Project No. G-36-611

Project Director: Dr. Richard DeMillo

Sponsor: Naval Air Development Center

Type Agreement: Contract No. F33657-82-C-2083, Delivery Order No. JCO3

Award Period: From 7/5/83 To 10/31/85 (Performance) (Reports)

Sponsor Amount: Estimated: $ 450,000

Cost Sharing Amount: $ Cost Sharing No:

Title: "Software Test and Evaluation Project, Phases III and IV"

ADMINISTRATIVE DATA

OCA Contact: John W. Burdette X4820

1) Sponsor Technical Contact:
   Mr. Fred Shocket
   Naval Air Development Center
   Warminster, PA 18974

2) Sponsor Admin/Contractual Matters:
   Mr. Thomas A. Bryant
   Office of Naval Research
   Resident Representative
   206 O'Keefe Building
   Georgia Institute of Technology
   Atlanta, GA 30332

Defense Priority Rating: DOC-9E

Military Security Classification:

RESTRICTIONS

See Attached Gov't Supplemental Information Sheet for Additional Requirements.

Travel: Foreign travel must have prior approval — Contact OCA in each case. Domestic travel requires sponsor approval where total will exceed greater of $500 or 125% of approved proposal budget category.

Equipment: Title vests with Government except those items costing $1,000 or less vest with GIT providing prior written approval to purchase received from Contracting Officer.

COMMENTS:


COPIES TO:

Procurement/EES Supply Services
Research Security Services
Reports Coordinator (OCA)
Research Communications (2)
GTRI Library Project File Other
GEORGIA INSTITUTE OF TECHNOLOGY
OFFICE OF CONTRACT ADMINISTRATION

SPONSORED PROJECT TERMINATION/CLOSEOUT SHEET

Date 5/21/86

Project No. C-36-611

School XXX ICS

Includes Subproject No.(s) A-3686, A-3975, A-4211

Project Director(s) Richard DeMillo

Sponsor Naval Air Development Center, Warminster, PA.

Title Software Test and Evaluation Project, Phases III and IV

Effective Completion Date: 10/31/85 (Performance) --- (Reports)

Contract Closeout Actions Remaining:

☐ None

☐ Final Invoice or Final Fiscal Report

☐ Closing Documents

☐ Final Report of Inventions - Patent Questionnaire sent to P.I.

☐ Govt. Property Inventory & Related Certificate

☐ Classified Material Certificate

☐ Other

The Project No. ____________________________ Continued by Project No. ____________________________

TO:

Acting Director
Research Administrative Network
Research Property Management
Purchasing
Agreement/EES Supply Services
Research Security Services
IT Coordinator [TCA] [TCA]
Services

Library
GTRC
Research Communications (2)
Project File
Other Jones, Embry

OCA 60 1028
THE SOFTWARE TEST AND EVALUATION PROJECT

PHASES III & IV

Quarterly Progress Report Number 1
05 July 1983 - 04 December 1983

06 December 1983

Supported by

Naval Air Development Center (NADC)
Contract: F33657-82-G-2083
GIT Project: G36-611

School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332
The Software Test and Evaluation Project: Phases III & IV
Quarterly Progress Report Number 1

Table of Contents

1. Introduction ................................................................. 1
2. Organization and Staffing ............................................. 2
3. Current Task Areas ...................................................... 2
4. Summary of Progress .................................................... 5
5. Travel Related to STEP .................................................. 13
6. Visitors ........................................................................ 16
7. Publications, Reports, and Briefings ............................... 17
8. Report Distribution ...................................................... 19
1. **INTRODUCTION**

This is the First Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

**Project Description**

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.

**Project Support**

The following contract provides support for this project.

**Title:** "Software Test and Evaluation Project, Phases III and IV"

**Funding Agency:** Naval Air Development Center (NADC)

**Contract Number:** F33657-82-G-2083

**GIT Project Number:** G36-611

**Principal Investigator:** Richard A. DeMillo
2. **ORGANIZATION AND STAFFING**

**Faculty**

DeMillo, Richard A. - Professor  
Gagliano, Ross A. - Senior Research Scientist  
Martin, Rhonda J. - Research Scientist II  
Passafiume, John F. - Senior Research Scientist  

**Staff**

Flinn, Perry B. - Research Scientist I  
Hutchins, Ronald R. - Research Technologist I  
Myrick, Dorothy K. - Clerk/Typist  
Richards, Esther E. - Administrative Secretary  
Seay, Brenda G. - Student Assistant  

**Students**

Bilsel, M. Sinasi - Graduate Research Assistant  
Brannen, James M. - Graduate Research Assistant  
Exel, Matija B. - Graduate Research Assistant  
Offutt, VI, A. Jefferson - Graduate Research Assistant  
Sycara-Cranski, Ekaterini P. - Graduate Research Assistant  

3. **CURRENT TASK AREAS**

The work to be performed has been organized into the task areas identified below.

**DDT&E Task Areas**

A. Software Test and Evaluation Guidebook

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition lifecycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. Software Test and Evaluation Model

Task 8 of the STEP Statement of Work states that "The effects of alternative software lifecycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software lifecycles. Examples of software lifecycles include the traditional waterfall lifecycles and the currently advocated evolutionary lifecycles.
STEP

C. NADC/Technology Demonstration

The NADC/Technology Demonstration is intended to provide a laboratory framework for carrying out most of the technology-based STEP activities. These activities are described in the STEP Statement of Work Tasks 3 - 9 and the Software Technology for Adaptable, Reliable Systems (STARS) Program Task Areas 5 - 7. STEP Tasks 3 - 9 are concerned with formulating guidance statements, assessing the effects of new policy guidance, gathering evidence which indicates the effectiveness of systematic program testing, developing a software T&E model, and developing a software T&E guidebook. STARS Program Task Areas 5 - 7 are concerned with providing a test technology baseline and identification procedures, providing demonstration and qualification procedures for test technologies, and providing functional requirements for Ada Programming Support Environment (APSE) test environments.

D. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the Software Test and Evaluation Guidebook.

E. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

* Ada is a registered trademark of the U.S. Government, Ada Joint Program Office
STARS Program Task Areas

F. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USDR&E). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

G. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USDR&E. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for Service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

H. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

I. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

J. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

K. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.

4. SUMMARY OF PROGRESS

A. Software Test and Evaluation Guidebook

During the acquisition lifecycle, a variety of organizations become involved with the software. The usual form which this involvement takes is that of reviewing documentation. The documents reviewed serve different purposes and vary in the amount of detail presented concerning the software. In addition, each organization has specific interests and objectives which must be satisfied through the review process.
For example, personnel in the program office may only be concerned that the testing performed demonstrates the capabilities of the software as required by the specifications. On the other hand, personnel in logistics organizations may be more concerned that testing demonstrates the maintainability or enhanceability of the software. Documents specifically concerned with testing range from the system level test document, the TEMP, through the lower level test plans and results, to the unit development folders which contain information pertaining to the testing of individual modules of code. Although program office personnel and logistics personnel may each review all of these document types, their evaluations would attempt to answer different questions.

The Software Test and Evaluation Guidebook is intended to support the various organizations during this review process. Checklists tailored to the concerns of the individual organizations are being developed to help ensure that information required by the STEP Phases I and II Recommendations is included, as appropriate, in each document type. The checklists will be supported by reference materials and "good examples" to further aid the reviewing organizations, as well as those involved in the development of the documents.

Efforts to date on the Software Test and Evaluation Guidebook have concentrated on two areas: (1) the determination of which organizations and document types require support, and (2) the development of a checklist for use by DDT&E in evaluating plans for software T&E. A strawman list of organizations and associated documents has been developed and will be refined shortly. This list will then be distributed for review by the STEP Technical Advisory Panel and the STEP Military Advisory Panel. A draft of the checklist for use when evaluating TEMP's has been presented to DDT&E. This draft is being expanded and will form the basis for a paper which is being written for publication in the International Test and Evaluation Association (ITEA) Journal.

In addition to the activities described above, guidebooks and other reference materials gathered during STEP Phases I and II are being reviewed to identify portions which may be relevant for inclusion in the Software Test and Evaluation Guidebook which will result from this effort.

B. Software Test and Evaluation Model

Efforts to date pertaining to this task area have centered on reviewing documentation describing the traditional software development lifecycles and its relationship to the Defense System Acquisition Review Council (DSARC) milestones. In addition, one of the STEP Team members recently directed a National Security Industrial Association (NSIA) Study which investigated the implications of evolutionary acquisition on the software development lifecycle. The results of that study are currently being briefed to selected organizations and will be documented in a final report in the near future. Those results will also serve as input to this task area.
C. NADC/Technology Demonstration

Current ideas in this area involve the use of NADC’s Facility for Automated Software Production (FASP), in concert with a project supported by that facility, to demonstrate the implementation of STEP planning recommendations. Technology required to support the plans developed would be inserted into FASP for application to the demonstration project. This technology would also be available to NADC for application on other projects which use FASP. Finally, the planning, technology identification, and technology insertion process would be documented such that it could be transitioned to other projects.

Two meetings have been held at NADC to discuss this task area. At the first meeting on 23 September 1983, STEP personnel gathered information and documentation on FASP and discussed the criteria to be applied when selecting a project for use in the technology demonstration. Selection criteria included the use of FASP, the status of the project (under development, deployed, etc.), the availability of documentation, and the extent to which the software was subjected to systematic testing. In addition, the process of gaining access to FASP was initiated.

At the second meeting on 06 - 07 October 1983, NADC personnel presented information on candidate projects and recommended the selection of the Light Airborne Multipurpose System (LAMPS). STEP personnel agreed that this project best satisfied the selection criteria discussed previously. In addition, alternative strategies for accessing FASP were discussed (long distance phone lines, dedicated lines, via ARPANET, etc.). Progress in this task area has been delayed while funds are located to cover NADC computer charges and costs related to NADC personnel support of STEP activities.

D. Good Examples

While initial plans were being formulated as to the best way to proceed, both literature and other materials were assembled on various systems, some of which were not analyzed during STEP Phases I and II. Criteria were then developed for system selection along with a compilation of possible candidate systems from the Army, Navy, and Air Force. A meeting was held in Washington on 18 August 1983 to discuss the list of systems and criteria for selection, etc. Using 7 selection criteria, an original set of approximately 65 systems was pared down to 6 systems, 2 from each Service. The selection process was aided by consultation with key Pentagon individuals, as well as contacts at the Defense Systems Management College, and selected development commands and project offices.
The work that proceeded from that point was to develop an approach to gathering data on the 6 systems by starting with the Data Gathering Guides which had been the basis of the data collection during STEP Phases I and II (ref. STEP Phases I and II Final Report, Volume 3: Software Test and Evaluation: Current Defense Practices Overview). The guides required modification since the earlier data gathering effort did not attempt to identify "good examples" nor assemble any basis for that "success". A compilation of system attributes and points of contact for each system was begun in order to facilitate site visits and information gathering.

Several meetings were held to work on task plans. The first meeting was with STEP personnel in Washington on 14 September 1983), the second with Air Force personnel at AFSC on 16 September 1983, and the third with Navy personnel at NADC on 07 October 1983. It became clear that an extensive document trace trace would be necessary starting with the TEMP and continuing through the Request for Proposal (RFP), the Statement of Work (SOW), software development and lifecycle management plans, and the various reports, reviews, and audits of the different levels of testing.

Two matrices have been assembled: the first details systems versus organizational levels (to include names and phone numbers of points of contact) and the second identifies specific items to be looked for. Additional sources of information have also been listed.

E. Risk Model

In the last quarter, research in this task area has concentrated on isolating those features of software T&E that distinguish its risk modelling problems from those encountered in other technologies. In summary, these features stem from the unavailability of commonly accepted quantitative and probabilistic interpretations of such characteristics as reliability, cost, and risk. The lack of accepted quantitative theories in these areas rules out many otherwise attractive risk modelling approaches.

In our initial work, we have focused on a branch of decision theory called "decision making under uncertainty." In this context, the term "uncertainty" means that the underlying probability distributions are either unknown or not meaningful.

