United States General Accounting Office

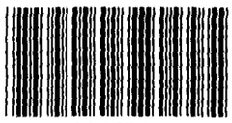
GAO

Report to the Chairman, Subcommittee on Projection Forces and Regional Defense, Committee on Armed Services, U.S. Senate

August 1991

SUBMARINE COMBAT SYSTEM

BSY-2 Development Risks Must Be Addressed and Production Schedule Reassessed



GAO/IMTEC-91-30

GAO

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

Information Management and
Technology Division

B-242594

August 22, 1991

The Honorable Edward M. Kennedy
Chairman, Subcommittee on Projection Forces
and Regional Defense
Committee on Armed Services
United States Senate

Dear Mr. Chairman:

The Navy's SSN-21 Seawolf nuclear attack submarine is designed to be the world's most advanced attack submarine. To perform its mission, the Seawolf will depend on its AN/BSY-2 (BSY-2) automated combat system to detect, classify, track, and launch weapons at enemy subsurface, surface, and land targets. The BSY-2 combat system is the largest Navy Ada[1] software development project ever undertaken, requiring over 2 million lines of new Ada computer code to be developed, integrated, and tested with another 1 million lines of code written in other programming languages. BSY-2 development is being driven by the Seawolf submarine schedule; the first Seawolf is to be delivered to the Navy in 1995. The Navy currently has three BSY-2 combat systems under development.

At your request, we identified technical risks in the combat system's software and hardware development, including design, testing and integration, and the use of independent verification and validation (IV&V). A detailed explanation of our objectives, scope, and methodology is contained in appendix I. Department of Defense comments are contained in appendix II and are discussed throughout the report.

Results in Brief

The risks that the Navy has allowed in the development of its BSY-2 combat system are serious and must be addressed. This combat system, which will largely dictate the effectiveness of the SSN-21 Seawolf attack submarine, is a mammoth software engineering challenge, costing $1.4 billion for development and production of just the first three systems and involving over 3 million lines of code.

In its endeavor to meet BSY-2 delivery schedules, tied closely to the submarine's delivery, the Navy is not following some sound management

[1] Ada is a relatively new computer language used by the Department of Defense that encourages the use of modern software development methods to reduce costs.

principles and practices, and is pushing forward not only with development of the first three systems but also for approval of three additional systems. By doing so, the Navy could find itself with combat systems that fall short of their promised capability and could cost millions to enhance.

In particular, the Navy has (1) allowed an already demanding BSY-2 development and testing schedule to be further compressed, (2) not ensured that important software design tasks are complete, (3) based system development and production approvals on incomplete test and evaluation results, (4) limited its visibility over a significant portion of software component testing, (5) not ensured that retesting of critical software components will be adequate, (6) not ensured that central guidance on unit-testing has been provided to programmers, (7) not addressed early indications of processor and capacity problems, and (8) just begun to assign IV&V the priority it deserves.

We believe that the Navy must address these risks and adjust the BSY-2 development approach as warranted. Further, the Congress should not fund any additional systems until the Secretary of Defense certifies that the BSY-2 combat system is sufficiently developed and tested.

This report and our recommendations specifically address the BSY-2 combat system for the Seawolf attack submarine. We are currently reviewing other aspects of the submarine.

Background

The successful development of a fully capable BSY-2 combat system is critical to the Seawolf achieving its mission requirements. The submarine is intended to counter the Soviet Union's new generation of quieter, more capable submarines by being quieter, deeper diving, and tactically faster than previous submarines while providing other enhanced capabilities. BSY-2 is being designed to (1) enable the submarine to detect and locate targets more quickly, (2) allow operators to perform multiple tasks and address multiple targets concurrently, and (3) ultimately reduce the time between detecting a threat and launching weapons. The BSY-2 combat system and the Seawolf submarine are being developed concurrently, but are managed by different Navy program offices and built by different contractors. Costs for development and production of the first three BSY-2 combat systems are estimated at $1.4 billion.

The first Seawolf submarine is scheduled for delivery in May 1995. In order to meet this timetable, the Navy contracted with the General Electric Company in March 1988 for full-scale development and low-rate initial production of three BSY-2 combat systems. Under the Navy's approach, BSY-2 development and production of three systems will be accomplished concurrently. The first combat system will be delivered for installation on the first submarine in November 1993 with all of the computer hardware and 86 percent of the needed software. The remaining software is to be completed by November 1994 for installation by January 1995. The second and third combat systems will be used to develop, test, and integrate this software and will eventually be installed on the second and third Seawolf submarines.

Before all software and hardware are developed in November 1994, the Navy plans to contract for up to three additional BSY-2 combat systems at an estimated cost of about $900 million. Navy officials stated it is necessary to procure these three additional systems before completed development of the first system, in order to field submarines as soon as possible to meet the advanced Soviet submarine threat.

Building BSY-2 Is a Very Difficult Software Engineering Challenge

Building BSY-2 is an inherently difficult job. Developing, integrating, and fully testing BSY-2's estimated 3.2 million lines of code is an immense undertaking, further complicated in that over 2 million lines of code will be written in the relatively new Ada programming language. According to the May 1990 Ada Joint Program Office survey of 529 Ada projects, the lines of code in the BSY-2 program are exceeded only by the Air Force's Advanced Tactical Fighter.

Because there are not many experienced Ada software engineers and programmers, contractor software personnel will need to be trained to obtain the skills necessary to develop the system. According to Defense and contractor officials, the shortage of Ada-experienced personnel is due to the relative newness of the language, which entails a new approach to software development.[2] To offset this shortfall, contractor staff are receiving Ada training from commercial vendors. However, classroom training cannot substitute for the experience gained from building an Ada system. Even with classroom training, Ada experts have stated that becoming truly productive in the use of Ada can take 4 to 6 months of experience.

[2]Ada facilitates the use of software engineering principles intended to decrease maintenance costs and improve reliability of software.

Defense recognizes that developing BSY-2 is a difficult software engineering challenge, but believes that the use of Ada does not require additional skills and experience. Defense cites the BSY-2 contractor's Ada training program and its development of detailed design documents as initial successes.

In the area of training, for example, Defense notes that personnel completing the contractor's Ada software training program consistently scored above the 75th percentile on a nationwide test. However, this is very misleading since these exam results reflect the third time the students took the same proficiency exam containing the same questions. Furthermore, the exam measures only the student's knowledge of coding programs in the Ada language, and not the student's knowledge of the software engineering principles supported by Ada. Expert software developers that we spoke with, including officials from the Navy's Ada Joint Program Office, stressed that Ada training requires instruction on the Ada programming language as well as the software engineering principles it supports. The wise use of these principles is key to the effective use of Ada.

Even with training, developers will need substantial experience before they become proficient and productive in the use of Ada. Defense's Ada Adoption Handbook states that developing a large Ada system without Ada-experienced personnel represents a substantial risk to successful development.

Tight Development Schedule Is Compressed Further by Delay in Critical Design Review

Early in the program, Navy and BSY-2 contractor officials recognized that development, integration, and testing of a combat system with over a hundred processors and millions of lines of computer code within the original 4-year development schedule would be difficult. The challenge has been exacerbated by a 1-year delay in the Critical Design Review (CDR).[3] According to contract requirements, CDR must be completed to establish a detailed design baseline for all hardware and software prior to the full initiation of software coding. As a result of this delay, the amount of time available for developing and testing the system has decreased about 25 percent, from the original schedule of over 4 years to just over 3 years.

[3]The purpose of CDR is to review the detailed design, test description, and operation and support documents with the contracting agency and to demonstrate to the agency that the detailed design satisfies requirements.

CDR was originally scheduled to be completed in August 1989, but was not considered complete by the Navy until over a year later, in September 1990. Contractor officials attributed the delay to an overly optimistic schedule—the contractor needed more time to finish required system engineering.

Some Design Tasks Not Completed at CDR

Navy and contractor officials stated that CDR was completed in September 1990, and subsequently the contractor began coding and hardware development. However, certain detailed design tasks, contained in Defense Standard 2167[4] and incorporated into the contract, were to be completed prior to CDR, but were still ongoing in November 1990. For example, detailed design descriptions for the software that will control the Seawolf's weapons, or the data base management system that will store critical navigation and target data used for launching these weapons were not completed in September 1990. These descriptions are critical because they provide a blueprint for programmers to code the system. Coding done prior to completion of these tasks will have to be modified if the code does not conform to the eventual detailed design description.

Standard 2167 also requires the review of component test descriptions at CDR, including test input data, expected output data and results, and criteria for evaluating test results. These reviews help provide assurance that the tests will be adequate and will meet Navy requirements. However, plans for reviewing test descriptions of software components at CDR, to ensure that they are consistent with the software test plan, have been dropped from the contractor's Master Test and Evaluation Plan. Contractor and BSY-2 program office officials could not explain to us why the planned reviews were dropped.

Defense does not believe that the delay in completing BSY-2's CDR adds to system development risk. Defense states that CDR was not a one-time event, but consisted of incremental document submissions and reviews, which had minimal impact on the overall development schedule. Defense suggests that extending the amount of time to complete CDR reduces technical and schedule risk by establishing a better understanding of design and implementation requirements.

