



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

National Security and
International Affairs Division

January 5, 1999

The Honorable William S. Cohen
The Secretary of Defense

Attention: DOD Office of the Inspector General
Deputy Director for GAO Affairs

Dear Mr. Secretary:

This is to advise you that we are initiating a review to examine the best practices for test and evaluation of a new product or weapon system. The assignment is one in a series of best practice reviews we have been requested to conduct for the Senate Armed Services Subcommittee on Acquisition and Technology. It is being performed under job code 707401.

The objective of the assignment is to determine whether best practices within the realm of test and evaluation offer opportunities to enhance the outcomes of DOD's weapon system acquisitions. This assignment will require work at the Office of the Secretary of Defense in Washington, D.C.; headquarters and various test facilities in the Army, Navy, and Air Force; and selected weapon system programs. We will contact defense contractors and commercial firms noted for their best practices in test and evaluation. We may also conduct work at NASA.

If you have any questions about this assignment, please call Mr. Paul Francis on (202)512-2811 or Rae Ann Sapp at (937)258-7908.

Sincerely yours,

Katherine V. Schinasi
Associate Director
Defense Acquisitions Issues

*Cluska Davis
Naval Inspector Gen (DC.)
(202) 433-4534*

The General Accounting Office is performing an ongoing body of work for the Senate Armed Services Committee, Subcommittee on Readiness and Management Support. This body of work is focussed on studying the best practices in commercial industry and government over a wide variety of topics with the intent of recommending improvements to weapon acquisitions. These topics have included earned value management, supplier relationships, transition to production, technology inclusion, and training. The objective of our current review is to study the best practices in testing and evaluation of commercial and military programs. We have visited several leading commercial firms and discussed the practices and methods they employ in their product testing. The next portion of our review entails studying the testing and evaluation practices of several military programs across the services.

We have selected the SLAM-ER as one of these programs. In general, our approach will be to capture the early test approach and history of the program during PDRR, understand lessons learned, and see how the subsequent EMD and operational testing reflects those lessons. The following questions were formulated to identify various events that have occurred in SLAM-ER's testing and evaluation process and we would like them to serve as a basis for discussion.

Test Approach and Techniques

1. Was modeling & simulation used in early stages? How was it used?
2. When did testing begin on subsystems and full system?
3. When did M&S begin?
4. Were revisions made to the test plan based on findings from M&S?
5. Is end-to-end simulation being used?
6. Were statistical techniques used to maximize knowledge capture (eg: design of experiments)? If so, please describe the techniques that were used and subsystems they were used on.
7. What benefits were derived from using these techniques?
8. Was there a non-penalty ("no-fault") environment used in testing?
9. Was hardware-in-the-loop testing used?
10. Was planned hardware and software ground testing completed prior to full-system testing? Was the ground testing adequate?
11. Was an incremental (step-up) testing process used? For example: component – subsystem – full-up testing.
12. Was the testing process event driven vs. schedule driven?
13. Who directed and managed the DT test program, the contractor or the government? What are the implications of contractor-led testing?
14. Were any key development test events reduced, delayed or dropped?
15. What subsystems became "long poles" in the tent?
16. What was the date of CDR and what was the number of ECO's submitted after CDR and after major flight tests?

Test Productivity

1. Was there sufficient time to analyze test results before additional testing was performed?

2. Were test objectives and criteria clear?
3. Did the test staff have authority to make changes to the program?
4. Were tests performed realistically in view of the anticipated envelope?
5. Were product releases managed to allow for efficient alignments between hardware and software?
6. Did test articles possess desired maturity/functionality? (See a. thru c. below.)
 - a. Was the software at the planned maturity level at the time of the test? What was the delta between the planned and actual s/w releases?
 - b. Was the hardware at the planned maturity at the time of the test?
 - c. Did the configuration tested match the end item?
7. Was a "closed loop" failure analysis process utilized for anomalies? Who tracked them?
8. To what extent were COTS parts/components used? What testing was needed on these items?

Effect of External Factors

1. Were testing assumptions (ground rules) realistic?
2. Were test plans (schedules) realistic?
3. Did the product exhibit design stability? For example: were there significant changes to the design after critical design review?
4. Did the program have funding stability?
5. Did the program exhibit stability in personnel?
6. Was there early and continuous customer involvement?
7. Were the mission requirements clear and stable?
8. Were test tools available and operating effectively?
9. What was the nature of external oversight on the DT test program? Was it effective?