These models take the form of two-person games. The "pay-off" entries in the game represent the utility of adopting a test strategy in light of operational characteristics of the integrated system. The utility functions for software T&E are determined by two parameters: cost and performance. The cost of a test is a weighted cost-of-test versus cost-of-failure function. The performance of a test is an index; the ordering of all possible indices for a test is consistent with a total ordering of the following conditional probabilities:

\[ \Pr\{\text{Failure } F \text{ occurs | test } T \text{ is passed}\} \]
Even though exact values of these distributions may not be available, most systematic software test methodologies allow consistent orderings of these indices.

The major issues to be resolved are: (1) the identification of additional parameters that affect test utility, (2) the development of optimal test selection strategies, and (3) the design of algorithms for test selection.

The ultimate goal is to provide an axiomatic characterisation of optimal selection strategies. Such characterisation allows a decision maker to quantify the risk at decision points: the decision risk is the difference between the utility of the candidate decision and the choice given by the optimal strategy.

This approach has the advantage of being automatable; decision quantities of this sort can be computed very quickly using algorithms that resemble the simplex algorithm for linear programs.

We also anticipate holding a workshop on the state-of-the-art for software risk modelling.

F. Technology Management

STARS Program documentation currently available to the STEP Team has been reviewed. During the initial meeting of the STEP Technical Advisory Panel, discussions were held as to the role of the proposed Acquisition Panel. Consultations will be held with STARS Program management in regard to this issue and a strawman management plan will be developed. This plan will be reviewed at a later meeting of the STEP Technical Advisory Panel. The strawman management plan will also be presented to the STEP Military Advisory Panel for their evaluation and comments.

G. Technology Insertion

Initial discussions of the activity to produce a software T&E management and operating plan for the SEI were held during the first meeting of the STEP Technical Advisory Panel on 08-09 November 1983. Future efforts are awaiting the promulgation of the SEI charter.

The definition of an initial procedure for identifying test tools, environments, and methodologies for Service qualification and the development of a baseline of software testing tools and technology has begun. The STEP Team has reduced the catalog of testing tools found in the STEP Phases I and II Final Report, Volume 2: Software Test and Evaluation: State-of-the-Art Overview to a set of 25 tools which are candidates for inclusion in the baseline. The criteria for selection addressed considerations such as availability, portability, and language. A collateral effort has been a search for CMS-2 tools that could be ported to FASP. This is an on-going effort.
Finally, as a result of discussions during the STEP Technical Advisory Panel meeting, selected panel members will be assigned the responsibility of investigating several issues concerning the development of procedures for demonstrating and qualifying testing tools for DoD use. This topic will be addressed at length at a later meeting of the STEP Technical Advisory Panel. In addition, a strawman procedure will be developed for review by both the Technical Advisory Panel and the STEP Military Advisory Panel.

H. APSE Test Environment

The progress on this task area for the past quarter is represented by a set of strawman "issues" that are prerequisites for discussions of integrated test environments. This work is being coordinated with the JSSEE Team under the STARS Program. STEP has formally established a liaison with the JSSEE Team and at least one member of the STEP Team attends all JSSEE Team meetings.

It is our view that the problem of designing an integrated software test environment reduces to the problem of designing an operating system that is optimized for testing. The issues that are currently under discussion include the following:

1. Lifecycle Scope: Are there technical limitations on a test environment that are imposed by lifecycle constraints?

2. Control: Should the control restrictions imposed by the testing process itself have a design impact on the test environment? It appears that the answer to this question is yes; we have therefore begun to call the test environment a "test sub-environment" or TSE.

3. Performance: What constraints do performance testing parameters place on the design of the TSE and the utilization of host-target configurations to be supported by the JSSEE?

4. Test Mode: What degree of testing is considered to be minimal in the JSSEE and how should additional test capabilities be integrated into lifecycle methodologies supported by the JSSEE?

5. Human Factors and Data Reduction: How should test status information be presented to the tester/user and what additional requirements does this place on the JSSEE?

These issues have been briefed to the JSSEE Team and will be refined during the next two quarters for delivery to the STARS Program as a draft preliminary design and requirements document.
I. Coordination with Related Efforts

The following reflects ongoing efforts with which STEP is currently coordinating.

**DT&E Commanders Conference:** The Development Test and Evaluation (DT&E) Commanders meet periodically to discuss common problems, benefit from one and others experience, and determine areas in which combined actions are necessary. STEP personnel have briefed this organization and are working with selected individuals towards the goal of solving, in the near term, problems being experienced with software T&E.

**Joint Logistics Commanders Initiatives:** The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development of Tri-Service standards for Software Development (DOD-STD-SDS) and Software Quality Assessment and Measurement (DOD-STD-SQAM). STEP personnel have been involved in the review of these documents and receive updates as they become available. In addition, the Joint Logistics Commanders recently sponsored their third Software Workshop, Orlando I (previous workshops were Monterey I & II). One member of the STEP Team was invited to participate in this workshop, providing input to a panel investigating IV&V.

**STARS Program Panels:** The STARS Program has established working groups and/or panels to investigate selected issues. STEP personnel have been active on both the Data Rights Working Group and the Software Engineering Institute Panel.

**JSSEE Team:** This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.

**Evaluation and Validation (E&V) Efforts:** Ada compilers and implementations of the Common APSE Interface Set (CAIS) will be required to undergo a standard evaluation and validation prior to being approved for DoD use. STEP is in the process of establishing a formal coordination link with these efforts. The goal of this coordination is to determine the relationship between E&V and T&E, and to influence the E&V process as necessitated by STEP recommendations.

**IEEE Computer Society Software Engineering Standards Subcommittee:** This group is in the process of developing a standard for unit testing. A member of the STEP Team attended the latest meeting of the group to brief them on our efforts and results. We will continue to evaluate the proposed standard, in terms of STEP recommendations, as it evolves.

**NSIA C² Software Development and Acquisition Study:** A STEP Team member recently directed an NSIA study for the Air Force Electronic Systems Division on C² Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on a software development effort. Results of this study will feed into the Software T&E Model task.
DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

J. Direct Support to DDT&E

At the request of DDT&E, STEP personnel recently evaluated the AMPE TEMP in terms of its treatment of software and the SUBACS TEMP in terms of its treatment of software reliability. Comments were forwarded to DDT&E for consideration.

K. Laboratory Support

Currently, STEP personnel are using a Fortune 32:16 System which is on consignment to the Georgia Institute of Technology from Fortune Systems. It was configured with the new Fortune operating system on 14 November 1983. The new operating system is an update to the previous Fortune version of Berkeley Unix 4.1 and contains several enhancements and additions. The Fortune System is directly connected to a Vax 11/780, a Wang word processor, and the local Georgia Tech network, Net-One.

A purchase order has been issued for another Fortune 32:16 System. Delivery will occur as soon as possible, although no delivery date is available at this time. This system will belong solely to STEP and will provide primary support for STEP workstations. Plans are to link this system to the current Fortune 32:16 to provide access to Net-One and the other machines.

The STEP Phases I and II Final Report is currently stored on Wang diskettes. For long-term retention and data manipulation purposes, these files need to be transferred to the Vax 11/780. This can be accomplished by first transferring the files from the Wang to the Fortune and then from the Fortune to the Vax. The capability of transferring files from the Fortune to the Vax exists today. However, the transfer of STEP documents from the Wang to the Fortune has been delayed by the absence of a 3780 bisynchronous communications protocol for the Fortune computer. The hardware required for this capability will be delivered with the new Fortune 32:16. The required software will be delivered soon thereafter.

In order to perform the work associated with the NADC/Technology Demonstration, it is necessary for STEP personnel to have access to the NADC Central Computer System, in general, and FASP, in particular. STEP has obtained an account on the NADC Cyber and identified several means of connecting to that system. Currently, we plan to access the NADC system through ARPANET via CSNET and over a modem from a local Vax 11/780 directly to the NADC system.
5. **TRAVEL RELATED TO STEP**

Date of Trip: 19 Jul 83  
Individual(s) Traveling: R. A. DeMillo & R. J. Martin  
Itinerary: The Pentagon, Washington, D.C.  
Contact: John Bolino (ODDT&E) & Samuel Redwine (STARS)  
Purpose: Discussions of STARS funding for STEP

Date of Trip: 20 - 21 Jul 83  
Individual(s) Traveling: R. A. DeMillo & R. J. Martin  
Itinerary: Tysons Corner, VA  
Contact: Charles Watt (DDDT&E/SNC31)  
Purpose: Attend DDT&E Program Review

Date of Trip: 22 Jul 83  
Individual(s) Traveling: R. A. DeMillo, R. A. Gagliano, & R. J. Martin  
Itinerary: The Pentagon, Washington, D.C.  
Contact: John Bolino (ODDT&E)  
Purpose: Brief STEP to Adm. Linder (DDT&E)

Date of Trip: 28 Jul 83  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: The Pentagon, Washington, D.C.  
Contact: John Bolino (ODDT&E)  
Purpose: Discussions of STEP Contract

Date of Trip: 01 - 03 Aug 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Boston, MA & Hanscom AFB, MA  
Contact: Gen. James W. Stansberry (Commander, AF/ESD)  
Purpose: Preparation for and Delivery of C2 Software Development and Acquisition Study Final Briefing

Date of Trip: 02 - 03 Aug 83  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: Los Alamos, NM  
Purpose: Brief Risk Models at Conference on The Measurement of Computer Software Performance

Date of Trip: 04 Aug 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Tysons Corner, VA  
Contact: Jacqueline Hamilton (BDM)  
Purpose: Coordination of BDM work on Functional Description of Acquisition T&E Process and STEP work on Software T&E Model

Date of Trip: 18 Aug 83  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: The Pentagon, Washington, D.C.  
Contact: Charles Watt (DDT&E/SNC31)  
Purpose: Discussions of Potential Linkages between Research Efforts
Date of Trip: 18 Aug 83
Individual(s) Traveling: R. A. Gagliano & R. J. Martin
Itinerary: The Pentagon, Washington, D.C.
Contact: John Bolino (ODDT&E)
Purpose: Discussions of STEP Work on Good Examples

Date of Trip: 23 Aug 83
Individual(s) Traveling: R. A. DeMillo & R. J. Martin
Itinerary: Naval Underwater Systems Center, Newport, RI
Purpose: Brief STEP at 14th Multi-Service DT&E Commanders Conference

Date of Trip: 13 - 15 Sep 83
Individual(s) Traveling: R. A. Gagliano
Itinerary: Washington, D.C.
Purpose: Attend Federal Computer Conference (MCCR Track)

Date of Trip: 14 - 15 Sep 83
Individual(s) Traveling: R. J. Martin & J. F. Passafiume
Itinerary: Washington, D.C.
Purpose: Attend Common APSE Interface Set (CAIS) Review

Date of Trip: 15 Sep 83
Individual(s) Traveling: R. A. DeMillo
Itinerary: Washington, D.C.
Purpose: Brief STEP at Federal Computer Conference

Date of Trip: 16 Sep 83
Individual(s) Traveling: R. A. Gagliano & R. J. Martin
Itinerary: Andrews AFB, MD
Contact: Maj. Ed Stevens & Capt. Lee Cooper (AFSC)
Purpose: Planning for Air Force Involvement in STEP

Date of Trip: 23 Sep 83
Individual(s) Traveling: R. J. Martin
Itinerary: NADC, Warminster, PA
Contact: Hank Stuebing & Fred Shocket
Purpose: Planning for STEP NADC/Technology Demonstration

Date of Trip: 28 Sep 83
Individual(s) Traveling: R. A. DeMillo
Itinerary: The Pentagon, Washington, D.C.
Contact: John Bolino (ODDT&E) & David Gelperin (IEEE)
Purpose: Discussions of IEEE Computer Society Software Engineering Standards Subcommittee Work on Unit Testing Standard

Date of Trip: 03 - 04 Oct 83
Individual(s) Traveling: R. A. Gagliano
Itinerary: Ft. Lee, VA
Purpose: Attend Army Operations Research Symposium on "The Integration of Modeling and Simulation with Testing"