[4]Department of Defense Standard 2167, Defense System Software Development, June 4, 1985.

We recognize that Defense needs to take whatever time necessary to complete CDR satisfactorily. However, the original schedule for developing, integrating, and testing the system, which Navy stated would have been difficult to meet, has been compressed. Since the original 6-1/2-year schedule for combined design and development of the combat system has not changed, but the design period has been extended, the amount of time allotted for development—coding, integrating, and testing the system—has decreased from the original 4 years to just over 3 years. Although Defense argues that CDR was conducted incrementally for various system components, the bulk of the design reviews were not completed until mid-1990 or later, instead of by August 1989 as originally scheduled. Furthermore, these reviews resulted in hundreds of comments on and revisions to the BSY-2 detailed design and support documents.

Finally, Defense commented that complete detail design descriptions and reviews of software test descriptions are not required at CDR, and that the BSY-2 data base design is considered on schedule with all required documentation delivered to the Navy. However, Defense's comments are not consistent with its agreements. The BSY-2 contract's statement of work requires the contractor, at CDR, to (1) establish "the complete, modular, lower-level design" for each software component, and (2) produce a software detailed design document for each software component.[5] Further, the contract provides for a review of the software test descriptions for each software component. In its comments, Defense states that detailed design documents for the weapons system were redelivered, but did not state whether they have been reviewed, completed, and accepted. While we sought clarification of Defense's statement and requested documentation to ascertain the status of the detailed design documents, Defense has yet to provide either.

Shortcomings in Testing and Evaluation Could Jeopardize Successful System Development

Adequate test plans and early system testing are critical for ensuring that problems are identified and corrected. However, in several areas testing and evaluation have either been reduced or will occur late. This significantly increases the risk that design flaws and errors will not be detected until very late in system development, when they will be considerably more difficult and expensive to correct.

[5]These provisions are contained in Defense Standard 2167, and are contractually required in appendix D of the statement of work.

Navy BSY-2 Development and Production Decisions Not Supported by Thorough Evaluation

To reduce risk and uncertainty before resources are committed, Defense policy[6] requires that systems undergo test and evaluation before moving into either full-scale development or low-rate initial production. To support these milestone decisions on BSY-2 the Navy used, as allowed by Defense policy, assessments to evaluate its potential operational effectiveness and suitability.[7] These assessments consisted of analyses of (1) computer modeling and simulation of the system; (2) program documents such as system requirements, engineering proposals, and design specifications; and (3) available subsystem test results.

At the request of the Chief of Naval Operations, the Navy's Operational Test and Evaluation Force (OPTEVFOR) attempted, in October 1987, to perform an assessment on BSY-2 before it proceeded into full-scale development and low-rate initial production. However, due to the lack of validated system models and simulations, OPTEVFOR was unable to estimate the potential operational effectiveness and suitability of BSY-2 against the threat for which it is being designed. Instead, OPTEVFOR was only able to determine that BSY-2 was potentially more effective than predecessor systems, but could not say how much more effective. Nonetheless, the Office of the Secretary of Defense and the Navy approved BSY-2 to proceed into full-scale development and low-rate initial production; the full-scale development contract was awarded and production of the first BSY-2 system authorized in 1988.

In March 1990, OPTEVFOR performed another assessment. The assessment consisted mainly of analyzing BSY-2 and the Seawolf in mission-representative scenarios using computer simulations to determine if the systems met the threat. OPTEVFOR reported that BSY-2 was potentially operationally effective, but was unable to project whether it would be potentially suitable for satisfactory use in the field. Nonetheless, this assessment was used to support the Navy's decision to request funding for producing the second and third BSY-2 systems in 1991.

The significance of proceeding with BSY-2 full-scale development and low-rate initial production without a complete assessment of potential operational effectiveness and suitability to support the milestone decisions cannot be overstated. Defense may spend hundreds of millions of

[6]Directive 5000.3, Test and Evaluation, Mar. 12, 1986.

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

dollars on a system that may never work or does not perform any better than existing systems. At best, costly changes and retrofitting may be necessary, as has happened with other Defense systems. For example, due to insufficient information on the F/A-18 system, OPTEVFOR was unable to assess the potential operational effectiveness of the system prior to low-rate initial production. However, low-rate initial production was nonetheless approved, and problems found during subsequent testing have required millions of dollars in changes and retrofitting.

Defense stated that BSY-2 testing performed to support development and production decisions is in full compliance with its established policies and procedures. Defense Directive 5000.3 requires "demonstration of a system's technical capabilities and its operational effectiveness and suitability" before advancing from one acquisition phase to another. To comply at this early stage of system design, the Chief of Naval Operations instructed OPTEVFOR to perform an assessment of BSY-2's potential operational effectiveness and suitability. But OPTEVFOR reported that it could not assess the system's potential operational effectiveness and suitability due to the lack of detailed information on the system. Nonetheless, the Navy proceeded into full-scale development and low-rate initial production.

Late Testing May Impair Navy's Early Visibility Into Contractor's Software Testing

For the BSY-2 effort, the Navy expressed its intentions to maintain early government visibility over the contractor's software component testing efforts. To achieve this early visibility and its benefits, the Navy included a contract requirement that government-witnessed tests of software components be conducted as early as possible, but no later than 3 months after the component's delivery to the contractor's test facility for integration. According to Navy officials, this visibility is intended to ensure that problems are identified and fixed as early in development as possible, when corrections are more cheaply and quickly made. Later changes are often difficult and expensive to make because hundreds of thousands of lines of code of different software components have already been integrated.

The Navy may not achieve its goal of early visibility over component testing. Eighteen, or about 15 percent, of the 121 major software components[8] will be integrated with other components before undergoing government-witnessed testing late in development. Government-witnessed

[8]BSY-2's 3.2 million lines of code have been divided into 121 major software components. Each of these major components consists of many subcomponents.

testing for these components is not planned to occur until the last year of system development, right before the system is delivered to the Navy. At this stage, code has already been written and integrated, negating much of the benefit of government-witnessed testing. Government-witnessed testing at this late point leaves the Navy little time and flexibility to identify specific problems and bring them to the attention of the contractor for resolution without affecting the delivery of the system.

Defense stated that the BSY-2 software development and testing approach has been developed to promote government visibility over contractor efforts. The Navy said the last 18 major software components will undergo lower-level testing (subcomponent) and will allow for Navy visibility. However, while the contract requires government-witnessed component acceptance testing,[9] no contract requirement exists to ensure visibility into subcomponent testing. While contractor officials stated that they will allow the Navy to witness subcomponent tests, they stipulated that if time runs short or development becomes difficult, they will no longer allow Navy to witness testing. If this happens, Navy's discovery of serious problems may be delayed.

Navy's Software Testing Approach Does Not Ensure Adequate Retesting of Major Components After Changes

The Navy's contract with the system developer requires that if 5 percent or more of the lines of code in any software component are changed after government-witnessed testing, the entire component must be retested. Navy officials stated that this 5-percent threshold is intended to ensure that significant changes are retested to verify that the changes do not impair the functions of the component. However, under this approach, 49 of the 121 software components could each have changes to 1,000 or more lines of code and not be retested before being integrated with hundreds of thousands of lines of code from other software components. With a system as complex and important as BSY-2, prudent management dictates that the amount of retesting should be based on the criticality of the software component being affected rather than an arbitrary percentage of the component's size.

Contractor officials stated that the 5-percent threshold was negotiated with the Navy because of the time and expense required for retesting. The officials, however, recognized the need to do additional retesting and stated they would perform some retesting on changes under the 5-percent threshold, but that the decision to retest would be based on the

[9]The software components are made up of several levels of subcomponents. A software component cannot undergo acceptance testing until all subcomponents have been tested and integrated.

contractor's judgment and made on a case-by-case basis. Thus, the Navy lacks adequate assurance that the contractor's retest efforts are sufficient to ensure that changes do not introduce adverse effects.

Our concern is not whether software changes will be tested, but whether the retesting is adequate to verify that other portions of the component are not adversely affected by the changes, particularly for those components that are critical to the performance of the BSY-2 system. Experience has shown that testing is frequently truncated when attempting to meet demanding schedules. BSY-2 has already had a 1-year slip in completing CDR. We are, therefore, concerned whether there will be sufficient time available for BSY-2 software integration and testing.

Central Guidance on Unit-Testing Lacking

Unit-testing of software is carried out by programmers as the code is developed to ensure that errors are identified as early as possible, when it is easiest and least expensive to correct them. According to software development experts, central guidance on unit-testing is important to ensure that programmers adequately test smaller software components before the components are integrated with others to form larger components with thousands of lines of code.

For BSY-2, central guidance on unit-testing is even more important given the compressed development schedule resulting from the 1-year delay in completing the CDR and the fact that many of the programmers will be writing Ada code for the first time. However, the Navy lacks assurance that the integrated units will be adequately tested before integration, since neither the contractor's Software Standards and Procedures Manual nor the contractor's Ada training program provides detailed procedures to be followed in carrying out unit-testing. The contractor's Software Development Plan leaves unit-testing procedures up to the discretion of the various development organizations. Without central guidance, the contractor and Navy lack assurance that unit-testing will be consistently and comprehensively conducted.

