) The GAO further concluded that the nature of operational assessments also shifts to emphasis from "determining" whether a system works to "predicting" whether the system will work and blurs the important distinctions between OT&E and DT&E. The GAO expressed strong concern that, in the long term, relying on data from non-independent sources and simulation models may compromise the OPTEVFOR independence. The GAO also concluded that the operational assessments may be used by persons who are

7 of 11

Now on pp. 1 and 6-9

not aware of the difference between OT&E and operational assessments. (The GAO noted that, while, OPTEVFOR reports describe limitations to the scope of the agency's work, these reports are often synopsisized into highly condensed reports, such as Congressional Data Sheets, were limitations of operational assessments may not be as evident.) (p.2, pp. 10-13/GAO Draft Report)

DoD RESPONSE: Partially concur. The DoD does not concur with the basic GAO conclusion that operational testing is being deemphasized. To the contrary, through the efforts of the Director, Operational Test and Evaluation, OT&E has been given an ever increasing emphasis throughout all phases of system acquisition and employment. The applicable DoD Directive 5000.3 stipulates that "in each DoD Component one major field OTA shall be responsible for planning and conducting all operational test and evaluation." The GAO is correct in stating that in the Navy, OPTEVFOR has that responsibility. However, the GAO conclusion that "the added responsibility for operational assessments makes OPTEVFOR dependent on the results of modeling, simulation, studies, and tests conducted by agencies or contractors with a strong interest in getting the system into production" is invalid. The use of modeling and simulation in test and evaluation is not new. The DoD Directive 5000.3 (page 7) states: "The use of modeling and simulation is encouraged to ensure the availability of operational effectiveness and suitability projections at milestone II." There has been no shift away from "determining" whether a system works to "predicting" whether the system will work, as the GAO states. The DoD approach has been to ascertain as much as possible, early in the program, concerning a system's ability to fulfill an operational need, progressively updating such independent operationally oriented assessments as more information becomes available. The GAO is correct in stating that all early assessments of system performance are to some degree predictive. However, the DoD, in general, and the DOT&E, in particular, have not advocated such assessments as a replacement for operational testing. To the contrary, such assessments fill a void by providing an operationally oriented perspective not previously available. The GAO conclusion that "relying on data from non-independent sources may in the long term compromise the independence of the operational test agency" is unfounded. An operational test agency (OTA) would not blindly use "data from non-independent sources and simulation models owned by other organizations which may not have been validated." Rather, the OTA would carry out evaluations that are unencumbered by the perspective of any developing agency. The key is for the independent OTA to be able to use all information available, at any point in time; discounting that information which is not considered to be credible, and using all germane techniques. It is entirely appropriate for an OTA to provide periodic Early Operational Assessments (EOAs) before OT&E information from actual field testing is available. The DOT&E will monitor and assist in

8 of 11

this process, as warranted, to ensure the independence and credibility of the OTA assessment. It is the DoD perspective that an OTA should have the authority to evaluate all pertinent information as it becomes available, regardless of the source, for input into periodic early operational assessments (EOAs), estimating the system's potential operational effectiveness and suitability until such time as actual field testing can be completed to ascertain the de facto capabilities of a system through the evaluation of actual field testing.

FINDING E: Transition To Operational Assessments. The GAO found that the DoD has not required the Navy to follow published DoD policy calling for OT&E before decisions are made to begin full-scale development or LRIP of new weapons systems. The GAO pointed out that, during the period FY 1985 through FY 1987, the Navy seldom conducted OT&E prior to decisions to begin full-scale development and it often delayed OT&E until the LRIP had been authorized. The GAO further found that, during the same period, the Navy conducted a limited number of operational assessments in support of early milestone decisions. The GAO reviewed 20 Milestone II and Milestone IIIA decisions covering 19 Navy systems and found that OT&E was not conducted before any of the ten full-scale development decisions and was conducted before only three of the ten LRIP decisions. (Emphasis Added) The GAO also noted the OPTEVFOR prepared operational assessments for three of the systems (that were not operationally tested), which supported initial production decisions. The GAO described these assessments as early efforts and based on such limited data that, in two of the cases, the OPTEVFOR could not project the potential systems' effectiveness or suitability. The GAO reported the Navy recognized the shortcomings in this initial approach to operational assessments and, during its (the GAO) review, adopted a new approach requiring other agencies to assist the OPTEVFOR. According to the GAO, this joint arrangement was dictated, in part, by the lack of OPTEVFOR capability to develop modeling and simulation data or to verify such data provided by other agencies. The GAO explained that, under the new approach, the OPTEVFOR has established a working relationship with the Center for Naval Analyses and has recommended that the Center be granted a charter similar to the OPTEVFOR. The GAO noted that the first operational assessment under the new arrangement began in 1988 (at the end of the GAO review) and the GAO was unable to evaluate the assessment. (p. 2, pp. 13-17/GAO Draft Report)

