

GAO

Testimony



138870

For Release
On Delivery
Expected at
9:30 a.m. EST
Friday,
June 16, 1989

Adequacy of Department of Defense Operational
Test and Evaluation

Statement of
Frank C. Conahan, Assistant Comptroller General
National Security and International Affairs
Division

Before the
Subcommittee on Federal Services, Post Office
and Civil Services, Committee on Governmental
Affairs - United States Senate



045713/138870

Mr. Chairman and Members of the Subcommittee:

We appreciate the opportunity to appear before you today to discuss our work concerning Department of Defense (DOD) operational testing and evaluation (OT&E) of major weapon systems.

Since 1970, we have issued over 50 reports dealing with the adequacy of operational testing and evaluation of weapon systems to demonstrate their capability to perform their intended missions. Testing has not been comprehensive, realistic or rigorous; nor has there been sufficient oversight of the testing and evaluation function by top management in DOD.

Our work has shown that during the later years of program development, weapons have historically experienced significant cost growth, schedule slippage and performance shortfalls. Sound and independent testing is needed if systems are to avoid costly redesign and modification after production or deployment. Similarly, we cannot afford to invest large sums in programs such as Aquila and DIVAD only to have them terminated because they do not work. While many have applauded the decisions to end such programs, they should never have progressed as far as they did.

Continual trade-offs are made between doing adequate operational testing and the consequent delays that may occur in fielding a system plus the high costs associated with keeping a production line open. Such factors as urgency of the requirement and the cost of building prototypes may, in the view of decision makers, outweigh the need to identify and correct performance shortcomings identified through operational testing and evaluation. Our fundamental objective is to assure that decision makers are fully informed on the impact of starting production when (1) sufficient OT&E results are not available or (2) test

results support the need to reduce risk by redesigning or modifying systems.

Today, I will discuss (1) inadequate oversight and reporting by the Office of the Director, Operational Test and Evaluation (DOT&E), (2) lack of realistic testing and adequate test resources, (3) concurrent development and production that has created an environment which is not conducive to thorough OT&E, and (4) the need to maintain OT&E independence, the evolving concept of early operational assessments and the absence of OT&E to support low-rate initial production decisions.

INADEQUATE OVERSIGHT AND REPORTING

The Congress has consistently shown a long-standing interest in the performance of major weapon systems and the adequacy and timeliness of OT&E. As early as 1971, the Congress enacted legislation requiring the Department of Defense to provide the Congress with data on the OT&E results of major weapon systems before committing major production dollars.

In 1983, Congress enacted additional legislation creating the Office of the Director of Operational Test and Evaluation (DOT&E). Prominent among the objectives of this Office were independent oversight, coordination of the military services' planning and execution of operational tests, independent evaluation of the results of operational tests, and objective reporting of test results to decision makers in the Department of Defense and the Congress.

In March 1987, we evaluated DOT&E's effectiveness in carrying out its oversight responsibilities. We reported, among other things, that DOT&E made contributions to OT&E activities, especially in the test planning area. The Office was responsible for improvement in Test and Evaluation Master Plans and detailed

operational test plans. However, we identified several deficiencies in DOT&E's effectiveness in carrying out its oversight activities. Specifically:

- DOT&E appeared to be making only limited numbers of actual on-site observations of operational tests,
- DOT&E's analysis of operational testing was primarily based on military service test reports with little assessment of actual test results, and
- DOT&E had not provided policy and procedural guidance or maintained reliable records on some of their principal activities.

In the official response to our report received in July 1987, DOD generally disagreed with the principal findings. We were surprised by the DOD response because we met with the Director of OT&E and his staff several times during the assignment and obtained confirmation of the accuracy of the information being developed. And, DOT&E officials acknowledged our reported findings in their comments on an advance copy of the draft report.

Recent follow-up on the current status of DOT&E efforts to resolve these problems indicate that corrective action has not been taken. For example, DOT&E has not established policy and procedures or maintained reliable records on observing tests, analyzing service test reports or preparing reports on systems prior to their entering full rate production. Also, DOT&E has not developed a formal system to document how DOT&E staff spend their time, nor do they believe one is necessary. We continue to believe that a formal system should be available to evaluate DOT&E's effectiveness in carrying out its principal activities.

In July 1988, we issued another report confirming many of the long-standing OT&E problems. The systems we reviewed included the Army's Helicopter Improvement Program and Aquila Remotely Piloted Vehicle, the Navy's Tomahawk Land Attack Missile and DDG-51 Destroyer and the Air Force's Imaging Infrared Maverick and Low Altitude Navigation and Targeting Infrared System for Night.

We evaluated (1) the methodological adequacy of OT&E under DOT&E oversight and (2) the quality of DOT&E dissemination of information to the Congress. We reported that there were significant problems and limitations in the planning, execution, realism, analysis, and reporting of OT&E by the service operational test agencies. Some of these problems and limitations were unavoidable due to time, resource, or safety constraints, although numerous others were not.