Defense recognizes the importance of unit-testing guidance, and stated that such guidance is provided in the Software Test Plan Style Guide and Software Development Plan. The guidance cited by Defense provides general test objectives for software developers to follow, but does not provide detailed procedures to be followed in conducting unit-testing. The cited guidance states that these objectives are only to be used where applicable and that the tests are to be conducted informally. For example, the guidance states that testing should use expected, as

well as extreme and erroneous values, but provides no detailed procedures on what specific tasks should be carried out during unit testing to test all such values. The objectives also state that each line of code in the unit should be executed during the test, but no tools and procedures for doing this are provided.

In its 1990 assessment of BSY-2 software development, the Institute for Defense Analyses (IDA) also expressed concern about the lack of detailed unit-testing guidance. IDA found that the contractor's general test objectives provided little direction on what activities should be performed during unit testing, and recommended that tools and techniques be established to ensure that unit tests were adequately conducted to detect errors as early as possible.

We maintain that detailed unit-testing guidance is needed to ensure that units will be appropriately and adequately tested prior to integration, especially since the amount of time remaining for coding, testing, and integration has been significantly reduced and many of the software developers will be newly trained in Ada. As IDA reported, without such detailed guidance there is a substantial risk that errors will be detected late in software development, will be costly to correct, and may severely affect BSY-2 development.

BSY-2 Could Experience System Performance Problems

Adequate system performance is critical to ensuring that the BSY-2 system provides all required functions and capabilities within time constraints specified by the Navy. However, even though it is still early in system development, there are already indications of potential performance problems. The contractor is experiencing problems in meeting start-up requirements in using a new standard Navy signal processor, and preliminary estimates indicate that processing capacity may be inadequate.

Enhanced Modular Signal Processor Does Not Meet Timing Requirements

The Enhanced Modular Signal Processor (EMSP) is one of the most critical components of the BSY-2 system in that it processes the bulk of the acoustic sensor data for transfer to other processors and operator displays. EMSP is being developed by another contractor as a standard signal processor for many Navy programs, and will be provided as government-furnished equipment to the BSY-2 contractor. EMSP is technologically complex. The Navy has recognized that EMSP poses considerable development risk, and stated in its 1988 Acquisition Plan that the risk would be mitigated in BSY-2 by using EMSP first in other Navy systems.

However, BSY-2 is now the first and only planned user of the initial version of EMSP, called SEM-B. A more advanced version of EMSP, called SEM-E, is scheduled for use in several Navy systems.

Currently, the SEM-B version of EMSP is experiencing development problems in meeting initialization (start-up) times for BSY-2 under several operational modes. For example, preliminary modeling estimates indicate that SEM-B takes at least 40 seconds to initialize, whereas Navy requires 20-second initialization. SEM-B's slow start-up time impairs BSY-2's ability to meet other critical system-wide initialization requirements. As a result, the start-up of the entire BSY-2 system under the self-protect mode[10] takes about 3 times longer than the Navy requires.

According to BSY-2 program officials, the upgraded SEM-E version of EMSP is planned to provide faster start-up, but is still under development and has not been proven to meet all BSY-2 requirements. Since SEM-E will not be available, the first three BSY-2 systems will use SEM-B at a cost of $45 million. The Navy plans to replace the SEM-B versions with SEM-E at the first three submarines' overhaul, about 6 years after they have been in operational use. The Navy plans to use SEM-E for additional BSY-2 systems.

Defense stated that use of EMSP in BSY-2 was required after the EMSP design was completed, and therefore the stringent BSY-2 start-up time requirements were not specified in the EMSP design. They added that through modifications the Navy has reduced the start-up times by a factor of four.

We recognize that various factors have contributed to this situation and that Navy has taken efforts to reduce the times, but we maintain that EMSP timing is a significant development risk. We reiterate, and the Navy does not dispute, that EMSP, even after modifications, still does not meet start-up requirements. This, in turn, could impair the ability of the entire BSY-2 system to meet critical system-wide start-up requirements. Due to the highly automated nature of the combat system, its ability to quickly restart after a failure is critical to accomplishing its mission.

[10]The self-protect mode consists of minimum defensive functions needed to protect the submarine from damage by hostile forces, but does not include all mission capabilities.

Estimates Show Processing Capacity Not Meeting Design Goals

The BSY-2 combat system needs sufficient spare processing capacity to facilitate growth to meet future work-load needs. The Navy has set standards[11] on spare capacity and maximum system utilization to avoid past problems with systems that could not process growing work loads. To ensure that the system meets these standards when delivered to the Navy, the contractor is estimating software component size and performance during design, and is using analytical models to predict how the system will behave when delivered. However, based on the analytical results thus far, over one-half of the processors do not meet the contractor's goal of not exceeding 50-percent spare capacity.

BSY-2 and the Seawolf submarine are intended to serve for decades, during which time multiple upgrades to the sensors and combat system will surely be developed and implemented. The Navy requires that delivered systems have at least 20-percent spare capacity (no more than 80-percent utilization), so that work-load growth can be accommodated. According to the Navy standards, failure to provide sufficient spare capacity can result in the delivery of systems that are too small or too slow, necessitating costly reprogramming or additional computer processors or upgrades.

To ensure that the system has sufficient capacity at the time BSY-2 is delivered to the Navy, the contractor's Systems Engineering Management Plan established as a design goal a 50-percent processor utilization limit throughout the development of BSY-2. Contractor officials stated this more stringent threshold reflects the uncertainty in early size and performance estimates, and in the results of analytical models. They explained that this conservative approach is a prudent way of ensuring adequate processing capacity at delivery. However, when CDR was completed in September 1990, processor utilization estimates in the contractor's Software Development Plan showed that over one-half of the processors were already exceeding the 50-percent threshold.

Moreover, based on the stringent physical and timing constraints and the complexity of the BSY-2 system, 20-percent spare capacity at delivery may be inadequate to accommodate change and growth in work load. Officials in Navy's Space and Naval Warfare Systems Command, who issue the standards, emphasized that the 20-percent spare capacity standard was a minimum, and that greater spare capacity should be provided under certain circumstances, such as when building particularly complex systems. BSY-2 is clearly an extremely complex system. Further,

[11] Department of Navy, Tactical Digital Standard (TADSTAND), July 2, 1980.

after the BSY-2 full-scale development contract was awarded, the Navy changed the standard to require 50-percent spare capacity at the time of delivery, instead of 20 percent. However, BSY-2 is subject to previous Navy standard to provide 20-percent spare capacity because the contract predated the change in standards, and the wisdom of requiring only 20-percent spare capacity has not been reassessed. If the delivered BSY-2 system contains processors with inadequate spare capacity, the Navy will be faced with expensive options, such as upgrading the processors or redesigning the system.

Defense stated that BSY-2 has sufficient spare capacity and that we have no basis to conclude that capacity may be insufficient, and that the current BSY-2 design will allow for 30 years of expected functional growth. Defense contends processing capacity is not a concern for BSY-2.

We do not conclude that BSY-2 reserve capacity will not meet expected requirements. Our concern is that even as early in the development process as CDR, processing capacity is not meeting the contractor's design goal to maintain 50-percent reserve capacity throughout system development—a goal that the contractor called prudent in light of uncertainties about the system's eventual size and performance when fully developed and the models used to project both. Documentation provided to us by the program office showed that at CDR completion, the majority of the BSY-2 data processors were below the contractor's goal of 50-percent estimated reserve capacity, bringing into question at this early stage of development the ability of each of the BSY-2's processors to achieve 20-percent reserve capacity at delivery. Further, since it is the contractor's goal to have capacity not exceed 50 percent throughout development, Defense's statement that it is expected and acceptable for loading limits to increase beyond this figure seems contradictory.

Defense claims an average of 63-percent overall spare capacity for all processors; it is unclear, however, how Defense computed this figure. Defense discusses extra spare processors, unused space in computer cabinets, and room for additional cabinets that could be used to provide extra capacity. However, Navy officials told us that not all of the processors may be able to use other processor's reserve capacity, since they may not be interconnected. Also, some processors may not be available to provide reserve capacity because they are already dedicated as back-up in case of processor failures in order to meet reliability requirements. Moreover, additional cabinets are intended to accommodate planned future functions and thus may not be available to provide spare capacity to handle system growth during development or other

unplanned growth. Our point was not to prove that capacity is insufficient, since the system has not yet been built, but only to show that reserve capacity goals for individual processors were largely not being met, which is a cause for concern.

Independent Verification and Validation Plan Developed Late

Effective IV&V planning used throughout system design and development is intended to significantly decrease development risks. Verification tasks help ensure the accuracy of the specifications, requirements, and design, while validation tasks are later performed to confirm that the software products perform efficiently and effectively and comply with requirements. IV&V is commonly used to minimize risks on critically important large-scale system developments, where loss of life or other catastrophic events could occur. IV&V provides an independent, third-party analysis of potential software problems throughout development.