Date of Trip: 05 Oct 83
Individual(s) Traveling: R. A. DeMillo & R. A. Gagliano
Itinerary: The Pentagon, Washington, D.C.
Contact: John Bolino & Don Greenlee (ODDT&E)
Purpose: Discussion of TEMP's
Date of Trip: 06 - 07 Oct 83  
Individual(s) Traveling: R. A. DeMillo, R. A. Gagliano, & R. J. Martin  
Itinerary: NADC, Warminster, PA  
Contact: Fred Shocket  
Purpose: Planning for NADC Involvement in STEP Phases III & IV

Date of Trip: 12 - 13 Oct 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: STARS Office, Arlington, VA  
Purpose: Attend JSSEE Team Meeting

Date of Trip: 17 Oct 83  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: The Pentagon, Washington, D.C.  
Contact: John Bolino (ODDT&E)  
Purpose: Brief STEP to Adm. Linder's Staff & Dr. Wade

Date of Trip: 18 Oct 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Norfolk, VA  
Purpose: Brief STEP at ITEA/Tidewater Chapter Meeting

Date of Trip: 21 Oct 83  
Individual(s) Traveling: R. A. DeMillo, R. A. Gagliano, & R. J. Martin  
Itinerary: Tysons Corner, VA  
Contact: Charles Watt (DDDT&E/SNC3I)  
Purpose: Attend DDT&E Roundtable

Date of Trip: 25 Oct 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Hartford, CT  
Purpose: Brief C² Software Development and Acquisition Study at AIAA Computers in Aerospace IV Conference

Date of Trip: 27 Oct 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Kansas City, MO  
Purpose: Brief C² Software Development and Acquisition Study at AFCEA Evolutionary Development of Command and Control Software, Military and Commercial Symposium

Date of Trip: 31 Oct 83 - 04 Nov 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Orlando, FL  
Purpose: Attend Joint Logistics Commanders Orlando I Software Workshop

Date of Trip: 01 Nov 83  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: Norfolk, VA  
Purpose: Brief STEP at 1983 Old Crow’s Association Conference

Date of Trip: 10 Nov 83  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: STARS Office, Arlington, VA  
Purpose: Brief STEP Work on APSE Test Environment at JSSEE Team Meeting
Date of Trip: 10 Nov 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: STARS Office, Arlington, VA  
Purpose: Attend JSSEE Team Meeting

Date of Trip: 16 - 17 Nov 83  
Individual(s) Traveling: R. A. Gagliano  
Itinerary: University of Texas, Arlington, TX  
Purpose: Brief STEP to IEEE Computer Society Software Engineering Standards Subcommittee

Date of Trip: 01 Dec 83  
Individual(s) Traveling: R. J. Martin  
Itinerary: Hanscom AFB, MA & Boston, MA  
Purpose: Brief C2 Software Development and Acquisition Study  
Gen. Stansberry's staff, AF/ESD Program Managers, and MITRE personnel

6. VISITORS

Date of Visit: 13 Jul 83  
Visitor(s): Charles Watt (DDDT&E/SNC3I)  
Contact: R. A. DeMillo, R. J. Martin, & R. A. Gagliano  
Purpose: Discussion of plans for STEP Phases III and IV

Date of Visit: 21 - 22 Sep 83  
Visitor(s): W. Michael McCracken (General Electric Company)  
Contact: R. A. DeMillo & R. J. Martin  
Purpose: Provide industry reaction to STEP Phases I and II Recommendations

Date of Visit: 07 Nov 83  
Visitor(s): John Bolino (ODDT&E)  
Contact: R. A. DeMillo & STEP Personnel  
Purpose: Prepare for STEP Technical Advisory Panel Meeting

Date of Visit: 08 - 09 Nov 83  
Visitor(s): STEP Technical Advisory Panel  
John Bolino (ODDT&E)  
John Bowen (Hughes Aircraft Company)  
Carolyn Gannon (General Research Corporation)  
David Gelperin (Software Quality Engineering)  
Lt. Col. Vance Mall (STARS Program Office)  
Michael McCracken (General Electric Company)  
Raymond Rubey (SoftTech, Inc.)  
Ralph SanAntonio (Dynamics Research Corporation)  
Fred Shocket (NADC)  
Marilyn Stewart (Logicon)  
Peter Wegner (Brown University, Dept. of Computer Science)  
Contact: R. A. DeMillo & STEP Personnel  
Purpose: Attend meeting of the STEP Technical Advisory Panel
7. PUBLICATIONS, REPORTS, AND BRIEFINGS

Publications


Reports


Briefings

Date of Briefing: 22 Jul 83
Briefer(s): R. A. DeMillo
Location: The Pentagon, Washington, D.C.
Audience: Adm. Linder (DDT&E)
Subject of Briefing: STEP

Date of Briefing: 02 - 03 Aug 83
Briefer(s): R. A. DeMillo
Location: Los Alamos, NM
Audience: Conference on the Measurement of Computer Software Performance
Subject of Briefing: Risk Models

Date of Briefing: 03 Aug 83
Briefer(s): R. J. Martin
Location: Hanscom AFB, MA
Audience: Gen. James W. Stansberry (Commander, AF/ESD)
Subject of Briefing: C² Software Development and Acquisition Study

Date of Briefing: 23 Aug 83
Briefer(s): R. A. DeMillo & R. J. Martin
Location: Naval Underwater Systems Center, Newport, RI
Audience: DT&E Commanders' Conference
Subject of Briefing: STEP

Date of Briefing: 15 Sep 83
Briefer(s): R. A. DeMillo
Location: Washington, D.C.
Audience: Federal Computer Conference (MCCR Track)
Subject of Briefing: STEP
Date of Briefing: 12 Oct 83
Briefer(s): R. J. Martin
Location: STARS Office, Arlington, VA
Audience: JSSEE Team
Subject of Briefing: STEP

Date of Briefing: 17 Oct 83
Briefer(s): R. A. DeMillo
Location: The Pentagon, Washington, D.C.
Audience: Adm. Linder's Staff & Dr. Wade
Subject of Briefing: STEP

Date of Briefing: 18 Oct 83
Briefer(s): R. J. Martin
Location: Norfolk, VA
Audience: ITEA/Tidewater Chapter
Subject of Briefing: STEP

Date of Briefing: 25 Oct 83
Briefer(s): R. J. Martin
Location: Hartford, CT
Audience: AIAA Computers in Aerospace IV Conference
Subject of Briefing: C² Software Development and Acquisition Study

Date of Briefing: 27 Oct 83
Briefer(s): R. J. Martin
Location: Kansas City, MO
Audience: AFCEA Evolutionary Development of Command and Control
Software, Military and Commercial Symposium
Subject of Briefing: C² Software Development and Acquisition Study

Date of Briefing: 01 Nov 83
Briefer(s): R. A. DeMillo
Location: Norfolk, VA
Audience: 1983 Old Crow's Association Conference
Subject of Briefing: STEP

Date of Briefing: 08 Nov 83
Briefer(s): R. A. Gagliano
Location: Georgia Institute of Technology, Atlanta, GA
Audience: STEP Technical Advisory Panel
Subject of Briefing: STEP

Date of Briefing: 10 Nov 83
Briefer(s): R. A. DeMillo
Location: STARS Office, Arlington, VA
Audience: JSSEE Team
Subject of Briefing: APSE Test Environment

Date of Briefing: 16 Nov 83
Briefer(s): R. A. Gagliano
Location: University of Texas, Arlington, TX
Audience: IEEE Computer Society Software Engineering Standards Subcommittee
Subject of Briefing: STEP
Date of Briefing: 01 Dec 83
Briefer(s): R. J. Martin
Location: Hanscom AFB, MA & Boston, MA
Audience: Gen. Stansberry's staff, AF/ESD Program Managers, & MITRE personnel
Subject of Briefing: C² Software Development and Acquisition Study

8. Report Distribution

The number of STEP Phases I and II Final Reports distributed during the first quarter are as follows.

Volume 1: Report and Recommendations ......................... 155
Volume 2: Software Test and Evaluation:
            State-of-the-Art Overview .......................... 163
Volume 3: Software Test and Evaluation:
            Current Defense Practices Overview ................. 151
Volume 4: Transcript of STEP Workshop, March 1982 * ............. 45
Volume 5: Report of Expert Panel on
            Software Test and Evaluation ............................ 156
Volume 6: Tactical Computer System Applicability Study ** ........ 11

* Volume 4 distribution is limited to a select audience.

** The distribution of Volume 6 is being handled by its author,
   Dr. James F. Leathrum of Clemson University.
THE SOFTWARE TEST AND EVALUATION PROJECT

PHASES III & IV

Quarterly Progress Report Number 2
04 December 1983 - 06 January 1983

06 January 1984

Supported by

Naval Air Development Center (NADC)
Contract: F33657-82-C-2083
GIT Project: G36-611

School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332
The Software Test and Evaluation Project: Phases III & IV
Quarterly Progress Report Number 2

Table of Contents

1. Introduction ................................................................. 1
2. Organization and Staffing ................................................. 2
3. Current Task Areas .......................................................... 2
4. Summary of Progress ...................................................... 5
5. Travel Related to STEP ................................................... 8
6. Visitors .............................................................................. 9
7. Publications, Reports, and Briefings ................................. 9
8. Report Distribution .......................................................... 10
1. INTRODUCTION

This is the Second Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, Development Organizations, Independent Verification and Validation (IV&V) Organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.

Project Support

The following contract provides support for this project.

Title: "Software Test and Evaluation Project, Phases III and IV"
Funding Agency: Naval Air Development Center (NADC)
Contract Number: F33657-82-C-2083
GIT Project Number: G36-611
Principal Investigator: Richard A. Dewille
2. ORGANIZATION AND STAFFING

Faculty

DeMillo, Richard A. - Professor  
Calviano, Ross A. - Senior Research Scientist  
Martin, Rhonda J. - Research Scientist II  
Passafiume, John F. - Senior Research Scientist

Staff

Myrick, Dorothy K. - Clerk/Typist  
Richards, Esther E. - Administrative Secretary  
Seay, Brenda G. - Student Assistant

Students

Pilsal, M. Sinasi - Graduate Research Assistant  
Brannen, James M. - Graduate Research Assistant  
Felix, Matija B. - Graduate Research Assistant  
Offutt, VI, A. Jefferson - Graduate Research Assistant

3. CURRENT TASK AREAS

The work to be performed has been organized into the task areas identified below.

DTRF Task Areas

A. Software Test and Evaluation Guidebook

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition lifecycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. Software Test and Evaluation Model

Task 8 of the STEP Statement of Work states that "The effects of alternative software lifecycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software lifecycles. Examples of software lifecycles include the traditional waterfall lifecycles and the currently advocated evolutionary lifecycles.
C. NADC/Technology Demonstration

The NADC/Technology Demonstration is intended to provide a laboratory framework for carrying out most of the technology-based STEP activities. These activities are described in the STEP Statement of Work Tasks 3 - 9 and the Software Technology for Adaptable, Reliable Systems (STARS) Program Task Areas 5 - 7. STEP Tasks 3 - 9 are concerned with formulating guidance statements, assessing the effects of new policy guidance, gathering evidence which indicates the effectiveness of systematic program testing, developing a software T&E model, and developing a software T&E guidebook. STARS Program Task Areas 5 - 7 are concerned with providing a test technology baseline and identification procedures, providing demonstration and qualification procedures for test technologies, and providing functional requirements for Ada* Programming Support Environment (APSE) test environments.

D. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the Software Test and Evaluation Guidebook.

F. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

* Ada is a registered trademark of the U.S. Government, Ada Joint Program Office
STARS Program Task Areas

F. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USD&RE). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

G. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USD&RE. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

H. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

I. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

J. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

K. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.

4. SUMMARY OF PROGRESS

A. Software Test and Evaluation Guidebook


B. Software Test and Evaluation Model

During the period of 05 December 1983 - 04 January 1984, the only work in this area was to continue briefing the NSIA study results.
C. NADC/Technology Demonstration

Progress in this task area is still awaiting the location of funds to cover NADC computer charges and costs related to NADC personnel support of STEP activities.

