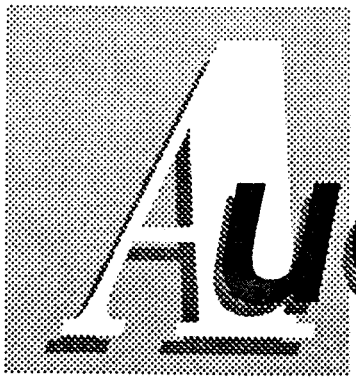
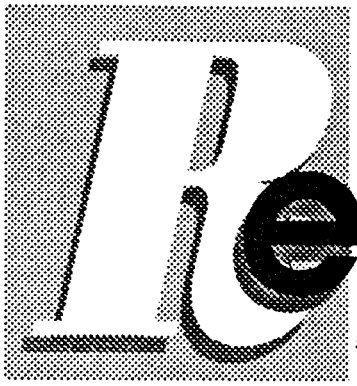




OFFICE OF THE INSPECTOR GENERAL

OPERATIONAL TESTING PERFORMED ON WEAPONS SYSTEMS

Report No. 96-107 May 6, 1996

Department of Defense

Additional Copies

To obtain additional copies of this audit report, contact the Secondary Reports Distribution Unit of the Analysis, Planning, and Technical Support Directorate at (703) 604-8937 (DSN 664-8937) or FAX (703) 604-8932.

Suggestions for Future Audits

To suggest ideas for or to request future audits, contact the Planning and Coordination Branch of the Analysis, Planning, and Technical Support Directorate at (703) 604-8939 (DSN 664-8939) or FAX (703) 604-8932. Ideas and requests can also be mailed to:

OAIG-AUD (ATTN: APTS Audit Suggestions)
Inspector General, Department of Defense
400 Army Navy Drive (Room 801)
Arlington, Virginia 22202-2884

Defense Hotline

To report fraud, waste, or abuse, contact the Defense Hotline by calling (800) 424-9098; by sending an electronic message to Hotline@DODIG.OSD.MIL; or by writing the Defense Hotline, The Pentagon, Washington, D.C. 20301-1900. The identity of each writer and caller is fully protected.

Acronyms

ASAS	All Source Analysis System
COI	Critical Operational Issue
DoDI	Department of Defense Instruction
DOT&E	Director, Operational Test and Evaluation
DTSE&E	Director, Test, Systems Engineering, and Evaluation
GSM	Ground Station Module
IPT	Integrated Product Team
JSTARS	Joint Surveillance Target Attack Radar System
LRIP	Low-Rate Initial Production
MILSTAR	Military Strategic and Tactical Relay
OSD	Office of the Secretary of Defense
OTA	Operational Test Agency
OT&E	Operational Test and Evaluation
SAE	Service Acquisition Executive
SAR	Selected Acquisition Report
T&E	Test and Evaluation
TEMP	Test and Evaluation Master Plan
TPWG	Test Plan Working Group

INSPECTOR GENERAL
DEPARTMENT OF DEFENSE
400 ARMY NAVY DRIVE
ARLINGTON, VIRGINIA 22202-2884

May 6, 1996

MEMORANDUM FOR UNDER SECRETARY OF DEFENSE FOR ACQUISITION AND TECHNOLOGY
UNDER SECRETARY OF DEFENSE (COMPTROLLER)
ASSISTANT SECRETARY OF THE NAVY (FINANCIAL MANAGEMENT AND COMPTROLLER)
ASSISTANT SECRETARY OF THE AIR FORCE (FINANCIAL MANAGEMENT AND COMPTROLLER)
DIRECTOR, OPERATIONAL TEST AND EVALUATION
AUDITOR GENERAL, DEPARTMENT OF THE ARMY

SUBJECT: Audit Report on Operational Testing Performed on Weapon Systems (Report No. 96-107)

We are providing this audit report for review and comment. We considered comments on a draft of this report in preparing the final report.

DoD Directive 7650.3 requires that audit recommendations be resolved promptly. The comments that we received from the Under Secretary of Defense for Acquisition and Technology; Director, Operational Test and Evaluation; and the Military Departments were partially responsive. As a result of management comments, we revised, renumbered, redirected, and deleted some draft recommendations and findings. See the end of each finding for the finding revision, recommendation changes, and responses required.

We ask that management provide comments in response to the final report by June 6, 1996. Comments must describe actions taken or planned and provide completion dates for those actions.

We appreciate the courtesies extended to the audit staff. Questions on the audit should be directed to Mr. Raymond A. Spencer, Audit Program Director, at (703) 604-9071 (DSN 664-9071) or Mr. Louis F. Schleuger, Acting Audit Project Manager, at (703) 604-9009 (DSN 664-9009). If management requests, we will provide a formal briefing on the audit results. See Appendix L for the report distribution. The inside back cover lists the audit team members.

Robert J. Lieberman

Robert J. Lieberman
Assistant Inspector General
for Auditing

Office of the Inspector General, DoD

Report No. 96-107 **May 6, 1996**
(Project No. 5AB-0031)

Operational Testing Performed on Weapon Systems

Executive Summary

Introduction. Operational testing is an integral part of the DoD weapon system acquisition process. Operational testing evaluates weapon systems under conditions as realistic as possible to determine the system's operational effectiveness and suitability for combat.

Audit Objectives. The overall audit objective was to determine whether the policies and procedures implemented by the Office of the Director, Operational Test and Evaluation, resulted in operational test plans that ensured the realistic portrayal of threat forces; the realism of test operation environments; the accuracy of the reporting by Military Departments' test agencies; and the accuracy of the Director, Operational Test and Evaluation, reports to Congress. We also evaluated the adequacy of the test agencies' management controls as they apply to the overall audit objective.

Audit Results. The policies and procedures used to develop and execute operational test plans were generally adequate for testing major Defense acquisition programs. Also, the policies and procedures for ensuring the realistic portrayal of threat forces were generally adequate. However, improvements are needed due to the following:

o the test planning process was not efficient, resulting in inefficient use of resources (Finding A);

o for at least 4 of 15 systems reviewed, the Military Departments conducted operational tests with significant system limitations, resulting in additional funding requirements for testing (Finding B);

o for 15 of 15 systems reviewed, the Military Departments did not adequately assess the impact of test limitations and the risk of fielding systems that will not counter the threats, resulting in test documents that did not provide adequate information about the impact of the limitations for decisionmakers (Finding C);

o for 16 of 65 systems reported, the Director, Operational Test and Evaluation, Annual Report to the Secretary of Defense and Congress did not provide complete and consistent assessments of the systems' performance, resulting in incomplete information for decisionmaking (Finding D); and

o for 4 of 12 systems reviewed requiring a Selected Acquisition Report to Congress, the Report did not provide a complete and accurate summary of the program status, resulting in an overly optimistic status reported (Finding E).

Implementing the recommendations in this report will improve the effectiveness and efficiency of DoD test planning and testing. Appendix J summarizes the potential benefits of the audit.

Summary of Recommendations. We recommend:

o applying the integrated product team concept to the test planning and execution process;

o developing a decisionmaking model to evaluate whether to proceed with operational testing when serious limitations are identified;

o approving test and evaluation master plans, test plans, and test reports after the impact of limitations are adequately addressed;

o reporting on all operational testing and providing complete and consistent information to decisionmakers; and

o preparing the Selected Acquisition Report to ensure responsible decisionmakers receive complete and accurate program status.

Management Comments. The Under Secretary of Defense for Acquisition and Technology; Director, Operational Test and Evaluation; and the Military Departments, in general, concurred with applying the integrated product team concept to the test planning and execution process. The Under Secretary of Defense for Acquisition and Technology nonconcurred with the need to develop a decisionmaking model to evaluate readiness to proceed with testing; however, he suggested the Defense Acquisition Deskbook would be an appropriate venue for criteria to make the determination. The Director, Operational Test and Evaluation, nonconcurred with approving the test plans and reports only when the impact of limitations are adequately addressed and he asserted the test plans and reports have not denied decisionmakers essential information. The Under Secretary of Defense (Comptroller) suggested redirecting the draft recommendation to budget adequate resources. The Director, Operational Test and Evaluation, partially concurred with reporting on all operational testing and will determine what operational test information on non-major systems should be reported. The Under Secretary of Defense for Acquisition and Technology stated that the Selected Acquisition Report is fully coordinated and represents the DoD position; however, Selected Acquisition Report reform initiatives may address issues raised in the report. See Part I for a discussion of management comments and Part III for the complete text of management comments.

Audit Response. The comments that we received were partially responsive. As a result of management comments, we revised, renumbered, redirected, and deleted some draft recommendations and findings. See the end of each finding for the finding revisions, recommendation changes, and additional responses requested.

We ask that the Under Secretary of Defense for Acquisition and Technology; the Under Secretary of Defense (Comptroller); the Director, Operational Test and Evaluation; the Assistant Secretary of the Army (Financial Management and Comptroller); the Assistant Secretary of the Navy (Financial Management and Comptroller); the Assistant Secretary of the Air Force (Financial Management and Comptroller); the Army Director, Test and Evaluation Management Agency; the Navy Director, Test and Evaluation and Technology Requirements; and the Air Force Director, Test and Evaluation Directorate, comment on unresolved recommendations in this final report by June 6, 1996.

Table of Contents

Executive Summary i

Part I - Audit Results

Audit Background 2
Audit Objectives 2
Finding A. Test and Evaluation Planning Process 4
Finding B. Readiness for Testing 16
Finding C. Test Threat and Environment 25
Finding D. Test Status Reporting 33
Finding E. Reporting Program Results 40

Part II - Additional Information

Appendix A. Scope and Methodology 48
Sample Selection 49
Management Control Program 49
Area Not Requiring Further Review 50
Appendix B. Summary of Prior Audits and Other Reviews 51
Appendix C. Operational Test Organizations and Agencies 54
Appendix D. Description of Weapon Systems 59
Appendix E. Test and Evaluation Master Plan Approval Time 62
Appendix F. Documents Reviewed, Approved, and Published by Director, Operational Test and Evaluation 63
Appendix G. Approved Performance Characteristics 65
Appendix H. Program Performance Summary 67
Appendix I. Management Comments on the Findings and Audit Responses 68
Appendix J. Summary of Potential Benefits Resulting From Audit 73
Appendix K. Organizations Visited or Contacted 74
Appendix L. Report Distribution 76

Part III - Management Comments

Under Secretary of Defense for Acquisition and Technology Comments 80
Director, Test, Systems Engineering, and Evaluation, Comments 83
Under Secretary of Defense (Comptroller) Comments 87
Director, Operational Test and Evaluation, Comments 88
Department of the Army Comments 98
Department of the Navy Comments 103
Department of the Air Force Comments 108

Part I - Audit Results

Audit Background

Test and evaluation is an integral part of the DoD weapon system acquisition process. Test and evaluation consists of two principal categories: development test and evaluation and operational test and evaluation (OT&E). The development test and evaluation assists in engineering design and development and verifies "attainment of technical performance specifications, objectives, and supportability." The OT&E is the field test, under realistic combat conditions, to determine the operational effectiveness and suitability of a system. Operational effectiveness and suitability are measured in terms of the system's ability to meet the user's requirements.

In 1971, the Deputy Secretary of Defense directed each Military Department to designate field commands to be responsible for OT&E. The Services (Army, Navy, Air Force, and Marine Corps) established operational test agencies (OTAs) to plan and conduct operational testing. They were chartered to be separate and distinct from the developing, procuring, and using commands. The OTAs were tasked to report directly to their respective Service Chief of Staff. The OTAs include the Army Operational Test and Evaluation Command, the Navy Operational Test and Evaluation Force, the Air Force Operational Test and Evaluation Center, and the Marine Corps Operational Test and Evaluation Activity.

In 1983, Congress enacted Public Law 98-94 establishing the position of the Director, Operational Test and Evaluation (DOT&E), as the principal OT&E authority within the Office of the Secretary of Defense. The DOT&E is to provide independent oversight, coordination, and evaluation of the Military Departments' planning and execution of operational tests and reports. The Director reports directly to the Secretary of Defense and has special reporting requirements to the Congress.

See Appendix C for a description and organizational diagram of the OT&E community.

Audit Objectives

We evaluated the procedures to develop and execute operational test plans. Specifically, we evaluated the policies and procedures implemented by the Office of the Director, Operational Test and Evaluation, to ensure the realistic portrayal of threat forces, the realism of test operation environments, the accuracy of the reporting by Military Departments' test agencies, and the accuracy of the Director's reports to Congress. We evaluated the management controls as they apply to the audit objectives. The policies and procedures for ensuring the realistic

portrayal of threat forces are generally adequate. See Appendix A for the audit scope and methodology. See Appendix B for a summary of prior coverage related to the audit objectives. See Appendix D for a description of the weapon systems included in this audit.

Finding A. Test and Evaluation Planning Process

The Military Departments' process for approving test and evaluation master plans was not efficient because the Military Departments used a sequential approval process instead of a more efficient integrated process. As a result, resources used for additional test and evaluation plan development and approval activities could be put to better use.

Test and Evaluation Planning

The primary planning document for test and evaluation is the Test and Evaluation Master Plan (TEMP). The system's program management office prepares the TEMP with operational test information provided by the Military Department's OTA. It translates the test concept into testable events and responsibilities to generate detailed test and evaluation plans. The TEMP is also used to ascertain schedule and resource implications associated with the test and evaluation program.

Statutory. United States Code, title 10, section 2399 (10 U.S.C. 2399), states that operational testing of a major Defense acquisition program may not be conducted until the DOT&E gives written approval to the adequacy of the test plans for operational test and evaluation to be conducted in connection with the program.

DoD Guidance. DoD Instruction (DoDI) 5000.2[1], "Defense Acquisition Management Policies and Procedures," part 8, "Test and Evaluation," February 23, 1991, implemented 10 U.S.C. 2399 and stated that a TEMP will be prepared for all acquisition programs. The Instruction also stated that the TEMP will be approved by DOT&E and the Director, Test, Systems Engineering, and Evaluation (DTSE&E), for acquisition category I programs and other acquisition programs designated for Office of the Secretary of Defense (OSD) test and evaluation oversight.

DoD 5000.2-M[1], "Defense Acquisition Management Documentation and Reports," part 7, "Test and Evaluation Master Plan," February 23, 1991, stated that the TEMP is to be approved within 45 days of its submission to DTSE&E and DOT&E. The Manual required seven mandatory approval authorities for the TEMP:

[1]The DoD 5000 series was revised as of March 15, 1996. The DoDI 5000.2 and DoD 5000.2-M used as criteria for this audit were current as of October 31, 1995. DoD guidance in the manual will be incorporated into the Defense Acquisition Deskbook.

o Program Manager,

o Program Executive Officer,

o Operational Test Agency,

o User's Representative (only one signature),

o DoD Component Acquisition Executive or milestone decision authority,

o the DTSE&E, and

o the DOT&E.

Military Department Implementation. The Military Departments' guidance states that the program manager should use a team to develop and maintain the TEMP. The Army calls its team the test integration working group and the Navy and the Air Force call their team a test planning working group (TPWG) [hereafter TPWG is used for all three working groups]. The Military Departments' guidance states that the TEMP originates in the applicable program office, is coordinated within the working group, and proceeds for approval through the Military Department acquisition chain of command. The Military Departments' guidance also states that the program manager should coordinate the TEMP for comments with all organizations represented on the TPWG and allow 30 days for the organizations' responses.

TEMP Process

The Military Departments' process for approval of TEMPs was not efficient. Inspector General, DoD, Inspection Report No. 91-INS-09, "Operational Test and Evaluation Within the Department of Defense," May 24, 1991, concluded that the TEMP process was not timely. The report stated that the DTSE&E (formerly Deputy Director, Defense Research and Engineering [Test and Evaluation]) and the DOT&E approval process averaged about 100 days and that the DTSE&E and the DOT&E recognized that 100 days was an excessive amount of time.

The DTSE&E conducted a study of TEMP approval times in response to concerns about approval duration. In addition, we reviewed TEMP approval times as a part of this audit. The DTSE&E and we concluded that approval time depends on the interactions within the Military Department development and approval chain. Table 1 shows significant delays within the Military Departments' approval chain. The TEMP approval times shown in both studies are not additive and do not include "informal" time spent staffing the TEMP within the Military Departments and OSD before the program manager's signature. We based our study on 14 TEMPs associated with audited programs, with program manager signature dates from June 1988 to December 1994.

OSD based its study on 58 TEMPs approved by OSD, with program manager signature dates from December 1992 through November 1994. Nine of the 14 TEMPs in our study predate the OSD study dates.

Table 1. Comparison of Test and Evaluation Master Plan Approval Times

	Average Number of Days After the Program Manager Signature for Signature of:				Average Number of Days After OSD Receipt for Signature of:	
Activity	PEO	OTA	USER	SAE	DTSE&E	DOT&E
IG Study	32	23	38	112	55	63
OSD Study	58	58	86	145	30	33

IG Inspector General, DoD
PEO Program Executive Officer
SAE Service Acquisition Executive

Comparing the results of our study with the more recent data of the DTSE&E study shows that the OSD processing times improved. However, the processing times within the Military Departments regressed during the same period. The average approval time within OSD decreased by a factor of 1.9. However, the average approval time within the Military Departments increased by a factor of 2.0. The approval times for each Military Department from the OSD and our studies are shown in Appendix E, Tables E-1 and E-2, respectively.

A review of program documents and interviews with OSD, OTA, and program office officials, combined with results of the OSD and Inspector General studies, indicated that these delays are caused by the Military Departments' sequential process for approving TEMPs.

Sequential Process

The Military Departments used a sequential instead of a more efficient integrated process to prepare, update, and approve the TEMPs. The program management office prepares a draft TEMP that is distributed for comment to TPWG members. At a later meeting, members generally review the TEMP to ensure the comments are incorporated into a revised draft. At another TPWG meeting, TPWG members refine the draft into a finished product.

TEMP Development. Lack of participation and empowerment hindered the TPWG process. Some organizational representatives, particularly those involved with program oversight, did not actively participate in the TPWG and simply waited for the TEMP to be forwarded for approval to voice additional test issues.

Participation. Some organizational representatives did not actively participate in the "call for comments" and arrived at the TPWG with new and additional comments that the working group was unprepared to handle. For example, the Military Strategic and Tactical Relay (MILSTAR) TPWG met six times from October 1990 through August 1993. According to the August 19, 1993, TPWG meeting minutes, a copy of the proposed TEMP changes was provided to the participants 12 days before the meeting so all organizations could review the submitted changes. Since the TPWG is meant to be a forum to resolve existing issues, valuable time was wasted resolving unexpected issues.

Table 2 illustrates lack of participation by key organizations during the MILSTAR TPWG meetings. In addition to these multi-Service TPWG meetings, each Military Department held its own working group meetings.

Table 2. MILSTAR Multi-Service Test Plan Working Group Participants

Required Approval Authority Office	Number of Meetings Attended
Program Office	6
PEO	0
SAE	0
User	6
OTA	6
DTSE&E	0
DOT&E	1

The DOT&E had an action officer at only one meeting. The Air Force Program Executive Officer, the Navy Program Executive Officer, and the Army Program Executive Officer attended none.

In addition to the six TPWG meetings, 35 Air Force meetings were held from January 1993 through June 1995 to plan the actual MILSTAR tests. Similar participation problems occurred.

Empowerment. Working group representatives are not given adequate authority to make program testing decisions. For example, a MILSTAR program office official stated that lack of participation and failure to empower TPWG members were problems. The official also stated that the lack of empowerment affects the organizational representatives' ability to exercise decision authority at the TPWG.

TEMP Approval. The TPWG members are to provide guidance to the program manager on program test and evaluation issues. However, without empowerment, the TPWG representatives cannot provide adequate and

timely guidance to the program manager, thus increasing delays in obtaining approval of the TEMP. Additionally, before the TEMP reaches the approval authority, additional organizational staffing was often required.

Hierarchical Approval. After development or update and TPWG agreement, the TEMP is ready for the program manager's approval. Upon the program manager's approval, the TEMP is forwarded for approval to the program executive officer, the OTA, the user, the DoD Component acquisition executive or milestone decision authority, and finally to DTSE&E and DOT&E, in turn. With a sequential process, one hierarchical level's approval is contingent on the approval of the preceding hierarchical level. In this sequential process, the TEMP is approved by officials who often have not been actively involved in TEMP development.

The program offices, the users, and DOT&E officials told us that, often, the TEMP awaits approval because an official is out of town, has higher priorities, or is unfamiliar with the program and its test plans.

In the Air Force, approval delays also occur because the Air Force requires a TEMP approval signature in addition to those that were required by DoDI 5000.2. In his November 14, 1994, Policy Memorandum 94A-020, the Assistant Secretary of the Air Force (Acquisition) directed the Air Force (Test and Evaluation) directorate to sign all acquisition category I, OSD oversight, and multi-Service or multi-agency TEMPs. In the Army, the Assistant Secretary of the Army (Research, Development, and Acquisition) delegated the signature authority to the Deputy Under Secretary of the Army (Operations Research), who is the Army T&E director. However, the new DoD 5000.2-R, March 15, 1996, requires the T&E director's signature in addition to that of the acquisition executive (the Military Department Assistant Secretaries) or Milestone Decision Authority.

Horizontal Approvals. After the program manager approves the TEMP, it is distributed for approval throughout the representative organizations for comment. For example, an Army Training and Doctrine Command official stated that the TEMP is concurrently staffed to a minimum of four different offices within the Command. Extensive staffing was generally not necessary and led to extraneous comments. Additionally, the official stated that the TEMP is delayed because representatives often placed a low priority on the TEMP.

The scope of the TEMP staffing depended on the program size and the organization responsible for the testing. For example, at the Air Force OTA, generally, parallel coordination exists among the test manager's support group, which includes such areas as analysis, effectiveness, suitability, policy, and human factors. Comments were compiled and forwarded to an analysis group for consideration. The changes were made to the TEMP and then forwarded to the OTA Commander for approval.

Another example of the extent of TEMP staffing is the Joint Surveillance Target Attack Radar System (JSTARS) TEMP. The program office sent the 30-page TEMP to OSD for approval. The TEMP was subsequently returned to the program office with 24 pages of comments from six organizations. These organizations were:

o the Office of the Under Secretary of Defense (Acquisition and Technology),

o the Office of the Assistant Secretary of Defense (Production and Logistics) (now the Deputy Undersecretary of Defense [Logistics]),

o the DOT&E,

o the Defense Intelligence Agency,

o the Joint Staff, and

o the Joint Interoperability Test Center.

Although these organizations may have introduced constructive comments, they should have been addressed within the TPWG by empowered representatives, thus putting direct costs in time and money and opportunity costs to better use.

Senior Management Delays. TEMP delays often occurred because of senior management demands. Program office officials stated that senior management often held the TEMP until program issues were resolved. For example, the Deputy Under Secretary of the Army (Operations Research) memorandum, April 26, 1995, stated that the JSTARS TEMP would not be forwarded to the Service Acquisition Executive for approval unless it addressed the "role in the ATACMS [Army Tactical Missile System] end-to-end evaluation." The memorandum contains specific language that demanded the following be added to the TEMP:

> The JSTARS MOTE [multi-service operational test and evaluation] will be used as one of several test events to support OEC's [Operational Evaluation Command] end-to-end evaluation of the ATACMS. Sensor missions, target sets, and possible inclusion of other Army systems will be coordinated so as to satisfy both the needs of the MOTE and the ATACMS end-to-end evaluation.

A second example is a DOT&E memorandum, November 12, 1992, which states that "In order to obtain OSD approval of the Joint STARS TEMP, the above actions must be accomplished." Holding on to TEMPs contributes to approval delays within the Military Departments as identified in both the OSD and Inspector General studies. These examples of delays illustrate raising issues at senior levels that should have been identified and resolved earlier in the TEMP development process.

The lack of participation and empowerment, hierarchical approval structure, extensive horizontal staffing, and senior management delays are all parts of the sequential approval process, which is time-consuming, inefficient, and unproductive. Improving the TEMP development, update, and approval process is necessary to enhance productivity and facilitate timely acquisition programs.

Effective Use of Resources

Resources used for additional test and evaluation plan development and approval activities could be put to better use. Staffing the TEMP for comments throughout approving organizations, instead of using an empowered TPWG, increases resource costs. Lack of participation by TPWG organizational representatives generally results in additional TPWG meetings to resolve issues raised outside the working group and increases opportunity costs.

The MILSTAR program office TPWG co-chairman estimated that for six TPWG meetings, the travel costs were $340,380. The official stated that had all representatives fully participated and been adequately empowered, two TPWG meetings would have been adequate to prepare and update the TEMP. Elimination of four meetings would have saved about $226,920 in travel costs. The actual costs for these meetings include travel and opportunity costs.

To improve test planning efficiency, the Army implemented guidance in Army Regulation 73-1, "Test and Evaluation Policy," February 27, 1995, that states the TPWG members should be empowered by their respective organizations to provide a position. Additionally, the Army policy memorandum, "Reengineering the Test and Evaluation Master Plan (TEMP) Review and Approval Process," March 18, 1996, established a test and evaluation integrated product team (IPT), to include the program Manager and Program Executive Officer, user, and OTA representatives. The policy requires that those IPT members be empowered to sign the TEMP. Also, the policy requires that the TEMP be signed by its TEMP approval authority within 30 days upon receipt of the TEMP from the test and evaluation IPT. Neither the Navy nor the Air Force have implemented similar guidance.

Integrated Product Team. On April 28, 1995, the Under Secretary of Defense for Acquisition and Technology issued a policy memorandum directing his staff and Component staff organizations to:

- o participate as members of an IPT or IPTs,
- o participate early on and in an on-going basis with the program office teams, and
- o resolve issues as they arise.

The policy memorandum also directs the program managers to utilize the experience of the OSD and Component staff organizations.

The IPTs are composed of representatives from all appropriate functional disciplines working together to identify and resolve issues and make sound and timely decisions.

In his May 10, 1995, policy memorandum, the Secretary of Defense directed the DoD Components to use IPTs to "enhance our ability to provide what the warfighter needs, when needed and at a cost that the Department can afford." The Secretary of Defense also directed that:

> The functional representatives assigned to the IPT at all levels must be empowered by their leadership to give good advice and counsel to the Program Manager. They must be able to speak for their superiors, the 'principals,' in the decision making process.

The IPT process is to replace the current TEMP development and approval sequential process. Implementation of IPTs at the program office level, *with empowered representatives*, is a crucial step to reap the benefits of the efficiency and effectiveness that IPTs offer. Without improving the test planning process, resources will continue to be used inefficiently and ineffectively.