Even though the BSY-2 program had been in full-scale development for 2-1/2 years and passed through four of the eight major program review milestones, an IV&V management plan was only developed in June 1990. The Navy had developed prior preliminary IV&V plans, but they were discarded. Thus, early BSY-2 development activities did not benefit from effective IV&V oversight.

The new management plan is a step in the right direction, as it describes an IV&V scope of effort consistent with federal IV&V guidance, such as the Federal Information Processing Standards and Air Force guidance. The effectiveness of the plan will depend on Navy's continued commitment to it and a willingness to act on the findings and recommendations of the IV&V effort. Further, as with any effort, adequate funding for the IV&V program is critical for its success. The level of IV&V funding is negotiated each year between the Naval Underwater Systems Center, which provides technical support to the BSY-2 program, and the BSY-2 program office. The Navy needs to ensure that if BSY-2 funding becomes tight or is required for other system development needs, IV&V tasks are not reduced.

Defense concurred that in the past an IV&V agent was lacking, but stated that verification and validation activities that either met or exceeded Federal guidance were nonetheless performed. Defense stated that an IV&V plan has now been developed and implemented. Finally, Defense disagreed with our statement that prior IV&V plans were discarded, instead maintaining that they were revised.

The objective of IV&V is to conduct an independent third-party analysis of potential system problems throughout development. Without an independent organization performing this function—which the Navy acknowledges it lacked for the first 2-1/2 years of BSY-2 development—the Navy has only minimal assurance that all problems and risks were effectively identified and assessed.

Navy had no IV&V plan before March 1989. At that time, we reported the lack of IV&V efforts,[12] and the Navy wrote a one page outline of an IV&V plan. Subsequently, the BSY-2 Program Office wrote a preliminary IV&V plan that lacked sufficient detail on IV&V functions and tasks. However, the Navy never approved the plan, which according to BSY-2 program officials, was discarded in favor of a different, detailed IV&V plan that was developed in June 1990. The 1990 plan includes principles and specific tasks expected in a detailed IV&V plan, such as analyzing code to verify adherence to coding standards and independently testing code. Continued Navy commitment to the IV&V program will greatly enhance the likelihood of BSY-2 program success.

Conclusions

Despite difficulties and delays encountered in building BSY-2, the Navy has not substantively altered its original approach or schedule. We believe that the Navy's approach to developing BSY-2 is flawed, which could result in the BSY-2 combat system not meeting its development schedule or satisfying all the mission requirements that the Navy has specified. Should the delivered combat system fail to meet these requirements, significant cost increases could be required to acquire the missing capability.

We recognize that every system development effort entails some risk. However, the potential impact of these risks on system performance and costs warrants additional Navy actions to ensure that the contractor proceeds with a prudent development approach that acknowledges these risks and begins to resolve them. In the past, the military services have experienced significant cost overruns, delays, and sometimes delivery of systems that do not provide significant performance improvements over the system they are to replace. Worse yet, the systems sometimes do not work at all and must be scrapped. In our opinion, the Navy must take action to minimize the risks associated with the three combat systems already under contract. Follow-on combat systems

[12]Submarine Combat System: Technical Challenges Confronting Navy's Seawolf AN/BSY-2 Development (GAO/IMTEC-89-35, Mar. 13, 1989).

should not be procured until the Secretary of Defense certifies that the initial BSY-2 system is sufficiently developed and tested.

Recommendations to the Secretary of Defense

We recommend that the Secretary of Defense direct the Secretary of Navy to (1) determine the impact of the risks we have identified on the performance, cost, and current delivery schedule for the combat system; (2) adjust its current development approach, as warranted; and (3) report the results of the assessment and planned actions to mitigate these risks to the House and Senate Armed Services and Appropriations Committees. We also recommend that the Secretary of Defense direct the Navy to ensure that central guidance on unit-testing is developed and issued for use by all BSY-2 programmers and software engineers.

Recommendation to the Congress

We recommend that the House and Senate Appropriations and Armed Services Committees not fund any additional systems beyond the first three until the Secretary of Defense certifies that the initial BSY-2 is sufficiently developed and tested and can provide assurance that the system will satisfy mission requirements.

Agency Comments and Our Evaluation

The Department of Defense disagreed with most of our findings and the recommendations contained in our draft report (see app. II). Defense stated that it understands the BSY-2 development risks, has taken strong action to counter any potential impact, and has conducted a thorough evaluation of the BSY-2 program and found the overall level of risk to be moderate.

We recognize that Defense has taken some management actions to lessen development risk, but significant development risks still exist in several areas and specific actions are needed. Given the potential impact of these risks, we disagree that the overall development risk is moderate. In fact, as part of the Defense Acquisition Board's January 1991 evaluation of BSY-2, the Office of the Director of Defense Research and Engineering concluded that development of BSY-2's 3.2 million lines of software code represented high-schedule risk.

Defense Views on GAO's Recommendations to the Secretary of Defense

Defense disagreed with our recommendations to assess the impact of the various risks and take corrective action. Defense stated the recommended actions were not necessary since they periodically examine the BSY-2 development status and monitor the risk areas to ensure BSY-2 meets

baseline thresholds. Defense also states that it is in full compliance with policies and procedures, and will be reporting on the risks to the Congress as required by section 2400 of Title 10 U.S.C. in the Low-Rate Initial Production report.

Although Defense claims that it examines and monitors BSY-2 risks in its management functions, we have found several areas where significant risks exist and believe that Defense needs to assess the impact of these risks and adjust the development approach as warranted. Defense's Low-Rate Initial Production report may not adequately discuss how Navy plans to mitigate these risks, since the purpose of the report is not to assess risks. As required by section 2400, Defense only has to report on the rate and quantity prescribed for low-rate initial production and considerations in establishing it, a test and evaluation master plan, and an acquisition strategy that includes the procurement objectives in terms of total quantity to be procured and annual production rates. Further, as discussed in this report, unit-testing guidance is inadequate and more detailed guidance needs to be provided.

Defense Views on GAO's Recommendations to the Congress

Defense commented that development and testing is in full compliance with congressional guidance and Defense policies and procedures. Defense stated that completing operational testing before funding further procurement would result in cost growths so large it could greatly exceed the cost of correcting any deficiencies identified during testing.

We have revised our recommendation that operational testing be completed before funding further procurement because of the significant cost impacts Defense claims it could have on the Seawolf. However, we continue to believe funding additional BSY-2 combat systems before they are sufficiently developed and tested introduces considerable risk that cannot be overlooked. Therefore, we believe that BSY-2 production should not be funded until the combat system is sufficiently developed and tested and the Secretary of Defense can provide assurance to the Congress that the system will satisfy all mission requirements.

As agreed with your office, unless you publicly announce the contents of this report earlier, we plan no further distribution until 15 days from the date of this letter. We will then send copies of this report to the Chairmen, Senate and House Committees on Appropriations; the Chairman, House Committee on Armed Services; and the Director, Office

of Management and Budget. Copies will also be made available to others upon request.

This report was prepared under the direction of Samuel W. Bowlin, Director, Defense and Security Information Systems, who can be reached at (202) 275-4649. Other major contributors are listed in appendix III.

Sincerely yours,

Ralph V. Carlone
Assistant Comptroller General

Contents

Abbreviations

CDR	critical design review
EMSP	enhanced modular signal processor
GAO	General Accounting Office
IDA	Institute for Defense Analyses
IMTEC	Information Management and Technology Division
IV&V	independent verification and validation
NSIAD	National Security and International Affairs Division
OPTEVFOR	Navy's Operational Test and Evaluation Force

Objectives, Scope, and Methodology

In response to a request from the Chairman, Subcommittee on Projection Forces and Regional Defense, Senate Armed Services Committee, and as a result of subsequent discussions with the Chairman's office, we evaluated the Navy's development of the SSN-21 Seawolf attack submarine's BSY-2 combat system. Our work focused on the following aspects of AN/BSY-2 software and hardware development: system design, testing and integration, and the use of independent verification and validation. We conducted our work at Navy, contractor, and subcontractor offices having responsibility for the BSY-2 program. Our primary work locations included the Naval Underwater Systems Center in Newport, Rhode Island; the Naval Sea Systems Command and the BSY-2 Program Office in Arlington, Virginia; and at the General Electric Company in Syracuse, New York.

We interviewed numerous Navy, contractor, and subcontractor officials and analyzed relevant BSY-2 documents, including the BSY-2 contract, critical design review data and reports, the System Design Document, the Software Development Plan, architecture modeling reports, the Master Test and Evaluation Plan, and various Navy and Defense policies and regulations. In addition, we interviewed recognized software development experts regarding technical issues associated with BSY-2 and the Ada programming language.

To obtain background information on BSY-2, we analyzed relevant Defense and GAO reports, including the January 1990 report by the Institute for Defense Analysis, Assessment of the Software Development Program for the AN/BSY-2 System; the March 1989 GAO report, Submarine Combat System: Technical Challenges Confronting Navy's Seawolf AN/BSY-2 Development (GAO/IMTEC-89-35); and the January 1990 GAO report, Navy Acquisition: Cost, Schedule, and Performance of New Submarine Combat Systems (GAO/NSIAD-90-72). For information about software and hardware design, we interviewed technical and management officials at General Electric responsible for systems engineering, software development, and testing and integration. We also interviewed Navy Underwater Systems Command and Naval Sea System officials on Navy's oversight roles and responsibilities for the BSY-2 program.