D. Good Examples

During this period there were three activities of interest: 1) A planning meeting took place to develop questions for interviews with cognizant personnel who were associated with the six systems selected for in-depth analysis. The data gathering guides of earlier phases were reviewed for potential benefit. 2) Calls were made to key individuals at agencies to find cognizant personnel, e.g., Fred McCoy at OTEA for M-issileminder and Firefinder. 3) Plans for interviews/site visits were revised to include a possible trip to Hurstville (Redstone Arsenal) in January 1984.

E. Risk Model

During this quarter, work continued on defining the mathematical model appropriate for software T&E. In addition, a bibliography construction project to compile a list of related research efforts began.

F. Technology Management

The project team has been investigating alternative existing organizations to fill this role in the event that the decision is made to forgo the establishment of the Acquisition Panel. These alternatives will be addressed at the next meeting of the STEP Technical Advisory Panel. It is anticipated that at this meeting a data-gathering plan for support of this task will be developed.

G. Technology Insertion

Work on this task to date has emphasized the development of a baseline for CMS-2 tools and the identification, from this baseline, of tools for possible insertion into FASP at NADC.
1. **APSE Test Environment**

During this quarter, a design team was assembled and began regular meetings to isolate requirements and propose preliminary designs of an APSE test environment. Initial issues lists were submitted by team members and used as a basis for future meetings.

2. **Coordination with Related Efforts**

The following reflects ongoing efforts with which STEP is currently coordinating.

**Joint Logistics Commanders Initiatives:** The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development of Tri-Service standards for Software Development (DOD-STD-SDS) and Software Quality Assessment and Measurement (DOD-STD-SQAM). STEP personnel have been involved in the review of these documents and receive updates as they become available. In addition, the Joint Logistics Commanders recently sponsored their third Software workshop, Orlando I (previous workshops were Monterey I & II). One member of the STF/Team was invited to participate in this workshop, providing input to a panel investigating IV&V.

**STARS Program Panels:** The STARS Program has established working groups and/or panels to investigate selected issues. STEP personnel have been active in these groups.

**JSSEE Team:** This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.

**Evaluation and Validation (E&V) Efforts:** Ada compilers and implementations of the Common APSE Interface Set (CAI-S) will be required to undergo a standard evaluation and validation prior to being approved for DoD use. STEP is in the process of establishing a formal coordination link with these efforts. The goal of this coordination is to determine the relationship between E&V and T&E, and to influence the E&V process as necessary by STEP recommendations.

**NSIA C2 Software Development and Acquisition Study:** A STEP Team member recently directed an NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on a software development
effort. Results of this study will feed into the Software T&E Model task.

DDT&E Research Initiatives: In addition to STFP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

K. Laboratory Support

As mentioned in the previous report, a second Fortune 32:16 system was ordered. It arrived on January 12th with a new disk drive controller which provides faster access to the disk. Although both Fortunes are networked with Georgia Tech's local Net-One, they are not connected together. At the moment, there is no lab available in which both machines can be replaced.

The communications equipment and software necessary to transport the STEP Phase I and II Final Report from the Wang to the Fortune System have not yet arrived but should shortly.

We have managed to connect, through the Fortune System, to the NADC system. Due to budgetary restrictions, we still have no account on their systems to actually use their facilities.

5. TRAVEL RELATED TO STEP

Date of Trip: 12 Dec 1983
Individual(s) Traveling: R. J. Martin
Itinerary: NADC, Warminster, PA
Contact: Fred Shockat
Purpose: Discussions of STFP Contract Audit

Date of Trip: 13 Dec 1983
Individual(s) Traveling: R. J. Martin
Itinerary: NADC, Warminster, PA
Purpose: Attend JSSEE Team Meeting
Date of Trip: 14 Dec 77
Individual(s) Traveling: R. J. Martin
Itinerary: Crystal City, VA
Purpose: Brief STEP and C2 software development and Acquisition Study at NSIA Software Committee Membership Meeting

Date of Trip: 7 Dec 83
Individual(s) Traveling: R. A. DeMillo
Itinerary: Monterey, CA
Purpose: To participate in Software Maintenance Workshop

6. VISITORS

Date of Visit: 19 Dec 83
Visitor(s): B. Zempolich (Naval Air Systems Command)
Contact: R. A. DeMillo, R. J. Martin
Purpose: Technical Discussion on STEP

7. PUBLICATIONS, REPORTS, AND BRIEFINGS

Reports

Briefings
Date of Briefing: 14 Dec 83
Briefer(s): R. J. Martin
Location: Crystal City, VA
Audience: NSIA Software Committee
Subject of Briefing: C2 Software Development and Acquisition Study

Date of Briefing: 14 Dec 83
Briefer(s): R. J. Martin
Location: Crystal City, VA
Audience: NSIA Software Committee
Subject of Briefing: STEP
8. Report Distribution

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations ......................... 155
Volume 2: Software Test and Evaluation:
           State-of-the-Art Overview .......................... 163
Volume 3: Software Test and Evaluation:
           Current Defense Practices Overview ................. 151
Volume 4: Transcript of STEP Workshop, March 1982* .......... 45
Volume 5: Report of Expert Panel on
           Software Test and Evaluation ........................ 156
Volume 6: Tactical Computer System Applicability Study** ...... 11

* Volume 4 distribution is limited to a selected audience.

** The distribution of Volume 6 is being handled by its author, Dr. James Leathrum of Clemson University.
THE SOFTWARE TEST AND EVALUATION PROJECT
PHASES III & IV

Supported by
NAVAL AIR DEVELOPMENT CENTER (NADC)
Contract: F33657-82-G-2083


OCTOBER 1984

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
THE SOFTWARE TEST AND EVALUATION PROJECT
PHASES III & IV

Quarterly Progress Report Number 3
06 January 1984 - 05 April 1984

19 Oct 1984

Supported by

Naval Air Development Center (NADC)
Contract: F33657-82-G-2083
GIT Project: G36-611

School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332
The Software Test and Evaluation Project: Phases III & IV
Quarterly Progress Report Number 3

Table of Contents

1. Introduction ....................................................... 1
2. Organization and Staffing ......................................... 3
3. Current Task Areas ................................................ 4
4. Summary of Progress ................................................ 8
5. Travel Related to STEP ............................................. 14
6. Visitors .............................................................. 16
7. Publications, Reports, and Briefings .............................. 17
8. Report Distribution .................................................. 19
1. INTRODUCTION

This is the Third Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.
Project Support

The following contract provides support for this project.

Title: "Software Test and Evaluation Project, Phases III and IV"
Funding Agency: Naval Air Development Center (NADC)
Contract Number: F33657-82-G-2083
GIT Project Number: G36-611
Principal Investigator: Richard A. DeMillo
2. ORGANIZATION AND STAFFING

Faculty

DeMillo, Richard A. - Professor
Gagliano, Ross A. - Senior Research Scientist
Martin, Rhonda J. - Research Scientist II
McCracken, W. Michael - Senior Research Engineer
Passafiume, John F. - Senior Research Scientist

Staff

Myrick, Dorothy K. - Clerk/Typist
Richards, Esther E. - Administrative Secretary
Seay, Brenda G. - Student Assistant

Students

Bilsel, M. Sinasi - Graduate Research Assistant
Offutt, VI, A. Jefferson - Graduate Research Assistant
3. CURRENT TASK AREAS

The work to be performed has been organized into the task areas identified below.

DPT&E Task Areas

A. Software Test and Evaluation Guidebook

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition lifecycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. Software Test and Evaluation Model

Task 8 of the STEP Statement of Work states that "The effects of alternative software lifecycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software lifecycles. Examples of software lifecycles include the traditional waterfall lifecycles and the currently advocated evolutionary lifecycles.
C. Technology Demonstration

The Technology Demonstration is intended to provide a laboratory framework for carrying out most of the technology-based STEP activities. These activities are described in the STEP Statement of Work Tasks 3 - 9 and the Software Technology for Adaptable, Reliable Systems (STARS) Program Task Areas 5 - 7. STEP Tasks 3 - 9 are concerned with formulating guidance statements, assessing the effects of new policy guidance, gathering evidence which indicates the effectiveness of systematic program testing, developing a software T&E model, and developing a software T&E guidebook. STARS Program Task Areas 5 - 7 are concerned with providing a test technology baseline and identification procedures, providing demonstration and qualification procedures for test technologies, and providing functional requirements for Ada (*) Programming Support Environment (APSE) test environments.

D. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the Software Test and Evaluation Guidebook.

E. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

(*) Ada is a registered trademark of the U.S. Government AJPO
STARS Program Task Areas

F. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USDAR&E). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

G. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USDAR&E. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

H. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

I. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals, and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

J. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

K. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.
4. SUMMARY OF PROGRESS

A. Software Test and Evaluation Guidebook

The primary accomplishment in this area during this reporting period was the completion of "Policy Recommendations for Software Test and Evaluation: System Level Test Issues" by R. A. DeMillo, R. A. Gagliano, R. J. Martin, and J. F. Passafiume, which was subsequently published in the Journal of Test and Evaluation, Volume V, Number 1, January 1984. The TEMP evaluation checklist, which was included as an appendix to the paper, will form the basis upon which the Software Test and Evaluation Guidebook will be built.

The evolution of the checklist into the guidebook was a topic of discussion at the Technical Advisory Panel Meeting in early April. First of all, the panel formulated questions critical to the assessment of adequate and suitable testing. A mapping of the list to the checklist was then constructed. In addition, the panel reviewed and commented on the checklist in terms of its format and ease of use.

B. Software Test and Evaluation Model

There was no activity in this task area during the reporting period.

C. Technology Demonstration

During this quarter, access to NADC's Facility for Automated Software Production (FASP) was established. LAMPS software documentation was received to allow preparation for the use of that program for technology demonstration purposes.

In addition, initial meetings were held with personnel at the Naval Weapons Center, China Lake, CA to investigate the feasibility of using the A7-E Program, which is employing the Parnas Methodology in its software development, for technology demonstration purposes. A literature search on the Parnas efforts, sponsored by the Naval Research Laboratory, was also completed during this quarter.
D. **Good Examples**

Planning continued to be the major activity related to this task during this reporting period. A strategy was developed for the contacts with the key personnel at the various sites of the six systems previously reported. The major results were: a plan for the set of telephone contacts, a pro-forma letter to be sent in advance to the contact persons, and a systematic data gathering procedure. Some assistance may be required from the Military Advisory Panel or DDT&E.

E. **Risk Model**

Work on risk modelling during this period has concentrated on verifying the applicability of a particular model to the software T&E problem and isolating problems for future study. As described in previous progress reports, the model under investigation is a tableaux model for decision-making under uncertainty. Key aspects of the model to be verified include choosing utility (risk) functions to measure the degree of desirability of outcomes of tests and deriving an appropriate concept of optimality with respect to these functions. The outline of the tableaux model was presented to the STEP Technical Advisory Panel during this period.

A survey of existing quality metrics was begun. The utility functions which have the most likelihood of providing meaningful results for software T&E are thought to be derivable from such metrics weighted by cost measurements. The question of what optimality means for software T&E decision-making is still open. Activity has also centered on the compilation of axioms and decision criteria that have proved useful in other applications of the tableux model.

Finally during this period, planning commenced for a workshop on software risk modelling. The STEP team has observed that there are relatively few efforts underway in this area and that they are geographically dispersed. The planned workshop will allow for an interchange of ideas related to software risk modelling, facilitate an evaluation of the state-of-the-art and provide for a cross-fertilization of ideas in this emerging field.
F. **Technology Management**

During this quarter, the project team developed a set of issues related to the management of software test and evaluation. This issues were derived from the 21 STEP recommendations and address the need to continue the current efforts by STEP. A draft software test and evaluation management plan will then be developed and presented to the Military Advisory Panel for their consideration.