The revised DoD 5000.2-R, March 15, 1996, addresses the use of empowered IPTs at the program office levels. However, the Regulation still requires a hierarchical approval process for the TEMP. Additionally, the draft "Rules of the Road, A Guide for Conducting Successful Integrated Product Teams," November 1995, recommends, but does not require, the program executive officer, Service acquisition executive or milestone decision authority, DTSE&E, and DOT&E be members of the working-level IPT. Representatives of the program executive officer, Service acquisition executive or milestone decision authority, DTSE&E, and DOT&E should be formal members of the program office's working-level IPT to expedite the approval process of test documents. To meet the intent of 10 U.S.C. 139, DOT&E should be a non-voting member so as to maintain his independence.

Summary

The process for developing, updating, and approving TEMPs needs improvement. We identified problems within the TEMP development, update, and approval process that can be resolved by implementing the IPT process at the TPWG level. If the recommendation to implement the IPT process at the TPWG level is adopted, management oversight and the time spent in test planning will be reduced. See Appendix J for the Summary of Benefits Resulting From Audit.

Management Comments on the Finding and Audit Response

DTSE&E commented on the results of the reviews of TEMP approval times. As a result, we modified our wording and the accompanying table to clarify our intent. In addition, the Army commented on their update policy on the use of test and evaluation IPTs in their comments to Recommendation A.1. As a result, we expanded our discussion of the Army's TEMP process.

Recommendations, Management Comments, and Audit Responses

Revised and Unresolved Recommendations. As a result of management comments, we revised draft Recommendations A.2. and A.3. by replacing the word "himself" with "a member of that office." We revised recommendation A.3. for DTSE&E to be a "non-voting" member of the test planning IPT. We ask that the Army, the Navy, and the Air Force provide additional comments on Recommendations A.1.b. and A.1.c. See Part III for the full text of management comments to the findings and recommendations.

A.1. We recommend that the Army Director, Test and Evaluation Management Agency; the Navy Director, Test and Evaluation and Technology Requirements; and the Air Force Director, Test and Evaluation Directorate:

a. Apply the integrated product team concept to the test planning and execution process.

Army and Navy Comments. The Army and the Navy concurred. The Army stated that test planning IPTs have been formed to develop the test and evaluation strategy and coordinate the TEMP. The Navy stated that the IPT concept has been institutionalized.

Audit Response. The Army and Navy comments meet the intent of our recommendation.

Air Force Comments. The Air Force concurred with the application of the IPT concept to the test planning and execution process. The Air Force stated it welcomes greater reliance on IPTs for test planning and is already using IPTs. The Air Force also stated that Recommendation A.1.a. should be directed to program managers, since they, not Air Force test and evaluation (T&E) directors, control the TEMP process.

Audit Response. The Air Force comments partially meet the intent of our recommendation. We addressed the recommendation to the Military Department T&E directors instead of the program managers because this application of the IPT concept should be clearly defined by and included in the

Military Department T&E planning guidance. We understand that the program managers are responsible for the establishment of the TPWGs and the development of the TEMPs. However, the Military Department T&E directors should make the determination and provide the guidance on the T&E participation of testing IPTs.

DTSE&E Comments. Although not required to comment, the DTSE&E concurred, stating that initiation of program IPTs at all levels should improve the TEMP-approval process.

b. Delegate approval authority for the Test and Evaluation Master Plan to the integrated product team.

Army, Navy, and Air Force Comments. The Army, the Navy, and the Air Force nonconcurred. The Army stated that for the approval of the TEMP, the appropriate signature authorities are the Service Test and Evaluation Director and the Service Acquisition Executive (SAE). The Navy stated that oversight by the appropriate Milestone Decision Authority is still required. It also stated that an empowered IPT is the proper forum for TEMP formulation and that the IPT process efficiency will improve if all IPT members are present and properly empowered. The Air Force stated that bypassing the Service T&E Director and the SAE would cede TEMP approval authority to OSD.

Audit Response. We ask the Army, the Navy, and the Air Force to reconsider their positions and provide additional comments in response to the final report. Personnel involved in the TEMP-approval process stated that a lack of empowered representatives at test planning IPT meetings delayed the TEMP updates and approvals. The predominant remedy discussed for this problem was for all approval authority representatives to approve and sign the TEMP upon agreement at the IPT meeting, thereby bypassing the time-consuming hierarchical approval structure.

Participation by all those involved, including the Service T&E Director and the SAE (or their representatives), in TEMP development and approval should eliminate the need for hierarchical approvals. The intent is not to eliminate the Milestone Decision Authority, or others, from the approval process. The intent is to have empowered functional representatives from the cognizant approval authority offices (to include the offices of the Service T&E Directors and SAEs) participate in discussions and exert approval and signature authority at the test planning IPT, where TEMP issues are raised. This delegation does not eliminate them from the approval process or cede their statutory responsibilities to OSD. It simply states that all TEMP concerns should be voiced and decisions made within the test planning IPT by empowered representatives of interested parties. This recommendation will allow the timely production and approval of a TEMP.

We commend the Army for empowering its test and evaluation IPT members, to include signing the TEMP. However, to fully meet the intent of our recommendation, empowered representatives from its Test and Evaluation Management Agency and Deputy Under Secretary of the Army (Operations Research) should also be members of its test and evaluation IPT. We also

commend the Air Force Operational Test and Evaluation Command for staffing its TEMPs in parallel and for using electronic processing when coordinating TEMPs. However, neither of these actions addresses or corrects the approval delays resulting from horizontal staffing within agencies, hierarchical approvals, and senior management delays.

DOT&E Comments. Although not required to comment, the Director partially concurred with the concept of empowered IPTs. The Director stated that he fully supported the IPT process, which is similar to the TPWG meetings that the test community has used for years. He also stated that the report failed to recognize the extent to which the IPT concept was used in test planning. However, he stated that implementation of this recommendation would preclude the Military Departments' T&E and Acquisition executives from performing a final check and approval on a TEMP.

Audit Response. We agree with DOT&E that the current, *unempowered*, implementation of IPTs is like the TPWGs that the test community used and are recognized by the audit report. Applying the IPT concept without incorporating empowered functional representatives reinforces the status quo in the TEMP development and approval process. Empowered representatives at the program office test planning IPT are key to this concept and will improve test planning and the TEMP-approval process.

c. Forward the Test and Evaluation Master Plan directly from the program manager to the Director, Test, Systems Engineering, and Evaluation, and the Director, Operational Test and Evaluation, for approval.

Army, Navy, and Air Force Comments. The Army, the Navy, and the Air Force nonconcurred. The Army stated that the TEMP should be forwarded through the Service T&E Director and the SAE before forwarding it to OSD. The Army stated that a centralized individual should be responsible for developing and promulgating T&E policy and procedures for the TEMP process and for interfacing with the DoD T&E and TEMP policy-makers. The Navy stated that the Service T&E directors have primary responsibility for the completeness and accuracy of the TEMP and must have the opportunity to review it. The Air Force stated that large portions of the authority of the Service T&E Director and the SAE would be ceded to OSD if these Service Chiefs were bypassed in the TEMP process.

Audit Response. If a functional representative from the approval authority office participates at the test planning IPT and is empowered to speak with regard to decisionmaking for the principals of that approval authority office, they need not forward the TEMP through the Military Department chain of command, since TEMP approval and signing would have already been exercised at the test planning IPT. We ask the Army, the Navy, and the Air Force to reconsider their positions and provide additional comments in their responses to the final report.

A.2. We recommend that the Director, Operational Test and Evaluation, include a member of that office as a formal, non-voting member of the program office integrated product team to expedite review of test documents.

DOT&E Comments. The Director partially concurred, stating that he is too senior to be a member of the program office IPT. The Director recommended that the word "himself" be changed to read "a member of his office."

Audit Response. We agree with the Director's comment and revised the recommendation accordingly.

Army, Navy, and Air Force Comments. Although not required to comment, the Army, the Navy, and the Air Force concurred. The Air Force also stated its primary concern was that OSD remains largely absent from the TPWG process and that absence contributed significantly to the length of the process.

A.3. We recommend that the Director, Test, Systems Engineering, and Evaluation, include a member of his office as a formal, non-voting member of the program office integrated product team to expedite his review of test documents.

DTSE&E Comments. The Director concurred, with modification, stating that DTSE&E should be a formal, but non-voting, member of the program office IPTs.

Audit Response. We agree with the Director's comment and revised the recommendation accordingly.

Army, Navy, and Air Force Comments. Although not required to comment, the Army, the Navy, and the Air Force concurred.

Finding B. Readiness for Testing

The Military Departments delivered 4 of the 15 systems we examined for operational testing that were not ready for testing. This situation occurred because a calendar schedule rather than system readiness often drove the start of testing. In addition, neither DoD nor the Military Departments have formalized criteria to help certification officials and operational testers decide whether to delay tests because of system limitations. Because the Military Departments often cannot conduct a complete system test, they incur additional costs to repeat a test or to test a system that should have been in the original evaluation.

Test Planning and Execution

Operational test and evaluation is a criterion for proceeding beyond low-rate initial production for major Defense programs. Because of this criterion, demands are made to complete operational test and evaluation quickly to complete milestones and continue production.

Statutory. United States Code, title 10, section 2399 (10 U.S.C. 2399), requires initial operational test and evaluation for major Defense acquisition programs. DOT&E must report on operational test and evaluation to the Secretary of Defense and Congress.

DoD Guidance. DoDI 5000.2 implemented 10 U.S.C. 2399 and required that operational test and evaluation:

- o verify that systems are operationally effective and suitable and
- o provide essential information in support of decisionmaking.

The Instruction directed that OT&E focus on verifying the operational effectiveness and suitability of the system under realistic combat conditions. Additional conditions included that, whenever possible, the test must be against threat-representative forces, employ typical users, and use production or production-representative articles.

The Instruction also required that test and evaluation planning must, at a minimum, address all system components that are critical to minimum acceptable operational performance requirements specified in the operational requirements document. Both test planning and test reporting must identify any test limitations. The developing agency must also formally certify the system as ready for the dedicated phase of operational test and evaluation. The March 15, 1996, revision to the instruction did not change these requirements.

DoDI 5000.2, part 11, section A, "Program Objectives and Baselines," defined exit criteria for acquisition programs as:

> the specific minimum requirements that must be satisfactorily demonstrated before an effort or program can progress further in the current acquisition phase or transition to the next acquisition phase. Failure to meet an exit criterion halts the progress of the system towards the next milestone decision point.

Military Department Implementation. The Military Departments developed procedures for planning, executing, and reporting test and evaluation activities, including certifying systems as ready for operational testing. The Military Departments' procedures prescribe identifying any test limitations both when certifying test readiness and when reporting test results.

Test Readiness and Completeness

The Military Departments were not consistently delivering systems for operational testing that were ready for testing. The Military Departments often noted deficiencies or limitations in their certifications for operational tests and test reports as required by guidance. However, tests went forward despite significant limitations, stated or not stated.

Army. The All Source Analysis System (ASAS) was not ready for operational testing, although it had been certified ready with no critical deficiencies. The Army completed initial OT&E testing on the system October 11, 1992, and concluded the system was neither operationally effective nor suitable.

Following the failed test, the Army did a 6-month corrective action plan and in April 1993 completed a demonstration of the modified system. After failing initial OT&E and not being able to determine effectiveness and suitability in the April 1993 test, the Army performed, at added cost, a second initial OT&E in August 1994.

More recently, the testing of the Army Light Ground Station Module (GSM) highlighted deficiencies with both the certification process and the evaluation of test results. On May 2, 1995, the Army Acquisition Executive approved Low-Rate Initial Production (LRIP) for 10 Light GSM units. Before the decision, the Army OTA concluded in an "Abbreviated Operational Assessment (AOA) of the Joint Surveillance Target Attack Radar System (Joint Stars) Light Ground Station Module (LGSM)," March 7, 1995, that the system had the "*potential* to be fully effective and suitable [emphasis added]." The assessment was based on previous test reports, including a Force Development Test completed by Operational Test and Evaluation Command on September 20, 1994. The assessment was done to validate attainment of the exit criteria for going into LRIP, which had been established for the Light GSM by the Defense Acquisition Board, July 23, 1993.

Table 3, taken from a DOT&E analysis, May 25, 1995, portrays significant disagreement concerning the evaluation results. Table 3 does not show that some Army Material Systems Analysis Activity and Operational Evaluation Command assessments were based on the Force Development Test, as well as on a demonstration of the Medium GSM completed January 18, 1995. The Operational Evaluation Command concluded from the Force Development Test that the exit criteria were not met for the "Sensor/GSM Interoperability" and portions of the multi-sensor display. However, the Operational Evaluation Command modified its assessment based on the Medium GSM demonstration.

Table 3. Light Ground Station Module Exit Criteria Test Assessment

Exit Criteria Parameters	PM	User	AMSAA	OEC	DOT&E
Met	13	12	6	5	2
Not Met	0	1	4	5	11
Insufficient Data	0	0	3	2	0
No Assessment	0	0	0	1	0

PM Program Manager
AMSAA Army Material Systems Analysis Activity
OEC Operational Evaluation Command

Navy. The Navy T-45 Aircraft Training System (T-45TS) program based its phased testing on test readiness at a point in time. Seven production lots of the T-45TS were purchased over 8 years before initial OT&E, which the Navy refers to as operational evaluation. Although operational testing identified deficiencies, the Navy continued to buy the T-45s and now all aircraft must be upgraded.

Table 4 shows the results of phased operational testing for the T-45. The numbers refer to the number of Critical Operational Issues (COIs) for a particular result. For example, for OT-IIA, the results for seven COIs was "partially met." Tests OT-IIA, OT-IIB, and OT-IIC were pre-initial OT&E tests. The OT-IIA test plan cited four limitations, OT-IIB test plan cited nine limitations, and OT-IIC cited 13 limitations. The Navy divided initial OT&E into two phases as depicted in Table 4 as OT-IID, Phase 1 and Phase 2.

Table 4. T-45TS Operational Testing Limitations and Results

Results by COI	OT-IIA	OT-IIB	OT-IIC	OT-IID Phase 1	OT-IID Phase 2
Partially met	7	13	15	15	0
Not tested	11	6	4	2	0
Unsatisfactory	1	0	0	0	1
Satisfactory	0	0	0	1	18

Air Force. The AC-130U Gunship Replacement Program illustrates how testing proceeded despite significant limitations. The readiness status for the items to be evaluated during the initial OT&E is shown in Table 5.

Table 5. Status of AC-130U Evaluation Areas Before Initial OT&E

Category	Satisfactory	Marginal
Sensors		X
Navigation	X	
Flight director and guidance		X
Airframe		X
Fire Control		X
Electronic warfare		X
Radio frequency management	X	
Communications		X
Performance	X	
Human factors		X
Reliability and maintainability		X
Mission support/planning		X

Masked in the summary is that the All Light Level Television sensor, which the TEMP identified as a key sensor for the Gunship, was limited for use for only training during initial OT&E. Additionally, the ammunition handling system for the 25-millimeter gun, a fire control system, was rated unsatisfactory and was also limited to training.

The primary testing of the AC-130U after certification was done from August through November 1994. One additional flight was done in January 1995. The Air Force OTA has not published a report as of October 16, 1995, because the report must be rewritten. While differences exist in the explanations concerning why the report must be rewritten, one official stated the test data does not

support the conclusions in the draft. One OTA official stated follow-on OT&E will be needed to validate performance of systems not adequately tested, either because of test design or system limitations.

Schedule Requirements and Decisionmaking Criteria

Calendar Schedule. The preceding examples confirm that systems enter operational testing before they are ready. Our audit did not identify a specific, universal cause; however, interviews with program offices, testers, and DoD officials provided clues as to why premature testing occurs. Systems that are not ready are tested because a calendar schedule rather than system readiness often drives the beginning of testing.

The causes that drive testing schedules are not new. Inspector General, DoD, Report 91-INS-09, "Inspection of Operational Test and Evaluation Within the Department of Defense," May 24, 1991, concluded that "operational test events are forced into acquisition timetables and compressed, as necessary, to accommodate rigid acquisition milestones. Often operational tests are limited, or evaluation and report writing periods are shortened, to the point where adequate time is not available for a full and objective assessment on the system tested."

While not specifically discussed in Military Departments' test plans and test reports, program office and OTA officials stated testing is done when significant limitations are present because of the general program schedule, the need to keep the program moving forward, and budget considerations.

Beyond the requirement to note any limitations, the guidance is silent on whether to proceed with the test when limitations exist. We have not found criteria by which Military Departments can consistently evaluate cost and schedule implications of the decision to test when limitations exist.

Certification Criteria. With the exception of the Air Force, criteria have not been formalized to help certification officials and operational testers in assessing whether to delay tests due to limitations. The Air Force Manual 63-119, "Certification of System Readiness for Dedicated Operational Test and Evaluation," February 1995, established a set of 33 templates as a part of a process for identifying and reducing risks associated with the transition from developmental T&E to operational T&E. The Manual also states that the templates cover "a broad range of subjects that have historically impacted systems transitioning from development T&E to dedicated operational T&E. Not all templates may apply to every program." A similar, tailorable, set of criteria would assist the other Military Departments in assessing the costs and benefits of testing with limitations.

The consequences of operationally testing systems that are not ready for the test vary depending on the system, whether the deficiencies were known before testing, and the type and extent of later testing.

Weapon System Readiness and Funding

Because of testing with limitations, the Military Departments often cannot conduct a complete system test, which results in additional cost when repeating a test or later testing a sub-system that should have been in the original evaluation.

Army. The Army paid for a second phase of OT&E on the ASAS, August 1994, 2 years after failing the first OT&E. The second test report, "Test and Evaluation Report (TER) for All Source Analysis System (ASAS) Block I initial Operational Test and Evaluation II (IOTE II), "December 16, 1994, concluded that the system was effective and suitable, although the system lacked both electronic countermeasures testing and an interactive scenario.

The August 1994 OT&E did not fully test one of four COIs: whether the Army can deploy and operate ASAS on the battlefield. The COI requires operation in an electronic countermeasures environment. The TEMP stated that the ability to operate in an electronic warfare environment was a major consideration and that the battlefield of the future will be "characterized by extensive use of offensive electronic warfare." We question the decision to report a system as "effective and suitable" when such a critical issue is not evaluated.

The September 1994 Force Development Test of the Light GSM failed to confirm that the system was effective and suitable. More testing was done to validate that the July 1993 exit criteria were met. In January 1995, the Army did a demonstration with a surrogate system to validate those exit criteria not met in the Force Development Test.

Navy. The Navy paid for an additional phase of operational testing for the T-45TS. The original plan had two phases of operational testing before initial OT&E. The test plan for the second phase states that the "deficiencies [those identified in the first test] led to airframe and powerplant changes to the T-45A and a program restructure with an additional phase of IOT&E [initial OT&E] included prior to T-45TS operational evaluation."

Air Force. Because of the inability to test some sub-systems and premature testing of others, additional testing at additional cost of the AC-130U is likely, but details were not available. One OTA official said future OT&E will be needed to evaluate the sub-systems not fully tested and the future improvements to unready systems that were tested.

Summary

The perceived need to keep the program moving forward and the absence of criteria to assist the Military Departments in deciding whether to test with limitations resulted in premature testing. The Military Departments need to

approval authority and establish a decision model, such as the Air Force System Design and Performance Deficiency Resolution Template, for assessing the cost and benefits of proceeding with testing with limitations. The Template requires an impact analysis and prioritization of known deficiencies and a plan for correcting and testing those deficiencies.

Recommendations, Management Comments, and Audit Responses

Revised and Unresolved Recommendations. As a result of management comments, we revised draft Recommendation B.2. to add "incorporating criteria" following "decisionmaking model," and that the criteria should be included in the Defense Acquisition Deskbook. We ask that the DTSE&E and Under Secretary of Defense for Acquisition and Technology provide additional comments on unresolved Recommendations B.1. and B.2. The Under Secretary of Defense for Acquisition and Technology, the DTSE&E, the DOT&E, the Army, and the Air Force commented on the finding. See Appendix I for a summary of their comments and audit responses.

B. 1. We recommend the Director, Test, Systems Engineering, and Evaluation, review and approve certification of readiness for operational testing for acquisition category I and designated oversight programs.

DTSE&E Comments. The Director nonconcurred, stating that inserting DTSE&E into the chain of command would compromise his objectivity and oversight responsibility and that the existing management structure is adequate.

Audit Response. We disagree that inserting the Director into the chain of command will compromise his objectivity and oversight responsibility. The Director is already in the chain of command to approve the TEMP, which does not compromise his objectivity and oversight responsibilities. We also disagree that the existing management structure is adequate to prevent premature operational testing. Based on the 15 programs we examined and discussions we held with operational test agencies and program officers, testing with significant limitations was a problem. We ask the Director to reconsider his position and provide additional comments in his response to the final report.

DOT&E, Army, Navy, and Air Force Comments. Although not required to comment, the DOT&E, Navy, and Air Force nonconcurred and the Army partially concurred. The DOT&E stated that the report should say that the Program Executive Officer is required to certify system readiness for operational testing and that we should make recommendations with respect to better implementing current regulations. He offered no specific suggestions as to how current regulations might be better implemented. The Army, the Navy, and the Air Force stated that moving the responsibility to DTSE&E adds another layer of oversight, adds no value, and will slow program execution schedules. The Navy also stated that DOT&E already performs readiness for operational test reviews.

Audit Response. We recognize that having DTSE&E in the certification approval process adds another "layer of review." However, the use of IPTs for the test readiness certification, with the DTSE&E office participating, will offset any delay. Also, the benefits of ensuring that systems are ready for testing and not arbitrarily sent to testing because of a calendar schedule outweighs the minimal delays that might occur. We agree with the DOT&E that the Program Executive Officer certifies system readiness for operational testing. We disagree with the Navy that DOT&E performs readiness for operational test reviews. The DOT&E is an observer at the Navy's Operational Test Readiness Reviews, not a formal member.

B. 2. We recommend the Under Secretary of Defense for Acquisition and Technology include a requirement in the Defense Acquisition Deskbook that the Army, the Navy, and the Air Force develop a decisionmaking model incorporating criteria to evaluate whether continue testing when serious limitations are identified before operational testing and to determine the cost/benefit of testing with limitations.

Under Secretary of Defense for Acquisition and Technology Comments. The Under Secretary nonconcurred, stating that the four programs presented in the draft report did not persuade him that a pervasive problem existed. Further, he stated that the proposed revision to DoDI 5000.2 captures most of the intent of this recommendation, and that the Defense Acquisition Deskbook would be the most appropriate place for the criteria.

Audit Response. We maintain that a significant problem does exist. The language in the revised DoD 5000.2-R, particularly the part requiring an impact analysis of unmet metrics before certification for operational tests, is directed at the same problem. We maintain that specific, known criteria or analytical procedures should be used to decide when to start testing with significant system limitations. Based on acquisition reform initiatives, we see the merit of addressing the criteria in the Defense Acquisition Deskbook and have revised the recommendation accordingly. We ask the Under Secretary to reconsider his position and provide additional comments on the revised recommendation in his response to the final report.

DOT&E, Army, Navy, and Air Force Comments. Although not required to comment, the DOT&E and the Air Force partially concurred, the Army concurred, and the Navy nonconcurred. The DOT&E stated that the Services' evaluation processes to proceed with operational testing are already in place and that he felt the policy should be better implemented. The Army stated it is currently using a decisionmaking model, which is in Army Regulation 73-1, "Test and Evaluation Policy," February 27, 1995. The Navy stated that its Operational Test Readiness Review process and each System Command's internal Operational Test Readiness Review checklist provides adequate risk assessment to the decisionmaker. It stated a decisionmaking model does not add value to its current process. The Air Force stated that the other Military Departments could use the Air Force's or a similar certification process. It stated that mandating "one-size-fits-all" decision criteria instead of basing criteria on individual program requirements would be counter-productive. Additionally, it stated that tests are proceeding on the basis of schedule rather

Additionally, it stated that tests are proceeding on the basis of schedule rather than systems readiness due to the influences of the program manager, user, and funding.

Audit Response. We agree with DOT&E that an evaluation process to proceed with operational testing is in place. However, the process lacks specific criteria for evaluating the costs, benefits, and future test requirements. We agree with the Air Force comments that clearly state why testing is often driven by schedule rather than performance. The Army and the Navy have processes that review the system readiness for operational testing. However, their systems lack specific criteria for evaluating the cost, benefits, and future test requirements when testing the limitations. The demand to test in spite of limitations is very powerful. The Army and the Navy need to enhance their processes to evaluate the overall readiness for testing a system. The Air Force has a well-defined process for decisionmaking that the Army and the Navy might consider when improving their processes. A need still exists for developing criteria for evaluating the cost, benefits, and future test requirements when testing with limitations. If the Defense Acquisition Deskbook contains the guidance and criteria and the Military Departments use the guidance, all the Military Departments' processes would be improved.

Finding C. Test Threat and Environment

The Military Departments were not adequately assessing the impact of test limitations or the risk of fielding systems that will not counter the identified threats for the 15 systems we examined. The Military Departments listed or labelled a limitation without explaining the cause or impact. As a result, the test documents did not provide decisionmakers with adequate information concerning the potential impact of test limitations.

Requirements for Test Realism

Statutory. United States Code, title 10, section 139 (10 U.S.C. 139), states that an operational test is a field test under realistic combat conditions against threat-representative forces.

DoD Guidance. DoDI 5000.2, part 8, "Test and Evaluation," February 23, 1991, stated that test and evaluation planning must address resource requirements such as targets, threat systems, validated threat simulators, or surrogates. The guidance required that test plans enumerate test limitations and that test reports identify significant test limitations and the impact of those limitations on meeting minimum operational performance requirements.

DoD Manual 5000.2-M, part 5, "System Threat Assessment Report," stated that the System Threat Assessment Report is the primary threat document. When the Military Departments write test plans, this document should be the reference for determining the threats to test against and the environment in which to test.

DoD Manual 5000.2-M, part 7, "Test and Evaluation Master Plan," required the TEMP to discuss test limitations. The TEMP should include "the impact of the test limitations on the ability to resolve critical operational issues and the ability to formulate conclusions regarding operational effectiveness and operational suitability."

Realism of Tests

Assessment of Limitations. Operational tests completed on the 15 audited programs had test realism limitations. The Military Departments did not adequately assess the impact of those test limitations on the ability to resolve COIs and formulate conclusions regarding operational effectiveness and suitability or the risk of fielding systems that may not meet the identified threats.