To obtain information on the Ada programming language, we interviewed various Ada experts; Ada training vendors; officials at Carnegie-Mellon University's Software Engineering Institute in Pittsburgh, Pennsylvania; and Defense's Ada Joint Program Office in Arlington, Virginia. We obtained guidelines on management of Ada-based systems from the

Armed Forces Communications and Electronics Association, and analyzed the prime contractor's Ada training program, the Ada Adoption Handbook, and case studies on Ada.

We discussed the use of IV&V and relevant Defense regulations with Defense officials, IV&V contractors, and the Naval Underwater Systems Center. We also analyzed BSY-2 IV&V plans. We conducted our review in accordance with generally accepted government auditing standards, between February 1990 and March 1991.

Comments From the Department of Defense

DIRECTOR OF DEFENSE RESEARCH AND ENGINEERING

WASHINGTON, DC 20301-3010

30 APR 1991

Mr. Ralph V. Carlone
Assistant Comptoller General
Information Management and Technology Division
U.S. General Accounting Office
Washington, DC 20548

Dear Mr. Carlone:

This is the Department of Defense (DoD) response to the General Accounting Office (GAO) draft report, "SUBMARINE COMBAT SYSTEM: Computer System Development Risks Must Be Addressed and Production Slowed," dated March 21, 1991 (GAO Code 510532/OSD Case 8640). The DoD only partially agrees with the report.

The GAO stated that "the risks that the Navy has allowed in the development of its BSY-2 combat system (sic) are serious and must be addressed." In fact, all of the "risks" identified in the draft report were raised by the GAO several years ago. During the past two years, numerous oversight organizations have evaluated those same specific risks and reported that the Department understands them and has taken strong action to counter any potential impact. The oversight reviews endorsed the strong actions taken by the Program Manager to maintain control over the risks.

As part of the process that led to a January 1991 Defense Acquisition Board review on the SEAWOLF submarine and the AN/BSY-2 Submarine Combat System, the Department of Defense conducted a thorough technical evaluation of the AN/BSY-2 program, focusing primarily on software development. The overall level of risk was determined to be moderate. The Defense Acquisition Board determined that the Navy was not "side-stepping some sound management principles and practices," but in fact is employing sound and prudent contractually-invoked risk management approaches in a disciplined software development environment incorporated in DoD-STD-2167.

Further, the GAO voices objections to the program concurrency. Both the Office of the Secretary of Defense and the Congress have recognized the unique requirements in the development of naval vessels (and military satellites) which, by their very nature, involve a high degree of concurrency. As a direct result of that recognition, the Congress passed Section 2400 of Title 10, U.S.C. which permitted exceptions for naval vessels and military satellites in determining the number of ships to be procured during the Low Rate Initial Production period. The BSY-2 system is being developed in full compliance with that legislation.

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

Sincerely,

Charles M. Herzfeld

Enclosure

**GAO DRAFT REPORT - DATED MARCH 21, 1991
(GAO CODE 510532) OSD CASE 8640**

"SUBMARINE COMBAT SYSTEM: COMPUTER SYSTEM DEVELOPMENT RISKS MUST BE ADDRESSED AND PRODUCTION SLOWED"

DEPARTMENT OF DEFENSE COMMENTS

* * * * *

FINDINGS

- **FINDING A: The Difficulty Of Building The AN/BSY-2.** The GAO reported that the Navy is currently developing the SSN-21 SEAWOLF, designed to be the world's most advanced nuclear attack submarine. The GAO explained that to perform its mission, the SEAWOLF will depend on the AN/BSY-2 automated combat system to detect, classify, track, and launch weapons at enemy subsurface, surface, and land targets. According to the GAO, the BSY-2 system is the largest Ada software development project ever undertaken, requiring over two million lines of new Ada computer code to be developed, integrated, and tested, with another one million lines of code written in other program languages.

  The GAO observed that building the BSY-2 is an inherently difficult job, further complicated by the use of Ada, (a relatively new programming language), for over two million lines of code. The GAO reported that, because it is still relatively new, there are not many experienced Ada software engineers and programmers. As a result, the GAO reported contractor software personnel will need to be trained to obtain the skill necessary to develop the system. The GAO found that contractor personnel are receiving Ada training from commercial vendors. The GAO also pointed out, however, that Ada experts have said, even with such training, becoming truly productive in the use of Ada can take 4 to 6 months of experience. The GAO further noted that the DoD Ada Adoption Handbook states that developing a large system in Ada without experienced personnel represents a substantial risk to successful development. The GAO concluded that building the BSY-2 is a very difficult software engineering challenge. (pp. 2-3, pp. 6-7/GAO Draft Report)

Enclosure

- **DoD Response: Partially concur.** The DoD concurs that developing the BSY-2 system is a difficult software engineering challenge. The DoD does not agree, however, with the GAO perception of the risk of using Ada, or the need for additional training.

  Ada does not entail a new approach to software development. It simply forces discipline in implementing recognized, sound software engineering principles. Personnel with previous programming experience can be expected to become quickly proficient in Ada. In fact, 25 to 30 percent of the BSY-2 software personnel actually had previous Ada programming experience. Moreover, the training program established for the BSY-2 has been complete since December 1989. A small number of new hires continue to be trained, as required, under the same training regimen. About 90 to 95 percent of the BSY-2 software engineers are now well past the 4 to 6 month experience point in using Ada mentioned by the GAO. (The Ada Adoption Handbook, cited by the GAO, is a technical publication of the Software Engineering Institute. It is not a formal DOD Military Handbook.)

  The training efforts to date have been successful. Strong points of the training program are that (1) the training included coding exercises derived from BSY-2 system examples, and (2) students consistently scored above the 75th national percentile on a widely administered test for programming proficiency (Psychometrics).

  The delivered software detail design documents consist almost exclusively of Ada or Ada Program Design Language. All Ada software detailed design documents submitted for Critical Design Review have been checked for syntactical correctness by automated tools and many have been compiled successfully, even though that is not required by the contract.

- **FINDING B: Delay In The Critical Design Review Has Further Impacted The Tight Development Schedule.** The GAO reported that early in the program, both Navy and contractor officials recognized that development, integration, and testing of the BSY-2, within the original 4-year development schedule, would be difficult. The GAO found that this has been exacerbated by a one-year delay in the critical design review, from August 1989 to September 1990. The GAO explained that contract requirements indicate that this review must be completed to establish a detailed design baseline for all hardware and software, prior to the full initiation of software coding. The GAO observed that as a result of this delay, the amount of time available for developing and testing the system has decreased about 25 percent, from over 4 to over 3 years.

The GAO also reported that, even though Navy and contractor officials have stated that the critical design review is now complete and the contractor has begun coding and hardware development, certain detailed design tasks, required by DoD policy to be completed prior to the review, are still ongoing. The GAO reported, for example, that detailed design descriptions are not yet complete for the software that will control the SEAWOLF weapons, or the data base management system that will store critical navigation and target data used for launching weapons. The GAO observed that these descriptions are critical, because they provide a blueprint for programmers to code the system. The GAO noted that coding done prior to completion of these tasks will have to be modified, if the code does not conform to the eventual detailed design.

The GAO further reported that DoD policy requires the review of component test descriptions at the critical design review, including test input data, expected output data, and criteria for evaluating test results, to provide assurance the tests will be adequate and meet requirements. The GAO found, however, that plans for reviewing test descriptions of software components at the critical design review have been dropped from the contractor Master Test and Evaluation Plan. The GAO concluded that the failure to complete all required critical design review tasks further increases the BSY-2 development risks. (p. 4, pp. 7-9/GAO Draft Report)

- **DoD Response:** **Nonconcur.** The BSY-2 development, integration, and test has always spanned the period from March 1988 to November 1994. It is and always has been a 6-1/2 year schedule, which has not changed since contract award. Further, development of the BSY-2 has been in full compliance with DoD development and testing requirements.

  The Critical Design Review evolved from a schedule driven activity to an event based activity. To this end it was not conducted as a discrete, one time event. When it became apparent that extra time was required in the design phase (to ensure that a sufficient level of detailed design was accomplished prior to committing to implementation of the chosen design), the review process was modified to permit incremental submittal and review. As portions of the program completed detailed design, they were submitted for review and approval. Approval allowed transition to implementation. The design and review processes specifically were structured to ensure that general design criteria and early critical portions of the program were completed first. The use of Ada programming language and the partitioned nature of the program permitted incremental submittal and review to occur without incurring risk. That approach was a prudent step, considering the extensive

design work that had to be contained in the Critical Design Review design documents resulting from the combined use of both Ada and DOD-STD-2167.