G. **Technology Insertion**

Work on the development of the baseline for CMS-2 tools continues. A list of the known tools will be developed and submitted to the Navy representatives for validation and comment. The team will then begin to compile a list of tools supporting the other military languages, namely, FORTRAN and Jovial. The Technical Advisory Panel was asked to participate in the development of a preliminary working paper on the validation procedures to be used for test tools.

H. **APSE Test Environment**

Work on the APSE test environment during this period began with the formation of a preliminary design team consisting of key STEP staff members and graduate students from Georgia Tech's School of Information and Computer Science. Preliminary to the derivation of functional requirements for the test environment, the team reviewed the usage scenarios that such an environment would be called upon to support. Of special importance to these early discussions was the relationship of the environment to the requirements imposed by DoD-STD-SDS.

A preliminary taxonomy of requirements based on a three-dimensional classification scheme was investigated. The first dimension included the lifecycle activity which imposed the requirement. The second took into account the nature of the requirement (e.g., reporting, data gathering, technological). The third dimension attempted a prioritization of the requirement (e.g., required by DoD-STD-SDS, highly desirable, desirable, low priority). The results of this classification will be used to select an initial set of functional capabilities.
A number of design considerations were also discussed. These included interfaces with and impact of native operating systems, approaches to efficiency and isolation in a test environment, and the feasibility of creating a subenvironment of the planned Joint Service Software Engineering Environment. Work was begun on a preliminary draft of the test environment requirements, with team members critiquing portions of the draft as they became available.

I. Coordination with Related Efforts

The following reflects ongoing efforts with which STEP is currently coordinating.

Joint Logistics Commanders Initiatives: The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development of Tri-Service standards for Software Development (DoD-STD-SDS) and Software Quality Assessment and Measurement (DoD-STD-SQAM). STEP personnel have been involved in the review of these documents and receive updates as they become available.

STARS Program Panels: The STARS Program has established working groups and/or panels to investigate selected issues. STEP personnel have been active in these groups.

JSSEE Team: This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.
Evaluation and Validation (E&V) Team: This Tri-Service Team is tasked with developing the technology necessary to evaluate the suitability and effectiveness of Ada Programming Support Environments (APSEs) and validate conformance to appropriate military standards. Maximum cooperation between the STEP and E&V tasks promises the following benefits: 1) The efforts of the E&V Task will allow the insertion of demonstrated risk reduction technology (in the form of support software which has been thoroughly evaluated and validated) into the acquisition cycle. 2) The testing tool qualification procedures currently being developed by STEP will be elaborated by the E&V Task and inserted into an environment where the standard operating procedures include the E&V of support software. 3) The functional requirements for APSE test environments currently being developed by STEP will be supported by a technology which ensures their implementation. Because of these opportunities, a STEP representative is now participating in the E&V task.

NSIA C2 Software Development and Acquisition Study: STEP Team members were recently involved in a NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on a software development effort. Results of this study will feed into the Software T&E Model task.

DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

J. Direct Support to DDT&E

Direct support activities, fell into two specific categories during this period. First, STEP provided support to the Product Engineering Support Office (PESO) through DDT&E for a software review of the Army's Patriot Air Defense System. Second, STEP undertook planning for a workshop on decision support systems to supplement the discussions begun at previous DDT&E Research Initiatives Roundtable meetings.
At the request of DDT&E and PESO, STEP staff members accompanied PESO staff to Patriot project offices at Raytheon Corporation (Prime Contractor) and Teledyne-Brown Corporation (V&V Contractor). The purpose of these visits was to aid in the assessment of Patriot software maturity -- specifically, whether or not the software was progressing toward stated RAM goals. During these visits, STEP staff members were briefed on the status of Patriot software with particular emphasis on planning for the Follow-On Evaluation III testing scheduled for mid-1984. STEP staff members were allowed to question engineers from the contractors' organizations as well as the Army's Patriot Program Office. A report summarizing the results of these visits is forthcoming.

The planning for a decision support system (DSS) workshop began at the direction of DDT&E. During meetings with researchers involved in various DDT&E initiatives, it was determined that many areas of mutual interest existed in the DSS arena. In order to share insights, problems and technical approaches, STEP staff members have arranged for a two-day workshop to be held at Georgia Tech and attended by DDT&E Research Initiative investigators. The subject matter of the workshop is planned to include an overview of DSS technology, the DSS requirements of DDT&E, and the hardware and software implications of those requirements. Ample time is planned for discussions aimed at crystallizing a plan of action from the workshop participants. The goal of the session is to produce a coordinated plan upon which DDT&E can act.

K. Laboratory Support

During this period, Fortune terminals were installed for Ross Gagliano and Mike McCracken. This allows them to access the Fortune system and Georgia Tech's other computing facilities. In addition, we have been provided access to the Facility for Automated Software Production (FASP) at NADC. We have also received advanced word processing software for installation on the Fortune system.
5. TRAVEL RELATED TO STEP

Date of Trip: 19 Jan 1984
Individual(s) Traveling: R. A. DeMillo
Itinerary: The Pentagon, Washington, D.C.
Contact: J. Bolino
Purpose: Discussion of Future STARS Funding/Tasks

Date of Trip: 26-27 Jan 1984
Individual(s) Traveling: R. A. DeMillo
Itinerary: Raytheon, Bedford, MA
Contact: B. Welch
Purpose: Review of Patriot Software

Date of Trip: 31 Jan 1984 - 01 Feb 1984
Individual(s) Traveling: R. J. Martin
Itinerary: STARS Office, Arlington, VA
Contact: H. Steubing
Purpose: Attend JSSEE Team Meeting

Date of Trip: 02 Feb 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin, R. A. Gagliano
Itinerary: BDM, Tysons Corner, VA
Contact: C. Watt
Purpose: Attend DDT&E Roundtable

Date of Trip: 03 Feb 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: Quantico, VA
Contact: Col. Cobbel
Purpose: Brief STEP

Date of Trip: 09 Feb 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: Teledyne Brown Engineering, Huntsville, AL
Contact: H. Welch
Purpose: Review Patriot Software

Date of Trip: 23 Feb 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: IDA, Alexandria, VA
Contact: J. Kramer
Purpose: Discussions of Relationship of E&V to STEP

Date of Trip: 27 Feb 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: AFOTEC, Kirtland AFB, NM
Contact: Col. Mueller
Purpose: Brief STEP
Date of Trip: 28 Feb 1984
Individual(s) Traveling: R. A. DeMillo
Itinerary: Camp Pendleton, CA
Contact: Col. Hudson
Purpose: Brief STEP

Date of Trip: 28-29 Feb 1984
Individual(s) Traveling: R. J. Martin
Itinerary: NOSC, San Diego, CA
Contact: H. Steubing/G. Myers
Purpose: Attend JSSEE Team Meeting

Date of Trip: 05 Mar 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: NWC, China Lake, CA
Contact: S. Fryer
Purpose: Brief STEP and Discuss Future Activities

Date of Trip: 07-08 Mar 1984
Individual(s) Traveling: R. J. Martin
Itinerary: Wright Patterson AFB, OH
Contact: G. Castor
Purpose: Attend E&V Meeting

Date of Trip: 27 Mar 1984
Individual(s) Traveling: R. J. Martin
Itinerary: IDA, Alexandria, VA
Contact: B. Bailey/J. Kramer
Purpose: Attend E&V Requirements Sub-group Meeting

Date of Trip: 29 Mar 1984
Individual(s) Traveling: R. A. DeMillo
Itinerary: The Pentagon, Washington, D.C.
Contact: J. Bolino
Purpose: Deliver STEP Monthly Oral Progress Report

Date of Trip: 02-06 April 1984
Individual(s) Traveling: R. J. Martin
Itinerary: Airlie, VA
Contact: G. Castor
Purpose: Attend E&V Workshop
6. **VISITORS**

Date of Visit: 04-05 April 1984  
Visitor(s): STEP Technical Advisory Panel  
  John Bowen (Hughes Aircraft Company)  
  Carolyn Gannon (General Research Corporation)  
  Richard Lipton (Princeton University)  
  Edward Miller, Jr. (Software Research Assoc.)  
  Capt. William Nelson (Brooks AFB, TX)  
  Fred Sayward (ITT)  
  Fred Shocket (Naval Air Development Center)  
Contact: R. A. DeMillo & STEP Personnel  
Purpose: Attend STEP Technical Advisory Panel meeting
7. PUBLICATIONS, REPORTS, AND BRIEFINGS

Publications


Reports


Briefings

Date of Briefing: 03 Feb 1984
Briefer(s): R. A. DeMillo
Location: Quantico, VA
Audience: Marine Corps OTEA
Subject of Briefing: STEP

Date of Briefing: 23 Feb 1984
Briefer(s): R. A. DeMillo
Location: IDA, Alexandria, VA
Audience: AJPO & E&V Team Representatives
Subject of Briefing: STEP and its Relationship to E&V

Date of Briefing: 27 Feb 1984
Briefer(s): R. A. DeMillo
Location: Kirtland AFB, NM
Audience: AFOTEC
Subject of Briefing: STEP

Date of Briefing: 28 Feb 1984
Briefer(s): R. A. DeMillo
Location: Camp Pendleton, CA
Audience: Marine Corps FCDSSA
Subject of Briefing: STEP

Date of Briefing: 05 Mar 1984
Briefer(s): R. A. DeMillo
Location: China Lake, CA
Audience: Naval Weapons Center Personnel
Subject of Briefing: STEP
Date of Briefing: 07 Mar 1984
Briefer(s): R. J. Martin
Location: Wright Patterson AFB, OH
Audience: E&V Team
Subject of Briefing: STEP

Date of Briefing: 19 Mar 1984
Briefer(s): R. A. DeMillo
Location: Atlanta, GA
Audience: GUIDE Meeting Attendees
Subject of Briefing: STEP
8. REPORT DISTRIBUTION

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations .................. 163

Volume 2: Software Test and Evaluation:
    State-of-the-Art Overview ......................... 171

Volume 3: Software Test and Evaluation:
    Current Defense Practices Overview ............ 158

Volume 4: Transcript of STEP Workshop, March 1982 (*) .. 45

Volume 5: Report of Expert Panel on
    Software Test and Evaluation .................. 163

Volume 6: Tactical Computer System
    Applicability Study (**) ......................... 11

(*) Volume 4 distribution is limited to a selected audience.

(**) The distribution of Volume 6 is being handled by its author, Dr. James Leathrum of Clemson University.
# Table of Contents

1. Introduction ......................................................... 1
2. Organization and Staffing ............................................. 3
3. Current Task Areas ..................................................... 4
4. Summary of Progress ..................................................... 8
5. Travel Related to STEP .................................................. 14
6. Visitors ................................................................. 16
7. Publications, Reports, and Briefings ................................. 17
8. Report Distribution ..................................................... 18
1. INTRODUCTION

This is the Fourth Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.
Project Support

The following contract provides support for this project.

Title: "Software Test and Evaluation Project, Phases III and IV"
Funding Agency: Naval Air Development Center (NADC)
Contract Number: F33657-82-G-2083
GIT Project Number: G36-611
Principal Investigator: Richard A. DeMillo
2. ORGANIZATION AND STAFFING

Faculty

DeMillo, Richard A. - Professor
Gagliano, Ross A. - Senior Research Scientist
Martin, Rhonda J. - Research Scientist II
McCracken, W. Michael - Senior Research Engineer
Passafiume, John F. - Senior Research Engineer

Staff

Hardy, Melinda - Student Assistant
Illingworth, Karen - Clerk/Typist
Myrick, Dorothy K. - Clerk/Typist
Perry, Rosalind - Clerk/Typist
Richards, Esther E. - Administrative Secretary
Seay, Brenda G. - Student Assistant

Students

Bilsel, M. Sinasi - Graduate Research Assistant
Offutt, VI, A. Jefferson - Graduate Research Assistant
3. **CURRENT TASK AREAS**

The work to be performed has been organized into the task areas identified below.