DoD RESPONSE: Partially Concur. The GAO conclusion that the DoD has not required the Navy to follow published DoD policy calling for OT&E before decisions are made to begin full-scale development is incorrect. Published DoD policy does not include the requirement ascribed by the GAO. The GAO

9 of 11

Now on pp. 1 and 6-9

See comments 1 & 7

apparently did not consider the definition of a "major defense acquisition program" in 10 USC 138 nor the applicability and scope of DoDD 5000.3. The DoD policy concerning milestone II for major acquisition programs does state: "Decisions to commit funds for long-lead items or Low Rate Initial Production (LRIP) must be supported by an OT&E assessment" (DoDD 5000.3, page 7). That directive also states that management of systems not designated as major defense acquisition programs shall be guided by the principles set forth in that directive. Neither the Packard Commission, nor 10 USC 138, nor any DoD directive calls for operational testing of a system before that system enters full-scale development. Operational as well as developmental assessments are provided at milestone meetings. Included in Milestone II (FSD) considerations are the results from prototyping and demonstration/validation. The Milestone II decision establishes more specific cost, schedule, and operational effectiveness and suitability goals and thresholds, including approval of the baseline agreement between the Defense Acquisition Executive, the Service Acquisition Executive, the Program Executive Officer and the Program Manager. (See DoDD 5000.2, September 1, 1987). Consequently, in most weapon systems a prototype suitable for operational testing cannot be feasibly and cost-effectively produced prior to the Full-Scale Development (FSD) decision. DoDD 5000.2 also stipulates that one of the primary considerations in Defense Acquisition Board Milestone III (Full-Rate Production) decisions is the results of completed operational test and evaluation. Procedures in these DoD directives were appropriately followed for the programs identified by the GAO. Both developmental test results and operational assessments and test results are considered in DoD acquisition decisions.

In Appendix I to the draft report, the GAO states that its objectives were "(1) to determine whether Navy weapon systems are being operationally tested before being approved for full-scale development and LRIP, as required by published DoD policy and (2) to determine reasons for any disparities between policy and practice." As stated, the GAO objectives appear to assume that DoD policy requires operational testing prior to Milestone II (approval to enter Full-Scale Development). This is not DoD policy. It is not possible to operationally field test a system that has not yet commenced full-scale development. It is this fact that inspired the use of EOAs to permit independent, objective, and candid assessments of a system's progress from an operationally oriented point of view to be available to decision makers throughout the acquisition process.

RECOMMENDATIONS

RECOMMENDATION 1: The Secretary of Defense take the necessary actions to eliminate the disparity between written OT&E policy and practice. (p. 17/GAO Draft Report)

Now on p. 10

10 of 11

DoD RESPONSE: Nonconcur. It is the Department's position that this GAO recommendation is unfounded and based on a misunderstanding of DoD policy. As stated in the DoD response to Finding E, there is no DoD policy which stipulates that operational testing must be conducted before a system enters full-scale development. The DoD policy does, however, require an approach to early and progressive assessments of operational capability, including realistic operational field testing before full-scale production may commence. The 1987 report by the Secretary of Defense to the Congress outlined a formalized approach to ensure that independent operationally oriented assessments are considered during all phases of system development and acquisition.

RECOMMENDATION 2: The Secretary of Defense assure that (1) the independence of the services' OT&E agencies is not compromised and (2) the basis for weapon system assessments is fully disclosed when the assessments are reported to the Congressional and the DoD decision makers. (p. 17/GAO Draft Report)

DoD RESPONSE: Partially Concur. The Department agrees that the independence of the Service OTAs should not be compromised. The Department does not, however, view OTA use of all available information to make early and independent assessments as a compromise of their independence. In fact, such early involvement only strengthens the credibility, utility, and independent voice of the Service OTAs throughout the acquisition process. The basis for every operational assessment is included in appropriate detail when reported to congressional and DoD decision makers. The assessments for non-major programs are identified at requisite decision-making levels. While the basis for each assessment is included at an appropriate level of detail, complete, voluminous details of all evaluations and assessments are not routinely included in reports to higher decision levels. However, such information is made available when requested on a case-by-case basis.