With regard to the quality of DOT&E's dissemination of information to the Congress, DOT&E reports we reviewed contained incomplete or inaccurate statements, and most contained both. We realize that some problems and limitations in OT&E cannot be avoided. However reports that omit, or do not accurately portray test results can create a misleading impression of weapon system performance; thus denying Congress complete and accurate information to make informed budgetary decisions. It can also allow systems to advance without the problems being corrected, and lead to costly redesign, modifications, and terminations.

REALISTIC TESTING AND TEST RESOURCES

We have issued numerous reports on DOD's need to improve the quantity and quality of test and evaluation resources. As early as 1975, we reported that the most troublesome problem in test resources was the lack of realistic targets. Other reports have also stressed the need for improved realism in testing and more

representative test resources. Realistic testing requires test resources--such as threat simulators and aerial targets--that duplicate, to the extent possible, the characteristics of the current and projected threat.

As a result of test resource shortfalls, major weapon systems have been deployed without having fully demonstrated their capabilities under representative combat conditions.

In 1983, in a major effort evaluating test resources, we found that insufficient support for testing and test resources within the DOD weapon development community was contributing to inadequate and, in some cases, nonexistent test resources. Our report focused on two types of test resources--electronic warfare threat simulators and aerial targets. It also discussed actions necessary to improve DOD's capability to support testing.

In December 1986, we issued a report which summarized historical problems on the adequacy of OT&E that we had reported since 1970. Further, we examined six deployed weapon systems and found, among other things, that test resources were often not available nor adequate.

In March 1988, we reported concerns in planning, organizing, and sponsoring test resources including our assessment of the Navy's long-term underwater test resources plan. Generally, we found that test resources had received relatively little management attention. Also, we found the Navy's long-term underwater test resources plan should have consolidated and prioritized all proposed test resource development and acquisition programs so that the most critical test resource development programs could be supported.

OSD has begun several initiatives to improve test capabilities and bring a central focus to test resource management. A Test

and Evaluation Committee, which reports to the Defense Acquisition Board, has brought a high level focus and emphasis on test resources. And through the Operational Test and Evaluation Capability Improvement Program, OSD, through DOT&E, is also addressing the problem of test realism. The Program attempts to improve operational test realism by ensuring that selected high priority weapon systems have the critical resources necessary to be properly tested in the next 2 to 3 years.

We have received congressional requests to review DOT&E's involvement in both the Capability Improvement Program and the Test and Evaluation Committee. In these efforts, the primary concern is not whether there is a need to identify and acquire test resources, but whether DOT&E's current role in this process is in consonance with its mission and charter.

As directed by the Conference Report on the Department of Defense (DOD) Appropriations Act for Fiscal Year 1989 (P.L. 100-463), we are evaluating whether the acquisition and management of resources under the Capability Improvement Program falls within the mission of DOT&E or whether it should be given to another DOD organization. We subsequently were requested to evaluate the Services' acceptance of the program, the documentation justifying the Program, whether the items being acquired are resulting in possible duplication, and the degree of coordination between DOT&E and the intelligence community.

The House Armed Services Committee has also requested that we examine DOT&E's role in similar areas, such as whether DOT&E's role as Chairman of the Test and Evaluation Committee is appropriate because of its involvement in both developmental and operational test resources.

THE IMPACT OF CONCURRENCY
ON PRODUCTION DECISIONS

DOD's policy on major weapon system acquisition stresses the importance of minimizing the time to develop, produce, and deploy major systems for use by the operational forces. In a concurrent program, production is started while development is still underway and operational testing has not been completed. In a nonconcurrent program, development is usually completed before production.

Concurrency is a technique to expedite the development and production of weapon systems. Concurrency must be well planned and controlled and adequate safeguards must be built into the program to minimize the risks. At the very least, these safeguards should provide for performance of at least some OT&E before production.

A major contributor to the problems that surfaced on the B-1B bomber was the heavy reliance on concurrency: the simultaneous development and production of a weapon system in which copy after copy is produced before all the bugs are worked out. Use of concurrency to produce sophisticated new weapons is dangerous when the technology is so advanced that the demands of prudent production cannot keep pace. In June 1985, we reported that in five concurrently developed and produced weapon systems--the Air Launch Cruise Missile, B-1B bomber, Sergeant York Air Defense Gun, F/A-18 aircraft, and the AGM-88A High Speed Antiradiation Missile--DOD did not obtain OT&E results critical to assessing mission performance before production start-up, even though it was initially planned that these test results would be available before making such decisions.

In January 1989, we reported on the concurrency in the Peacekeeper Rail Garrison Missile System and our concerns over the overlap between testing, production and deployment.