DOT&E. The "FY 1994 Annual Report of the Director of Operational Test and Evaluation," February 1995, discusses limitations for the reported programs. For example, according to the test activity for the Navy T-45TS, "The operational testing was carried out in as realistic an operational environment as could be achieved within the constraints of available test ranges, resources, and safety." However, the impact of these constraints on test conclusions and the risk of wrongly concluding the system was effective and suitable were not discussed.

The three beyond LRIP reports included with the FY 1994 Annual Report listed 21 limitations. The reports were for the Rolling Airframe Missile Combat System; E-2C Update Development Program, Group II; and the M1A2 Abrams Tank System. Of the 21 limitations, 13 discussed the impact of the limitation on the ability to resolve the COIs and only 4 discussed the impact of the limitation on the ability to formulate conclusions regarding operational effectiveness and suitability.

Military Departments. The Military Departments were reporting test limitations in their planning documents and reports of test results. However, the Military Departments' OTA test reports did not present as comprehensive an analysis of the impact of test limitations as that of the DOT&E beyond LRIP reports.

This reporting weakness is not new. Inspector General, DoD, Report 91-INS-09, "Inspection of Operational Test and Evaluation within the Department of Defense," May 24, 1991, states that "In each of the eight reports reviewed, we found terse discussions of the limitations to the scope of testing, with little or no discussion of the impact thereof."

Test Realism Limitations. The Military Departments performed operational tests despite threat and test environment limitations. The limitations were not always acknowledged in the test certifications, but were at least identified in the test report. Limitations for 3 of the 15 programs illustrate shortcomings in testing in a realistic environment against representative threats.

Army. The ASAS program failed to test in a realistic combat environment. During the August 1994 OT&E, the COI on whether the Army could deploy and operate the system on the battlefield was not fully evaluated. One measure of performance for the COI assesses the capability of the system to operate in an electronic countermeasures environment (See Finding B). The TEMP declared that the battlefield of the future will be "characterized by extensive use of offensive electronic warfare." The Army did not explain the reason for the electronic countermeasures limitation in the TEMP, the test plan, or the test report.

The July 1993 initial OT&E report, "Test and Evaluation Report (TER) for All Source Analysis System (ASAS) Block I Initial Operation Test and Evaluation (IOT&E)," states that "Only ASAS communications are susceptible to electronic countermeasures. . . . MSE [multiple subscriber equipment] is the primary communications for ASAS. MSE to date [July 1993] is also the only communications system which has been procured and fielded without extensive

ECM [electronic countermeasures] testing." The December 1994 OT&E report, "Test and Evaluation Report for the All Source Analysis System (ASAS)," states that the failure to test in an electronic countermeasures environment "prevented an evaluation of how well ASAS can work through the jamming of its communications."

The 1994 test plan said the system was not directly effected by not testing in an electronic countermeasures environment because countermeasures would disrupt the multiple subscriber equipment communications path rather than the system. The plan states that testing of the system's ability to "establish and maintain interfaces, both internally and externally, will be accomplished by a series of subtests," which would use 13 communications links. Nine of 13 links would use the multiple subscriber equipment. Because of the importance of the multiple subscriber equipment to ASAS communications, the Army should have tested the system in an electronic countermeasures environment.

Navy. Operational testing of the Navy MK-48 torpedo upgrade proceeded despite significant test limitations. The test plans recognized shortcomings in the availability of appropriate targets, the test environment, countermeasures availability, safety considerations, and tactic restrictions that limited operational testing. The test plan did not always explain the cause of the shortcomings. Further, modeling and simulation capability that might have ameliorated the limitations were not available. The test plan for OT-IIIA states that "Dynamic modeling and simulation capabilities, which would provide for assessment of the MK-48 ADCAP [Advanced Capability] torpedo in sophisticated CM [countermeasures], counterfire, and acoustic environment do not currently exist."

Air Force. The recent operational testing of the F-16 High-Speed Anti-Radiation Missile targeting system illustrates not testing against representative threats in an adequate number of realistic threat environments. The F-16 TEMP states that "a complete array of threat surface and air weapons systems will not be available" for OT&E. It states that "Several emitter [electronic signal generators] simulators will be used to simulate the signals of threats otherwise unavailable for field testing. The emitters available on the test ranges cannot duplicate the real world in either variety or density."

Additionally, the F-16 operational test plan states that "The majority of the testing will be accomplished in the high desert environment of the USAFWTC [U.S. Air Force Weapons and Tactics Center] ranges." The plan states, "Results may differ from those obtained in other environmental conditions of the world." Testing of the system advanced despite both threat and environmental limitations. Although the Air Force recognized the possibility of different results in another environment, the Air Force should have tested in an environment representative of the Western major regional conflict scenario in the Defense Planning Guide.

Reporting Test Limitations

While the Military Department test plans and reports list test limitations, coverage of the impact and cause for each limitation was less complete.

Reporting Limitation Assessments. Determining the cause of the limitation is hampered because the Military Departments list or label a limitation without explaining the cause. Our review of testing documents for the ASAS, MK-48, and F-16 programs is instructive. The ASAS test report, December 1994, explained the cause for only 33 percent of the limitations for that test. The MK-48 test report for OT-IIIA, November 20, 1991, explained the cause for 86 percent of the cited limitations. However, for OT-IIIC, February 7, 1995, the test plan gave a cause for only 43 percent of the listed limitations. The F-16 overall TEMP, October 14, 1992, explained 57 percent of the limitations while the F-16 TEMP Annex for the missile targeting system did not explain the cause of the single [only listed one] noted limitation. Even those percentages overstate, to some extent, the reporting of causes because the explanations describe a category-like threat, security, or range limitation without further discussion.

Army. The ASAS initial OT&E II test and evaluation report cites nine limitations. For one limitation, the report states that testing with live sensors was not possible because of costs, but then states that not testing with live sensors would have no impact. However, the capability of the system to pass information to the live sensors and the capability and reliability of the sensors to transmit information could not be evaluated.

A second limitation was deferral of electronic countermeasures testing without stating how long. The impact was that a measure of performance concerning whether the system can deploy and operate on the battlefield in an electronic countermeasures environment could not be evaluated. The related measure of effectiveness criteria was that commanders and their staffs must render a positive assessment of the capability to operate in an electronic countermeasures environment. The test report did not state why the test was deferred or the impact.

Navy. The MK-48 test plans for OT-IIIA, OT-IIIB, and OT-IIIC illustrate the lack of information provided as to the cause of a test limitation. The three separate test plans cite 19 limitations with safety cited most often as the cause as shown in Table 6. For 7 of 19 limitations, or 37 percent, the Navy labeled but did not actually describe the cause, for example, listing a threat limitation without describing the cause for the limitation. For 2 of 19 limitations, the Navy did not even label the cause.

Table 6. MK-48 Reasons for Test Limitations

Test	Number of Limitations	Safety	Other	Not Stated
OT-IIIA	5	2	0	3
OT-IIIB	7	3	1	3
OT-IIIC	7	2	1	4

Air Force. The F-16 operational test plan states that "The majority of the testing will be accomplished in the high desert environment of the USAFWTC ranges because of cost constraints [emphasis added]." The second limitation in the test plan, concerning the availability of threat surface and air weapons, was noted without explanation.

Additionally, the JSTARS TEMP, September 1, 1993, states test range limitations will affect the ability to operationally test the system adequately. The main limitations are the non-availability of ranges simultaneously possessing overall electronic counter-countermeasures testing capability, large quantities of ground targets, range and target instrumentation, and operational terrain and weather factors.

Resources. The FY 1994 Annual Report and DOT&E and Military Department officials offered explanations for the test limitations as discussed below. Test limitations occur because the Military Departments lacked funding to procure the threat systems, an adequate quantity of systems necessary to conduct a realistic test, and the opportunity to acquire the appropriate threat system.

OSD management uses the Foreign Material Acquisition/Foreign Material Evaluation Program to acquire threat systems. Also, OSD management uses the Central Test and Evaluation Improvement Program as a tool for funding high-priority, multi-Service DoD test and evaluation requirements, such as modelling and simulation development and expanding the test and evaluation technology base for emerging technologies. However, DoD managers have curtailed test resource expenditures over the past few years. For example, the Department's FY 1995 test and evaluation resource investment is about 78 percent of the FY 1990 investment.

With the change from a global threat to numerous regional ones and the insertion of newer technologies into already developed and fielded threat systems, the variety of threat systems has increased. As a result, DOT&E officials stated that costs to procure the varied threat systems has increased and, with reduced funding, the quantity that can be procured for operational tests is more limited. The DOT&E officials also stated that threat systems are not always available to acquire. Therefore, to mitigate the limitations of realistic operational environment, DoD builds surrogates based on estimates of how the system may perform.

Reporting Obstacles. The failure to address the impact of test limitations adequately has not changed materially since 1991. The May 24, 1991, Inspector General report states that "when questioned about the rationale for the omission [of a discussion of test limitations], the test directors stated their informal guidance was to present test results as positively as possible and to avoid any negative statements that might serve as ammunition for opponents of the weapons system tested."

Contemporary concerns about reporting limitations and their impact can be inferred from an Army briefing on the ATCCS III [Army Tactical Command and Control System], July 28, 1994. Three briefing charts cite United States Army Operational Test and Evaluation Command Memorandum 73-1, May 11, 1994, that states policy for release of test information. For release to acquisition team members, Army staff, DOT&E, General Accounting Office, and other officials, adherence to specific procedures was required to include a need to know and approval by the commanding general of Operational Test and Evaluation Command.

Uncertainty and Risk

While citing limitations and general causes, the Military Departments are not explicit to either the cause or the impact of the limitations.

Assessment. The Military Departments do not adequately report the cause or impact of test limitations. Also, the Military Departments' test documents are not providing essential information for assessing acquisition risk and for decisionmaking. Military Departments are required to report the limitations and the impact of the limitations on the ability to resolve the COIs and the capability to formulate conclusions regarding operational effectiveness and suitability. The following examples illustrate the lack of information on the reported limitations and their impact.

Army. The ASAS test plans and reports do not adequately address the impact of the test limitations. The December 1995 test report listed nine limitations. For two of these limitations (22 percent), the report addressed the impact on the COIs, but not the impact on the ability to formulate a conclusion regarding operational effectiveness and suitability.

Navy. The Navy planning and reporting for MK-48 limitations was similar to that of the Army. The test report, however, implies that the Navy was aware of the requirement to assess the impact of the limitations because the report states that the limitations "neither impacted the ability to resolve critical operational issues (COI) nor precluded the formation of conclusions regarding operational effectiveness or operational suitability." The report did not discuss the impact of the individual limitation. In the February 1995 test plan, the general comment that limitations did not impact the ability to resolve COIs and form conclusions on effectiveness and suitability was included; however, the impact of one of the seven individual limitations was addressed.

Air Force. The F-16 TEMP mirrors the approach of the Army and Navy. Neither the general F-16 TEMP nor the annex for the missile targeting system discussed the impact of the limitations on resolving COIs or developing conclusions for effectiveness and suitability.

System Effectiveness. The Military Departments are fielding systems with unknown effectiveness against identified threats. While the preceding examples do not prove that the fielded systems cannot meet requirements in combat, the purpose of operational testing is to demonstrate that they can. Testing with limitations increases the risk of incorrectly concluding that a system is effective and suitable.

Failure to test a system designed to communicate with other battlefield systems in an electronic countermeasures environment (JSTARS and ASAS) introduces substantial risk in forming conclusions as to effectiveness of the system in combat. Not testing a system against the full range of expected targets and in an environment not representative of expected combat conditions introduces uncertainty as to the system's effectiveness and suitability. Mitigating some threat limitations is the capability to use simulators instead of actual threat targets. Testing the F-16 missile targeting system only in the desert, although it has a worldwide operational mission requirement, introduces a high risk on its effectiveness in other climates.

Summary

Test documents and DOT&E and Military Department officials often cited test constraints caused by the lack of threat systems, range availability, and safety as bases of test limitations. Although those long-standing resource shortfalls are well documented and understood, the cause and impact of test limitations need to be in test documents so decisionmakers can make informed decisions. Because test plans are being approved without adequate analysis of the cause and impact of test limitations, decisionmakers cannot make informed decisions for obtaining those needed resources or to conclude whether a system is effective and suitable.

Recommendations, Management Comments, and Audit Responses

Deleted and Renumbered Recommendations. As a result of management comments, we deleted draft Recommendation C.1.a. and C.1.b. and renumbered C.2. to C. The DOT&E, Army, Navy, and Air Force commented on the finding. See Appendix I for a summary of their comments and audit responses.

C. We recommend that the Director, Operational Test and Evaluation, establish a policy to approve Test and Evaluation Master Plans or Test Plans only when they address the specific cause for each limitation and the probable impact of the limitation on test conclusions, both for the evaluation of critical operational issues and the ability to conclude whether the system is effective and suitable.

DOT&E Comments. The Director nonconcurred, stating that although the Military Departments have not always reported all effects of test limitations, DOT&E has no indication that this limitation has denied decisionmakers essential information.

Audit Response. Although the Director nonconcurred with the recommendation, the revised DoD 5000.2-R, March 15, 1996, meets the intent of our recommendation The DoD 5000.2-R requires the Military Departments to report the impact of limitations:

> Discuss the test limitations including threat realism, resource availability, limited operational (military, climatic, nuclear, etc.) environments, limited support environment, maturity of the tested system, safety, etc., that may impact the resolution of affected critical operational issues. Indicate the impact of the test limitations on the ability to resolve critical operational issues and the ability to formulate conclusions regarding operational effectiveness and operational suitability.

Better analysis and reporting of limitations in TEMPs, test plans, test reports, certification messages, and certification briefings will help decisionmakers evaluate whether to test with limitations.

Army, Navy, and Air Force Comments. Although not required to comment, the Army concurred, the Air Force partially concurred, and the Navy nonconcurred. The Army stated that it was already reporting the impact of limitations. The Navy stated that a nonapproval policy by DOT&E would reduce efficiency; however, the proper use of the IPT process will preclude the need for a nonapproval policy. The Air Force stated that it describes limitations when and where they increase understanding of the test and its results.

Audit Response. We disagree with the Navy about the potential inefficiency of a DOT&E nonapproval policy; the DOT&E already reviews TEMPs and test plans. However, we do agree with the Navy in the expectation that the proper use of IPTs will obviate the need for DOT&E disapprovals of TEMPs and test plans. Although the Air Force stated that it describes limitations when they feel it is necessary, we maintain that the assessments of the impact of the limitations are necessary at all times to ensure a full and complete understanding of the testing and the related results.

Finding D. Test Status Reporting

The FY 1994 Annual Report of the Director, Operational Test and Evaluation (the Annual Report), did not include results of operational testing performed on nonmajor programs that were not selected for oversight. Also, for 16 of 65 of the reported systems, the report did not present a complete and consistent assessment of the system's performance demonstrated in operational tests because DOT&E lacks resources to monitor, review, and report all DoD operational testing. The DOT&E also lacked adequate internal guidance and training for its action officers. As a result, the Secretary of Defense, the Under Secretary of Defense for Acquisition and Technology, and Congress did not get complete and consistent operational test information for decisionmaking.

Test Reporting Policy

The DOT&E is to provide the Secretary of Defense and Congress an unbiased insight into the operational effectiveness and suitability of new systems and major modifications to existing systems. Each year, the DOT&E issues his Annual Report to the Secretary of Defense, Under Secretary of Defense for Acquisition and Technology, and Congress summarizing the operational testing performed for DoD oversight programs for the fiscal year. Oversight programs are acquisition category I programs and other programs (nonmajor) selected due to their relative importance and sensitivity.

Statute. The Director is required by 10 U.S.C. 139 to monitor and review all operational test and evaluation in DoD and to summarize the OT&E activities of the DoD during the preceding fiscal year. The law also requires DOT&E to have sufficient professional staff to implement the duties and responsibilities of the Director.

DoD Guidance. DoDI 5000.2 implemented 10 U.S.C. 139 and also stated that the Director will prepare an annual report summarizing all OT&E activities within the DoD during the preceding fiscal year.

Although the guidance stated that DOT&E is to report on all OT&E activities, the FY 1994 Annual Report, February 1995, states that DOT&E is responsible for reporting the operational test results only for major Defense acquisition programs and nonmajor systems with DOT&E oversight.

Annual Test Report

Report Content. The FY 1994 Annual Report omits results of operational testing performed on nonmajor programs not selected for DOT&E oversight. The report summarized the OT&E activity for 59 of the 189 DoD major programs and other designated oversight programs. Neither 10 U.S.C. 139 nor DoDI 5000.2 stated that the DOT&E reports should be limited to an "oversight" listing or exclude nonmajor programs.

Operational testing and evaluation and reporting issues also occur in nonmajor programs that do not have DOT&E operational testing oversight. For example, the AC-130U Special Operation Forces Gunship was on the DOT&E oversight listing for live fire testing, but not operational testing. Major OT&E issues occurred as discussed in Finding B.

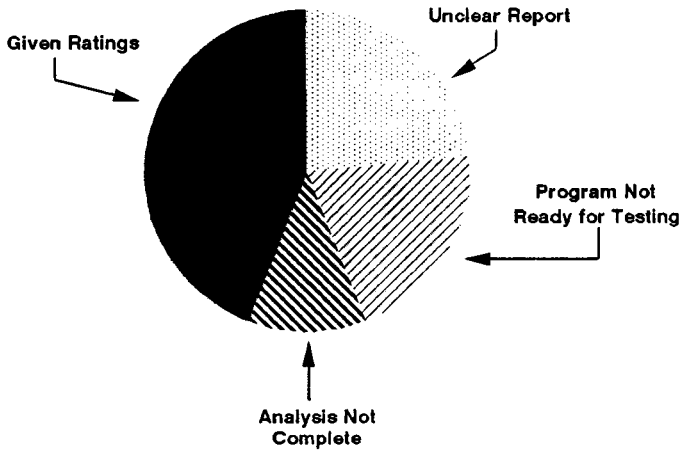
Additionally, the Naval Audit Service reviewed eight nonmajor programs with total projected costs of $2.7 billion. The Navy's subsequent report (Appendix B) concluded that operational tests and test results were not adequately documented and that the test results were not given adequate consideration in production decisions. The Inspector General, DoD, Report 92-079, "Operational Test and Evaluation of Nonmajor Systems," April 17, 1992, concluded that OT&E was inappropriately limited or omitted for 8 of 17 systems that did not have the DOT&E oversight.

Congress has not given DOT&E a waiver or other legislative relief from its OT&E monitoring, reviewing, and reporting responsibilities. A DOT&E official stated that the decision not to monitor, review, and report all operational testing in the Annual Report was a conscious, but informal, decision based on the lack of resources.

Clarity of the Annual Report. The Annual Report does not always completely and consistently assess the systems' operational testing performance. According to 10 U.S.C. 2399, the Annual Report is to describe the status of test and evaluation activities in comparison to the TEMP for the systems covered in the report. The status of the test, the source of the analysis, and the overall rating of the system by DOT&E were not always discernible from the report's systems summaries.

For example, the FY 1994 Annual Report states that the Navy OTA determined that the AN/SQQ-89(V)6 Antisubmarine Warfare Combat System was operationally effective and suitable. However, DOT&E did not state whether the system was operationally effective and suitable. This ambiguity leaves the reader uncertain as to whether DOT&E agreed with the Navy assessment.

Our review of the FY 1994 Annual Report showed that 23.5 percent of the systems' summaries were not clear, as shown in the figure.



Status of Test and Evaluation Reported

Both the report content and clarity issues result from a lack of sufficient resources for the DOT&E to meet the reporting intent of the congressionally directed oversight of OT&E. Current legislation and policy require earlier and continuing operational assessments, more explicit and implicit reporting requirements, and the Live Fire Test responsibilities added to the DOT&E workload.

Oversight Activity and Resources

DOT&E lacked resources to monitor, review, and report on all operational testing performed in the DoD, as required by 10 U.S.C. 139. In addition, the Federal Acquisition Streamlining Act of 1994 transferred the responsibility for oversight of the Live Fire Test and Evaluation from the DTSE&E to the DOT&E along with a requirement to submit an unclassified annual report concurrent with the classified one to include live fire test results.

Activity. The DOT&E FY 1994 review activities for the 189 oversight programs included:

- o approving 54 TEMPs;
- o approving 49 operational test plans;

o preparing and submitting numerous reports to the Defense Acquisition Board;

o publishing three beyond LRIP reports;

o reviewing the planning, conducting, and evaluating operational test activities;

o meeting with Military Department OTAs, program officials, private-sector organizations, and academia; and

o providing information to the Defense Acquisition Board principals, the Secretary and Deputy Secretary of Defense, the Under Secretary of Defense for Acquisition and Technology, the Military Departments, and Congress.

Documents that the DOT&E action officers review, approve, and produce are illustrated in Appendix F.

Resource Requirements. The DOT&E currently employs 50 people, divided primarily into two groups: 26 action officers and 24 policy and support personnel. Each action officer is responsible for monitoring, reviewing, and reporting as few as 2 and as many as 21 individual programs, to include live fire test. Because of DOT&E workload and special needs, in FY 1994, the DOT&E contracted with the Institute for Defense Analysis (the Institute), a Federally Funded Research and Development Center, for approximately 62 staff years at a cost of $10 million to assist in monitoring, reviewing, and reporting the OT&E. The use of the Institute is consistent with the Office of Federal Procurement Policy Letter 84-1 and Federal Acquisition Regulations that allow the use of federally funded research and development centers for special needs that in-house resources cannot meet. In the future, DOT&E, because of an increased workload, will rely even more on the Institute for support. In 1990, DOT&E concluded that the then current staff of 48 could not do the work in-house and "do the job right."

Action Officers' Guidance and Training. Of the 26 action officers, 12 are active duty military officers who are assigned from operating units to DOT&E. These officers generally do not have acquisition or test and evaluation experience.

The DOT&E internal guidance and training for its action officers are limited. The action officers use an internal policy guide, the "New Assistant's Guide," that provides general, but limited, guidance. DOT&E management also developed a self-help video so the action officers can learn their new work responsibilities. Additionally, new action officers to DOT&E receive "on-the-job" training instead of formal training for monitoring, reviewing, and reporting operational testing for their assigned programs. The Defense Acquisition University offers courses on the basic acquisition process and test and evaluation, which would benefit the productivity of DOT&E new action officers and DOT&E.

The policy guide summarizes how to review test plans and the resources available to accomplish this task. However, the guide does not provide details for reporting the operational test results in the annual report.

New action officers work for a short time with outgoing action officers and DOT&E management to be trained on-the-job. The action officers develop an informal network with other action officers for assistance.

The report summarization process is verbally passed from the previous action officer to the new action officer. Also, prior annual reports are used as "guides" for the current annual report. Completeness and consistency of reporting were inconsistent and selectively accomplished and dependent more on the action officer assigned.

Completeness and Consistency of the Annual Report

The Secretary of Defense, the Under Secretary of Defense for Acquisition and Technology, and Congress are not getting complete and consistent operational test information on systems for decisionmaking, which lessens the credibility of DOT&E products. The "New Assistant's Guide" states that reports are the most important product of the Office and that the reporting is to be objective and complete. However, the objectiveness of the FY 1994 Annual Report summary for the Combat Service Support Control System is questionable.

The summary states that the DOT&E had not evaluated the test data from the system's September 1994 initial OT&E. The initial OT&E was to support the Army Systems Acquisition Review Council full-rate production (Milestone III) decision scheduled for April 1995. Although the report said that the DOT&E had not evaluated the test data, it states, "We are aware of no significant issues which would prevent the 1995 Milestone III decision." At the same time the Annual Report was published (February 1995), the Army OTA published its test report for the system's initial OT&E that concluded the system demonstrated neither operational effectiveness nor operational suitability.

The DOT&E statement did not clearly present the system's performance, which prevented DOT&E from presenting an objective, unbiased overview of the system. The "New Assistant's Guide" states that the action officers must "ensure that the operational effectiveness and suitability of weapon systems are tested adequately, evaluated objectively, and reported independently to acquisition decision makers."

Summary

DOT&E must maintain credible products on which Department acquisition decisions are based and provide an unbiased insight of system performance, to include the nonmajor systems, to the DoD senior management and Congress. Producing credible products may be accomplished with a well-constituted oversight program that has the appropriate professional staffing level to "do the job right" and definite standards and specific checks.

Recommendations, Management Comments, and Audit Responses

Unresolved Recommendations. We ask that DOT&E provide additional comments on unresolved Recommendation D.2. The DOT&E and the Air Force commented on the finding. See Appendix I for a summary of their comments and audit responses.

D.1. We recommend that the Under Secretary of Defense (Comptroller), provide the Director, Operational Test and Evaluation, the necessary funds for increased DoD civilian staff years.

Under Secretary of Defense (Comptroller) Comments. The Under Secretary of Defense (Comptroller) nonconcurred with the draft recommendation addressed to him, stating that the Director, Washington Headquarters Services, determines the funding and staffing levels of all OSD staff elements.

DOT&E Comments. Although not required to comment, the Director concurred, stating that he would work with the Comptroller to achieve increased civilian staff years.

Audit Response. The Under Secretary of Defense (Comptroller) comments were not responsive. However, the intent of the recommendation will be met if the Director seeks staff augmentation through the DoD budget process.

D.2. We recommend that the Director, Operational Test and Evaluation:

a. Conduct a new staffing-requirements study to determine the appropriate mix of in-house DoD staffing and contractor support requirements needed to perform the Director, Operational Test and Evaluation, mission.

DOT&E Comments. The Director concurred, stating that he would conduct the staffing and contractor support study.

Audit Response. The Director's proposed action meets the intent of our recommendation. We ask that he provide the effective date for the planned action in his response to the final report.

b. Provide the Director, Operational Test and Evaluation, staff formal guidance for writing the annual report and for ensuring that the reports are complete and consistent.

DOT&E Comments. The Director concurred, stating that he issued guidance on September 29, 1995, for writing the Annual Report and for ensuring that the reports are complete and consistent.

Audit Response. We examined the September 29, 1995, guidance issued by the Director; it did not address the intent of our recommendation. The guidance does reference 10 U.S.C. 2399 and 10 U.S.C. 2366 reporting requirements. However, to fully meet the intent of our recommendation, it should state what is to be included and should include a checklist for action officers to use to ensure a complete, consistent, and accurate report. We would be happy to assist the Director in developing such a checklist as an aid for action officers. We ask the Director to reconsider his response and provide additional comments to the final report.

c. Provide new action officers with formal training on the acquisition process, including test and evaluation.

DOT&E Comments. The Director concurred, stating that he initiated the development of a proactive training plan for action officers, which will include milestones for monitoring its execution.