**DDT&E Task Areas**

**A. Software Test and Evaluation Guidebook**

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition lifecycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

**B. Software Test and Evaluation Model**

Task 8 of the STEP Statement of Work states that "The effects of alternative software lifecycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software lifecycles. Examples of software lifecycles include the traditional waterfall lifecycles and the currently advocated evolutionary lifecycles.
C. Technology Demonstration

The Technology Demonstration is intended to provide a laboratory framework for carrying out most of the technology-based STEP activities. These activities are described in the STEP Statement of Work Tasks 3 - 9 and the Software Technology for Adaptable, Reliable Systems (STARS) Program Task Areas 5 - 7. STEP Tasks 3 - 9 are concerned with formulating guidance statements, assessing the effects of new policy guidance, gathering evidence which indicates the effectiveness of systematic program testing, developing a software T&E model, and developing a software T&E guidebook. STARS Program Task Areas 5 - 7 are concerned with providing a test technology baseline and identification procedures, providing demonstration and qualification procedures for test technologies, and providing functional requirements for Ada (*) Programming Support Environment (APSE) test environments.

D. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the software T&E guidebook.

E. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

(*) Ada is a registered trademark of the U.S. Government AJPO
STARS Program Task Areas

F. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USD R&E). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

G. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USD R&E. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

H. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

I. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to:
1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals, and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

J. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

K. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.
4. **SUMMARY OF PROGRESS**

A. **Software Test and Evaluation Guidebook**

Work began on compiling the first draft of the Guidebook. In the interests of user convenience, the decision was made to produce the guidebook in stages, the first of which will address the TEMP Checklist described in previous reports. The initial plans for this guidebook include separate sections for checklists, explanatory notes, glossaries and selected readings. Subsequent guides will deal with other review aspects of software T&E.

B. **Software Test and Evaluation Model**

During this quarter, a format for the T&E model was arrived at. The goal of the model is to aid DoD decision-makers in relating acquisition milestones to the expected maturity of software systems and subsystems. Accordingly, the model will consist of a matrix which makes this traceability explicit. The form of the matrix is shown below:

```
<table>
<thead>
<tr>
<th>Acquisition Phases and Milestones</th>
</tr>
</thead>
<tbody>
<tr>
<td>Milestone</td>
</tr>
<tr>
<td>Phase 0</td>
</tr>
<tr>
<td>Lifecycle A</td>
</tr>
<tr>
<td>Lifecycle B</td>
</tr>
<tr>
<td>Lifecycle C</td>
</tr>
<tr>
<td>Lifecycle D</td>
</tr>
</tbody>
</table>
```

The entries of the matrix will be succinct descriptions of software status versus acquisition milestones for each of the software lifecycles A, B, C, D, ... The initial portion of this research will concentrate on examining existing systems for examples of software developments that correspond to these lifecycles.
C. **Technology Demonstration**

Technology demonstration-related efforts this quarter centered on trying to make contact with B. Zempolich of NAVAIR to discuss possible funding for work with the A7-E Program. Further progress in this area is awaiting the location of such funding.

D. **Good Examples**

Efforts this quarter centered on further development of the data gathering procedure that will be used to acquire the needed evidence of good software testing examples. The core of the data gathering procedure consists of a detailed set of questions targeted to determine the extent and success of software testing performed on the program. The questions are organized in a hierarchical manner to allow tracing of the test requirements from the Test and Evaluation Master Plan (TEMP) to the actual development and/or operational test plans, procedures, and results. The questions are derived such that their answers should be able to be substantiated by actual documentation of testing activities or events.

The first site visit is planned for 06 July 1984 with the Missleminder (AN/TSQ-73) Program Office at the U.S. Army Missile Command in Huntsville, AL. An additional objective for this visit, besides that of gathering good examples, is to evaluate the data gathering procedure in terms of its capabilities with respect to (1) efficiently extracting the data necessary to determine if a good example of software testing exists, and (2) locating the substantiating documentation.

E. **Risk Model**

The only reportable activity regarding risk modelling during this period was the continued planning for a 1985 symposium on software risk modelling. Participants in this workshop will include software professionals with related interest as well as decision theorists who may be able to contribute their expertise to the decision problems that are to be encompassed by the STEP risk model.
F. **Technology Management**

During this quarter, the STEP team, in concert with the STEP Military Advisory Panel (MAP), reviewed the set of issues and proposed solutions at a meeting in Washington, D.C. The MAP members were asked to provide detailed comments from which the team would then prepare a draft management plan. Comments were received from two of the members of the Panel during the reporting period.

G. **Technology Insertion**

Work continues on the refinement of the baseline for CMS-2 tools. A list of known tools has been developed and submitted to the Navy representatives to the STEP Military Advisory Panel for comment. Comments are still outstanding at the end of the reporting period. The team also began to develop a list of contacts for use when compiling the lists of tools for the other military languages. We plan to hire another individual to participate in this task.

H. **APSE Test Environment**

The major result of this period's work on APSE test environments was the preparation and on-schedule delivery of a draft requirements document. The requirements span both general and specific requirements for functional capabilities.

The perceived need for integrity, reliability and integration have led to the concept of a subenvironment -- that is, a closed, self-contained entity to house the software T&E capabilities that are recommended for APSE inclusion. In addition to the general requirements, the following requirements are significant:

1. **Test Building Capabilities:** The user/tester should have the functional capability to construct useful test sets/case/scenarios within the subenvironment.

2. **Test Evaluation Capabilities:** The user/tester should have the functional capability to assess the effectiveness of his test data by coverage analysis, instrumentation, and error-based methods, as well as by combinations of methods that can be tailored to specific testing tasks.
3. Test Support Capabilities: Sufficient run-time support functions should be provided to allow the user/tester to carry out test activities with minimal overhead and minor impact on the performance of the system for other users. These capabilities should range from simple test harnesses to such features as automatic data capture monitors.

The draft requirements document will be briefed to appropriate organizations and coordination will continue with the STARS office to provide effective feedback for future revisions.

I. **Coordination with Related Efforts**

The following reflects ongoing efforts with which STEP is currently coordinating.

**Joint Logistics Commanders Initiatives:** The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development of Tri-Service standards for Software Development (DoD-STD-SDS) and Software Quality Assessment and Measurement (DoD-STD-SQAM). STEP personnel have been involved in the review of these documents and receive updates as they become available.

**JSSEE Team:** This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.
Evaluation and Validation (E&V) Team: This Tri-Service Team is tasked with developing the technology necessary to evaluate the suitability and effectiveness of Ada Programming Support Environments (APSEs) and validate conformance to appropriate military standards. Maximum cooperation between the STEP and E&V tasks promises the following benefits:

1) The efforts of the E&V Task will allow the insertion of demonstrated risk reduction technology (in the form of support software which has been thoroughly evaluated and validated) into the acquisition cycle. 2) The testing tool qualification procedures currently being developed by STEP will be elaborated by the E&V Task and inserted into an environment where the standard operating procedures include the E&V of support software. 3) The functional requirements for APSE test environments currently being developed by STEP will be supported by a technology which ensures their implementation. Because of these opportunities, a STEP representative is now participating in the E&V task.

NSIA C2 Software Development and Acquisition Study: STEP Team members were recently involved in a NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on a software development effort. Results of this study will feed into the Software T&E Model task.

DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.
J. Direct Support to DDT&E

The major support activity undertaken during this period was an expansion of the STEP role in evaluating the software in the Patriot System. Members of the STEP team accompanied the acting Direct, Defense T&E to the Patriot Program Office in Huntsville. After being briefed on the status of Patriot and following a meeting with the Chief Scientist, OTEA, STEP was requested to provide direct support to OTEA in evaluating the results of FOE III. This activity will be conducted through a separate contract via the U. S. Army, during which STEP personnel will be available in Atlanta, Washington and White Sands Missile Range to aid in the conduct of this test.

K. Laboratory Support

Our major laboratory concern this quarter has been with regard to maintenance of our Fortune equipment. After realizing that we had no maintenance agreement for the Fortune 32:16, we were able to work out an agreement with Fortune whereby they will provide free maintenance until 30 Sep 1984. At that time, we will need to purchase a maintenance agreement.

In addition, during this quarter, one of our Fortune terminals and two Fortune keyboards were stolen. We were able to get replacement equipment and have since improved security in our lab.

Finally, we have obtained bisynchronous software and hardware for use when moving documents between the Fortune system and our Wang word processing workstation.
5. **TRAVEL RELATED TO STEP**

Date of Trip: 09 Apr 1984  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: The Pentagon, Washington, D.C.  
Contact: J. Bolino  
Purpose: Contract Discussions

Date of Trip: 18 Apr 1984  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: STARS Office, Arlington, VA  
Contact: H. Steubing  
Purpose: Attend JSSEE Team Meeting

Date of Trip: 18-19 Apr 1984  
Individual(s) Traveling: R. J. Martin  
Itinerary: STARS Office, Arlington, VA  
Contact: H. Steubing  
Purpose: Attend JSSEE Team Meeting

Date of Trip: 23 Apr 1984  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: Pittsburgh, PA  
Contact: P. Henderson  
Purpose: Attend ACM SIGSOFT/SIGPLAN Symposium on Practical Software Development Environments

Date of Trip: 24-26 Apr 1984  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: The Pentagon, Washington, D.C.  
Contact: D. Greenlee  
Purpose: Deliver STEP Monthly Oral Progress Report

Date of Trip: 25-26 Apr 1984  
Individual(s) Traveling: R. J. Martin, W. M. McCracken  
Itinerary: The Pentagon, Washington, D.C.  
Contact: D. Greenlee  
Purpose: Deliver STEP Monthly Oral Progress Report

Date of Trip: 16-18 May 1984  
Individual(s) Traveling: R. A. DeMillo, R. J. Martin, W. M. McCracken, J. F. Passafiume  
Itinerary: The Pentagon, Washington, D.C.  
Contact: D. Greenlee  
Purpose: Attend STEP Military Advisory Panel Meeting

Date of Trip: 17-18 May 1984  
Individual(s) Traveling: R. A. Gagliano  
Itinerary: The Pentagon, Washington, D.C.  
Contact: D. Greenlee  
Purpose: Attend STEP Military Advisory Panel Meeting
Date of Trip: 05–08 Jun 1984  
Individual(s) Traveling: R. J. Martin  
Itinerary: Wright Patterson AFB, OH  
Contact: G. Castor  
Purpose: Attend E&V Meeting

Date of Trip: 06 Jun 1984  
Individual(s) Traveling: R. A. DeMillo, W. M. McCracken  
Itinerary: Huntsville, AL  
Contact: H. Welch  
Purpose: Visit Patriot Program Office w/C. Watt

Date of Trip: 14 Jun 1984  
Individual(s) Traveling: R. A. DeMillo, R. J. Martin  
Itinerary: The Pentagon, Washington, D.C.  
Contact: C. Watt  
Purpose: Discussions of Patriot Program w/C. Watt & P. Dickinson
6. VISITORS

Date of Visit: 14 Jun 1984
Visitor(s): C. Watt
Contact: R. A. DeMillo, W. M. McCracken
Purpose: Discussions of Patriot Program
7. PUBLICATIONS, REPORTS, AND BRIEFINGS

Publications


Reports


Briefings

None during this reporting period.
8. REPORT DISTRIBUTION

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations ..................... 170
Volume 2: Software Test and Evaluation:
         State-of-the-Art Overview .................. 178
Volume 3: Software Test and Evaluation:
         Current Defense Practices Overview ........ 165
Volume 4: Transcript of STEP Workshop, March 1982 (*) .. 45
Volume 5: Report of Expert Panel on
         Software Test and Evaluation ............... 169
Volume 6: Tactical Computer System
         Applicability Study (**) .................... 11

(*) Volume 4 distribution is limited to a selected audience.