While the incremental approach had the effect of rescheduling the completion date of the Critical Design Review, the impact on the overall development schedule has been minimal. In fact, such a delay typically reduces technical risk by establishing a better understanding of design and implementation requirements. The reduction in technical risk also typically promotes a reduction in schedule risk due to: (1) better definition of requirements, which leads to an improved understanding of testing requirements and, therefore, the test program becomes more structured and focused on verifying performance achievement rather than establishing performance objectives; (2) a disciplined system integration that benefits from clearly defined interfaces; and (3) a net reduction in the number of software errors, failures and faults that could be encountered during integration and test, due to the increased design effort prior to Critical Design Review. In fact, the number of lines of new code to be developed for the BSY-2 has been very stable and actually has decreased. That is a very good trend in a widely recognized indicator of software design maturity. It is also well recognized that additional time spent in the detailed design phase will result in fewer problems in the ensuing integration and test phases.

Contrary to the GAO statements, the DoD-STD-2167 is not a Defense policy document. It is a part of the contracting mechanisms that provide tasking to the contractor. The DoD-STD-2167 is designed to be tailored by the contracting agency to meet the specific needs of the program. It does not require all detailed design descriptions to be complete for Critical Design Review.

Despite the above fact, by the close of the Critical Design Review (September 30, 1990), all BSY-2 Tactical System Software Detail Design Documents were shipped to the Navy. Two Software Detail Design Documents were redelivered on October 9, 1990, with greater levels of detail (10 days after the close of the Critical Design Review). The redelivered Software Detail Design Documents were for the conventional weapon Computer System Configuration Item. The redelivery did not, however, impact the implementation schedule for the weapon Computer System Configuration Items or cause any recoding efforts. Final versions of the Software Detail Design Documents will be delivered with the Software Product Specification, as specified by the contract.

The Database Design Documents for each database site in the tactical system were delivered on time to meet the Critical Design Review milestone. The database design effort has no delinquent documentation and is on schedule. To accommodate the complexity of the BSY-2 system, the scope and content of the BSY-2 Database Design Documents exceed the DOD-STD-2167 guidelines.

The GAO may be referring to the lack of a physical database design in the Database Design Documents delivered for Critical Design Review. The Database Design Documents contain sufficient detail in the area of storage and size requirements, access methods to allow implementation to proceed without undue risk. The details of the physical database design are not required by applications developers when a relational database is used, as is the case for the BSY-2.

Defense policy does not require a review of detailed test descriptions at Critical Design Review. Prior to program tailoring, the DoD-STD-2167 may recommend test plans be reviewed. Software test plans for all Computer Software Configuration Items were reviewed at both the Preliminary Design Review and the Critical Design Review, per DOD-STD-2167.

The Software Test Design document Contract Data Requirements List schedule was revised by agreement between the contractor and the Navy to require delivery of the documents 135 days prior to the start of the respective thread tests. The BSY-2 test schedule is phased over a period of greater than 3 years. Development and review of the documents closer to actual testing will enhance the quality of the Software Test Design and ensure the availability Software Test Design developer for successive tasks, such as the Software Test Procedures development and performance testing. The described approach provides continuity of personnel and reduces schedule risk.

- **FINDING C: Development And Production Decisions Not Supported By Thorough Evaluation.** The GAO reported that, to reduce risk and uncertainty before resources are committed, DoD policy requires that systems undergo test and evaluation before moving into either full-scale development or low-rate initial production. In the case of the BSY-2, the GAO found that, to support these milestones, the Navy used--as allowed by DoD policy--assessments to evaluate the potential operational effectiveness and suitability. The GAO, however, identified several problems with the BSY-2 assessments.

The GAO reported, for example, that in October 1987, the Navy Operational Test and Evaluation Force attempted to perform an assessment on the BSY-2, before it proceeded into full-scale development and low-rate initial production. According to the GAO, however, the Navy Operational Test and Evaluation Force was unable to estimate the potential effectiveness and suitability of the BSY-2 against the design threat, due to the lack of validated system models--instead, it could only determine the BSY-2 was potentially more effective than predecessor systems. Nonetheless, the GAO found that, in 1988, the DoD approved the BSY-2 to proceed into full-scale development and low-rate initial production.

As a second example, the GAO found that the Navy Operational Test and Evaluation Force performed another assessment in March 1990, based mainly on the use of computer simulations. According to the GAO, the Navy Operational Test and Evaluation Force reported the BSY-2 was potentially operationally effective, but it could not project whether the BSY-2 would be potentially suitable for satisfactory use in the field. The GAO found that this assessment was then used to support the Navy decision to request funding for producing the second and third BSY-2 systems in 1991.

The GAO observed that proceeding with BSY-2 full-scale development and low-rate production, without a complete assessment of potential operational effectiveness and suitability to support milestone decisions, cannot be over-stated. The GAO concluded that this could result in the DoD spending hundreds of millions of dollars on a system that may never work, or does not perform any better than existing systems. The GAO also concluded that, at best, it may require costly changes and retrofitting, as happened on the F/A-18 aircraft. The GAO further concluded that these test and evaluation shortcomings should be addressed to better ensure successful system development. (p. 4, pp. 9-12, pp. 21-22/GAO Draft Report)

- **DoD Response:** **Nonconcur.** While the GAO description of the BSY-2 test and evaluation efforts to date is generally correct, the DoD does not agree that DoD development and production decisions are not supported by thorough evaluation. As indicated by the GAO, testing done and assessments performed are in full compliance with established DoD policies and procedures. The GAO report lists no specific test reduction or elimination, so it is impossible to address the criticism more specifically. There has been no reduction in testing and planned Operational Test and Evaluation is on schedule.

The fidelity of an Operational Test and Evaluation assessment is directly related to the maturity of the development system--i.e., until the system has completed development there will always be some system aspects that operational Test and Evaluation cannot assess with a high degree of certainty. That is why decisions are supported by both developmental Test and Evaluation and operational Test and Evaluation. Developmental Test and Evaluation provides the quantitative evidence of progress towards the satisfaction of systems requirements that complements the operational Test and Evaluation assessments. Such complementary Test and Evaluation assessments are especially valuable prior to the Milestone III decision, and were used for the AN/BSY-2 decision process.

The procedures followed for the AN/BSY-2 program by both the Office of the Secretary of Defense and the Navy in evaluating operational effectiveness and operational suitability are consistent with established DoD directives and guidance. The March 1990 assessment conducted by the Commander, Operational Test and Evaluation Force and independently evaluated by Office of the Director, Operational Test and Evaluation, was an extensive effort directed by Under Secretary of Defense (Acquisition) to highlight the strengths and weaknesses of the program as early as possible. That assessment demonstrated a commitment to identify critical issues as early as possible to enhance program stability. It was aimed at identifying those high risk areas in need of management attention. To assign fault to the management of the program because the Operational Test and Evaluation Force was unable to evaluate the program to any greater degree is inappropriate in that the program was in its very early stages and no operational testing had yet occurred. There is no operational test data upon which either the Commander, Operational Test and Evaluation Force or the Director, Operational Test and Evaluation, could make a determination of operational effectiveness or suitability beyond that which has occurred.

Because of the unique nature of shipbuilding, the Congress recognized that a "complete assessment" cannot be made for naval vessels and passed legislation permitting exceptions in establishing Low Rate Initial Production. The SSN 21 program, of which the BSY-2 system is an integral part, is being executed within the guidance provided by the Congress. It was in that light that the recent Defense Acquisition Board decision was made--i.e. determining the program delivered an acceptable balance of cost, risk, and programmatic needs.

The goal of completing operational testing before funding further procurement would result in cost growths so large

(due to the costs of shutting down and re-starting production lines being spread over so few systems), it could greatly exceed the cost of correcting any deficiencies identified during operational testing. To minimize the number of deficiencies that will not be identified until operational testing, the Navy conducted a vigorous program of critical component testing to ensure that key technologies will perform as required. In addition, the Navy will be conducting extensive preliminary testing to ensure subsystems will operate together, including a series of fully integrated shore based tests designed to stress both operators and machines.

- **FINDING D: Late Testing Of Many Software Components.** The GAO reported that, to achieve early visibility over the contractor software component testing efforts, the Navy included a requirement in the contract for Government-witnessed tests of software components. According to the GAO, these tests were to be conducted as early as possible, but no later than three months after the delivery of components to the contractor test facility for integration. The GAO reported that the goal is to ensure that problems are identified and corrected as early in development as possible, when corrections are more cheaply and quickly made.

  The GAO observed, however, that the Navy may not be able to achieve its goal of early visibility over component testing. The GAO explained that about 18 of the 121 major software components will be integrated with other components, before undergoing Government-witnessed testing. According to the GAO, government-witnessed testing for these components is not planned until the last year of system development. The GAO observed that by then, code will have already been written and integrated, negating much of the benefit of government-witnessed testing--little time will be left to identify problems and resolve them, without affecting the delivery of the system. The GAO concluded that such late testing of software components impairs the early visibility of the Navy into contractor software testing efforts. The GAO also concluded that this is another indication that additional Navy actions are needed to ensure the contractor proceeds with a prudent development approach. (p. 4, pp. 12-13, pp. 21-22/GAO Draft Report)

- **DoD Response: Nonconcur.** The BSY-2 software development and testing approach has been developed to promote Government visibility of contractor efforts. The contract requirement for Government-witnessed software tests referred to by the GAO is for Computer Software Configuration Item acceptance testing. A Computer Software Configuration Item

is composed of many Computer Software Components--which, in turn, are composed of many units. Units are developed and tested prior to integration with other units to form Computer Software Components. Computer Software Components are tested and integrated incrementally to form a Computer Software Configuration Item. Thus, before a Computer Software Configuration Item is presented for acceptance testing, multiple levels of test and integration already have been performed.