Audit Response. The Director's action meets the intent of our recommendation. We ask the Director to provide us the estimated completion date for the plan and the general contents of the plan, to include acquisition training when warranted, in his response to the final report.

d. Include a summary of the operational testing performed on nonmajor weapon systems in the Director's annual report.

DOT&E Comments. The Director concurred, stating he would work with congressional staffs, OSD offices, and the Military Departments to determine what operational test information on nonmajor weapon systems will be in the annual report.

Audit Response. The Director's proposed action meets the intent of our recommendation. However, he did not provide the date of when this action will be done. We ask that the Director provide estimated completion dates for these actions in his response to the final report.

Finding E. Reporting Program Results

The Selected Acquisition Report (SAR) does not present a complete and accurate assessment of the system's performance limitations identified during testing for 4 of 12 systems we examined. The program manager prepares the report, which does not always present an objective assessment of the system's performance limitations identified during operational testing. As a result, the SAR can present an overly optimistic assessment of systems to decisionmakers.

Program Status Reporting Policy

The decisionmakers in the acquisition chain can effectively control a program only when they are informed of emerging problems and issues, including test results. A management reporting system that contains consistent and reliable data on the status of the programs can keep those decisionmakers informed.

Statute. United States Code, title 10, section 2432 (10 U.S.C. 2432), requires the Secretary of Defense to submit the SAR to Congress at the end of each year. The SAR provides the status on the current major Defense acquisition programs; summarizes key program cost, schedule, and performance technical baseline information; and program variance analysis relative to the baseline. Data in the report are DoD Component estimates based upon test data; technical information; and contractor cost, schedule, and performance data.

DoD Guidance. The DoDI 5000.2, part 11, section D, "Periodic Program Status Reports and Required Certifications," implemented 10 U.S.C. 2432 and required the program manager to prepare the SAR. The Instruction also required the Under Secretary of Defense for Acquisition and Technology in coordination with other OSD offices (to include DOT&E) to review and approve the SAR, then forward it to Congress.

The DoD Manual 5000.2-M, part 17, "Selected Acquisition Report," provided specific guidance for preparing the SAR.

o Section 7, "Program Highlights," required a narrative summary of significant accomplishments and developments to include program highlights from inception to the present. Specifically, program highlights include the status of testing and an assessment of the extent to which the system is expected to satisfy its current mission requirements, identifying any areas where it may fall short.

o Section 10, "Performance Characteristics," required a list of quantifiable system performance characteristics that are the primary indicators of technical achievement of engineering objectives and thresholds and the system's operational capability to accomplish the mission. The performance characteristics should be representative of the characteristics that will be subject

to contractor development and Government operational tests to evaluate the system's effectiveness. Additionally, for each data element being reported, the "value (relative to the objective and/or thresholds) actually achieved in the latest development or operational" test is to be entered.

Test Status Reported

The SAR does not always present a complete and accurate assessment of the program's performance limitations identified during testing. We identified incomplete and inaccurate assessments for system performance and test results in the SAR for 4 of 12, or 33 percent, of the program SARs we reviewed. The four programs are the Army Combat Service Support Control System, JSTARS GSM, ASAS, and the Air Force Sensor Fused Weapon.

Completeness of Performance Characteristics Assessments. The system's demonstrated performance characteristics reported in the SAR were not always complete.

Army. The Combat Service Support Control System SAR, December 31, 1994[2], Section 10, as shown in Appendix G, did not include the latest test data. The SAR Section 7 states that a Limited Users Test was successfully completed in November 1993. Neither those test results nor any other test results, which were available, are shown in the "Performance Demonstrated" column. The column shows all performance elements as "TBD (to be demonstrated)." The SAR does state that the performance characteristics of "TBD" will be changed to reflect actual test results after the release of the test reports from the initial OT&E, which is scheduled for late 1996.

The JSTARS GSM SAR, December 31, 1994, Section 10, includes the performance objectives and thresholds for evaluating the system's effectiveness. The SAR sent to Congress did not include the latest test data for the Light GSM. The "Demonstrated Performance" column shows "TBD" for 13 of 25 performance elements. Data for 7 of the 13 "TBD" elements were available from the September 1994 Force Development Test and could have been included.

Air Force. The Sensor Fuzed Weapon SAR, December 31, 1994, Section 10, performance characteristics were not consistent with the program's acquisition program baseline. Section 10 did not provide detailed information that the system's effectiveness is reduced when the delivery envelope increases from low to high altitudes. The system's reduced effectiveness is discussed in the Inspector General, DoD, Report No. 95-271, "Acquisition of the Sensor Fuzed Weapon," June 30, 1995 (Appendix B). The report states that Section 10 performance characteristics are not consistent with program documents. The

[2]The Selected Acquisition Reports are dated December 31, 1994, but changes were accepted through March 8, 1995.

report also states that the "Demonstrated Performance" column and "Program Manager's Estimate" column footnotes do state that the system's multiple kills effectiveness is degraded at higher altitudes. However, the footnotes do not define the delivery altitude regime when the effectiveness degrades and to what degree.

Accuracy of SAR Program Highlights Assessments. The assessment of the program's performance limitations identified during testing are not always presented completely or accurately in the program highlights.

Army. The JSTARS GSM SAR, December 31, 1994, Section 7, states that the system's Force Development Test was completed September 1994. Section 7 also states that the Light GSM exceeded operational availability requirements, successfully interfaced with the unmanned aerial vehicle, and will satisfy mission requirements. Those statements are not supported by either the demonstrated performance characteristics reported in Section 10 or the Army OTA and DOT&E assessments of test results shown in Table 3 of Finding B.

The Army OTA assessment of the September 1994 Force Development Test, "Abbreviated Operational Assessment (AOA) of the Joint Surveillance Target Attack Radar System (Joint STARS) Light Ground Station Module (LGSM)," March 7, 1995, concluded that the Light GSM did not meet all LRIP exit criteria. Exit criteria not met included system operational availability and simultaneous receipt of JSTARS and unmanned aerial vehicle data. The assessment, unlike the SAR, concluded that the JSTARS and unmanned aerial vehicle interface had not been adequately tested. The OTA assessment of the Force Development Test was based on the "Army Material Systems Analysis Activity Interim Draft Independent Evaluation Report," January 20, 1995, and "Army Test and Experimentation Command Force Development Test and Experimentation Test Report," January 1995.

Air Force. The Sensor Fuzed Weapon SAR, December 31, 1994, Section 7, states that the system will satisfy all mission requirements. In contrast, Section 10 states that the system has "inherent performance degradation at higher altitude." Section 7 does not discuss that the operational tests do not validate the multiple kills per pass requirement across the delivery envelope (200 feet through 20,000 feet). It fails to note that the second phase of operational testing is limited to low altitudes due to the performance degradations for medium and high altitudes. Additionally, Section 7 does not discuss that the system may not satisfy all mission requirements.

In contrast, the FY 1994 Annual Report states that the DOT&E overall analysis of OT&E indicates that the Sensor Fuzed Weapon has the "potential" to be operationally effective and suitable. The report also said that the system's effectiveness decreases as release altitude, dive angle, and time of flight increase because of adverse effects of wind conditions, weapon dispersion, and aim point uncertainties. The report states that some type of compensated munition dispenser is necessary for the system to achieve optimum kills per pass performance when employed from medium and high altitudes.

Information in the SAR should be consistent with test results and provide an unbiased assessment of the system's performance.

Acceptance of the Selected Acquisition Report

The program manager prepares the SAR, which does not always present an objective assessment of the system's performance limitations identified during operational testing. Those who are reviewing and approving the SAR are not ensuring that the report is objective, complete, and accurate.

DOT&E is responsible for reviewing the test status reported in both Sections 7 and 10. Although required to review the SAR, DOT&E officials stated that often they either do only a cursory review or do not review it at all because of the workload and resource shortages.

Additionally, an Under Secretary of Defense for Acquisition and Technology official stated that, although OSD officials review and approve the SAR, changes are seldom made unless the assessment of the system's status is significantly different from that of the program manager. The official stated that when the assessments differ significantly, the OSD officials try to persuade the program manager to make the appropriate changes. The program manager's perspective is often overly optimistic because the program's performance is driven by cost and schedule. As a result, the information in the SAR reflects the program manager's perspective more than that of the Department of Defense.

Although DoDI 5000.2 assigns the responsibility to prepare the SAR to the program manager, by law it is the Secretary of Defense's report and should present the DoD position, not the program manager's. The SAR should provide accurate acquisition management information to decisionmakers so they can make informed decisions.

Status for Decisionmakers

The SAR is sometimes misleading when the system has not met its performance requirements.

Army. The JSTARS GSM SAR, December 31, 1994, Section 7, states that, "based on successful testing," the Army System Acquisition Review Council approved LRIP for 12 Medium GSMs in May 1993. It also states that the system will satisfy mission requirements. Section 10 shows that the Medium GSM meets or exceeds all of its performance characteristics.

In contrast, the Army OTA test report "Operational Assessment (OA) of the Medium Ground Station Module (MGSM) of the Joint Surveillance Target

Attack Radar System (Joint STARS)," July 2, 1993, concluded that the Medium GSM demonstrated the "potential" to be effective and suitable, even though the software "lacks robustness and reliability, and limits mission performance."

Additionally, the FY 1994 Annual Report states that the Medium GSM demonstrated the "potential" to be "effective in the surveillance mission but not in the targeting mission due to failure to provide accurate target location." The report also states that the Medium GSM "was not considered suitable due to low reliability."

Although the Medium GSM was assessed as only partially effective and not suitable, the Army procured 12 units. The GSMs are to be deployed with contingency forces and used as training equipment. Additionally, on May 2, 1995, the Army Acquisition Executive approved LRIP for 10 Light GSMs, even though the Light GSM had not met its exit criteria for LRIP as discussed in Finding B.

The ASAS SAR, December 31, 1994, Section 7, states that the ASAS Block I operational effectiveness was successfully demonstrated at the Technical Test and Operational Demonstration in April and May 1993. The system is "expected to satisfy mission requirements."

In contrast, the Army Operational Evaluation Command test report, "Abbreviated Operational Assessment (AOA) of the All Source Analysis System (ASAS) Block I (Modified)," May 26, 1993, concluded that the operational demonstration was not conducted under realistic operational conditions and that the determination whether the system was operationally effective and suitable could not be made.

Additionally, the DOT&E FY 1994 Annual Report states that before the unfavorable initial OT&E in October 1992, about 95 percent of the systems' components were procured to equip nine divisions and two corps. The report did not assess the 1993 Test Development and Operational Demonstration.

Air Force. The Sensor Fused Weapon has not demonstrated operational effectiveness and suitability for all mission requirements; however, the weapon is being placed in inventory. The Sensor Fuzed Weapon SAR, December 31, 1994, Section 7, states that a limited number of weapons have been placed in inventory and are available for the user.

Summary

To make informed decisions appropriately, acquisition decisionmakers need accurate status of the program to be able to analyze that status quickly and effectively. A program summary, similar to one used in the Defense Acquisition Executive Summary, could be in the SAR for the decisionmakers.

The summary assesses the entire program, not just the Selected Acquisition Report Baseline like the SAR. The Defense Acquisition Executive Summary is illustrated in Appendix H.

Recommendations, Management Comments, and Audit Responses

Unresolved Recommendations. We ask that the Under Secretary of Defense for Acquisition and Technology provide comments on unresolved Recommendations E.1. and E.2. The Under Secretary of Defense for Acquisition and Technology commented on the finding. See Appendix I for a summary of the Under Secretary's comments and audit response.

E. We recommend that the Under Secretary of Defense for Acquisition and Technology:

1. Present the Department of Defense perspective in the Selected Acquisition Report.

2. Include a summary in the Selected Acquisition Report similar to the one used in the Defense Acquisition Executive Summary and include the Director, Operational Test and Evaluation, assessment of the performance characteristics and test and evaluation.

Under Secretary of Defense for Acquisition and Technology. The Under Secretary stated that the SARs are fully coordinated with OSD staff, incorporate the appropriate comments, and, therefore, represent the DoD position. He stated that his office "will reiterate the DoD policy of accurate and timely reporting of Demonstrated Performance information in the upcoming release of December 1995 SAR guidance." Additionally, he stated that a working group on SAR reform has been established and that a proposal under review would replace the "Program Highlights" section with an Executive Summary similar to that used in the Defense Acquisition Executive Summary report. These changes are planned to be incorporated in the December 1996 SARs.

Audit Response. The Under Secretary's comments partially meet the intent of our recommendations. We acknowledge that the SARs are coordinated with the OSD. However, in spite of the coordination, SARs are not always complete and accurate. The SAR reform working group is a welcome initiative and we suggest our recommendations be staffed through that group. We ask the Under Secretary to provide additional comments in his response to the final report.

DOT&E Comments. Although not requested to comment, the DOT&E concurred with Recommendation E.1. and supported Recommendation E.2.

Part II - Additional Information

Appendix A. Scope and Methodology

Scope

We limited our scope to programs that had DOT&E oversight. We selected 15 systems that were either operationally tested from FYs 1989 through 1995 or had operational testing scheduled. A variety of systems in various stages of operational testing were in our sample, including systems with combined development and operational testing and joint Service (Army, Navy, Air Force, and Marine Corps) systems. The sample systems are described in Appendix D.

Methodology

We conducted this economy and efficiency audit from April through October 1995 in accordance with auditing standards issued by the Comptroller General of the United States, as implemented by the Inspector General, DoD, and included such tests of management controls as were deemed necessary.

We reviewed data from May 1987 through November 1995. We evaluated the policies and procedures of the DOT&E and the Military Departments for developing and executing test plans. We compared and analyzed the actual planning, execution, and reporting for selected systems against the stated criteria to determine whether test plans are properly developed and whether the systems are ready for operational testing. The audit did not rely on computer-generated data. We interviewed OSD and Military Department officials. The organizations contacted or visited are listed in Appendix K.

We evaluated the FY 1994 Annual Report of the Director, Operational Test and Evaluation, submitted to the Secretary of Defense, Office of the Under Secretary for Acquisition and Technology, and Congress. We compared and analyzed the report to the systems' requirements, test plans, and test results reported by the Military Departments to determine the accuracy of the DOT&E and the Military Departments' reports to Congress.

We evaluated the DoD Selected Acquisition Reports submitted to Congress. We compared and analyzed the Selected Acquisition Reports to the systems' requirements, test plans, and test results reported by the DOT&E and the Military Departments to determine the accuracy of the Military Departments' reports to Congress.

Sample Selection

We only selected systems with DOT&E oversight for our sample. We excluded the special access programs (classified) and Major Automated Information System Review Council systems.

The Military Departments test data lists were reviewed to determine which systems were operationally tested from FYs 1989 through 1995 or were scheduled to be tested in FY 1995. Also, the DoD audit coverage was reviewed to determine which audits related to our sample and objectives had not had extensive audit coverage.

We judgmentally selected systems based on the test data list and audit coverage, the type of system, and joint Service programs. A variety of systems in various stages of operational testing, including systems with combined development and operational testing, was desired. Observing an operational test was desired.

The DOT&E Master Oversight List contained 189 programs: Ballistic Missile Defense Organization (9), Army (54), Navy (64), Air Force (43), and other DoD Components (19). Of those programs, we selected 15: 5 Army, 5 Navy, and 5 Air Force.

Management Control Program

DoD Directive 5010.38, "Internal Management Control Program," April 14, 1987, requires DoD organizations to implement a comprehensive system of management controls that provides reasonable assurance that programs are operating as intended and to evaluate the adequacy of the controls.

Scope of Review of Management Control Program. We evaluated management controls applicable to the critical program management elements of the Military Department operational test agencies. In assessing those controls, we evaluated plans and procedures, vulnerability assessments, written policies and procedures, and management-initiated reviews. We also reviewed the results of the Military Departments' self-evaluation of those management controls. We did not assess management controls or self-evaluations at the Office of the Director, Operational Test and Evaluation.

Adequacy of Management Controls. We found the controls to be generally adequate. However, material management control weaknesses were identified as defined by DoD Directive 5010.38.

o The Military Departments lack criteria, such as a decisionmaking model, to evaluate whether to proceed with testing when serious limitations are identified before operational testing and to determine the cost/benefit of testing with limitations. (Finding B, Recommendation B.2.)

o The DOT&E lacks guidance for preparing his annual report for the DoD senior management and Congress, and the DOT&E action officers lack formal training on the acquisition process and operational testing. (Finding D, Recommendations D.2.b. and D.2.c.)

Although we could not quantify the potential monetary benefits associated with management's implementation of the recommendations, we did identify other potential benefits. See Appendix J for a summary of the potential benefits resulting from the audit. A copy of the final report will be provided to the senior official responsible for management controls in the Office of the Under Secretary of Defense for Acquisition and Technology; the Office of the Director, Operational Test and Evaluation; and each Military Department.

Adequacy of Management's Self-Evaluation. The Military Departments did not identify the area of testing with limitations as an assessable unit and, therefore, did not identify or report the material management control weakness identified by the audit. However, the Army Test and Experimentation Command was developing a new management control plan that included test management, test planning, test execution, and test reporting. The Navy Operational Test and Evaluation Force determined that the functional areas of plans and reports were both assigned a high level of risk and were scheduled for review during FY 1995. However, the Operational Test and Evaluation Force had not performed the applicable management control testing. Additionally, the Air Force has established research, development, test, and evaluation as a category of internal management controls, but has not addressed the specific area related to our finding.

Area Not Requiring Further Review

We reviewed the systems' operational requirements documents, system threat assessment reports, and test and evaluation plans to determine whether threat realism was adequately addressed for the systems. The systems' test and evaluation plans addressed the threats, which the Defense Intelligence Agency approved. We discussed the DOT&E process for ensuring the realistic portrayal of threat forces with the DOT&E action officers, operational test agencies, and the systems' program office test officers. The policies and procedures for ensuring the realistic portrayal of threat forces are generally adequate.

Appendix B. Summary of Prior Audits and Other Reviews

General Accounting Office

General Accounting Office, NSIAD Report 95-172R (OSD Case No. 9951), "Production of Joint STARS LGSM," May 1995, concluded that the Light Ground Station Module (Light GSM) has not met its exit criteria to enter LRIP. The report recommended that the Army postpone the system's LRIP contract award until DoD has assurance that the system demonstrates its ability to meet all exit criteria established by DoD for proceeding to LRIP. Management comments were not in the report.

General Accounting Office, NSIAD Report 95-18 (OSD Case No. 9725), "Weapons Acquisition, Low-Rate Initial Production Used to Buy Weapon Systems Prematurely," November 1994, concluded that current legislation and DoD acquisition policies permit LRIP to start before any OT&E is conducted resulting in unsatisfactory weapons and deployment of substandard systems to combat forces. The report recommended that the Secretary of Defense require programs to plan, buy prototypes for, and conduct realistic testing to certify that systems clearly demonstrate the potential to fully meet all minimum acceptable requirements before entering into LRIP; require those programs not required to test prototypes to instead test all key subsystems in an operational environment before entry into LRIP; and adopt the recommendations made by the Inspector General, DoD, regarding controls over the start and continuation of LRIP. Management did not concur with the recommendations, which are in resolution.

General Accounting Office, NSIAD Report 94-51 (OSD Case No. 9517), "Battlefield Automation, Premature Acquisition of the Army's Combat Service Support Control System," February 1994, concluded that the Army planned to procure Combat Service Support Control System equipment before operational testing. The report recommended that the Army defer procurement of the system's computers until the system successfully completes operational testing that demonstrates its military effectiveness and automated data exchange among and between the Army Tactical Command and Control System control segments. Also, the report recommended that the Army use the system's existing equipment to meet operational testing requirements. Management actions were considered responsive to the recommendations.

General Accounting Office, NSIAD Report 91-46 (OSD Case No. 8441), "T-45 Training System, Navy Should Reduce Risks Before Procuring More Aircraft," December 1990, concluded that the Navy planned to complete flight testing and initial OT&E after procurement commitments even though the T-45 aircraft design had not been proven effective or suitable for its mission. The report recommended that the authorization for the Navy to proceed with procurement

of additional T-45 aircraft be withheld until operational testing demonstrates that the system is effective and suitable for its mission. Management actions were considered responsive to the recommendations.

General Accounting Office, NSIAD Report 90-107 (OSD Case No. 8341), "Weapons Testing, DoD Needs To Plan and Conduct More Timely Operational Tests and Evaluation," May 1990, concluded that weapons systems are starting production before OT&E because DoD is not assuring that OT&E is planned and conducted before production. The report recommended that the DoD acquisition and testing directives clearly establish the need for OT&E before production start-up, define when OT&E must occur and when it may be appropriate for decisionmakers to rely on operational assessments, require the Services to plan for and conduct earlier OT&E, and require system or subsystem prototypes be built where practical and that these prototypes be operationally tested before production start-up. Management comments were not in the report.

Inspector General, Department of Defense

Report No. 95-271, "Acquisition of the Sensor Fuzed Weapon," June 30, 1995, concluded that the Sensor Fuzed Weapon operational tests did not validate the acquisition program baseline key performance parameters due to inconsistent parameters among program documents, the system's Cost and Operational Effectiveness Analysis did not provide adequate decisionmaking, and the acquisition objectives for the system were questionable. The report recommended that the Air Force revise key program documentation to clarify the key performance parameter; update the system's Cost and Operational Effectiveness Analysis to incorporate changes in operational concepts, available alternatives, and costs; and revise the system's Selected Acquisition Report to ensure Congress is accurately informed on the program's status. Management actions were considered responsive to the recommendations.

Report No. 94-014, "Low-Rate Initial Production in Major Defense Acquisition Programs," November 9, 1993, concluded that programs entered LRIP and full-rate production without completing some prerequisites in design, testing, and preparation for production. The report recommended that the DoD acquisition guidance be revised to provide additional internal controls for assessing the readiness of programs to enter LRIP and to limit the number of LRIP units produced to the minimum quantity necessary to support initial OT&E and production base considerations. Management partially concurred with the recommendations, which are still in resolution.

Report No. 93-087, "Review of the All Source Analysis System as a Part of the Audit of the Effectiveness of the Defense Acquisition Board Review Process--FY 1993," April 20, 1993, concluded that the Defense Acquisition Board process was not effective for the system and testing planned and conducted were inadequate to support a decision to field the Block I system or award the engineering and manufacturing development contract for the Block II system.

The report recommended a Defense Acquisition Board Milestone II review of the Program and that funding be withheld until a Milestone II acquisition decision memorandum was issued. Management actions were considered responsive to the recommendations.

Report No. 92-079, "Operational Test and Evaluation of Nonmajor Systems," April 17, 1992, concluded that OT&E was inappropriately limited or omitted resulting in systems procured with serious operational deficiencies. The report recommended that OT&E be conducted on limited procurement-urgent systems and that follow-on OT&E be conducted on one system developed under the foreign comparative testing program. Management concurred with the recommendations.

Inspection Report No. 91-INS-09, "Operational Test and Evaluation Within the Department of Defense," May 24, 1991, concluded that although OT&E adds value to the acquisition process, it does not have the major impact the Congress intended on acquisition milestone decisions. Instead of using OT&E results to delay or halt production of systems with questionable effectiveness or suitability, acquisition executives use the results to continue the test-fix-test scenario begun during developmental testing. The report recommended changes in legislation, organization, policy, and procedures to improve OT&E. Management action was considered responsive to the recommendations.

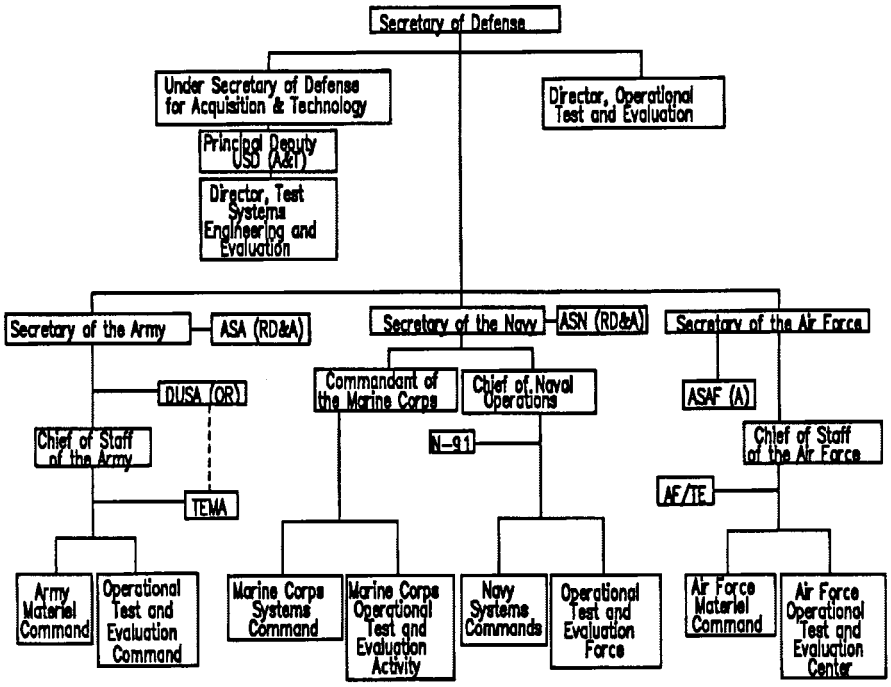
Navy

Naval Audit Service Report No. 061-C-91, "Operational Test and Evaluation of Nonmajor Systems," September 30, 1991, concluded that operational tests were not adequately documented, test results were not adequately documented, contractors were monitoring and participating in the testing, Navy program decision authorities used LRIP approvals to authorize piecemeal procurement of major portions of system requirements, and OT&E results were not adequately considered in production decisions. The report recommended changes in policy and procedures to improve OT&E. Management action was considered responsive to the recommendations.

Appendix C. Operational Test Organizations and Agencies

Office of the Secretary of Defense

Congress enacted Public Law 98-94 in 1983 establishing the position of the Director, Operational Test and Evaluation (DOT&E). The DOT&E is the principal OT&E authority within the OSD senior management structure and is chartered to provide independent oversight, coordination, and evaluation of the Military Departments' planning and execution of operational tests and reports. The Director is the principal advisor to the Secretary of Defense and the Under Secretary of Defense for Acquisition and Technology for OT&E matters. The Director communicates directly with the Secretary and must submit a report to Congress to meet statutory operational and live fire test reporting requirements. The DOT&E responsibility to Congress is to provide unbiased insight into operational effectiveness, suitability, lethality, and survivability of new systems and, in some cases, upgrades to fielded systems. The DOT&E approves test and evaluation master plans as well as operational and live fire test plans for designated systems. See Figure C-1.