(**) The distribution of Volume 6 is being handled by its author, Dr. James Leathrum of Clemson University.
QUARTERLY PROGRESS REPORT NO. 5
GIT Project G36-811

THE SOFTWARE TEST AND EVALUATION PROJECT
PHASES III & IV

Supported by

NAVAL AIR DEVELOPMENT CENTER (NADC)
Contract: F33657-82-G-2083


FEBRUARY 1985

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
THE SOFTWARE TEST AND EVALUATION PROJECT

PHASES III & IV

Quarterly Progress Report Number 5
06 July 1984 – 05 October 1984

01 Feb 1985

Supported by

Naval Air Development Center (NADC)
Contract: F33657-82-G-2083
GIT Project: G36-611

School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332
The Software Test and Evaluation Project: Phases III & IV
Quarterly Progress Report Number 5

Table of Contents

1. Introduction .......................................................... 1
2. Organization and Staffing ........................................ 3
3. Current Task Areas .................................................. 4
4. Summary of Progress ............................................... 8
5. Travel Related to STEP .......................................... 13
6. Visitors ............................................................... 16
7. Publications, Reports, and Briefings ......................... 17
8. Report Distribution ............................................... 18
1. INTRODUCTION

This is the Fifth Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.
Project Support

The following contract provides support for this project.

Title:  "Software Test and Evaluation Project, Phases III and IV"
Funding Agency:  Naval Air Development Center (NADC)
Contract Number:  F33657-82-G-2083
GIT Project Number:  G36-611
Principal Investigator:  Richard A. DeMillo
2. ORGANIZATION AND STAFFING

Faculty

DeMillo, Richard A. - Professor (50%)
Gagliano, Ross A. - Senior Research Scientist (67%)
Grover Jeffrey L. - Research Scientist II (17%)
Martin, Rhonda J. - Research Scientist II (50%)
McCracken, W. Michael - Senior Research Engineer (33%)
Passafiume, John F. - Senior Research Engineer (42%)

Note: During this reporting period, faculty personnel levels of effort were reduced due to assignments in other areas. Actual levels of effort expended on STEP-related activities are reported in parentheses above.

Staff

Hardy, Melinda - Student Assistant
Hinds, Lindsay - Student Assistant
Perry, Rosalind - Clerk/Typist
Richards, Esther E. - Administrative Secretary
Richliew, Ann K. - Clerk/Typist
Seay, Brenda G. - Student Assistant
Wandrick, Greg - Student Assistant

Students

Bilsel, M. Sinasi - Graduate Research Assistant
Harder, Rita M. - Graduate Research Assistant
Offutt, VI, A. Jefferson - Graduate Research Assistant
3. CURRENT TASK AREAS

The work to be performed has been organized into the task areas identified below.

DDT&E Task Areas

A. Software Test and Evaluation Guidebook

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition life cycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. Software Test and Evaluation Model

Task 8 of the STEP Statement of Work states that "The effects of alternative software life cycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software life cycles. Examples of software life cycles include the traditional waterfall life cycles and the currently advocated evolutionary life cycles.
C. Technology Demonstration

The Technology Demonstration is intended to provide a laboratory framework for carrying out most of the technology-based STEP activities. These activities are described in the STEP Statement of Work Tasks 3 - 9 and the Software Technology for Adaptable, Reliable Systems (STARS) Program Task Areas 5 - 7. STEP Tasks 3 - 9 are concerned with formulating guidance statements, assessing the effects of new policy guidance, gathering evidence which indicates the effectiveness of systematic program testing, developing a software T&E model, and developing a software T&E guidebook. STARS Program Task Areas 5 - 7 are concerned with providing a test technology baseline and identification procedures, providing demonstration and qualification procedures for test technologies, and providing functional requirements for Ada (*) Programming Support Environment (APSE) test environments.

D. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the software T&E guidebook.

E. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

(*) Ada is a registered trademark of the U.S. Government AJPO
STARS Program Task Areas

F. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USD R&E). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

G. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USD R&E. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

H. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

I. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals, and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

J. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

K. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.
4. SUMMARY OF PROGRESS

A. Software Test and Evaluation Guidebook

A draft guidebook for evaluation of software testing in Test and Evaluation Master plans was completed and delivered to DDT&E during this period. This draft is intended to be the first volume of a multi-part "Software Test and Evaluation Guidebook". Copies of the draft were also circulated to the members of the STEP Military Advisory Panel and others for comment. As planned, the guidebook includes separate sections for checklists, explanatory notes, glossaries and selected readings in a suitable format. Initial discussions were held with DDT&E to determine a briefing schedule for the draft document and a schedule for completing the reviewing and revision process.

Planning for subsequent volumes of the Guidebook also commenced during this period.

B. Software Test and Evaluation Model

Work continued on refining the format of the software T&E model as discussed in the previous quarterly report.

C. Technology Demonstration

Progress in this area is awaiting the location of funding.
D. Good Examples

Efforts this quarter were concentrated on organizing the ITEA Fall Symposium special session on good examples of software testing. Contacts were made with government and industry personnel for the purpose of soliciting papers associated with this subject. The session (the fourth day of the ITEA Fall Symposium) will be organized into four sections with an additional section at the end of the day for a panel discussion.

The first section will be "Software T&E Policy and Guidance" and will discuss the forthcoming revisions to DoD 5000.3 and the draft Software T&E Guidebook for use by the Office of the Director Defense Test and Evaluation. The second session will consist of presentations of software test management activities at the US Army Test and Evaluation Command and the US Air Force Operational Test and Evaluation Command. The third session will consist of presentations of the application of software test tools on recent DoD programs. The final presentation will be a review and status report of MIL-STD-SDS. The panel will consist of some of the participants and will be chaired by Dr. DeMillo.

One problem encountered when organizing the session concerned obtaining sponsor approval for contractor presentations. This resulted in a key speaker being unable to give a talk on the industry perspective of software testing on a DoD project.

E. Risk Model

Work continued on the mathematical basis for the STEP risk modelling work. In addition, discussions were held with cognizant personnel of the Air Force Operational Test and Evaluation Command concerning their work on software maintainability risk assessment. After presentation of the outline of the STEP model to AFOTEC personnel and engineers from BDM corporation, AFOTEC presented the metrics and assessment parameters that are being considered for inclusion in the maintainability model. It was concluded that there are possible relationships between the models that will be explored in the coming months.
F. **Technology Management**

During this reporting period, a draft management plan was developed based on the assumption that the proposed Embedded Computing Resources Acquisition Policy Panel (ECSAPP) will address issues concerning software test and evaluation. The draft management plan was circulated within the STEP Team for review and efforts are currently underway to develop a second draft of the plan.

G. **Technology Insertion**

During this period, a new project team member was assigned primary responsibility for the testing tools baseline and identification task. After completing a brief orientation period, this individual developed a plan for the completion of the task. Expected deliverables include a catalog of software testing tools and a "case study" methodology for identifying tools. Tools currently excluded from further consideration include those with inappropriate target languages, those which are not really test tools or are not in use, and those for which we have been unable to obtain information. Tool search efforts conducted thus far have concentrated on the identification of testing tools which support the Navy standard language, CMS-2. Future near term efforts will be dedicated to the development of a draft of the final version of the CMS-2 Baseline Report.

H. **APSE Test Environment**

There were three major activities concerning the APSE Test Environment that took place during this period. First, the process of gathering and integrating comments on the draft requirements document delivered during the last period began. Second, the draft requirements were briefed and discussed at the July meeting of the Joint Service Software Engineering Environment team meeting in Monterey. Finally, at the suggestion of several reviewers a reorganization of the draft requirements document was begun. While the reorganization will not affect the requirements, it will improve the readability of the document.
I. Coordination with Related Efforts

The following reflects ongoing efforts with which STEP is currently coordinating.

**Joint Logistics Commanders Initiatives:** The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development of Tri-Service standards for Software Development (DoD-STD-SDS) and Software Quality Assessment and Measurement (DoD-STD-SQAM). STEP personnel have been involved in the review of these documents and receive updates as they become available.

**JSSEE Team:** This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.

**Evaluation and Validation (E&V) Team:** This Tri-Service Team is tasked with developing the technology necessary to evaluate the suitability and effectiveness of Ada Programming Support Environments (APSEs) and validate conformance to appropriate military standards. Maximum cooperation between the STEP and E&V tasks promises the following benefits: 1) The efforts of the E&V Task will allow the insertion of demonstrated risk reduction technology (in the form of support software which has been thoroughly evaluated and validated) into the acquisition cycle. 2) The testing tool qualification procedures currently being developed by STEP will be elaborated by the E&V Task and inserted into an environment where the standard operating procedures include the E&V of support software. 3) The functional requirements for APSE test environments currently being developed by STEP will be supported by a technology which ensures their implementation. Because of these opportunities, a STEP representative is now participating in the E&V task.

**NSIA C2 Software Development and Acquisition Study:** STEP Team members were recently involved in a NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on a software development effort. Results of this study will feed into the Software T&E Model task.
DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

J. Direct Support to DDT&E

The major support activity during this period was the commencement of STEP participation in Follow-on Evaluation III of the Patriot System at White Sands Missile Range, New Mexico. Several visits to OTEA offices at White Sands were made. Details of the STEP review of software in this operational test will be made in reports and briefings to OTEA following the completion of the test.

K. Laboratory Support

During this quarter, STEP hired a student assistant to help in the lab. His duties include backing up our accounts on the Fortune system, managing the machine, helping STEP personnel with problems concerning the use of the Fortune, and writing small programs.

During this quarter, we also discovered that STEP was spending much more money using Georgia Tech's computer resources than anticipated. This was partially due to the fact that, in the past, the charging procedures had not been fully understood.
5. **TRAVEL RELATED TO STEP**

**Date of Trip:** 06 Jul 1984  
**Individual(s) Traveling:** R. A. Gagliano, R. J. Martin  
**Itinerary:** Huntsville, AL  
**Contact:** R. Roberts  
**Purpose:** Good Examples Task Discussions

**Date of Trip:** 16-19 Jul 1984  
**Individual(s) Traveling:** R. A. DeMillo, R. J. Martin  
**Itinerary:** Naval Postgraduate School, Monterey, CA  
**Contact:** B. Zempolich & H. Stuebing  
**Purpose:** Attend Navy Software Engineering Research & Development Technology Review and JSSEE Meeting

**Date of Trip:** 20 Jul 1984  
**Individual(s) Traveling:** R. A. DeMillo, R. J. Martin  
**Itinerary:** Albuquerque, NM  
**Contact:** LTC R. Cline  
**Purpose:** Risk Model Discussions with AFOTEC & BDM

**Date of Trip:** 09 Aug 1984  
**Individual(s) Traveling:** R. J. Martin, W. M. McCracken  
**Itinerary:** The Pentagon, Washington, D.C.  
**Contact:** C. Morgan  
**Purpose:** Contract Discussions ref. STARS funding

**Date of Trip:** 14-15 Aug 1984  
**Individual(s) Traveling:** R. J. Martin  
**Itinerary:** Rome, NY  
**Contact:** T. Robbins  
**Purpose:** Attend STARS Measurement DID's Workshop

**Date of Trip:** 15-17 Aug 1984  
**Individual(s) Traveling:** R. J. Martin  
**Itinerary:** Rome Air Development Center, Rome, NY  
**Contact:** H. Stuebing  
**Purpose:** Attend JSSEE Meeting

**Date of Trip:** 15-17 Aug 1984  
**Individual(s) Traveling:** W. M. McCracken  
**Itinerary:** White Sands Missile Range, NM  
**Contact:** L. Pickering  
**Purpose:** Patriot Program Discussions

**Date of Trip:** 29-31 Aug 1984  
**Individual(s) Traveling:** R. A. DeMillo, R. J. Martin, W. M. McCracken  
**Itinerary:** White Sands Missile Range, NM  
**Contact:** L. Pickering  
**Purpose:** Patriot Program Discussions
Date of Trip: 05-06 Sep 1984
Individual(s) Traveling: R. J. Martin
Itinerary: Wright Patterson AFB, Dayton, OH
Contact: V. Castor
Purpose: Attend E&V Meeting

Date of Trip: 11 Sep 1984
Individual(s) Traveling: W. M. McCracken
Itinerary: White Sands Missile Range, NM
Contact: L. Pickering
Purpose: Patriot Program Discussions

Date of Trip: 11-12 Sep 1984
Individual(s) Traveling: J. L. Grover
Itinerary: White Sands Missile Range, NM
Contact: L. Pickering
Purpose: Patriot Program Discussions

Date of Trip: 17-20 Sep 1984
Individual(s) Traveling: R. J. Martin
Itinerary: Santa Barbara, CA
Contact: H. Hart
Purpose: Attend ACM/Ada TEC Future Ada Environments Workshop

Date of Trip: 18-19 Sep 1984
Individual(s) Traveling: W. M. McCracken
Itinerary: Camp Pendelton, CA
Contact: Col. E. Kline
Purpose: Attend 16th Multi-Service Development Test and Evaluation Commanders' Conference

Date of Trip: 19 Sep 1984
Individual(s) Traveling: R. J. Martin
Itinerary: Camp Pendelton, CA
Contact: Col. E. Kline
Purpose: Attend 16th Multi-Service Development Test and Evaluation Commanders' Conference

Date of Trip: 02 Oct 1984
Individual(s) Traveling: W. M. McCracken
Itinerary: Hanscom AFB, MA
Contact: H. Lyness
Purpose: Attend AF/ESD-NSIA Meeting

Date of Trip: 02-03 Oct 1984
Individual(s) Traveling: R. J. Martin
Itinerary: NADC, Warminster, PA
Contact: H. Stuebing
Purpose: Attend JSSEE Meeting
Date of Trip: 04 Oct 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: The Pentagon, Washington, D.C.
Contact: D. Greenlee
Purpose: Contract Discussions
6. VISITORS

Date of Visit: 13 Sep 84
Visitor(s): D. Greenlee
Contact: R. A. DeMillo and STEP personnel
Purpose: Oral Progress Report and Contract Discussions
7. **PUBLICATIONS, REPORTS, AND BRIEFINGS**

**Publications**

None during this reporting period.