Others are also concerned about concurrency. For example, in August 1988, the CBO published a study which addressed cost growth experienced in concurrent programs. Unit costs for the PATRIOT Missile, a highly concurrent program, were cited to be over 250 percent of what was originally planned.

Until concurrent programs are managed better, including adequate operational testing, DOD will continue to face the prospect of more programs that will not achieve advertised performance levels and also be subject to major cost growth.

NEED TO MAINTAIN INDEPENDENCE IN OT&E

From time to time DOT&E has attempted to reorganize or redefine the test and evaluation functions. In early 1987, we opposed a DOT&E proposal to place the test and evaluation functions within the Office of the Under Secretary of Defense for Acquisition. At that time, we also expressed concerns over DOT&E's plans to consolidate developmental and operational test functions under its purview.

In response to our March 1987 report and DOD's attempt to realign test and evaluation activities, Congress reacted by adopting legislation which stated:

"The Director may not be assigned any responsibility for developmental test and evaluation, other than the provision of advice to officials responsible for such testing." (National Defense Authorization Act, FY 1988 and 1989, section 801)

In the Conference Report 100-446 dated November 17, 1987, the conferees agreed that responsibility for operational testing is separate and distinct from functions associated with developmental testing. The conferees further agreed that the Secretary of Defense should refrain from any realignment or new arrangement of test and evaluation activities until it could be debated by the Congress.

In September 1987, DOD attempted to redefine the test and evaluation functions. Traditionally, the portion of initial OT&E conducted during the acquisition process prior to the decision to proceed to full-rate production was accomplished using a prototype, preproduction article, or production-representative test article. The Secretary of Defense endorsed a realignment of test and evaluation activities which describe initial OT&E as a tool to provide insights about the potential operational worth of a system throughout its acquisition life. This concept could take advantage of any test results and may use simulation, modeling, and paper analysis to develop assessments. DOT&E officials stated that this definition generally describes an early operational assessment.

As of January 1989, it appears the Director, Operational Test and Evaluation has continued to realign the test and evaluation activities by requiring the use of "early operational assessments". The purpose of early operational assessments is to assess the ability of a system to satisfy operational issues and identify areas of risk by using available information. They provide a process to assure decision makers that the system is evolving toward readiness for operational test and evaluation, and that as many as possible operational effectiveness or suitability shortfalls have been identified and corrected.

In our May 8, 1989, report we pointed out, among other things, that because operational assessments rely on information from nonindependent sources, the important distinctions between actual operational test and evaluation and operational assessments can be blurred. We also pointed out that military service operational test and evaluation agencies are conducting operational assessments and have concerns that in performing such assessments, their independence may be compromised by a role that requires them to predict weapon system performance on a basis of data provided by contractors and agencies responsible for developing the weapon systems. Development and OT&E results are important in the acquisition process, but they have different purposes, use different criteria and are conducted under different conditions. Further, the OT&E agencies believe that misunderstandings may arise over the nature and extent of testing actually performed on a weapon system.

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 can be expensive and time consuming to build and often are not available prior to the full-scale development or low-rate initial production decisions.

DOD acquisition directives generally call for OT&E input at all major acquisition decision points, including full-scale development and low-rate initial production. We reported in May 1989, that OT&E was not conducted for ten full scale development decisions and was only conducted for three of ten low-rate initial production decisions. For three of the systems that have not been operationally tested, the Navy prepared operational assessments to support initial production decisions. Two of these assessments were so limited that the Navy's operational testing office could not project the systems' potential effectiveness or suitability.

Operational assessments early in the acquisition process, particularly when test hardware is not available, fill a void and offer useful information to decision makers. Our concern, however, is the tendency to rely on these assessments as a primary basis for deciding on whether to initiate or accelerate production. It should be clearly understood that these assessments cannot and should not take the place of actual OT&E before production.

The DOD Inspector General (IG) has been reviewing the adequacy of OT&E for transition into production. The DOD IG found that United States Code, Title 10, Section 2403, "Contract Guarantee" defines "mature full scale production" as the manufacture of all units of a weapon system after the initial production quantity or manufacture of the first 10 percent of the eventual production, whichever is less. DOD IG reports issued on the C-17 and SINCGARS programs illustrate the absence of OT&E, yet significant quantities had already been produced. For example, the DOD IG's application of "the 10-percent rule" to an Air Force planned purchase of 210 C-17 aircraft indicated that a production request of 40 aircraft, which is far in excess of 10 percent of 210 aircraft, represented a "de facto" full-rate production decision.

We do not know what the low-rate initial production rate should be, but we strongly support the need for having some OT&E results available early in the acquisition process before production commitments are made. We plan to explore this whole area in future work.

- - - - -

This concludes my prepared testimony. We will be pleased to answer any questions you might have. Thank you.