AF/TE	Air Force (Test and Evaluation)
ASAF(A)	Assistant Secretary of the Air Force (Acquisition)
ASA (RD&A)	Assistant Secretary of the Army (Research, Development,and Acquisition)
ASN (RD&A)	Assistant Secretary of the Navy (Research, Development, and Acquisition)
DUSA (OR)	Deputy Under Secretary of the Army (Operations Research)
N-91	Directorate of Navy T&E and Technology Requirements
TEMA	Test and Evaluation Management Agency
USD(A&T)	Under Secretary of Defense for Acquistion and Technology

Figure C-1. Department of Defense Test and Evaluation Organization Chart

Army

The Army established its operational test agency (OTA) September 25, 1972, as a field operating agency assigned to the Office of the Chief of Staff. The agency, Operational Test and Evaluation Command, performs OT&E on selected major systems. The agency combines the evaluation function performed by the Operational Evaluation Command (OEC) and the operational testing function performed by the Test and Experimentation Command (TEXCOM). The Army Test and Evaluation Management Agency is responsible for developing and monitoring test policy. The Army Test and Evaluation Management Agency is under direction of the Deputy Under Secretary of the Army for Operations Research and administers OT&E staff functions for the Army Chief of Staff. See Figure C-2.

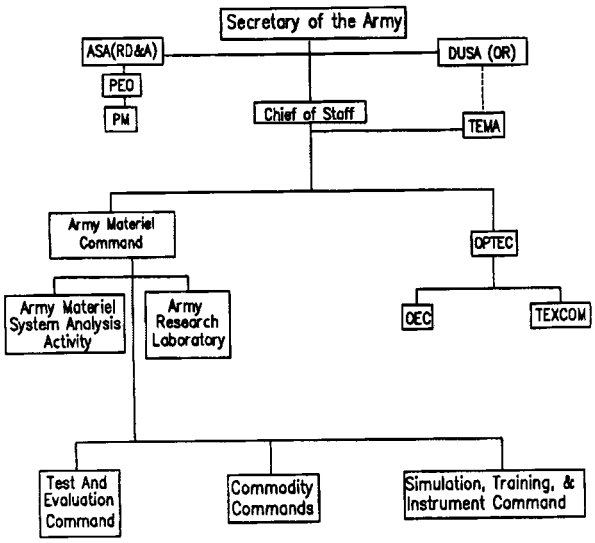


Figure C-2. Army Test and Evaluation Organization Chart

Navy

The Navy designated the Operational Test and Evaluation Forces as the independent test agency for OT&E in 1971. The Operational Test and Evaluation Forces is a field operating agency that reports to the Chief of Naval Operations. The Chief of Naval Operations is responsible for ensuring the adequacy of the overall test and evaluation program for the Navy. The Director, Test and Evaluation and Technology Requirements, establishes the Navy test and evaluation policy and guidance. The Test and Evaluation Division (OP-912) provides staff support for OT&E. See Figure C-3.

The Marine Corps established the Marine Corps Operational Test and Evaluation Activity in 1978. The activity is responsible for all OT&E in the Marine Corps. Marine Corps organizations do not provide staff level support equivalent to similar organizations in the Army, the Navy, and the Air Force.

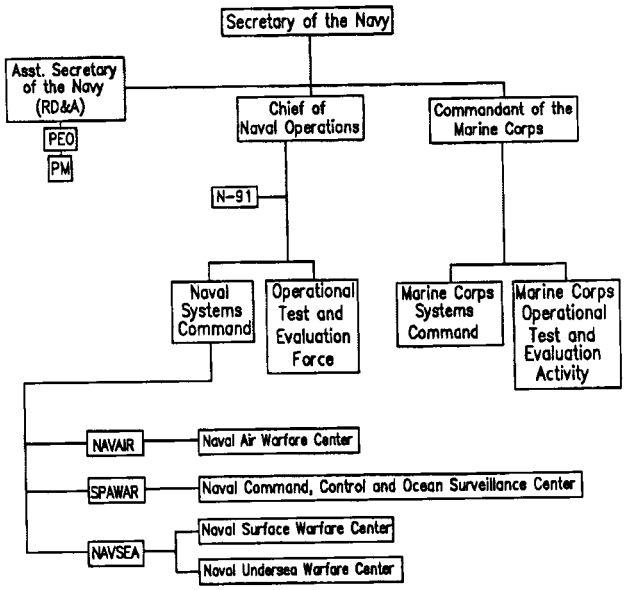


Figure C-3. Navy Test and Evaluation Organization Chart

Air Force

The Air Force designated the Air Force Test and Evaluation Center as the focal point for all OT&E in 1974. The Air Force Test and Evaluation Center is a field operating agency reporting to the Chief of Staff of the Air Force for all OT&E matters. The operational commands augment and support the agency in planning and conducting OT&E. Agency personnel conduct OT&E on all Air Force major Defense systems and monitor OT&E for non-major systems conducted by the operational commands. The Deputy Chief of Staff for Plans and Operations is responsible for supporting and coordinating the OT&E activities of the agency. The agency works with the Air Staff through the Air Force Test and Evaluation (AF/TE) Directorate. See Figure C-4.

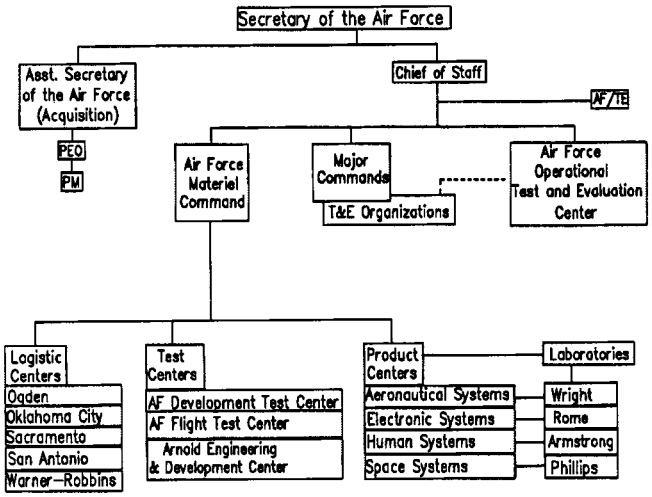


Figure C-4. Air Force Test and Evaluation Organization Chart

Appendix D. Description of Weapon Systems

Army

All Source Analysis System (ASAS). The ASAS is a ground-based mobile intelligence processor designed to provide automated support to Army combat commanders. The system will provide commanders a timely and comprehensive understanding of opposing force deployments, capabilities, and potential action.

Combat Service Support Control System. The system is a logistics command and control unit. The system will manage combat service support information for planning, coordinating, controlling, and executing logistic functions of sustainability and reconstitution of combat forces.

Joint Surveillance Target Attack Radar System (JSTARS) Ground Station Module (GSM). The JSTARS consists of an Air Force E-8 aircraft, Army ground station, and a data link between the two. The GSM receives, stores, and displays the surveillance and targeting data from the aircraft; provides tool for processing the data for situation development and targeting; and provides for the dissemination of information to users. The GSM will be configured on a high-mobility, multi-purpose wheeled vehicle (Light GSM) and a military standard 5-ton truck (Medium GSM). The JSTARS is a joint Army and Air Force (lead) program.

Secure Mobile Anti-Jam Reliable Tactical Terminal. The system is a MILSTAR, extremely high frequency, multi-channel, communications terminal. The system will provide the Army a mobile secure data and voice communications capability from the MILSTAR satellites and will provide connectivity to other systems. The system will be configured on a high-mobility, multi-purpose wheeled vehicle and will provide communications while unattended or from remote locations.

Single Channel Anti-Jam Manportable Terminal. The system is a MILSTAR, extremely high frequency, single-channel, man-portable, communications terminal. The system provides the Army a secure data and voice communications capability from the MILSTAR satellites. The system will give small combat units increased range for command and control communications.

Navy

Antisubmarine Warfare Combat System (AN/SQQ-89). The antisubmarine warfare combat system combines improved sensors, weapon control systems, and advanced acoustics data processing and display. The system provides long-range detection, tracking, localization, and correlation of subsurface and surface contacts.

Fixed Distributed System. The antisubmarine warfare surveillance system is designed to detect, classify, localize, and track submarines. The surveillance system employs seabed acoustic sensors distributed over large ocean areas.

MK-48 Advanced Capabilities Torpedo. The submarine-launched torpedo is designed for long- and short-range engagement of both submarines and surface targets. The torpedo was improved to include digital guidance and control systems and speed, depth, and range capabilities.

T-45 Training System (T-45TS). The aircraft is intended to provide intermediate and advanced student jet flight training. The trainer is a tandem-seat, lightweight, single-engine aircraft modified for carrier operations. The trainer includes a heads-up display and a weapons delivery capability for training.

V-22 Osprey. The V-22 Osprey is a tilt-rotor, vertical takeoff and landing aircraft. The tilt-rotor design combines the vertical flight capabilities of a helicopter with the speed and range of a turboprop airplane. The aircraft is being designed to meet the amphibious and vertical assault needs of the Marine Corps and the long-range extraction capabilities and to supplement U.S. Special Operations Forces needs of the Air Force. The V-22 is a joint Navy (lead), Air Force, and Marine Corps program.

Air Force

AC-130U Aircraft. The aircraft is designed to provide greater firepower for close air support and air interdiction missions for the U. S. Special Operations Forces. The AC-130U is a new production AC-130 aircraft modified to a sidefiring gunship configuration. The aircraft is integrated with a sophisticated fire control system with 105 millimeter, 40 millimeter, and 25 millimeter guns; armament; sensors; and electronic warfare systems. The AC-130U replaces the AC-130A aircraft.

F-16D Block 50. The Block 50 is an upgrade to the F-16 aircraft that includes integration of separately developed subsystems for more capabilities. The upgrade includes the missile carriage, targeting, electronic warfare, propulsion, and avionics subsystems. The missile carriage and targeting subsystems include the High Speed Antiradiation Missile and its targeting system.

Joint Surveillance Target Attack Radar System. The JSTARS consists of an Air Force E-8 aircraft, an Army ground station, and a data link between the two. The aircraft is host for the side-looking, electronically scanned, multi-mode radar; operations and control; communications; and self-defense subsystems. The airborne radar system is designed to provide wide area, near-real-time surveillance and targeting information on moving and stationary ground targets, slow-moving rotary and fixed-wing aircraft, and rotating antennas.

Military Strategic and Tactical Relay. The satellite system provides secure global communications. The system consists of a constellation of satellites for worldwide communications; multi-Service ground, airborne, and shipborne terminals for connectivity; and a mission control center with constellation control. The MILSTAR is an Air Force-led, joint Services program.

Sensor Fuzed Weapon. The weapon is a 1,000 pound class cluster bomb unit, designed to provide multiple kills per pass against armored and support vehicle combat formations. The bomb consists of 10 submunitions that contain four infrared heat-seeking projectiles each, for a total of 40 projectiles. Each projectile can independently detect targets and fire a high-explosive warhead that is designed to penetrate armored targets.

Appendix E. Test and Evaluation Master Plan Approval Time

Table E-1. OSD Comparison of the Military Departments' Test and Evaluation Master Plan Approval Times

	Average Number of Days After Program Manager Signature for Signature of:				Average Number of Days After OSD Receipt for Signature of:	
Service	PEO	OTA	USER	SAE	DTSE&E	DOT&E
Army (18)	39	107	97	159	40	44
Navy (28)	58	37	90	139	21	24
Air Force (12)	87	43	63	140	37	40
Overall (58)	**58**	**58**	**86**	**145**	**30**	**33**

Note: The sample is based on 58 TEMPs approved by OSD with program manager signature dates from December 1992 through November 1994.

PEO Program Executive Office
SAE Service Acquisition Executive

Table E-2. Inspector General, DoD, Comparison of the Military Departments' Test and Evaluation Master Plan Approval Times

	Average Number of Days After Program Manager Signature for Signature of:				Average Number of Days After OSD Receipt for Signature of:	
Service	PEO	OTA	USER	SAE	DTSE&E	DOT&E
Army (4)	5	72	29	117	64	71
Navy (6)	36	21	85	118	57	67
Air Force (4)	54	-21*	-24*	133	36	42
Overall (14)	**32**	**23**	**38**	**112**	**55**	**63**

Note: The sample is based on 14 TEMPs approved by OSD with program manager signature dates from June 1988 through December 1994.

*The negative anomaly occurs because both the user and OTA officials approved the MILSTAR terminal TEMP during a TPWG meeting and that date was used for the final TEMP. The user and the OTA approved the TEMP 99 days and 119 days, respectively, before the program manager.

Appendix F. Documents Reviewed, Approved, and Published by Director, Operational Test and Evaluation

Documents	Review	Approve	Produce	Assess/ Comment
Annual Report			X	
Beyond Low-Rate Initial Production Report			X	
Central Test and Evaluation Improvement Program	X			X
Congressional data sheets	X			X
Contract award reports	X			X
Cost and Operational Effectiveness Analysis	X			
Crossbow Program	X			X
Defense Acquisition Blue Book Input			X	
Defense Acquisition Executive Summary	X			X
Determination for number of articles of low-rate initial production needed for initial operational testing			X	
DOT&E budget			X	
Foreign Comparative Test Program	X			X
General Accounting Office audit reports	X			X
Independent test concepts			X	
Inspector General, DoD, audit reports	X			X

Documents Reviewed, Approved, and Published by Director, Operational Test and Evaluation (Cont'd)

Documents	Review	Approve	Produce	Assess/ Comment
Live Fire Test Plan	X	X		
Live Fire Test Report			X	
Mission Needs Statements	X			
Office automation programs		X		
Operational Requirements Document	X			
Operational test events				X
Operational test plan	X	X		
Operational test report	X			X
Selected Acquisition Reports	X			X
System Threat Assessment Report	X			
Test and Evaluation Master Plan	X	X		
Test Failure Review Board				X

Appendix G. Approved Performance Characteristics

Combat Service Support Control System

Performance Characteristics	Development Estimate	Approved Program Objective	Approved Program Threshold	Demonstrated Performance*	Current Estimate
Operational Temperature (degrees Fahrenheit)	0±120	0±120	+40±95	TBD	+40±95
Relative Humidity (percent)	10-80	10-80	10-80	TBD	10-80
Portability (number person carry)	2	1	2	TBD	2
Equipment Set-up/ Tear-down (hours)	<=0.5	<=0.5	<=0.5	TBD	<=.5
MTBOMF (hours) ACCS Hardware	220	220	220	TBD	220
ACCS CHS Hardware/Software	210	140	140	TBD	140
Automatic Message Handling User Responsiveness Display 24 Lines (seconds)	1.0	0.7	5.0	TBD	1.0
Scroll (line/seconds)	20	28	20	TBD	20
Error Feedback (seconds)	1.0	0.7	1.0	TBD	1.0
User Help Request (seconds)	3.0	2.1	3.0	TBD	3.0
Auto-message handling Speed-in (seconds)	10/500	7/500	10/500	TBD	10/500
Speed-out (seconds)	10/1,000	7/1,000	10/1,000	TBD	10/1,000
Message Transmit and Receipt 24 hour USMTF Transmission	334	477	334	TBD	334
24 hour Receipt and Processing (million characters)	6.9	9.86	6.9	TBD	6.9
(STAMIS messages)	4,400	6,286	4,400	TBD	4,400

Combat Service Support Control System (Cont'd)

Performance Characteristics	Development Estimate	Approved Program Objective / Threshold		Demonstrated Performance	Current Estimate
Capable of Update (every x hours)	3	2	3	TBD	3
Process All Information Received (within x hours)	3	2	3	TBD	3
On-Line Query Response Time (seconds/minutes)	5/180	5/7	2/3	TBD	5-180 seconds
Local Data File Update Response Time (seconds/minutes) (seconds)	5/180	5/7	5/15	TBD	5-180 seconds

*Demonstrated performance characteristics of TBD will be changed to reflect actual test results after the release of the test reports from the Combat Service Support Control System IOT&E.

ACCS	Army Command and Control System
CHS	Common Hardware and Software
MTBOMF	Mean Time Between Operational Maintenance Failures
STAMIS	Standard Army Management Information System
TBD	To Be Demonstrated
USMTF	United States Message Test Format

Appendix H. Program Performance Summary

An example of the Defense Acquisition Executive Summary, Section 2, Assessments, is below. As discussed in Finding E, a similar program summary could be in the SAR.

Assessments		
(Program Name)	Report Date:	Class:
Program Assessment Indicators		Assessment*
Performance Characteristics		G
Test and Evaluation		Y
Logistics Requirements and Readiness Objective		R
Cost		G
Funding		G
Schedule		G
Contracts		Y
Production		R
Management Structure		Y

*The summary uses a color-coded rating system to indicate the assessed status of the program.

Green - all aspects of the program are progressing satisfactorily.

Yellow - a potential or actual problem has occurred that impairs progress against one or more segments of the program.

Red - a major problem has occurred that seriously impedes successful accomplishment of one or more major objectives.

Advisory (Green or Yellow) - the program is generally progressing satisfactorily, but some problem has occurred or is anticipated that will require additional work.

Appendix I. Management Comments on the Findings and Audit Responses

Finding B

Under Secretary of Defense for Acquisition and Technology Comments. The Under Secretary expressed confusion regarding the discussion of the AC-130U aircraft and our recommended correction. He also questioned whether the Template was used in the case of the AC-130U program and, if so, why we would recommend DoD-wide use of the Template if it did not work for the AC-130U. Additionally, he expressed confusion about our limiting Recommendation B.2. to just the Military Departments.

Audit Response. The Air Force stated that it used an early version of the templates in the certification process for the AC-130U and in their development of that process. The templates did identify the deficiencies. However, the Air Force tested the AC-130U despite known deficiencies because the program office ran out of money to fix the deficiencies. Further, the Air Force stated that the current implementation of the process has provided them with "excellent results in OT&E." We limited our recommendation to the Military Departments because our audit was limited to programs managed by the Army, the Navy, and the Air Force. However, if the Under Secretary desires to include the Defense Agencies in the requirement, we would certainly support that decision.

DOT&E Comments. The Director stated that the finding did not recognize the extent to which the Services use decisionmaking models and cited the Army's formal readiness reviews chaired by the Commander of Operational Test and Evaluation Command as an example.

Audit Response. We agree that formal readiness reviews are required to certify systems as ready for operational testing. The Navy and the Air Force also do formal reviews, although in those Services the commander of the operational test agency does not chair the reviews. The readiness reviews are a process for certifying readiness rather than a decisionmaking model that has specific, known, criteria or analytical procedures to decide whether to start testing with significant limitations.

Army Comments. The Army partially agreed that it was delivering systems for operational tests that were not ready. However, it disagreed that it lacks criteria to help certification officials and operational testers decide whether to delay tests. It stated that Army Regulation 73-1, "Test and

Evaluation Policy," February 27, 1995, paragraph 3-12, "Waivers of Approved Testing," and paragraph 3-13, "Delay, Suspension, or Termination of Testing," provide such criteria.

Audit Response. Paragraph 3-12 does not provide standards for judging whether to delay testing. We commend the Army for the latest change in paragraph 3-13. The prior edition of the regulation contained general rules for suspending tests. The latest edition extends those rules to the delay of testing. Our evaluation of the regulation deduced four broad rules that state testing can be delayed or suspended: if the problems with the system will affect data validity; if there is little chance of attaining critical technical parameters; if there is little chance of satisfying critical operational criteria; and if significant safety problems exist. Those rules provide a start for a decision model; explicit criteria must be added. For example, comparing the costs of delaying the test to retesting or additional testing should be added.

Air Force Comments. The Air Force disagreed that there was no "formalized criteria to help certification officials and operational testers decide whether to delay tests because of system limitations." The Air Force stated it had a well-defined process for assessing the system readiness to begin dedicated operational test and evaluation. It stated its process does not levy bureaucratic delays or burdens on the program, but does require exit criteria and informed judgment by decisionmakers before systems go to the next phase of testing. The Air Force operational test agency also addressed the process aspect of the Air Force certification templates: "it is not possible to develop individual, specific criteria for all programs."

Audit Response. We agree that the Air Force template system is a well-defined process, albeit one that needs to be strengthened by adding criteria. Our draft report stated that the Air Force model has significant value in helping decisionmakers determine readiness for testing and favorably cited the deficiency resolution template from Air Force Manual 63-119, "Certification of System Readiness for Dedicated Operational Test and Evaluation," February 1995. Those decisionmaking templates should include specific decision criteria in addition to broad rules. For example, the System Design and Performance Deficiency Resolution Template states that "dedicated OT&E can be completed as planned and results will not be invalidated due to deferred deficiencies." This criterion is similar to the Army's rules cited above.

Finding C

Revision to the Finding. As a result of comments received from the DOT&E and the Air Force, we deleted the statement concerning the risk for test of the F-16 missile targeting system. The DOT&E stated that past experience indicates that climate is not a significant performance factor. The Air Force stated that a lack of testing in a particular environment cannot introduce wither a new risk or a high risk since all systems have some inherent amount of risk.

DOT&E Comments. The Director partially concurred with the finding. He strongly agreed with the need for adequate threat systems. He disagreed that test documents were not providing essential information for assessing acquisition risk and for decisionmaking; that is, the impact of test limitations is not adequately reported. He stated that in most cases a statement of the limitation is sufficient to understand its impact.

Audit Response. A statement of the limitation is not sufficient to understand the impact. For example, the F-16 Follow-on OT&E plan, August 1994, stated that "A complete array of threat surface and air weapon systems will not be available for this evaluation." The plan did not state the impact of those limitations. DoDI 5000.2 required information on the impact of limitations because the information is essential for assessing test risk. Further, the DoDI did not say the statement alone is sufficient. Additionally, DoDI 5000.2 required a discussion of limitations in the TEMP and the impact of the limitations on the ability to resolve critical operational issues and the ability to formulate conclusions regarding operational effectiveness and operational suitability. The March 1996 DoD 5000.2-R has not changed these requirements.

Army Comments. The Army stated that long-standing resource shortfalls cause most test limitations and that the Army decides its priorities and executes them within available resources. It stated it is complying with the requirement to analyze and report on the impact of limitations.

Air Force Comments. The Air Force also stated that the test limitations were due to long-standing resource shortfalls and that it routinely analyzes and reports on all significant limitations. It recommended that the statement concerning testing the F-16 missile targeting system in the desert as "introducing a high risk on its effectiveness in other climates" should be changed to read "fails to reduce uncertainty for decision makers about system effectiveness in non-desert climates."

Audit Response. While we agree many test limitations are the result of long-standing resource shortfalls, this shortfall is what needs to be corrected. We do not agree that the Army and the Air Force are adequately complying with the requirement to report the impact of limitations on test results. As a result of the Air Force comments, we revised the statement concerning risk for test of the F-16 missile targeting system.

Finding D

DOT&E Comments. The Director stated that, in general, he agreed with the audit results. However, he cautioned that his annual report is not designed to be a decisionmaking tool for individual weapon systems. Also, he stated that the law does not explicitly require the annual report to contain the same level of analysis on every OT&E activity conducted within the DoD. He stated that his assessment of the AN/SQQ-89(V)(6) supported the Navy OTA fundamental

conclusions about the system's operational effectiveness and suitability. Additionally, he stated that our example of the 1994 Annual Report assessment of Combat Service Support Control System contained a partial factual error.

Audit Response. Although the annual report was not designed to be a decisionmaking tool for individual weapon systems, it is an independent source of information available to decisionmakers. It is submitted to the Secretary of Defense, the Under Secretary of Defense for Acquisition and Technology, and Congress.

AN/SQQ-89(V)6. We agree that the Director concluded different results than the Navy for several areas of the Navy's OTA report on the AN/SQQ-89(V)6. The Director's quoting the Navy OTA report did not cause the ambiguity concerning the AN/SQQ-89(V)6 section of the annual report; the cause was the lack of his explicit conclusions about the operational effectiveness and suitability of the system.

Combat Service Support Control System. We disagree that the Combat Service Support Control System statements in the report contained a partial factual error. We agree that the cut-off date for the data for the annual report was September 30, 1994, and that it was published in February 1995. However, we maintain that the Director should not have provided unsubstantiated assurance, such as, "We are aware of no significant issues which would prevent the 1995 Milestone III decision," before the completion of the analysis.

Air Force Comments. Although not required to comment, the Air Force stated that the audit report "overlooked" 10 U.S.C. 2399, paragraph (b)(2), that requires the Director to analyze the results of the OT&E conducted for each major Defense acquisition program. The Air Force stated that this reference is the Title 10 definition of what the Director's annual report should contain.

Audit Response. We disagree that the content of the Director's annual report is defined by 10 U.S.C. 2399, paragraph (b)(2). This paragraph discusses the Beyond Low-Rate Initial Production Report (also issued by the Director), not his annual report. The law requires the Director to summarize all operational test and evaluation in the DoD and does not limit the reporting requirement to major Defense acquisition programs.

Finding E

Under Secretary of Defense for Acquisition and Technology Comments. The Under Secretary stated that, although the DoD Manual 5000.2-M requires reporting the test results starting with advanced development testing, the program managers are reluctant to report test results until after a full evaluation. Additionally, the evaluations may not be completed in time to be in a SAR.

Army GSM. The Under Secretary stated that although the Army GSM Force Development Test was done in September 1994, the evaluations of the test were not available in time to be included in the December 1994 SAR.

Sensor Fuzed Weapon. The Under Secretary disagreed with our statement that the Sensor Fuzed Weapon had not met operational effectiveness and suitability requirements.

Audit Response. We agree that program managers are reluctant to report test results that have not been fully evaluated and that the evaluation is not complete until the issuance of a report from the OTA.

Army GSM. We agree that the test assessment for the Army GSM was issued by the Operational Test and Evaluation Command, the Army's OTA, on March 7, 1995. However, the testing was completed in September 1994, well before the end of the reporting period. The advance submissions of the SARs were to be provided by March 1, 1995. Additionally, test reports from both the Army Material Systems Analysis Activity and the Test and Experimentation Command, a command subordinate to the OTA, were both issued in January 1995, well before the advance submission date. Therefore, this information was available to the program office in time to be incorporated into the SAR. The comments in Section 7 of the SAR and the test information in Section 10 were not reasonable in light of the information the OTA released within the deadline for advance submissions.

Sensor Fuzed Weapon. We contend that the weapon has not demonstrated operational effectiveness and suitability for all mission requirements. The system's requirement is for an average number of mobility kills per pass across the delivery envelope (200 feet through 20,000 feet). However, the system requires enhancements to improve its performance to meet this requirement. Additionally, the second phase of operational testing is limited to low altitudes due to its performance degradations for medium and high altitudes.