There are three to four levels of testing witnessed by the Navy prior to a Computer Software Configuration Item acceptance test. Although the contract calls for witnessing Computer Software Configuration Item, Critical Item, and System Design Certification testing, the prime contractor developed a Thread methodology to integrate and test the system on an incremental basis. The Navy has full access to the process. The Navy also has access to Software Development Files, which contain unit and component test results, and to the Program Trouble Reports, which define problems and resolution. System level requirements will be demonstrated for the Government early enough to allow sufficient time to correct any problems.

A Computer Software Configuration Item cannot be presented for acceptance until its components (Units, Computer Software Components) have been tested and integrated. It is the formal acceptance of the last 18 Computer Software Configuration Items that occurs in 1994. All lower level tests will have occurred earlier and all allow for Navy visibility. The current approach represents a prudent and widely accepted method to schedule development and acceptance testing.

- **FINDING E: Navy's Software Testing Approach does not Ensure that Major Components are Adequately Retested After Changes.** The GAO reported that the Navy contract with the system developer requires that, if five percent or more of lines of code in any software component is changed after government-witnessed testing, the entire component must be retested. The GAO noted that Navy officials said this is intended to ensure that significant changes are retested to verify that the changes do not impair functions of the component. The GAO observed, however, that with a system as complex and important as the BSY-2, prudent management dictates that the amount of retesting be based on the criticality of the software component being affected, rather than an arbitrary percentage of the component's size.

  According to the GAO, contractor officials said the five percent threshold was negotiated based on time and cost considerations for retesting. The GAO reported that the

contractor officials also recognized the need to do additional retesting under the five percent threshold, but said the decision to retest would be based on contractor judgment and made on a case-by-case basis. The GAO concluded, therefore, that the Navy lacks adequate assurance that the contractor retest efforts are sufficient to ensure that significant changes do not introduce adverse effects. The GAO concluded that this is another indication that additional Navy actions are needed to ensure the contractor proceeds with a prudent development approach. (p. 4, pp. 13-14, pp. 21-22/GAO Draft Report)

- **DoD Response: Nonconcur.** The Navy has more than adequate assurance that retest efforts are sufficient to ensure the effects of changes are both characterized fully and minimized. The 5 percent change threshold applies to Computer Software Configuration Items that have already undergone acceptance testing. (See the DoD response to Finding D.) The threshold was established during source selection in 1987, to establish an accurate, common basis for evaluating competitive proposals. This threshold requires retest of the entire Computer Software Configuration Item if the change is greater than 5 percent.

  All software changes, no matter how small, will be retested multiple times (from two to as many as 13) at various levels, prior to system acceptance. Both the Navy and the prime contractor use criticality analysis to determine the amount of regression needed to ensure confidence that a particular change introduces no adverse effects. That is prudent engineering practice.

  The software engineering principles inherent in the use of Ada (e.g., data abstraction, information hiding, strong typing) also serve to mitigate cascading effects of changes. The use of Navy test time provides another level of assurance that all changes will be evaluated fully for impact. The use of the test time is also prioritized by criticality analysis.

- **FINDING F: Central Guidance On Unit-Testing Lacking.** The GAO explained that unit-testing of software is carried out by programmers as the code is developed to ensure errors are identified as early as possible. According to the GAO, software development experts said central guidance on unit-testing is important to ensure that programmers test smaller software components adequately and consistently, before they are integrated with others. The GAO observed that central guidance for the BSY-2 is even more important, given the compressed development schedule and the fact that many of the programmers will be using Ada for the first

time. The GAO found, however, that contractor developed guidance lacks detailed procedures to be followed in carrying out unit-testing. The GAO concluded, therefore, that without central guidance, the Navy lacks assurance that unit-testing will be consistently and comprehensively conducted. The GAO also concluded that this is another indication that additional Navy actions are needed to ensure the contractor proceeds with a prudent development approach. (p. 4, pp. 14-15, pp. 21-22/GAO Draft Report)

- **DoD Response:** Nonconcur. While the GAO has correctly identified the importance of unit testing guidance, the GAO has not accurately described the guidance developed for the BSY-2 system.

  Program level direction for Unit Testing exists in the following Program Level Documentation. Section 3.1 of the prime contractor Software Test Plan Style Guide states specific requirements of Unit testing to include:

  - testing using nominal, upper and lower boundary, high and low out-of-bounds, and erroneous input values;
  - testing for error detection and proper error recovery, including appropriate error messages;
  - testing all executable statements and branches; and
  - testing each output option for correct formatting of output data and command signals.

  The Software Development Plan in Section 5.3 Table XXXII contains specific criteria to be verified by the Software Development and Software Quality Assurance organizations for both the Unit Test Cases and Unit Test Procedures/Test Results. These criteria include:

  - coverage of the unit requirements;
  - consistence with design documentation;
  - compliance with contractual requirements;
  - adherence to required format/content;
  - traceability to Section 4.2 of the Software Test Plan;
  - internal consistency;
  - understandability; and
  - completeness.

In general, the guidelines exceed the minimum acceptable level by requiring action items for such items as excessive cyclomatic complexity. (See Software Development Plan Section 5.1.6.)

- **FINDING G: Indications That The BSY-2 Could Experience Performance Problems.** The GAO found that, even though it is still early in development, there are already indications of potential performance problems. As one example, the GAO cited the Enhanced Modular Signal Processor. The GAO explained that the processor is one of the most critical components of the BSY-2 system, is technologically complex, and poses considerable development risk. The GAO found that currently, the signal processor is experiencing problems in meeting start-up times under several operational modes. The GAO noted that a more advanced version of the processor is under development, but will not be available initially for the first three BSY-2 systems.

  As another example, the GAO reported that estimates indicate BSY-2 processing capacity may be insufficient. The GAO explained that sufficient spare processing capacity is needed to facilitate growth in order to meet future workload needs, and the Navy has set standards on spare capacity and maximum system utilization to avoid past problems. The GAO reported, however, that analytical results thus far indicate the delivered system may not provide sufficient spare capacity. The GAO explained that, to meet the 20 percent spare capacity standard established by the Navy, contractor plans state that the processor estimates should show spare capacity of 50 percent at the critical design review. The GAO found, however, that when the critical design review was completed in September 1990, estimates showed that over one-half of the processors were already exceeding the 50 percent threshold. In addition, the GAO pointed out that the constraints and complexity of the BSY-2 may mean that the 20 percent spare capacity may be inadequate to accommodate change and growth in workload. The GAO observed that, if a BSY-2 system with inadequate spare capacity is delivered, it may be difficult or impossible to increase capacity by adding hardware, due to severe space constraints on submarines. The GAO concluded that these indications of BSY-2 performance problems are additional reasons why the Navy needs to take action to ensure a prudent development approach and to minimize risks. (p. 4, pp. 15-20/GAO Draft Report)

- **DoD Response: Partially concur.** The DoD does not agree that the two examples cited by the GAO are early indications of potential performance problems.

The use of Enhanced Modular Signal Processor in BSY-2 was congressionally mandated after the Enhanced Modular Signal Processor design was completed. The design did not account for BSY-2 instantiation time requirement. The Navy has funded development of modifications to reduce Enhanced Modular Signal Processor instantiation times and, so far, these times have already improved by a factor of 4. The Navy continues to address the timing concern. The prime contractor, at the urging of the Navy, designed the AN/BSY-2 system to incorporate the use of Enhanced Modular Signal Processors, while minimizing the impact of the specific Enhanced Modular Signal Processor shortfall.

With regard to the second example, there is no factual basis from which to conclude that the delivered system may not provide sufficient spare capacity. The open architecture of the BSY-2 Combat System allows for ease of growth. The tremendous spare capacity that already exists in the developed hardware is sufficient, so processing capacity is absolutely not a concern for the BSY-2, present or future. As currently designed, the BSY-2 will be delivered with enough spare capacity to accommodate 30 years of expected functional growth.

The System Engineering Management Plan states that computer processor estimates should show 50 percent spare capacity entering Critical Design Review-- which was, in fact, enforced. At the time of Critical Design Review exit, much more design detail was available and the estimates were based upon Computer Software Configuration Item Program Design Language. It was expected and acceptable for some of the loading estimates to increase.

The System Engineering Management Plan (P. 2-145) also states that "a design goal for limiting Central Processing Unit loading per processor to 50 percent will remain even up to the delivery of the AN/BSY-2 system."