11 of 11

Now on p. 10

The following are GAO's comments on the Department of Defense's letter dated December 12, 1988.

GAO Comments

1. We agree that title 10 U.S.C. 138 does not call for operational testing before a full-scale development decision is made, and our report states that OT&E is not legislatively required before this decision point. DOD stated that operational testing before milestone II is neither possible nor required by published policy. The operational testing policy delineated in DOD Directive 5000.3 does provide, as it should, some degree of flexibility in performing early operational testing. However, when viewed in the context of both the DOD directive that establishes overall acquisition policy (DOD Directive 5000.1), as well as the Navy instructions that implement DOD Directive 5000.3, it is clear that DOD's published policy calls for some operational testing before full-scale development. We recognize that the requirement is not absolute, and we do not advocate that it should be. Under the concept of tailored acquisition strategies, more compelling needs may override the need for early operational testing.

Directive 5000.3 specifically—and, we believe, correctly—emphasizes that decisions “to commit funds for long-lead items or Low Rate Initial Production...must be supported by an OT&E assessment;” but, it is important to note that the low-rate initial production decision can be made, and sometimes is made, at the same time full-scale development is authorized. The directive states that “OT&E shall be structured to provide inputs at each decision point, including major milestones,” that “OT&E shall take place as early as possible,” and that operational test agencies shall ensure “that OT&E is effectively planned and accomplished during all acquisition phases.”

2. DOD stated that our characterization of OT&E and DT&E was not entirely valid because it was not usually possible to insert operational tests into an actual combat environment. Thus, DOD concluded that we had erred in stating that OT&E can only be conducted in a combat environment. Our draft report stated that OT&E is to be conducted in a “realistic” combat environment, which we believe is an accurate paraphrase of “realistic combat conditions” and “realistic conditions” prescribed for OT&E.

3. DOD said the phrase “DT&E is controlled by program managers” incorrectly implies that DT&E is conducted without oversight from the Office of the Secretary of Defense. We recognize the role of the Deputy Director, Defense Research and Engineering (Test and Evaluation), in providing oversight for DT&E and have revised our report accordingly.

However, DT&E is controlled by program managers and may be conducted by contractors, user commands, or other parties with a vested interest in the system under development, whereas independent agencies within each service maintain control of operational testing.

4. Our draft report may have been unclear concerning our characterization of the concepts outlined in the Secretary of Defense's 1987 report. We agree that the intent of the proposal contained in the Secretary's report was to enhance the process of operational evaluation of weapon systems by providing operationally oriented information where none existed previously. This, we agree, is a step forward, and we have modified our report accordingly to more clearly articulate this position.

5. Contrary to DOD's statement, the Packard Commission specifically recommended that operational testing occur during the early phase of development. The Commission's final report, dated June 30, 1986, states:

"Operational tests should be combined with developmental tests of the prototypes to uncover operational as well as technical deficiencies before a decision is made to proceed with full-scale development."

Appendix A to the Commission's final report further states:

"The proper use of operational testing is critical to improving the operations performance of new weapons. We recommend that operational testing begin early in advanced development and continue through full-scale development, using prototype hardware."

6. DOD incorrectly stated that our report concludes that testing of prototypes was not addressed in either the Director's 1986 report to the Secretary of Defense or the Secretary's 1987 report to the Armed Services Committees. Our report states that both of these reports cite the positive nature of the Packard Commission's recommendation regarding testing of prototypes. However, the section of the Secretary's report that redefined OT&E to include "operational assessments" omitted any reference to preproduction OT&E of prototypes.

7. In addition to DOD directives and Navy instructions, OPTEVFOR's "Operational Test Director Guide" also stresses the importance of the OT&E phase the Navy calls OT-I—that is, OT&E in support of milestone II decisions that authorize full-scale development. The guide states that "OT-I is extremely important—it is usually required, and it usually requires

hands-on operation... by fleet personnel." Furthermore, according to the guide, the greatest opportunity for OPTEVFOR to influence the design, performance, and survivability of future fleet equipment is as a result of thorough OT-I before milestone II.

Major Contributors to This Report

**National Security and
International Affairs
Division Washington,
D.C.**

John D'Esopo, Assistant Director, Navy Issues, (202) 275-6504

**Norfolk Regional
Office**

Edward W. States, Regional Management Representative
James B. Marshall, Evaluator-in-Charge
William W. McComb, Jr., Evaluator