Appendix J. Summary of Potential Benefits Resulting From Audit

Recommendation Reference	Description of Benefit	Type of Benefit
A.1., A.2., A.3.	Economy and Efficiency. Reduces the time spent in test planning.	Nonmonetary.
B.1., B.2.	Management Controls. Avoids the expenditure of resources for testing weapon systems prematurely.	Nonmonetary.
C.	Program Results. Avoids fielding ineffective weapon systems.	Nonmonetary.
C., D.2.d	Program Results. Provides the additional information necessary for adequate evaluation.	Nonmonetary.
D.1., D.2.a.	Program Results. Provides identification of additional resources necessary for the performance of the mission.	Nonmonetary.
D.2.b., D.2.c.	Management Controls. Assists staff in the successful completion of their functions.	Nonmonetary.
E.1., E.2.	Economy and Efficiency. Provides unbiased information for efficient allocation of resources.	Nonmonetary.

Appendix K. Organizations Visited or Contacted

Office of the Secretary of Defense

Office of the Under Secretary of Defense for Acquisition and Technology, Washington, DC
- Director, Acquisition Program Integration, Washington, DC
- Director, Tactical Systems, Washington, DC
- Director, Test, Systems Engineering, and Evaluation, Washington, DC

Director, Operational Test and Evaluation, Washington, DC

Department of the Army

Army Materiel Command, Alexandria, VA
Communications and Electronic Command, Fort Monmouth, NJ
Operational Test and Evaluation Command, Alexandria, VA
- Operational Evaluation Command, Alexandria, VA
- Test and Experimentation Command, Fort Hood, TX

Training and Doctrine Command, Fort Monroe, VA
Program Executive Office, Command and Control Systems, Fort Monmouth, NJ
- Combat Service Support Control System Program Office, Fort Belvoir, VA
- Intelligence Fusion Program Office, McLean, VA
- Military Strategic and Tactical Relay Program Office, Fort Monmouth, NJ

Test and Evaluation Management Agency, Washington, DC
Army Audit Agency, Alexandria, VA

Department of the Navy

Office of the Chief of Naval Operations, Arlington, VA
- Operational Test and Evaluation Force, Norfolk, VA
- Navy Air Systems Command, Arlington, VA
  - Naval Undergraduate Flight Training Systems Program Office, Arlington, VA
  - V-22 Osprey Program Office, Arlington, VA
- Navy Sea Systems Command, Arlington, VA
  - Program Executive Office Undersea Warfare, Arlington, VA
  - MK-48 Advanced Capabilities Torpedo Program Office, Arlington, VA
- Space and Naval Warfare Systems Command, Arlington, VA
  - Fixed Distributed System Program Office, Arlington, VA
  - Military Strategic and Tactical Relay Program Office, Arlington, VA

Naval Audit Service, Arlington, VA

Department of the Air Force

Office of the Assistant Secretary of the Air Force (Acquisition), Washington, DC
- Office of the Deputy Assistant Secretary of the Air Force (Management Policy and Program Integration), Washington, DC
- Director, Fighter, C2, and Weapons Programs, Washington, DC
- Director, Long Range Power Projection, Special Operation Forces, Airlift, and Training Programs, Washington, DC
- Program Executive Office Organization, Washington, DC

Office of the Assistant Secretary of the Air Force (Financial Management and Comptroller), Washington, DC
Air Force Test and Evaluation Directorate, Washington, DC
Air Combat Command, Langley Air Force Base, VA
Air Force Materiel Command, Aeronautical Systems Center, Wright-Patterson Air Force Base, OH
- AC-130U Gunship System Program Office, Wright-Patterson Air Force Base, OH
- F-16 Program Office, Wright-Patterson Air Force Base, OH
- Sensor Fuzed Weapon Program Office, Eglin Air Force Base, FL

Air Force Space Command, Los Angeles Air Force Base, CA
- Air Force Space and Missile Center, Los Angeles Air Force Base, CA
- 4th Space Operations Squadron, Falcon Air Force Base, CO

Air Force Electronic Systems Center, Hanscom Air Force Base, MA
- Joint Surveillance Attack Radar System Program Office, Hanscom Air Force Base, MA
- Military Strategic and Tactical Relay Program Office, Hanscom Air Force Base, MA

Air Force Operational Test and Evaluation Center, Kirtland Air Force Base, NM
- Air Force Test and Evaluation Center, Detachment 2, Eglin Air Force Base, FL
- Air Force Test and Evaluation Center, Detachment 4, Peterson Air Force Base, CO

Other Government Organizations

Ballistic Missile Defense Organization, Washington, DC
U.S. Special Operations Command, McDill Air Force Base, FL
- Air Force Special Operations Command (18th Test Squadron), Hurlburt Field, FL
- 4th Special Operations Squadron, Hurlburt Field, FL

Defense Logistics Agency, Alexandria, VA
Defense Plant Representative Office, Textron Defense Systems, Wilmington, MA

Appendix L. Report Distribution

Office of the Secretary of Defense

Under Secretary of Defense for Acquisition and Technology
  Director, Acquisition Program Integration
  Director, Tests, Systems Engineering and Evaluation
  Director, Defense Logistics Studies Information Exchange
Under Secretary of Defense (Comptroller)
  Deputy Chief Financial Officer
  Deputy Comptroller (Program/Budget)
Deputy Under Secretary of Defense (Acquisition Reform)
Director, Operational Test and Evaluation
Assistant to the Secretary of Defense (Public Affairs)
Director, Washington Headquarters Service

Department of the Army

Assistant Secretary of the Army (Financial Management and Comptroller)
Assistant Secretary of the Army (Research, Development and Acquisition)
Auditor General, Department of the Army
Director, Test and Evaluation Management Agency

Department of the Navy

Assistant Secretary of the Navy (Financial Management and Comptroller)
Assistant Secretary of the Navy (Research, Development and Acquisition)
Auditor General, Department of the Navy
Director, Test and Evaluation and Technology Requirements

Department of the Air Force

Assistant Secretary of the Air Force (Acquisition)
Assistant Secretary of the Air Force (Financial Management and Comptroller)
Auditor General, Department of the Air Force
Director, Test and Evaluation Directorate

Other Defense Organizations

Ballistic Missile Defense Organization
Director, Defense Contract Audit Agency
Director, Defense Logistics Agency
Director, National Security Agency
Inspector General, National Security Agency

Non-Defense Organizations

Office of Management and Budget
Technical Information Center, National Security and International Affairs Division, General Accounting Office
Chairman and ranking minority member of each of the following congressional committees and subcommittees:
Senate Committee on Appropriations
Senate Subcommittee on Defense, Committee on Appropriations
Senate Committee on Armed Services
Senate Committee on Governmental Affairs
House Committee on Appropriations
House Subcommittee on National Security, Committee on Appropriations
House Committee on Government Reform and Oversight
House Subcommittee on National Security, International Affairs, and Criminal Justice, Committee on Government Reform and Oversight
House Committee on National Security

Part III - Management Comments

Under Secretary of Defense for Acquisition and Technology Comments

OFFICE OF THE UNDER SECRETARY OF DEFENSE

3000 DEFENSE PENTAGON
WASHINGTON DC 20301-3000

ACQUISITION AND TECHNOLOGY

29 JAN 1996

MEMORANDUM FOR DIRECTOR, ACQUISITION MANAGEMENT DIRECTORATE, INSPECTOR GENERAL

SUBJECT: DODIG Draft Report on "Operational Testing Performance on Weapon Systems" Report No. 5AB-0031

I am responding for the USD(A&T) on your draft report of November 24, 1995. Recommendation B-2 on page 19 recommends that DoDI 5000.2 be revised "to require the Army, the Navy, and the Air Force to develop a decision-making model to evaluate whether to proceed with testing when serious limitations are identified before operational testing and to determine the cost/benefit of testing with limitation."

Generally, I resist changing policy and procedures, until I fully understand that there is a real widespread, and recurring problem to be fixed. Your draft mentions only four programs so I am not persuaded that there is a pervasive problem. I am also confused by your finding on the Air Force AC-103U program which states that its testing proceeded despite significant limitations and notes an inability to test some of the subsystems and premature testing of others. Yet your fix to this problem (p-19) is to direct the Services use a decision model "such as the Air Force System Design and Performance Deficiency Resolution Template". Did the AC-103U use that template? It must have as it was the only Air Force program you report on. Why would I recommend use of a template DoD-wide if it didn't work for the Air Force? I am also confused why you would limit your recommendation to the three Services.

As you know we are finalizing the new DoDD 5000.1 and DoDI 5000.2 at this time but that effort is generally a streamlining effort and an attempt to reduce the mandatory burden on the PM. Your recommendation would add to the mandatory burden. I remain unconvinced that a decision model needs to be added to DoD policy regarding testing when limitations are identified. If you can make a stronger case for the Air Force decision model we can incorporate it in the electronic "acquisition deskbook" at a later time. The 5000 working group has recommended the following addition to the policy for "Certification of readiness for Operational Test and Evaluation" (part 3.4.3 of DoDI 5000.2):

The developing agency shall also provide software maturity criteria necessary for certification of operational test. Risk management

metrics, measures, indicators, and associated thresholds shall include cost, schedule, requirements traceability, and fault profile. A mission impact analysis of unmet metrics shall be completed before certification for operational tests.

The policy still requires the developing agency formally certify that the system is ready for the next OT phase and the additional wording captures most of the intent of your recommendation B-2.

In general, mandating specific models are anti-reform and don't have widespread support. We prefer the alternative of IPTs and relevant discussions among the acquisition professionals rather than cookbook solutions.

I. N. Blickstein
Director
Acquisition Program
Integration

cc: USD(A&T)
PDUSD(A&T)
D, OT&E
D, TSE&E
DD, API/PM

OFFICE OF THE UNDER SECRETARY OF DEFENSE

3000 DEFENSE PENTAGON
WASHINGTON DC 20301-3000

ACQUISITION AND TECHNOLOGY

05 FEB 1996

MEMORANDUM FOR DoD INSPECTOR GENERAL

SUBJECT: Draft DoDIG Report, "Operational Testing Performed on Weapon Systems"

Thank you for the opportunity to comment on the subject draft report. The following comments pertain to Finding E, "Reporting Program Results."

DoD Selected Acquisition Report (SAR) policy (DoD 5000.2-M (Part 17)) requires the reporting of Demonstrated Performance information starting with the results of advanced development testing. However, there has been a natural reluctance on the part of the Program Manager to delay the reporting of test results until after a careful evaluation has been completed by the government testing activities. For example, the Force Development Test for JSTARS GSM was completed in September 1994. The raw results of this test could easily have been reported in the December 1994 SAR. However, evaluations of that test were not completed by the Army's Operational Test Agency and the DOT&E until March 1995 and May 1995, respectively, and would not have been available in time to be included in the December 1994 SAR. These evaluations are important to ensure the validity of the test. Raw test data can be deceptive at times. Nevertheless, we will reiterate the DoD policy of accurate and timely reporting of Demonstrated Performance information in the upcoming release of December 1995 SAR guidance.

Regarding your comment that "Sensor Fused Weapon has not demonstrated operational effectiveness and suitability for all mission requirements...." (pp 39), a series of fact-finding discussions and a mediation meeting were held including representatives of the Air Force, OUSD(A&T), and the OAIG-AUD. Agreement was made that the referenced SFW 1994 SAR was consistent with System Operational Requirements Document (SORD) requirements in that there was no explicit requirement for mobility and firepower kills at medium and high altitude deliveries. Rather, the SORD requirement is for an average number of mobility kills per pass. Agreement was also made that the SFW APB would be updated and the 1995 SFW SAR would be consistent with the SORD and APB. Accordingly, the report is not correct regarding SFW not meeting operational effectiveness and suitability requirements.

With respect to your findings (E.1 and E.2), the SAR is fully coordinated with the OSD staff and all appropriate comments are incorporated prior to forwarding the report to the Congress. Therefore, the SAR represents a coordinated DoD position on the current status of major acquisition programs. In accordance with direction from the USD(A&T), a DoD working group has been established to develop and implement SAR reform. One of the proposals under consideration would replace the present SAR Program Highlights with an Executive Summary (similar to the DAES) This and other streamlining/improvements are planned to be incorporated in the December 1996 SAR.

Sam S. Smith

FOR I. N. Blickstein
Director, Acquisition
Program Integration

Director, Test, Systems Engineering, and Evaluation, Comments

OFFICE OF THE UNDER SECRETARY OF DEFENSE

3000 DEFENSE PENTAGON
WASHINGTON DC 20301-3000

ACQUISITION AND TECHNOLOGY

22 JAN 1996

MEMORANDUM FOR DIRECTOR, ACQUISITION MANAGEMENT DIRECTORATE
OFFICE OF THE INSPECTOR GENERAL

THROUGH: DIRECTOR, CONGRESSIONAL ACTIONS & INTERNAL REPORTS 23JAN 1996

SUBJECT: Audit Report on Operational Testing Performed on Weapon Systems (Project No. 5AB-0031)

This office has reviewed the subject report and provides the following comments to the recommendations and areas cited below:

Recommendation for Corrective Action (A.3, page 12): "We recommend that the Director Test, Systems Engineering and Evaluation, include himself as a formal member of the program office integrated product team to expedite his review of test documents." Concur with modification.

Well before the advent of integrated product teams, this office has encouraged active participation in program office Test and Evaluation Working Groups (TEWG). The TEWG forum generally includes the aspects of test planning, test integration, test conduct, and test reporting. It is acknowledged that in the past some program offices did not desire OSD participation in their TEWGs which often resulted in Test and Evaluation Master Plans (TEMP) being reviewed for the first time at presentation for OSD approval with the attendant problems that mostly could have been solved in advance. The current USD(A&T) emphasis on integrated product teams has resulted in more open T&E forums and a spirit of cooperation and teamwork. Additionally, we participate in Systems Engineering integrated product teams in order to apply systems engineering discipline into development and testing. Through these forums, we are participating in the drafting and early review of TEMPs, test integration, systems integration, test conduct, and test results. This office intends to continue this active role through formal membership in integrated product teams addressing testing, systems engineering, and requirements generation. DTSE&E action officers are empowered in these integrated product teams to make decisions and resolve issues.

Final Report Reference

Revised Recommendation A.3

The recommendation should be modified to read that the DTSE&E should be a formal, non-voting member of the program office integrated product teams. The integrated product team prepares a TEMP for the Program Manager's signature. Frequently, there are issues that cannot be resolved by the Program Manager and must be taken up at a higher level. Although DTSE&E action officers must participate fully in these integrated product teams to attempt to resolve these issues in a timely manner, DTSE&E must preserve its ability to approve or disapprove TEMPs when necessary.

Recommendation for Corrective Action (B.1, page 19): "We recommend the Director Test, Systems Engineering and Evaluation, review and approve certification of readiness for operational testing for acquisition category I and designated oversight programs." Non concur.

Inserting DTSE&E into the chain of command would compromise this office's objectivity and oversight responsibility. It is the responsibility of the service development agency to adequately test the weapon system in accordance with the OSD approved TEMP and then to certify its readiness to enter operational test and evaluation (OT&E). The Services' Program Executive Officers and Acquisition Executives must provide adequate funding and schedule to conduct the testing in the TEMP. It is also the responsibility of the Service's Operational Test Agency to accept the system into OT&E. The existing management structure is adequate for this purpose, when there is effective accountability.

TEMP Process (page 5): Management Comment.

Historically, the approval of TEMPs has been driven by elements outside of the TEMP development and approval process. Since the TEMP is the comprehensive top level document that provides the roadmap to evaluate system maturity and readiness to advance, it may be delayed until other programmatic issues are resolved. These issues can include such areas as: Definitive operational requirements and/or threat descriptions; Formulation of measurable evaluation criteria (MOEs & MOPs); Early identification of resources necessary and available to perform adequate realistic testing; Program stability (minimized funding issues, technical problems, restructuring, rebaselining, etc.); and Harmonization of key documents (ORD, COEA, STAR, ASR, APB, etc.). Although the T&E community has worked closely together in T&E Working Groups, its ability to finalize a TEMP is frequently hindered by these other issues -- many of which they have no control over. Approval of a TEMP is dependent on approval of the

ORD, COEA, and STAR. Consequently, the TEMP process is unfairly criticized when in fact it is delayed by the inability to resolve other programmatic issues and becomes an indicator of deeper problems. The recent initiation of program level integrated product teams at all levels should facilitate the development of all these documents and thus improve the TEMP approval process.

The DoD IG study selected 14 TEMPs in the June 1988 to December 1994 time frame as its basis for TEMP approval times. This is a significantly smaller and misleading data sample than the 58 TEMPs in the OSD study. Additional data (attached) from 1995 bringing the total to 82 TEMPs approved from January 1994 to January 1996 validates the earlier OSD statistics.

John A. Burt
Director, Test, Systems
Engineering & Evaluation

Attachment

SUMMARY OF TEMP STATISTICS JANUARY 1994 TO JANUARY 1996

	Average no. of days after PM signature for signature of:				Average no. of days after OSD receipt for signature of:	
	PEO	OTA	USER	SAE	DTSE&E	DOT&E
NAVY (41)	51	34	83	127	27	28
AIR FORCE (19)	71	38	62	133	28	32
ARMY (22)	39	122	100	186	42	44
OVERALL (82 TEMPS)	53	54	81	141	30	32

Under Secretary of Defense (Comptroller) Comments

OFFICE OF THE UNDER SECRETARY OF DEFENSE
1100 DEFENSE PENTAGON
WASHINGTON, DC 20301-1100

COMPTROLLER
(Program/Budget)

FEB 16 1996

MEMORANDUM FOR DIRECTOR, ACQUISITION MANAGEMENT DIRECTORATE

SUBJECT: Audit Report on Operational Testing Performed on Weapons Systems
(Project No. 5AB-0031)

I would like to recommend that you revise your Recommendation for Corrective Action D.1. This recommendation would have the Under Secretary of Defense (Comptroller) prepare alternative decision documents for Defense Resource Board (DRB) consideration to provide the Director, Operational Test and Evaluation, the necessary funds for increased DoD civilian staff years.

The Director, Operational Test and Evaluation organization is a Office of the Secretary of Defense staff office. The DRB does not determine either the number of military or civilian personnel or the funding levels to support OSD staff offices. That is a responsibility of the Director, Washington Headquarters Services (WHS), who determines the funding and staffing levels of all OSD staff elements.

Presumably, the Director, WHS has established what is considered to be an appropriate resource and personnel level for the operation of the OTE directorate. It is noted that that the Congress imposes an overall management headquarters ceiling for the OSD staff that the Director, WHS must stay within. Director, OTE staffing levels can not be considered in isolation..

I would suggest that the recommendation be revised to recommend that the Director, OTE, if he believes that the directorate has insufficient resources, request additional resources from the Director, WHS. That request should include a rationale that supports specific funding and civilian personnel increases required, which are not included in the draft audit report.

Ronald G. Garant
Director for Investment

Director, Operational Test and Evaluation, Comments

OFFICE OF THE SECRETARY OF DEFENSE
WASHINGTON, DC 20301-1700

OPERATIONAL TEST
AND EVALUATION

26 FEB 1996

MEMORANDUM FOR THE DEPARTMENT OF DEFENSE,
INSPECTOR GENERAL

Subject: Audit Report on Operational Testing Performed on Weapon Systems (Project No. 5AB-0031)

Thank you for the opportunity to comment on your draft Audit Report on Operational Testing Performed on Weapon Systems (Project No. 5AB-0031). We appreciate the effort that you put into this report. We are committed to the performance of appropriate operational test and evaluation (OT&E) so that we can be certain that our warfighters are being supplied with effective and suitable weapon systems. Your report helps all involved in OT&E to achieve that goal.

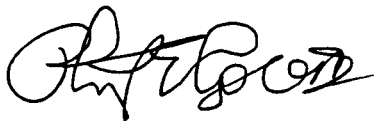
We are providing general comments on the draft audit report in this letter, and more detailed comments are contained in the attachment. In general, we agree with the audit results as described in the Executive Summary. However, we would caution that the Annual Report of the Director of Operational Test and Evaluation (DOT&E) is not designed to be a decision making tool for acquisition decisions on individual weapon systems. Our Annual Report is primarily a report on the results of recent OT&E and LFT&E and the state of this testing within the Department of Defense. Decisions on the future of developing weapon systems are made using other documentation in formalized processes.

In general, we agree with the summary of recommendations. The report does not seem to recognize the extent to which the Integrated Product Team (IPT) concept is used in test planning. All major DoD programs are using IPTs, and DOT&E supports this process, but your report recommendations on page 12 do not seem to recognize this. We fully support the IPT process and find that in general it is quite similar to the Test Working Group and Test Planning Working Group meetings which have been used by the test community for years.

With respect to operational test readiness, the report could indicate that by regulation, DoDI 5000.2 requires the PEO to certify readiness of a system for operational testing. The Services each have a process to evaluate whether to proceed with operational testing when serious limitations are identified. For example, in the Army, the Commander OPTEC holds a formal review to determine the readiness of a system for operational testing.

The report recommends that adequate resources are needed "to acquire actual threat systems and for approving Test and Evaluation Master Plans (TEMP) test plans and test reports after the impact of limitations are adequately addressed." We strongly agree with this. Resources for realistic threat systems are a special problem and we have reported on this regularly as noted in your report. We will continue to work with the Services to improve the quality of threat systems used in operational testing.

The Report recommends that adequate resources are needed "to report on all operational testing and to provide complete and consistent information to decision making." We agree with this. The report states "DOT&E lacked resources to monitor, review, and report on all operational testing performances in the DoD, as required by title 10 USC 139." This is true. Given the pressure on financial and personnel resources it is impractical that we will be able to report on all operational testing within the DoD. We will work with the Director of Administration and Management, and with the Military Services, to find the correct balance.



Philip E. Coyle
Director

Attachment

Detailed Comments on the IG Report

Page 5 (Test and Evaluation Planning)

Please add DOT&E to the end of the sentence at the top of the page so that it reads "... is to be approved 45 days of its submission to DTSE&E and DOT&E." This change is required to reflect that TEMPS are submitted for approval to DTSE&E and DOT&E jointly.

Page 7 (TEMP Development)

In discussing the Office of the DOT&E participation in MILSTAR, the audit makes the assumption that attendance at TPWGs is the measure of office involvement and impact on the OT&E of a program. This is not the case. During the period in question, representatives of this office participated in over 27 test planning meetings or actual test events on the Milstar program. During these planning meetings, the approach to the OT&E was discussed with the OTA and issues resolved. The OTA's representative was thus empowered to present the entire OT&E community's position at the TPWGs and DOT&E personnel only needed to attend meetings where unresolved issues remained.

Page 8 (Empowerment)

The office of OT&E emphasizes the principle of empowerment and our action officers are instructed that they fully represent the office in IPT, TPWG and other test meetings. The grade and experience levels of IPT participants varies widely. The amount of empowerment that a Service agency is willing to give correlates with the grade and experience level of the participants. Because not all agencies have a sufficient quantity of experienced personnel at an appropriate grade level at all times, full empowerment will be difficult to achieve. Additionally, final system approval must be at the agency-head level. As stated elsewhere in this document, agency heads may have knowledge of mitigating circumstances that would change an otherwise apparently rational decision.

Page 8 (Hierarchical Approval)

One of the methods for shortening the approval process in the hierarchical situation is to require that the participants keep their commanders informed of the issues involved. In that way, when the TEMP arrives for signature approval, the number of surprise issues is reduced and approval can be granted faster.

Page 9 (Senior Management Delays)

Recommend that the word "hostage" be removed from this paragraph. The concept of not approving a TEMP until it is adequate to properly evaluate the system effectiveness and suitability, fulfills one of the DOT&E missions. However, the term "hostage" and its negative implications are not appropriate in the context of this paragraph and do not describe the actual situation. Instead of simply disapproving a TEMP and requiring it to be re-coordinated through a long sequential process, DOT&E simplifies and shortens the process by stating changes, which, in its view, are

required for TEMP approval. Once changes are agreed to, often in a few change pages, the TEMP can be approved without extensive staffing, thus making the process shorter and faster.

Specifically, in the case of Joint STARS, the measure of effectiveness had been removed from the TEMP. Approving it without an agreed to definition of how testing will determine Joint STARS operational effectiveness would almost surely result in a test program that could not properly support acquisition decision making. Although the use of IPTs (properly empowered) is a useful method to facilitate agreement on test adequacy issues, TEMP disapproval should not be restricted when the Service-approved TEMP is determined to be inadequate by DOT&E.

Pages 10 and 11 (Effective Use of Resources)

The report does not seem to recognize the extent to which the IPT concept is used in test planning. All major DoD programs are using IPTs, and often a program will have ten or more concurrent IPTs. One program actually has more than 20 different IPTs. DOT&E endeavors to attend and support all IPTs applicable to testing. IPTs are held in many areas which affect testing including cost, requirements, and schedule, but many IPTs will not involve test issues.

We are encouraging more discipline in the IPT process (for example, the use of agendas known in advance) to make the best use of these meetings. We fully support the IPT process and find that in general it is quite similar to the TIWG and TPWG meetings that have been used by the test community for years.

As a practical matter, DOT&E is not staffed, nor should it be staffed to attend all IPT meetings. Additionally, the last sentence of the last full paragraph on page 11 should be changed to read: "To meet the intent of 10 USC. 139, *the Office of the* DOT&E should be a non-voting member so as to maintain his independence." As stated elsewhere in these comments, the DOT&E is not personally a participant in these meetings. However, he can and does appoint a representative who is empowered to represent the office.

Page 12 (Recommendations for Corrective Action)

We agree that the TEMP development and approval process is inefficient and support the recommendations to streamline the process. However, streamlining will require substantial reengineering by the Services since over 80% of the time is spent there. Also, as noted in the report, DOT&E has reduced its time for TEMP review to roughly 30 days, well within the 45-day limit required by DoD Manual 5000.2M. We have analyzed the review time for the most recent fiscal year and DOT&E TEMP review times average 32 days.

Page 12 (Recommendations for Corrective Action -- Recommendation A.1)

We partially concur with the concept of empowered IPTs, when there is unanimous agreement. However, implementation of this recommendation would preclude the Military Department T&E and Acquisition executives from performing a final check and approval on a TEMP. While action officers can be 'empowered' to speak for the T&E executives, the action officer is not privy to all of the information of the T&E executives and therefore could miss an important concept in his TEMP approval. We recommend a more reasonable approach be taken such as a much shorter time limit for Service executive approvals and tacit approvals by the Service executive if no action is taken within a reasonable time limit.