Currently, the total utilized capacity of the BSY-2 data processors is 36.9 percent for Central Processing Unit loading and 12.1 percent for memory, for a reserve capacity of 63 percent Central Processing Unit loading and 88 percent memory. Per TADSTAND definition, the estimate includes resources from installed spares, plus growth capacity by installation of processors (or memory) that require no backplane modification for use. All dedicated backup processors (not spares) are included in the utilization estimates and are not counted as spare for reserve capacity purpose. The utilized capacity estimates, therefore, are extremely conservative.

At the present time, only 5 percent of the data processing computers require examination to see if some of the spare

capacity needs to be allocated to provide extra utilization resources. In addition, spare (empty) slots in the major units represent 23 percent of all available module slots. That percentage does not include the fact that most units also have one or two empty (spare) drawers available; each of which represents 135 module slots.

There are also four entire cabinet footprints allocated to the BSY-2 Combat System, which are currently reserved for growth (not being used). Utilization of the footprints would more than double the current BSY-2 data processing capabilities.

- **FINDING H: Need For Management Commitment To Independent Verification And Validation.** The GAO explained that independent verification and validation is a commonly used approach to minimize risks on large scale development efforts, and also provides an independent, third party analysis of potential software problems throughout development. The GAO found, however, that even though the BSY-2 program had been in full-scale development for two and one-half years and passed through four of the eight major program review milestones, an independent verification and validation plan was only developed in June 1990.

  The GAO concluded that the new BSY-2 management plan is a step in the right direction, since it describes an independent verification and validation effort that is consistent with Federal guidance. The GAO observed, however, that the effectiveness of the plan will depend on the commitment of Navy management to the plan, and the willingness of management to act on the findings and recommendations of the independent verification and validation effort. In addition, the GAO observed that the funding for independent verification and validation may not receive the priority it deserves. The GAO concluded that, without management and financial commitment to independent verification and validation (which has been lacking over the last two and one-half years), system development risks will be increased. (p. 4, pp. 20-21/GAO Draft Report)

- **DoD Response: Partially concur.** The Navy has always recognized the importance of a Verification and Validation function that assures the quality and compliance of the AN/BSY-2 development with the Fleet requirements. That recognition has, in fact, been the driver in establishing and maintaining a well-funded Verification and Validation function from the inception of the AN/BSY-2 program. The Verification and Validation function was an integral component of the Navy team that established the contractual and technical requirements for the AN/BSY-2 Full Scale

Development and continued to perform as an effective and efficient oversight function during the initial phases of Full Scale Development.

Likewise, the Navy recognized the importance of maintaining programmatic objectivity within the Navy team as the AN/BSY-2 development proceeded into the critical phases of software development and its early test and integration. Consistent with that recognition, the Navy established a separate organization charged with providing the Navy team with an Independent Verification and Validation of the AN/BSY-2 Software Development. The objectives and responsibilities of the Independent Verification and Validation function are defined formally in Management and Implementation plans, and are in full execution.

The Navy past commitment to the Verification and Validation and now Independent Verification and Validation support is a matter of record. The Navy future management and financial commitments are accounted for clearly in the out-year budget submittals and reflect a significant Verification and Validation/Independent Verification and Validation activity, which extends beyond the Navy operational tests of the AN/BSY-2 system.

Initial, formal Independent Verification and Validation plans were revised, not discarded. Although formal Independent Verification and Validation was not in place through the program design review milestone, Independent Verification and Validation activities have been performed since program inception. Activities conducted through Preliminary Design review meet or exceed the intent of Federal guidance (the Air Force, the National Bureau of Standards, National Institute of Standards and Technology, and the Institute of Electrical and Electronics Engineers) commonly applied for Independent Verification and Validation. However, they were conducted by the Technical Direction Agent vice a specific Independent Verification and Validation agent. The plans, policies, and procedures applied are considered equivalent to the products that would be generated and utilized by an independent agent. In addition, the methodology applied by the Technical Direction Agent was based on extensive lessons learned and, in a number of key areas, included verification activities beyond those normally conducted. Notable examples include: (1) supplementing the review of individual configuration items with "string" reviews, which evaluate functionality and interdependencies by grouping Configuration Items; (2) top level design reviews on certain key areas systems level capabilities, such as Performance Monitoring and Fault Localization; (3) weapons snapshot and system reconfigurability. In addition, other key areas of the

system were reviewed, including:

- o database management;
- o combat system services;
- o non-development items;
- o system level functions review; and
- o man-machine interface operability.

The Technical Direction Agent managed the Verification and Validation activities utilizing a comprehensive set of plans, policies, and procedures that provided an audit trail for each milestone. A detailed plan was developed mutually with the contractor for each milestone, which included a set of guidelines for implementation, handbooks for each milestone that identified Navy/Contractor roles and responsibilities, the products to be reviewed, entrance criteria, review criteria, and exit criteria for each product.

* * * * *

RECOMMENDATIONS TO THE DEPARTMENT OF DEFENSE

o **RECOMMENDATION 1:** The GAO recommended that the Secretary of Defense direct the Secretary of the Navy to:

(1) -- assess the risks associated with the compressed schedule, incomplete critical design review tasks, testing and analysis shortcomings, early indications of potential system performance problems, and the delays in developing an independent verification and validation program;

(2) -- determine their collective impact on the performance, cost, and current delivery schedule for the combat system;

(3) -- adjust the current development approach, as is warranted; and

(4) -- report the results of the assessment and planned actions to the House and Senate Armed Services and Appropriations Committees. (p. 22/GAO Draft Report)

- **DoD Response: Nonconcur.** It is the DoD position the recommended actions are not necessary. In full compliance with congressional direction and DoD policies and procedures, the Department periodically examines development status of the BSY-2 system, and constantly monitors the risk areas to ensure the BSY-2 meets baseline thresholds (cost, schedule, and performance). Corrective action will be taken, as appropriate, to address any problem identified. With regard to item (4), the information will be transmitted to the Congress as required by Section 2400, Title 10 U.S.C. in the Low Rate Initial Production report.

- **RECOMMENDATION 2:** The GAO recommended that the Secretary of Defense direct the Navy to ensure that:

  -- central guidance on unit testing is developed and issued for use by all BSY-2 programmers and software engineers; and

  -- independent verification and validation is treated as a high priority task and adequate funding is provided. (p. 23/GAO Draft Report)

- **DoD Response: Partially concur.** While the DoD agrees with the importance of the cited principles, adequate unit testing guidance is already in place and priority has been assigned to independent verification and validation (see the DoD responses to Findings F and H). Accordingly, no additional action is required.

* * * * *

RECOMMENDATION TO THE CONGRESS

- **RECOMMENDATION 3:** The GAO recommended that the Congress withhold funding for any additional systems beyond the first three, until the BSY-2 is fully developed and thorough operational testing demonstrates that the system satisfies all mission requirements. (p. 23/GAO Draft Report)

- **DoD Response: Nonconcur.** Development and testing of the BSY-2 is in full compliance with congressional guidance and DoD policy and procedures. In January 1991, the Defense Acquisition Board reviewed the BSY-2 system and authorized the Navy to contract for the remaining ten SSN-21s and associated AN/BSY-2 systems for a total of 11 SSN-21's and 15 AN/BSY-2 systems to be procured during Low Rate Initial Production, as allowed by Section 2400 of Title 10 U.S.C. The decision will be reported to the Congress in the Low Rate Initial Production report.

Major Contributors to This Report

Information Management and Technology Division, Washington, D.C.	Michael T. Blair Jr., Assistant Director Andrew Patchan Jr., Evaluator-in-Charge Dr. Rona B. Stillman, Chief Scientist Edward R. Tekeley, Technical Adviser Gwendolyn A. Dittmer, Computer Scientist Heather A. Winand, Computer Scientist
Boston Regional Office	Jeffrey V. Rose, Regional Management Representative Edmund L. Kelley, Jr., Senior Evaluator

Related GAO Products

Submarine Combat System: Status of Selected Technical Risks in the BSY-2 Development (GAO/IMTEC-91-46BR, May 24, 1991).

Navy Ships: Concurrency Within the SSN-21 Program (GAO/NSIAD-90-297, Sept. 28, 1990).

Submarine Technology: Transition Plans Needed to Realize Gains From DOD Advanced Research (GAO/IMTEC-90-21, Feb. 14, 1990).

Navy Acquisition: Cost, Schedule, and Performance of New Submarine Combat Systems (GAO/NSIAD-90-72, Jan. 31, 1990).

Navy Weapons Testing: Defense Policy on Early Operational Testing (GAO/NSIAD-89-98, May 8, 1989).

Submarine Combat System: Technical Challenges Confronting Navy's Seawolf AN/BSY-2 Development (GAO/IMTEC-89-35, Mar. 13, 1989).

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

Ordering Information

The first five copies of each GAO report are free. Additional copies are $2 each. Orders should be sent to the following address, accompanied by a check or money order made out to the Superintendent of Documents, when necessary. Orders for 100 or more copies to be mailed to a single address are discounted 25 percent.

U.S. General Accounting Office
P.O. Box 6015
Gaithersburg, MD 20877

Orders may also be placed by calling (202) 275-6241.

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

Official Business
Penalty for Private Use $300

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