2

Final Report Reference

Page 12 (Recommendations for Corrective Action -- Recommendations A.2 and A.3)

Revised Recommendation A.2 and A.3

We partially concur with recommendations A.2 and A.3. The Director of OT&E and the Director of TSE&E are far too senior to be members of IPTs. Recommend that the word "himself" be changed to read "a member of his office." This will support the Directors appointing proper representatives for each IPT as they do now.

There are too many IPTs for OSD to be a member of each one. The Rules of the Road for IPTs call for the Overarching IPT to determine what Working IPTs will be formed and who shall participate on them. As a member of the Overarching IPTs, DOT&E Deputy Directors can and do ensure that DOT&E staff are included on the Test IPTs. The issue of voting or not-voting membership will be decided as each Test IPT determines its own operating rules. Voting status is not a significant factor in effective DOT&E participation.

Page 13 (Finding B. Readiness for Testing)

Basically, the determination of whether or not a system is ready for testing is a Service responsibility. By regulation, DoDI 5000.2 requires the developing agency to certify readiness of a system for operational testing. The Director OT&E has been urging earlier involvement by operational testers so as to derive their insights and experience earlier in the acquisition process.

Additionally, this comment does not seem to recognize the extent to which the Services use decision making models to evaluate readiness for operational testing. For example, in the Army, the Commander of OPTEC holds a formal readiness review of each system before proceeding with operational testing. This includes a review of any limitations due to resources, environmental considerations, or other reasons.

Page 17 (Schedule Requirements and Decisionmaking Criteria)

The report states that "... the TEMPs and test plans do not explain why testing begins on systems that are not ready." This is because the purpose of TEMPs and test plans is to provide guidance for testing a system to determine its status, not to explain the status of the system before the test. Also, of necessity, test plans are crafted relatively early in a new system's development schedule.

Page 19 (Recommendations for Corrective Action)

Recommendations B.1 and B.2 could take us back to a narrow pass / fail approach to testing. If and when systems fail, we will not hesitate to report it, but we are concerned that an overly formalized readiness review could thwart the progress we are making in involving operational testing early in the acquisition process. The operational test agencies participate in the test readiness reviews and are often a source of test readiness concerns.

We non-concur with recommendation B.1. With respect to operational test readiness, the report could indicate that by regulation, DoDI 5000.2 requires the PEO to certify readiness of a system for operational testing. The Services each have a process to evaluate whether to proceed with operational testing when serious limitations are identified. For example, in the Army the Comander OPTEC holds a formal review to determine the readiness of a system for operational testing. The IG recommendations should focus on better implementing current regulations. If it

3

becomes necessary to make significant change, DOT&E, with our early involvement, could provide the desired review.

We partially concur with recommendation B.2. The policy to use a decision making process is in place. However, it should be better implemented. The development of a decision-making model which takes into account test limitations and cost-benefit analysis may not be practical. There are too many factors that go into such a decision and trying to develop such a model may be counterproductive.

Page 22 (DOT&E.)

The draft report stated that for the T-45TS, the B-LRIP report did not discuss the impact of the constraints on test conclusions or on the risk of wrongly concluding that the system was effective and suitable. That is true. The reason that these constraints were not discussed was that, in the instant case, there were no effects on test conclusions and that there was insignificant risk of wrongly concluding that T-45TS was effective and suitable.

Page 23 (Resources and Reporting -- Third paragraph)

This paragraph appears to be a mixture of apples and oranges. Section C of the Draft Audit Report is entitled "Test Threat and Environment." However, the CTEIP program, which DOT&E enthusiastically supports, does not purchase foreign materiel. Additionally, T&E resource investment encompasses much more that just foreign material. Therefore, the comment about T&E resource investment is not completely relevant to the section topic.

While OSD does use the Foreign Material Acquisition / Foreign Material Evaluation Program to acquire threat systems, it is the Foreign Material Program Review Board (FMPRB) which determines the requirements for foreign material. In fact, DOT&E was instrumental in the creation of a Test and Evaluation Subcommittee to the FMPRB in an attempt to ensure that the T&E community has access to as much of the foreign material as possible.

Page 26 (Third paragraph)

The draft report states that "Testing the F-16 missile targeting system only in the desert, although it has a worldwide operational mission requirement, introduces a high risk on its effectiveness in other climates." DOT&E fully supports testing in realistic climates and environments. However, the experience from other HARM missile targeting systems indicates that climate is not a significant factor in the performance of these systems. Therefore, we recommend that this statement be deleted or that a better example be used.

Page 26 (Fourth paragraph)

The draft report states that "The Military Departments' test documents are not providing essential information for assessing acquisition risk and for decision making." The report contains no basis for this statement. Although the Military Departments have not always reported all the effects of test limitations, generally this has not denied decision makers essential information. In most cases the statement of the limitation itself is sufficient to convey its impact. Recommend that the comment be better substantiated.

Page 27 (3rd Paragraph which starts "Test documents . . . ")

The draft report states "Plans are being approved without adequate analysis of the cause and effect of test limitation." The fact that the analysis of the cause and the impact of test limitations are not documented in the approval of test documents should not imply that that analysis is not being performed or that its results are not being given full consideration in the test plan approval process. Test officials usually have far greater insight into the test planning process than is outlined in the test plan. They will understand the operational implications of the planned tests activities as well as the test limitations. In many cases, the cause of the test limitation is a requirement for funding which, after analysis, has been determined to be impractical or unreasonable to correct. When there is an impact due to test limitations it is documented in test reports.

Page 27 (Recommendations for Corrective Action)

We partially concur with recommendation C.1. However, it may not be practical or even possible to acquire sufficient copies of all of the different threat systems. DOT&E has been very active in attempting to obtain needed threat resources for testing.

The OSD Central Test and Evaluation Improvement Program (CTEIP) has promoted, with success, joint resource programs to enhance DoD's ability to conduct tests that would not have happened without it. DOT&E is an active participant in CTEIP and has consistently worked toward greater CTEIP funding.

DOT&E, in coordination with DTSE&E, initiated the actions to charter the Test and Evaluation Subcommittee of the Foreign Material Review Board and is presently the co-chair of that subcommittee. This subcommittee is designed to insure that the testing community needs are considered in the Foreign Material Acquisition / Foreign Material Evaluation Program. This subcommittee is expected to greatly assist the testing community's access to foreign material.

We non-concur with recommendation C.2. The forced inclusion of a specific cause for each test limitation and the probable impact of the limitation on the test conclusions would provide very little net value to test plans. It would turn each test plan into an accounting document which would require more effort to generate, add little to the understanding of the test itself, and ultimately result in fewer limitations being revealed because of the added workload to list them.

Page 28 (DoD Guidance)

While the DoD guidance requires that the Annual Report summarize all OT&E activities occurring during the past fiscal year, it does not explicitly require the same level of analysis for every activity. Producing such in-depth analyses for all systems has not been possible given staffing and resource contraints.

Page 29 (Report Content)

The Annual Report does not report on special access programs because they are covered in a separate report.

The Federal Acquisition Streamlining Act (FASA) moved oversight of the Live Fire Test and Evaluation (LFT&E) from USD(A&T) to DOT&E. Inclusion of LFT&E activities in the DOT&E Annual Report was a new requirement brought into being by the FASA legislation

passed just at the end of FY1994. Actual transfer of responsibility occurred in FY1995. Hence, LFT&E was not included in the FY1994 Annual Report. It will be included in the DOT&E Annual Report beginning with the FY1995 Annual Report. The FY1995 Annual Report will include both a synopsis of LFT&E Systems under oversight as well as commentary on LFT&E issues on a system-by-system basis.

Page 29 (Clarity of the Annual Report)

Please note that the comment in the first paragraph relative to 10 USC. 2399 only applies to designated systems.

The draft report states that because DOT&E quotes the Navy OTA report conclusions regarding the AN/SQQ-89(V)6 effectiveness and suitability, that there was ambiguity as to whether DOT&E agreed with the Navy assessment. A careful reading of this assessment section shows that DOT&E questioned several areas of the Navy OTA report, with different interpretations of the results. Overall, however, the DOT&E assessment was positive and supported the Navy OTA's fundamental conclusions about the systems' operational effectiveness and operational suitability.

Page 31 (Resource Requirements)

The second to last sentence beginning "In the future, ..." is incorrect. We are limited in our use of IDA by ceiling restrictions and money constraints. Also the last sentence is misleading because it ignores the issues of skills and training available in-house vice from an FFRDC raised in 1990.

Page 32 (Top paragraph)

The DOT&E action officers, both military and civilian, are very senior individuals normally with over 20 years of related hands-on experience. Many of the action officers have had previous R&D assignments which generally include an assignment at one of the Service Operational Test and Evaluation agencies (OTAs). The previous experience of these action officers provides them with sufficient background knowledge such that straightforward on-the-job training allows them to become very productive in their jobs. In the few cases where additional schooling is requested, it is provided. Additionally, as is stated in the draft IG report, The "New Assistant's Guide" is required reading and the self-help video is required viewing for all new action officers.

However, the IG Audit Report caused us to review our training procedures and we have determined that a more aggressive training approach is warranted. Therefore, we have initiated the development of a proactive training plan for our action officers. This plan will include milestones for monitoring its execution.

Page 32 (Second and Fourth paragraphs)

The purpose of the "New Assistant's Guide" is not to prepare action officers for the preparation of the Annual Report. Therefore, it does not contain information on the preparation of the Annual Report. However, we have prepared formal guidance for writing Annual Report

submissions and for ensuring that the reports are complete and consistent. This guidance was given to all of the action officers on September 29, 1995.

Page 32 (Completeness and Consistency of the Annual Report)

The Director of Operational Test and Evaluation Annual Report is not a decision making tool for individual weapon systems. Rather it is a report on the results of recent OT&E and LFT&E and the state of this testing within the Department of Defense. Decisions on the future of developing weapon systems are made using other documentation in formalized processes.

The second paragraph has a partial factual error. That is, while the FY 1994 Annual Report was published in February 1995, the cut-off date for information was September 30, 1994. Since the CSSCS's initial OT&E was conducted in September 1994, its test results were still undergoing evaluation by both the Army and DOT&E until a date after the Annual Report was written. As a general matter, our policy has been to not release test conclusions until the director has signed the final B-LRIP. In the particular CSSCS case, the Annual Report went to the printer prior to the Army reaching a conclusion that the MSIII would be delayed.

Page 33 (Recommendations for Corrective Action)

We concur with recommendation D.1 and will work with the comptroller to achieve increased civilian staff years. However, we also recognize that the Congress has mandated that DoD plan for reductions over the next 5 years.

We partially concur with recommendation D.2.

We will prepare the recommended staffing and contractor support study mentioned in D.2.a.

We have already taken appropriate actions to comply with D.2.b.

We are assessing the current action officer training program and will ensure that those new action officers who require formal training will receive all training possible within the scope of the training and time budgets as recommended by D.2.c.

DOT&E will work with congressional staffs, OSD offices and military departments to determine what operational test information on non-major programs is needed in the DOT&E Annual Report. Once this information is identified, DOT&E will determine what resources are required to develop and include this information and will request those resources.

Page 37 (Acceptance of the Selected Acquisition Report)

Although DOT&E is responsible for reviewing the test status reported in Sections 7 and 10 of the SAR, this document is primarily a program manager's report. DOT&E does not have the resources to obtain the necessary information to perform an independent, detailed review of each SAR. Basically, the program manager has the information needed to create the report and unless there has been a major finding by DOT&E as a result of an operational test, DOT&E action officers have little independent basis to perform a detailed review of the document.

Page 39 (Recommendations for Corrective Action)

We concur with recommendation E.1, and would support recommendation E.2 in concert with the USDA&T and his responsibilities as stated in recommendation E.1.

Page 49 (Office of the Secretary of Defense)

This paragraph does not include the fact that the DOT&E also now has statutory responsibility to report not only on operational effectiveness and suitability but also on system survivability and lethality. Recommend that following as a replacement for the last three sentences starting with "The Director reports directly . . . ."

> "The Director communicates directly with the Secretary and must submit a report to Congress to meet statutory operational and live fire test reporting requirements. The DOT&E responsibility to Congress is to provide unbiased insight into operational effectiveness, suitability, lethality, and survivability of new systems and, in some cases, upgrades to fielded systems. The DOT&E approves test and evaluation master plans as well as operational and live fire test plans for designated systems."

8

Department of the Army Comments

DEPARTMENT OF THE ARMY
OFFICE OF THE UNDER SECRETARY
WASHINGTON, D.C. 20310-0102

24 January 1996

SAUS-OR

MEMORANDUM FOR DIRECTOR, ACQUISITION MANAGEMENT DIRECTORATE, DEPARTMENT OF DEFENSE INSPECTOR GENERAL

SUBJECT: Audit Report on Operational Testing Performed on Weapon Systems (Project No. 5AB-0031)

Reference memorandum, 24 November 1995, subject as above.

This office has reviewed the draft subject report. Our response is attached at the enclosure.

Please direct any questions to Dr. John Foulkes, (703)695-8995.

Walter W. Hollis
Deputy Under Secretary of the Army
(Operations Research)

Enclosure

FINDING A. The Military Departments process for approving test and evaluation master plans was not efficient because the Military Departments used a sequential approval process instead of a more efficient integrated process. As a result, resources used for additional test and evaluation plan development and approval activities could be put to better use. **CONCUR WITH FINDING.**

Recommendation A.1. We recommend that the Army Director, Test and Evaluation Management Agency; the Navy Director, Test and Evaluation and Technology Requirements; and the Air Force Director, Test and Evaluation Directorate:

a. Apply the integrated product team concept to the test planning and execution process.

CONCUR. In support of Army programs which have instituted the integrated product team (IPT) process, Test and Evaluation (T&E) IPTs have been formed to develop the T&E strategy for the program and to develop and coordinate the Test and Evaluation Master Plan (TEMP). In addition, the Army currently is staffing draft policy and procedures for major programs which requires no more than 30 working days between signature of the T&E IPT-coordinated TEMP by the Project Manager and the signature by the Army TEMP approval authority (Deputy Under Secretary of the Army (Operations Research) (DUSA(OR))).

CORRECTIVE ACTION TAKEN. The Army currently is staffing draft policy and procedures for major programs which leverages the IPT process and which requires no more than 30 working days between signature of the T&E IPT-coordinated TEMP by the Project Manager and the signature by the Army TEMP approval authority (DUSA(OR)).

b. Delegate approval authority for the Test and Evaluation Master Plan to the integrated product team.

NONCONCUR. The Army believes it must remain cognizant of the status and content of all T&E strategies being developed for major programs. In addition, a centralized individual should be responsible for developing and promulgating T&E policy and procedures for the TEMP process and for interfacing with the Department of Defense T&E and TEMP policy makers. The Army believes that the appropriate individual for this is the Service T&E Director. As stated above, the Army has developed draft procedures which minimizes the time from PM signature to that of the Army T&E Principal, the DUSA(OR). Also note that the 19 January 1996 draft of DoD Instruction 5000.2 indicates that the Service approval authority for major program TEMPs (which currently is the Service Acquisition Executive) will be the Service T&E Director.

1

c. Forward the Test and Evaluation Master Plan directly from the Program Manager to the Director, Test , Systems Engineering, and Evaluation, and the Director, Operational Test and Evaluation.

NONCONCUR. The Test and Evaluation Master Plan should be forwarded through the Service T&E Director prior to forwarding to the Office of the Secretary of Defense per discussion in paragraph b. above.

RECOMMENDATION A.2. We recommend that the Director, Operational Test and Evaluation, include himself as a formal, non-voting member of the program office integrated product team to expedite his review of test documents.

RECOMMENDATION A.3. We recommend that the Director, Operational Test and Evaluation, include himself as a formal member of the program office integrated product team to expedite his review of test documents.

CONCUR with both recommendations A.2 and A.3, however the Army believes that empowered representatives of these two OSD offices should suffice as members of the IPT.

FINDING B. The Military Departments were delivering systems for operational testing that were not ready for testing. This situation occurred because a calendar schedule rather than system readiness often drove the start of testing. In addition, neither DoD nor the Military Departments have formalized criteria to help certification officials and operational testers decide whether to delay tests because of system limitations. Because the Military Departments often cannot conduct a complete system test, they incur additional costs to repeat a test or to test a system that should have been in the original evaluation.

PARTIALLY CONCUR WITH FINDING. Concur with the part of the finding that states that the Army was delivering systems for operational testing that were not ready. In the case of one of the cited systems (ASAS), the Project Manager and the operational tester were aware that ASAS was not ready, however the system proceeded since the system had been under development for several years without any user testing and user testing was necessary for continued development. Nonconcur with the part of the finding that states that the Army has no criteria to help certification officials and operational testers decide whether to delay tests. Army Regulation 73-1 (AR 73-1), 27 February 1995, paragraphs 3-12 and 3-13 provide such criteria.

RECOMMENDATION B.1. We recommend the Director, Test, Systems Engineering, and Evaluation, review and approve certification of readiness for operational testing for acquisition category I and designated oversight programs.

PARTIALLY CONCUR. Although this may improve the readiness of a particular system to enter operational testing, the added layer of oversight appears to contradict the IPT approach to efficient test planning as prescribed in Finding A.

CORRECTIVE ACTION TAKEN: None, action is directed to DoD.

RECOMMENDATION B.2. We recommend the Under Secretary of Defense for Acquisition and Technology revise Department of Defense Instruction 5000.2, "Test and Evaluation," to require the Army, Navy, and the Air Force to develop a decisionmaking model to evaluate whether to proceed with testing when serious limitations are identified before operational testing and to determine the cost/benefit of testing with limitations. **CONCUR.**

CORRECTIVE ACTION TAKEN: None, action is directed to DoD, however the Army will ensure that DoD is aware of the Army policy in AR 73-1.

FINDING C. The Military Departments' operational tests were not always performed in a realistic combat environment. Further, the Military Departments were neither adequately assessing the impact of test limitations nor the risk of fielding systems that will not counter the identified threats. Test limitations occurred because the Military Departments lack funding to procure the threat systems, have an inadequate quantity of threat systems necessary to conduct a realistic test, and lack opportunity to acquire the appropriate threat system. Additionally, range availability and safety are often cited as limitations. Test planning and report documents did not adequately report the impact of test limitations. As a result, the Military Departments fielded systems with uncertain effectiveness against identified threats. Further, test documents did not provide decisionmakers with adequate information concerning the potential impact of test limitations.

PARTIALLY CONCUR WITH FINDING. Most test limitations are caused by long-standing resource shortfalls, and are well documented and understood. Each Military Department determines its priorities and executes them according to the resources available.

RECOMMENDATION C.1. We recommend that the Assistant Secretary of the Army (Financial Management and Comptroller),the Assistant Secretary of the Navy (Financial Management and Comptroller), and the Assistant Secretary of the Air Force (Financial Management and Comptroller):

a. Budget adequate resources to acquire actual threat systems to overcome threat limitations.

PARTIALLY CONCUR. The ASA(FM&C) does not set the priorities for funding in the Army. The resourcing of test and evaluation requirements is a collaborative effort among several Headquarters, Department of the Army offices. In this continuing environment of severely constrained resources, the ASA(FM&C) cannot guarantee full funding of threat systems. That decision depends on prioritization consensus among the key players, with ultimate approval authority by the Army leadership.

CORRECTIVE ACTION TAKEN: The ASA(FM&C) can and will raise the audit recommendation for careful consideration in future resource allocation forums.

b. Acquire validated, verified, and accredited threat surrogates, models, or simulators for those threats that are not available.

PARTIALLY CONCUR. We concur with the recommendation, and have been executing it, however the ASA(FM&C) is not the office of responsibility.

CORRECTIVE ACTION TAKEN: The Army participates in the tri-service CROSSBOW office, which provides oversight and funding for tri-service threat simulator programs. Army-unique threat systems are acquired by the Program Manager for Instrumentation, Targets, and Threat Simulators, with oversight provided by the Test and Evaluation Management Agency.

RECOMMENDATION C.2. We recommend that the Director, Operational Test and Evaluation, establish a policy to approve Test and Evaluation Master Plans or Test Plans only when they address the specific cause for each limitation and the probable impact of the limitation on test conclusions, both for the evaluation of critical operational issues and the ability to conclude whether the system is effective and suitable.

CONCUR. It is the Army's belief that this is already been done.

CORRECTIVE ACTION TAKEN: None, action is directed to DoD.

FINDING D.
FINDING E.

No opinion, as these findings and associated recommendations are directed to DoD.

4

Department of the Navy Comments

DEPARTMENT OF THE NAVY
OFFICE OF THE ASSISTANT SECRETARY
RESEARCH, DEVELOPMENT AND ACQUISITION
1000 NAVY PENTAGON
WASHINGTON DC 20350-1000

MAR 8

MEMORANDUM FOR THE DEPARTMENT OF DEFENSE ASSISTANT INSPECTOR GENERAL FOR AUDITING

Subj: DRAFT REPORT ON THE AUDIT OF OPERATIONAL TESTING PERFORMED ON WEAPONS SYSTEMS (PROJECT NO. 5AB-0031) - ACTION MEMORANDUM

Ref: (a) DODIG memo of 24 Nov 95

Encl: (1) DoN Response to Draft Audit Report

I am responding to the draft audit report forwarded by reference (a) concerning the operational testing performed on weapons systems.

The Department of the Navy response is provided at enclosure (1). We generally agree with the draft report findings and recommendations with five noted exceptions. As outlined in the enclosed comments, many of the issues have already been resolved. The Department's current practices are significantly different from those studied and incorporate Integrated Product Team tenets and practices.

W. C. Bowes
Vice Admiral, U.S. Navy
Principal Deputy

Attachment To Memorandum For The Department of Defense Assistant Inspector General for Auditing

DEPARTMENT OF THE NAVY RESPONSE
to
DODIG Draft Report no. 5AB-0031 of 24 Nov 95
on
OPERATIONAL TESTING PERFORMED ON WEAPON SYSTEMS

Finding A - COMMENTS TO RECOMMENDATIONS

DODIG recommends that:

A1. The Navy Director, Test and Evaluation and Technology Requirements:

a. Apply the integrated product team concept to the test planning and execution process.

b. Delegate approval authority for the Test and Evaluation Master Plan to the integrated product team.

c. Forward the Test and Evaluation Master Plan directly from the program manager to the Director, Test, Systems Engineering, and Evaluation and the Director, Operational Test and Evaluation, for approval.

Reply: a. Concur. The IPT concept has been institutionalized in SECNAVINST 5420.188D and OPNAVINST 5000.42D.

b. Do not concur. An empowered IPT is the proper forum for TEMP formulation, but the IPT process will only improve efficiency if **all** IPT members are present and properly empowered. Appropriate review by principals and senior managers above the IPT member provides significant value added through bringing to bear increased insight and experience and providing the appropriate checks and balances. Oversight by the appropriate Milestone Decision Authority is still required.

c. Do not concur. Service Chiefs have the sole responsibility for conduct of operational testing and reporting. Service T&E Directors have primary responsibility for completeness and accuracy of TEMP and must have the opportunity to review the document prior to submission to DOT&E/DTSE&E.

A2. The Director, Operational Test and Evaluation, include himself as a formal, non-voting member of the program office integrated product team to expedite his review of test documents.

Enclosure (1)

Final Report Reference

Reply: Concur with this concept; however, an empowered representative of DOT&E is sufficient.

Revised Recommendation A.2

A3. The Director, Test, Systems Engineering, and Evaluation include himself as a formal member of the program office integrated product team to expedite his review of test documents.

Reply: Concur with this concept; however, an empowered representative of DTSE&E is sufficient.

Revised Recommendation A.3

Finding B - COMMENTS TO RECOMMENDATIONS

B1. Recommend the Director, Test, Systems Engineering, and Evaluation, review and approve certification of readiness for operational testing for acquisition category I and designated oversight programs.

Reply: Do not concur. DOT&E already performs readiness for operational test review and is an IPT member of the Navy SYSCOM/PEO Operational Test Readiness Review (OTRR) Board. Moving responsibility to DTSE&E presents no value added.

B2. Recommend the Under Secretary of Defense for Acquisition and Technology revise Department of Defense Instruction 5000.2, "Test and Evaluation," to require the Army, the Navy, and the Air Force to develop a decisionmaking model to evaluate whether to proceed with testing when serious limitations are identified before operational testing and to determine the cost/benefit of testing with limitations.

Reply: Do not concur. Development of a decision-making model is unnecessary. The Navy SYSCOM/PEO IPT Operational Test Readiness Review (OTRR) process and each SYSCOM's internal OTRR checklist provide adequate risk assessment to the decision maker. Institutionalized over the last two years, the Navy OTRR process corrects the deficiencies identified in the dated sample group OSD (Dec 92-Nov 94) and IG (Jun 88-Dec 94) studies. A decision-making model does not provide added value to the current process. This reply addresses the management control weakness discussed in Appendix A of the basic report.

2

Finding C - COMMENTS TO RECOMMENDATIONS

C1. Recommend that the Assistant Secretary of the Army (Financial Management and Comptroller), the Assistant Secretary of the Navy (Financial Management and Comptroller), and the Assistant Secretary of the Air Force (Financial Management and Comptroller):

a. Budget adequate resources to acquire actual threat systems to overcome threat limitations.

b. Acquire validated, verified, and accredited threat surrogates, models or simulators for those threats that are not available.

Reply: a. Concur. However, because of fiscal limitations, the MDA and Program Manager must balance requirements, testing limitations and cost into designing the testing program. In many instances actual threat systems are not available or too expensive to be considered, while some surrogates or simulators provide adequate and/or better threat representation. Acquiring actual threats where fiscally reasonable is each program's goal.

b. Concur. Ongoing programs exist to exploit capabilities of modeling and simulation to adequately simulate threats.

C2. Recommend that the Director, Operational Test and Evaluation, establish a policy to approve Test and Evaluation Master Plans or Test Plans only when they address the specific cause for each limitation and the probable impact of the limitation on test conclusions, both for the evaluation of critical operational issues and the ability to conclude whether the system is effective and suitable.

Reply: Do not concur. Recommendation for DOT&E policy on TEMP approval where test limitations exist reduces efficiency. Emphasis is better placed on TEMP developers to include more detailed discussions of test limitations and impact on testing. Proper use of the IPT process will preclude the need for DOT&E to establish a non-approval policy.

Findings D and E - Findings D and E are not germane to Navy Test and Evaluation Requirements.

3

ASN(RDA) TEMP PAT RESULTS

29 AUGUST 95

>CURRENT TEMP PROCESS IS VALUE-ADDED AND REQUIRED

>INSTITUTIONALIZE THE TEST PLAN WORKING GROUP (TPWG) IPT PROCESS

>INCORPORATE TPWG IPT PROCESS INTO OPNAV INSTRUCTION

Slide 16

Enclosure (2)

Department of the Air Force Comments

DEPARTMENT OF THE AIR FORCE
WASHINGTON, DC

Office of the Assistant Secretary

25 January 1996

MEMORANDUM FOR ASSISTANT INSPECTOR GENERAL FOR AUDITING
OFFICE OF THE INSPECTOR GENERAL
DEPARTMENT OF DEFENSE

FROM: SAF/FMPF

SUBJECT: Air Force Comments to DoD(IG) Draft Report on Operational Testing Performed on Weapon Systems (Project No. 5AB-0031)

The attached comments are provided in response to your 24 November 1995 memorandum requesting the Assistant Secretary of the Air Force (Financial Management and Comptroller) (SAF/FM) provide Air Force comments on the subject report. Attachment 1 presents the Air Force Test and Evaluation Directorate comments. Attachment 2 is the SAF/FM response to recommendation C.1. The SAF/FM response was provided through the Test and Evaluation Directorate.

VAUGHN E. SCHLUNZ
Director for Audit
Liaison and Followup
(Financial Management)

Attachments:
1. AF/TE Memo, 24 Jan 1996
2. SAF/FMBI Memo, 23 Jan 1996

DEPARTMENT OF THE AIR FORCE
HEADQUARTERS UNITED STATES AIR FORCE
WASHINGTON DC

24 JAN 1996

MEMORANDUM FOR SAF/FMP

FROM: AF/TE

SUBJECT: Review of DoD (IG) Audit Report, "Operational Testing Performed on Weapon Systems," Project No. 5AB-0031

We appreciate the opportunity to review and comment on the subject report. Although some findings and recommendations are correct, others are inconsistent with the basic principles of test and evaluation (T&E) and how it is managed. The report's lack of in-depth analysis leads to some unsupported generalizations. Our comments, combined with those from the Air Force Operational Test and Evaluation Center (AFOTEC), offer a more complete picture. We have the following comments:

Finding A. Test and Evaluation Planning Process

a. On page 8, the fifth paragraph tells only part of the story. The requirement for AF/TE to sign the TEMP aligns Air Force policy with the other Services' policies. The Navy requires its T&E Executive to sign TEMPs prior to the Service Acquisition Executive (SAE). The Army SAE delegated TEMP approval authority to the Army T&E Executive, so only one signature is required. As the Service functional representatives for all T&E matters, the T&E Executives must be key elements in the TEMP approval process. They ensure continuity and quality in the Service T&E processes.

b. **Recommendations: Partially concur.** While the Air Force welcomes greater reliance on integrated product teams (IPT) for test planning, the fact is we are already using IPTs. The structure of the test planning working group (TPWG) is in essence an IPT. Our primary concern is that OSD remains largely absent from the TPWG process and contributes significantly to the length of that process as shown in Tables 1 and 2. Also, we do not support the delegation of TEMP approval authority above Service level. The Service Chiefs have statutory responsibility for equipping and training their forces. Large portions of that authority would be ceded to OSD if the Service Chiefs were bypassed in the TEMP process.

Finding B. Readiness for Testing

a. The finding is incorrect when it states there are no "formalized criteria to help certification officials and operational testers decide whether to delay tests because of system limitations." The Air Force published Air Force Manual (AFMAN) 63-119, *Certification of System Readiness for Dedicated Operational Test and Evaluation*," in February 1995. A well-defined process for assessing the readiness of systems to enter dedicated OT&E was developed which did not levy bureaucratic delays or burdens on the program, but did require exit criteria and informed judgment by decision makers before systems proceeded to the next phase of testing. The Air Force certification process was field tested for a year before final publication, and has been a true success story.

Final Report Reference

b. On page 14, "Test Readiness and Completeness," first paragraph. It is not Air Force policy to begin OT&E when there is significant risk the program will fail OT&E. The goal of every development program is to reduce risks to prudent levels, not eliminate all risk. Some tests must go forward with known risks, but only when those risks and limitations are unavoidable and manageable. The Air Force certification process helps decision makers assess those risks.

c. We disagree with the statement on page 17, third paragraph, that "criteria have not been formalized to help certification officials and operational testers in assessing whether to delay tests because of limitations." We also disagree with the statement on page 18, first paragraph, that "the guidance is silent on whether to proceed with the test when limitations exist." AFMAN 63-119 provides ample guidance for decision makers to assess readiness for OT&E while retaining ample flexibility for program managers. The key criteria are, "Will the system pass OT&E?" These criteria are evaluated by comparing known system capabilities and limitations to the system requirements.

d. On page 16, the AC-130U Gunship Replacement Program is a poor example for illustrating how the Air Force certification process works. The AC-130U was one of the first systems certified during early field testing of our new process. During the DoD IG's initial interviews, we pointed out the C-17 and Milstar programs were much better examples of how the certification process works today, and those programs should be reviewed instead of the AC-130U. These programs produced excellent results in OT&E, largely as a result of the new certification process.

e. **Recommendations: Partially concur.**

(1). **Recommendation B.1: Non-concur.** The proposal adds another layer of review and approval to the short, streamlined acquisition chain directed in DoD 5000.1, Part 1. It usurps the PEO's authority over assigned programs, places review and approval further from the most knowledgeable sources, and will slow program execution schedules. Acquisition reform trends throughout government are pushing decision authority down to the lowest practical levels.

(2) **Recommendation B.2: Partially concur.** The Air Force has implemented a highly credible certification process. Recommend the other Services use this process or one similar to it. However, mandating "one-size-fits-all" decision criteria not based on individual program requirements would be counter-productive.

Finding C. Test Threat and Environment

Revised Finding C.

a. On page 26, "Uncertainty and Risk," second paragraph, the report describes testing of the F-16 missile targeting system only in the desert as "introducing a high risk on its effectiveness in other climates." Change the phrase to read, "fails to reduce uncertainty for decision makers about system effectiveness in non-desert climates." The lack of testing in specific environments cannot introduce new risk, let alone "high" risk, since the system has an inherent amount of risk.

b. We are concerned that Finding C's overall thrust misses the mark about the amount of detail needed when describing test limitations and the impacts on test reporting.

(1) On page 23, the Military Departments are criticized for not presenting "as comprehensive an analysis of the impact of test limitations as that of the DOT&E beyond LRIP reports." On page 24, the report states, "test planning and report documents do not adequately report the impact of test limitations." The Finding concludes with the recommendation to "address the specific cause for each limitation and the probable impact of the limitation on test conclusions." In general, most test limitations are due to long-standing resource shortfalls, are well documented and understood, and need no further elaboration. It is Air Force policy to always describe limitations when and where they increase understanding of the test and its results.

(2) We believe the report is too focused on attaining near-perfect test environments rather than the accurate characterization of system capabilities and limitations. For any test, the number of limitations is potentially unlimited, and testers find work-arounds for most of them. The Air Force routinely performs analysis and reports on all significant limitations which may impact test results. However, if test plans and reports do not discuss all test limitations, it is because testers believe they have acceptable, balanced programs and are able to reach credible conclusions in spite of resource constraints. Nonetheless, the subject report advocates the analysis of all test limitations without regard to the value added to the total test effort. More detailed explanations and analyses always add cost and time to test programs. Therefore, the TEMP and test plan approval processes will be slowed even further.

c. **Recommendations: Partially concur**. While we fully support increased budget authority for test resources and threat systems, the reality is that funding continues to decline despite our best efforts. The primary reason for the acquisition of threat systems, as well as our analysis capabilities, must remain focused on characterizing the capabilities and limitations of our systems, and not become preoccupied with limitations to test.

Finding D. Test Status Reporting. We non-concur with Recommendation D.2.d which expands the scope of DOT&E's Annual Report to include nonmajor programs. There are hundreds of nonmajor programs, most of which are successfully executed without OSD oversight. The added reporting requirements would further dilute the already limited time DOT&E staffers devote to major programs. TEMP and test plan approval times for major programs would increase as a result.

Finding E. Reporting Program Results. Concur.

The points of contact for this report are Lt Col Bill Sullivan and Mr. Chuck Triska, DSN 225-0900.

HOWARD W. LEAF
Lt Gen, USAF (Retired)
Director, Test and Evaluation

Attachment:
AFOTEC/CC Memo

DEPARTMENT OF THE AIR FORCE
HEADQUARTERS AIR FORCE OPERATIONAL TEST AND EVALUATION CENTER
KIRTLAND AIR FORCE BASE, NEW MEXICO

MEMORANDUM FOR AF/TE

2 5 JAN 1996

FROM: HQ AFOTEC/CC

SUBJECT: Comments, (Draft) DoD IG Report: Operational Testing Performed on Weapons Systems (Project No. 5AB-0031)

1. I have reviewed the subject audit report. Overall, I believe the report needs more insight from the tester point-of-view, because it misses certain important concepts about testing. Some conclusions are reached based on only a few programs, and I feel that an inadequate amount of in-depth analysis was done to discover if there were other factors within the program which may have caused the problems noted. I believe the auditors misunderstood or inadvertently stretched the inputs made by various program officials. Specific comments on the findings follow.

2. **Finding A, Test and Evaluation Planning Process (TEMP Approval).**

a. The report criticizes the "Military Departments" for using a sequential process to approve a TEMP, instead of empowering their members. It also cites the use of parallel staffing in gaining TEMP approvals. Unfortunately, the report misses the point about why the military staffs paperwork to different offices. It is because there is no single source of experience and information available. For example, AFOTEC staffs TEMPs in parallel in order to reach the widest experience base possible in the shortest amount of time. Plans are exposed to the needed expertise; comments are compiled; and then both are forwarded to the Commander for approval. Even this audit report was staffed this way to allow for a wide range of comments. AFOTEC has a wide cross-section of operational experts and analysts. Managers use these experts to gain the greatest insight possible to avoid costly mistakes in the future. Even if cross-directorate coordination takes longer, this time is better spent early, rather than later correcting errors. We suggest other military departments adopt similar parallel staffing processes. Using electronic processes when coordinating TEMPs among their various agencies will also help expedite coordination. These two suggestions will accelerate the approval process without losing the valuable perspective of the resident experts.

b. The audit report implies that empowered representatives sent to meetings can accomplish more in less time than is possible within an entire organization. This is not necessarily true. If a meeting were called and each organization was assigned one empowered member, that member could not just arrive, make decisions, and generate changes instantly. They would have to spend time and effort researching issues before arriving at optimal positions. The example given by the report, the JSTARS TEMP, is an excellent example of the inadequacy and hazards of one person coordination. This TEMP was sent to OSD and six subsequent organizations for comments. The results were 24

pages of changes. A single member could never have provided the insight and depth needed by that TEMP. There are other problems inherent with staffing at this lower level, including lack of expertise, overtasking, and a tendency to create subgroups to address specific problems. It is difficult to prove, but parallel staffing often results in a quicker set of comments. A single empowered decision maker may work faster, but faster does not necessarily equate to better.

c. **Senior Management Delays.** The audit report points out exactly why empowered representatives at meetings cannot approve TEMPs. The TEMP, especially for ACAT I programs, must be reviewed and approved at the highest levels due to the importance of the programs. TEMPs are guiding documents for programs. It is imperative that they be done right, and done right the first time whenever possible.

d. Finally, the report ignores a very important point. The AF has a process team that works TEMPs and their approvals: the Test Plan Working Group (TPWG). The DoD IG observed that executive level support at MILSTAR TPWG meetings was nonexistent. The report then suggests that DoD adopt an Integrated Process Team (IPT) with empowered attendees as a cure to speed the TEMP process. They do not explain how a MILSTAR IPT would be better attended or more empowered than the original TPWG was. An Air Force (AF) TPWG is almost exactly what is suggested by this report, but renaming the TPWG will not increase attendance or participation.

e. **Recommendations: Non-Concur.** The DoD IG recommendations do not offer any solutions. The TEMP is a direct reflection and the responsibility of the Program Manager (PM). The PM forms the TPWG to construct the TEMP. The Service T&E Directors (AF/TE in our case) then review TEMPs to ensure compliance with T&E policy and instructions. The first recommendation to use an IPT process should be directed at Program Managers. They control the TEMP process, not the Service T&E Directors. The second recommendation, sending the TEMP directly to senior level managers, is being done already. We recommend that organizations adopt parallel staffing and electronic dissemination to reduce review time.

3. **Finding B, Readiness for Testing (Testing with Deficiencies).**

a. The audit report gives examples of systems that began OT&E before they were ready. It questions why the TEMP and test plans do not reflect this situation. It concludes that systems enter test because programs are driven by calendar schedules instead of progress schedules. The report further states that this finding is not new, and it quotes a previous IG report that reached the same conclusion. AFOTEC understands this situation is a problem and uses AFMAN 63-119 to certify a system for operational testing. We work very strenuously to ensure systems are the best operationally representative system available with no significant deficiencies to impact our testing. AFOTEC also understands the problems programs encounter when trying to reach this level of capability.

Final Report Reference

b. Calendar event schedules (as opposed to progress event schedules) are driven by pressures from many sources, including the PM, funding, and the user. A PM's main motivation is getting the best product to the field for the cost and amount of time given. They are evaluated on the performance of their programs and are constantly juggling priorities. Calendar events mark program performance, but do not necessarily reflect system performance. Sometimes PMs sense that a system that passes a calendar event, even if it is an incomplete system, is a better measure of success than delaying the event and fixing the system. Funding is also a source of pressure. (Note: the AC-130U program mentioned in the audit report went to test simply because the program office ran out of funding to fix the remaining deficiencies. Stopping test would not have stopped fielding the system.) Fixing deficiencies costs time and money. Also, funding is tied by Congress to a time period; program delays risk getting programs canceled. A PM often feels that cancellation is a reflection of poor management performance, rather than poor system or contractor performance. Most PMs feel cancellation should be avoided at all costs, even at the risk of fielding a system which doesn't meet the user's requirements. Finally, the user is another source of pressure. There are times when they prefer having a partially capable system now, rather than a fully capable system later. A good example is a system that will fill an urgent need: an ECM system against a new threat. If the user requirement is for the system to be 80% effective, but the system is only 50% effective, the system would be reported as not meeting requirements. That does not mean the user will not want to acquire the system. That lesser capability is better than the current capability, which is zero (it can always be improved in the future). In sum, many conflicting pressures (the PM, funding, and the user need) lead to schedules being driven by time instead of performance.

c. The audit report suggests that testing should be stopped to fix problems, but it does not recognize the risks faced by a program manager and the user if this happens. For the PM, stopping testing to fix deficiencies may result in termination of the program. At the very least, delays will cost more time and money to complete the system. For the user, stopping testing and terminating a program represents a risk to meeting their need. Experience has proven that if a new development effort is begun (because an earlier system was canceled), it will take even more time to field the replacement system.

Revised Recommendation B.2

d. The audit report states that "we have not found criteria by which Military Departments can consistently evaluate cost and schedule implications of the decision to test when limitations exist." This is because it is not possible to develop individual, specific criteria for all programs. Each program is reviewed based on its individual merits to determine the tradeoffs associated with testing a system that has deficiencies. To aid this process, the AF has developed Certification Templates, AFMAN 63-119. The templates provide a process for PMs to certify a system for operational test. Each program is evaluated based on programmatic issues and the criticality of deficiencies. Using these templates, the OTA and the PM decide whether to certify a system (begin OT&E) or not. This "negotiation" process leaves the exact process flexible, allowing it to be adapted to a wide range of programs. This process empowers test managers and PMs to make value judgments on the readiness of a system for operational testing.

e. Finally, the audit report does not understand the wide range of testing available to decision makers. For example, neither the OT&E report for the AC-130U, nor the F-16 Block 50D/HTS programs supported an acquisition or procurement decision. The testing was intended to provide the user with an evaluation of the capability of the system, at that particular point in its acquisition cycle. The actual decision to procure had already been made, due to the needs of the user. Critical decision points occur many times within the acquisition process, for reasons that relate to program specific or user specific problems. Testing is invaluable to this sometimes convoluted process.

f. **Recommendations: Partially Concur.** Adding DTSE&E as approval authority for OT&E certification will extend the time needed to certify a system for OT. We believe this would be an unnecessary addition to the process. The AF already has an OT certification process, contained in AFMAN 63-119. It has been used successfully on a number of programs. We believe the other services should examine our process to see if it can fulfill their needs for a certification process. This action would satisfy the second part of the DoD IG recommendation: "to develop a decision making model to evaluate whether to proceed with testing when serious limitations are identified..."

4. **Finding C, Test Threat and Environment (Lack of Realism).**

a. This finding states that Final Reports and Test Plans address some test limitations, but do not assess the impact of the limitations when rating the COIs. The audit report also questions why testing is accomplished with known limitations. Testing with limitations is not a new problem. It is a fact of life. The report offers no viable solutions to its observations.

b. The audit report details limitations relating to threats and threat systems. It states that limited funding prevents the military from obtaining new threat systems, and it implies the military has "an inadequate quantity" of threats. If "inadequate quantity" means not having all possible threat systems, there will never be enough funding, nor will our adversaries ever be cooperative in providing us with their newest equipment. The services have a process to prioritize and acquire the needed threat systems. In addition, AFOTEC has proposed an Electronic Combat Test Process to aid in acquiring, validating, verifying, and accrediting threat systems. The audit report fails to mention this, or discuss other military efforts being made to address threat shortfalls with (less costly) modeling and simulation (M&S). M&S will not solve all shortfalls, but it may help the military re-prioritize their threat needs.

c. The audit report states that plans and reports do not adequately report the impact of test limitations, and it gives several examples. Although DoD, AF, and AFOTEC Instructions and training already emphasize the importance of detailing and reporting test limitations, emphasis can be added to ensure that limitations and their impacts on evaluating various systems are clearly stated. We suggest adding this emphasis is a more appropriate recommendation.

d. It is difficult for users and testers to explain every operational and training limitation that occurs in accomplishing their mission. The audit report does not understand the subjectiveness of these limitations, or why the test plans may not be able to explain the limitations. The audit report (paraphrasing an Army evaluation report) states that "testing with live sensors was not possible because of costs, but then states that not testing with live sensors would have no impact." The audit report next states that "the capability of the system to pass information to the live sensors and the capability and reliability of the sensors to transmit information could not be evaluated." There is not enough information, in the context in which it was given, to determine the validity of either statement. There may not have been a requirement to evaluate the sensors themselves. Many programs have information that is obvious to the testers and the users who will employ the system, but is not easily understood by persons unfamiliar with the system or user need. It may have been a more useful comment to state that the report was not written clearly, rather than state that the system was inadequately tested/evaluated. The auditors need to recognize that explanations of the military missions are available from program documents other than system test plans and reports.

e. The audit report addresses "Reporting Obstacles," and it suggests that policy about data release restricts information about system and test limitations. The reason for the data release policy is that some information is not clearly understood, and it can easily be misinterpreted. Test managers must exercise extreme care when releasing data that can be misunderstood. Also, the report suggests that test managers use "informal guidance ... to present test results as positively as possible..." At AFOTEC, instructions and training stress that reports must be objective. They must present as accurate a picture as possible, but they must not advocate a system in any way. Wholesale conclusions about an organization's policy based upon the frailties of a few is an unfair representation of that policy.

f. The audit report states that "tests are occurring despite threat and test environment limitations." A more correct statement would be "Every test occurs despite threat and test environment limitations and is done in the most credible and realistic manner given the real world limitations." It is not possible to test without limitations. Not all systems are testable in every intended environment. Testers accomplish the most realistic test possible, using the most current user information and the best resources available. Any criticism beyond that is program specific and is not related to policy or instruction issues.

g. **Uncertainty and Risk.** In this section, the audit report reiterates some earlier points. The statements made about F-16 testing in a desert environment are true, but they reveal no problems which can be fixed. If more test ranges and threat emitters were available in a variety of climatic environments, the OTA community would have embraced them and used them. In the meantime, the test community will continue to evaluate systems using the best environments and threats available.

h. **Recommendations: Non-Concur.** The audit report recommends that the Service Secretaries "budget adequate resources to acquire threat systems to overcome threat limitations." Test investment and infrastructure needs are put forth in the planning, programming and budgeting system. They have not received funding because other budget priorities have overridden the need for threat systems.

i. In addition (mentioned by the body of the audit report, but not mentioned in the recommendations), AFOTEC and AF Instructions emphasize the importance of detailing test limitations and their impact on assessing the effectiveness and suitability of systems. Many times test plans and reports do not show impact from test limitations because the testers believe they can make a viable report of a system's effectiveness and suitability based on the testing they can accomplish. OTAs attempt to ensure reports are as clear and understandable as possible. Sometimes clarifications with respect to various limitations only obfuscate the issue.

j. The last recommendation, that DOT&E ensure the TEMPs and test plans detail limitations and their impacts, is better done at a level lower than DOT&E. If the time required for DOT&E to approve a TEMP is already a problem, adding more approval criteria to the TEMP process will only slow the process down. We believe adding emphasis to test plan instructions, to ensure limitations are detailed and impacts are explained, is a better approach to the problem. It would place the emphasis where it needs to be placed: writing the test plans, not approving them.

5. **Finding D, Test Status Reporting (DOT&E Reports).**

a. Overall, this finding impacts DOT&E and does not affect AFOTEC. We have no comments on the recommendations. However, there are a few important points about inferring that DOT&E should monitor and review *all* OT&E programs, not just those that are major or are designated as DOT&E oversight programs.

b. This audit report finding states: "The FY 1994 (DOT&E) Annual Report did not include results of operational testing performed on non-major programs which were not selected for oversight." This is a correct statement. The DOT&E Annual Report is intended to report on major and oversight programs only. The audit report paraphrases Title 10 U.S.C., section 139, paragraph (b)(3): the Director shall monitor and review all operational test and evaluation in the Department of Defense. Unfortunately, they overlooked the Title 10 definition of what the DOT&E annual report should contain: section 2399, paragraph (b)(2) states "The Director shall analyze the results of the operational test and evaluation conducted for each ***major*** defense acquisition program" (emphasis added). DOT&E monitors and reviews all ot&e, but only reports on major systems and those designated for DOT&E oversight.

c. If DOT&E were to report on all programs it would tremendously increase their workload. This would add needless oversight to *hundreds* of smaller programs currently managed by OTAs and Using Commands. Adding these programs to the DOT&E

Oversight list would result in major bottlenecks. DOT&E would by duplicating functions already performed by OTAs and the Using Commands. If there is any motivation to accelerate the acquisition process, adding additional layers of oversight does not seem prudent, necessary, or useful.

d. **Recommendations: No Comment.**

6. **Finding E, Reporting Program Results (SAR).**

a. This section of the audit report states that Selected Acquisition Reports (SARs), submitted by OSD but written by PMs, do not accurately reflect the test results. It gives several examples of program SARs that contradict test reports, or do not accurately reflect the OT test reports. We cannot verify the allegation based on the few facts cited in the audit report. Although we agree with the recommendations, we have noted some possible errors in the audit report.

b. Before any OT&E report is complete, the effectiveness and suitability of a weapon system is subject to individual judgment. If a system is performing well in some areas, and poorly in others, the question of its effectiveness and suitability reflects the "perspective" of the individual. If the SAR is prepared by the Program Manager, his/her judgment and perspective is going to be influenced by *program* performance. This perspective may account for the difference between the SAR and the DOT&E report, especially if they are produced at different times.

c. The audit report considers a difference between SARs and OTA reports as significant, but this is not always true. The audit report quotes the Army JSTARS GSM SAR: "the system will satisfy mission requirements." It then quotes the OTA Operational Assessment (OA) which concluded that the GSM demonstrated the "potential" to be effective and suitable, despite some limitations in performance. OAs are often conducted to see if a program is ready for OT&E. When an OTA conducts an OA they limit the scope of their report to "potential" effectiveness and suitability. They may base their findings on previous development testing, limited testing, and use of non-representative or properly functional articles. OAs report on program progress, not on system performance. The differences between the SARs and OTA OAs are understandable to those familiar with these documents.

d. The audit report further observed that the Army procured additional units despite not meeting exit criteria for LRIP. This could be a reflection of user need, rather than system performance. If the user need is great enough, or the system represents a significant gain in capability even when performing to less than requirements, the user and decision makers may approve production or deployment.

e. **Recommendations: Concur.** OTAs make great efforts to accurately report system effectiveness and suitability. If this effort is not reported in the SAR, then changes should

be made to ensure accurate or more comprehensive reporting. However, we would urge caution when using OAs or other early reports for definitive conclusions.

7. Overall, we feel that the report has some good points to offer. However, it lacks credibility based on several misunderstandings of test and evaluation and the organizations that accomplish T&E. We welcome any response to these comments. Replies should be directed to our POC for this, Major Doug Forbes at HQ AFOTEC/XRX. He can be reached at DSN 246-5242, commercial 505-846-5242, or FORBESD@P3.AFOTEC.AF.MIL.

GEORGE B. HARRISON, MGen, USAF
Commander

DEPARTMENT OF THE AIR FORCE
WASHINGTON, DC

Office of the Assistant Secretary

23 January 1996

MEMORANDUM FOR AF/TEP

FROM: SAF/FMBI

SUBJECT: DoD (IG) Audit of Operational Testing Performed on Weapon Systems (Project No. 5AB-0031)

Having reviewed the DoD (IG) Audit, we would like the following comments be included in your response to SAF/FMPF. In particular, we have commented on section C.1 under Recommendations for Corrective Action which appears on page 27 of the report.

The Air Force Budget is a balance of many competing resources. Commitment to readiness and sustainibility, a strong modernization program, support of approved force structure, and provisions for quality of life programs must be financed within constrained levels of topline. Adequate funding of threat systems is one of the competing requirements. Through the Air Force corporate resource allocation process, budgeted funding levels for these systems reflect current Air Force priorities.

We regret our delay in providing input to your office and appreciate your willingness to include our comments in your response. Please direct any questions to me at 614-4996.

MICHAEL J. NOVEL
Acting Director Budget Investment
Deputy Assistant Secretary (Budget)

Attachment:
SAF/FMPF Memo, 25 Jan 95

Audit Team Members

This report was prepared by the Acquisition Management Directorate, Office of the Assistant Inspector General for Auditing, DoD.

Donald E. Reed
Patricia A. Brannin
Raymond A. Spencer
Louis F. Schleuger
Richard L. Collier
Harold R. Tollefson
Susanne M. Williams
Thomas P. Byers
Mary Ann Hourclé
Stacey L. Solomone