**Reports**


**Briefings**

Date of Briefing: 16 Jul 1984  
Briefer(s): R. A. DeMillo  
Location: Naval Postgraduate School, Monterey, CA  
Audience: JSSEE Team  
Subject of Briefing: Functional Requirements for Software Test and Evaluation

Date of Briefing: 19 Jul 1984  
Briefer(s): R. A. DeMillo  
Location: Naval Postgraduate School, Monterey, CA  
Audience: Navy Software Engineering Research & Development Technology Review  
Subject of Briefing: STEP

Date of Briefing: 19 Sep 1984  
Briefer(s): R. J. Martin, W. M. McCracken  
Location: Camp Pendelton, CA  
Audience: 16th Multi-Service Development Test and Evaluation Commanders' Conference  
Subject of Briefing: STEP
8. **REPORT DISTRIBUTION**

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations ......................... 180
Volume 2: Software Test and Evaluation:
    State-of-the-Art Overview .............................. 189
Volume 3: Software Test and Evaluation:
    Current Defense Practices Overview .............. 175
Volume 4: Transcript of STEP Workshop, March 1982 (*) .. 45
Volume 5: Report of Expert Panel on
    Software Test and Evaluation ....................... 178
Volume 6: Tactical Computer System
    Applicability Study (**) ............................. 11

(*) Volume 4 distribution is limited to a selected audience.

(**) The distribution of Volume 6 is being handled by its author, Dr. James Leathrum of Clemson University.
Quarterly Progress Report No. 6
GIT Project G38-611

THE SOFTWARE TEST AND EVALUATION PROJECT
PHASES III & IV

Supported by
NAVAL AIR DEVELOPMENT CENTER (NADC)
Contract: F33657-82-G-2083


March 1986

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
THE SOFTWARE TEST AND EVALUATION PROJECT

PHASES III & IV

Quarterly Progress Report Number 6
06 October 1984 - 05 January 1985
12 March 1985

Supported by

Naval Air Development Center (NADC)
Contract: F33567-82-G-2083
GIT Project: G36-611

Army Institute for Research in Management
Information and Computer Science (AIRMICS)
Contract: BOA DAAK70-79-D-0087, D. O. #0020
GIT Project: G36-623

School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332
Table of Contents

1. Introduction ................................................................. 1
2. Organization and Staffing ............................................... 3
3. Current Task Areas ......................................................... 4
4. Summary of Progress ...................................................... 8
5. Travel Related to STEP .................................................. 14
6. Visitors ............................................................... 15
7. Publications, Reports, and Briefings ......................... 16
8. Report Distribution ....................................................... 17
1. INTRODUCTION

This is the Sixth Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.
Project Support

The following contracts provide support for this project.

Title: "Software Test and Evaluation Project, Phases III and IV"
Funding Agency: Naval Air Development Center (NADC)
Contract Number: F33657-82-G-2083
GIT Project Number: G36-611
Principal Investigator: Richard A. DeMillo

Title: "Software Test and Evaluation Project Support"
Funding Agency: Army Institute for Research in Management Information and Computer Science (AIRMICS)
Contract Number: BOA DAAK70-79-D-0087, D. O. #0020
GIT Project Number: G36-623
Principal Investigator: Richard A. DeMillo
2. ORGANIZATION AND STAFFING

Faculty

DeMillo, Richard A. - Professor
Martin, Rhonda J. - Research Scientist II
McCracken, W. Michael - Senior Research Engineer
Passafiume, John F. - Senior Research Engineer

Staff

Hinds, Lindsay - Student Assistant
Richards, Esther E. - Administrative Secretary
Richliew, Ann K. - Clerk/Typist
Wandrick, Greg - Student Assistant

Students

Bilsel, M. Sinasi - Graduate Research Assistant
Harder, Rita M. - Graduate Research Assistant
Offutt, VI, A. Jefferson - Graduate Research Assistant
3. CURRENT TASK AREAS

The work to be performed has been organized into the task areas identified below.

DDT&E Task Areas

A. Software Test and Evaluation Guidebook

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition life cycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. Software Test and Evaluation Model

Task 8 of the STEP Statement of Work states that "The effects of alternative software life cycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software life cycles. Examples of software life cycles include the traditional waterfall life cycles and the currently advocated evolutionary life cycles.
C. Technology Demonstration

The Technology Demonstration task area is intended to provide a laboratory framework for carrying out most of the technology-based STEP activities. These activities are described in the STEP Statement of Work Tasks 3 - 9 and the Software Technology for Adaptable, Reliable Systems (STARS) Program Task Areas 5 - 7. STEP Tasks 3 - 9 are concerned with formulating guidance statements, assessing the effects of new policy guidance, gathering evidence which indicates the effectiveness of systematic program testing, developing a software T&E model, and developing a software T&E guidebook. STARS Program Task Areas 5 - 7 are concerned with providing a test technology baseline and identification procedures, providing demonstration and qualification procedures for test technologies, and providing functional requirements for Ada (*) Programming Support Environment (APSE) test environments.

D. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the software T&E guidebook.

E. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

(*) Ada is a registered trademark of the U.S. Government AJPO
STARS Program Task Areas

F. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USDR&E). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

G. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USDR&E. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

H. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

I. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals, and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

J. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

K. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.
4. **SUMMARY OF PROGRESS**

A. **Software Test and Evaluation Guidebook**

Circulation and solicitation of comments on the draft guidebook was a principal activity during this period. Major briefings of the guidebook included briefings to the STEP Military Advisory Panel (MAP), the National Security Industrial Association Software Committee, and the attendees of the ITEA Fall Symposium. In addition to a number of technical and stylistic comments received, the MAP recommended a number of changes to improve the readability of the document. The issue of how to structure remaining volumes of the guidebook was also discussed by the MAP. After examining MAP recommendations, STEP has decided that the guidebook will appear in two volumes; the first is a slightly revised version of the current draft volume. The second volume will provide information and guidelines to managers and testers which will aid them in structuring test programs to conform to the guidelines in volume one.

Work began on structuring the format of the second volume and revising the first volume for final review and approval by DDT&E.

B. **Software Test and Evaluation Model**

Work continued on refining the format of the software T&E model as discussed in the previous quarterly report.

The outline of the model was briefed to the STEP Technical Advisory Panel (TAP), and comments were solicited from TAP members concerning the model. These comments are currently under review.

C. **Technology Demonstration**

Progress in this area is awaiting the location of funding.
D. **Good Examples**

Efforts this quarter were concentrated in two areas: conducting the Good Examples Session of the ITEA Fall Symposium and replanning the Good Examples Task Area.

The ITEA Good Examples Session was well attended and the presentations were well received. The panel session was stimulated by some comments initiated by Dr. DeMillo and the audience participated in the ensuing discussions. The proceedings of the Good Example Session will be published as a STEP report and will be distributed to the normal STEP distribution as well as the attendees of the ITEA Symposium.

The replanning of the Good Examples Task Area consisted primarily of two activities. The first activity was concerned with reviewing the programs that had been selected to be used for input to this task. The review was prompted as a result of an initial visit to one of the selected military program offices and the difficulties encountered when trying to locate the program's corporate memory. This issue caused us to reexamine the criteria used in selecting the programs for this task and to include the requirement that the program be at a point in its life cycle such that its corporate memory is still in place. A modified set of programs was generated and a preliminary list of these was discussed with the STEP staff. The list is not complete and will be updated as a part of the task area effort. The second activity was concerned with trying to regain lost ground in this area and to attempt to generate an achievable schedule that meets the overall STEP milestones. Based on the current set of known programs and as evidenced by preliminary contacts, the overall milestones for STEP will be met as supported by this task.
E. Risk Model

Significant progress was made during this period on refining the form of the risk model to accommodate the special problems of software T&E. A notable development was the discovery of a methodology by which system-level test requirements can be treated in the model. The resulting structure appears to be a promising one for basing decision-making methodologies on. A number of theoretical problems arose in this process (e.g., how to associate non-statistical information from software tests with the statistical requirements of operational tests) that need to be clarified. In addition, work continues on methods for structuring the test selection process within the model in such a way that useful decisions criteria can be formulated. Preliminary results on both of these efforts are expected during the next reporting period.

F. Technology Management

During this quarter, the STEP team obtained inputs and suggestions from representatives of the STARS program office. The major recommendations centered on the choice of an agency/activity which would be the most appropriate to perform the management task for software T&E. The team plans to prepare a list of areas of concern which would be appropriately monitored by each of the two competing entities - namely the Embedded Computer System Acquisition Policy Panel (ESCAPP) and the Defense Computer Resources Board (DCRB).

G. Technology Insertion

A draft CMS-2 tools baseline report has been prepared and is under review by members of the STEP team. Efforts are underway to expand this baseline to include additional CMS-2 tools, as well as to initiate actions to obtain information on tools residing within the other Services' activities.
H. APSE Test Environment

Assessment of the comments on the draft requirements document continued during this period. These will be the basis of the final revision of the document, expected early in the next reporting period. Issues relating to the draft requirements were presented to the STEP Technical Advisory Panel. Special emphasis was placed on the usage scenarios and potential sources of risk in the requirements derived thus far. These discussions will also be useful in preparing the final requirements document.

In addition, the question of feasibility for the likely designs that meet the recommended requirements was addressed. The approach adopted was to begin preliminary prototyping efforts to assess feasibility. Since an existing test tool at Georgia Tech has many of the utilities that will be required in the test environment, it was decided to use this tool as a basis for experimentation. Simultaneously, work has begun on surveying existing environment technologies for architectural approaches that will support the test environment requirements.

I. Coordination with Related Efforts

The following reflects ongoing efforts with which STEP is currently coordinating.

**Joint Logistics Commanders Initiatives:** The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development of Tri-Service standards for Software Development (DoD-STD-SDS) and Software Quality (DoD-STD-SQS). STEP personnel have been involved in the review of these documents and receive updates as they become available.

**JSSEE Team:** This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.
Evaluation and Validation (E&V) Team: This Tri-Service Team is tasked with developing the technology necessary to evaluate the suitability and effectiveness of Ada Programming Support Environments (APSEs) and validate conformance to appropriate military standards. Maximum cooperation between the STEP and E&V tasks promises the following benefits: 1) The efforts of the E&V Task will allow the insertion of demonstrated risk reduction technology (in the form of support software which has been thoroughly evaluated and validated) into the acquisition cycle. 2) The testing tool qualification procedures currently being developed by STEP will be elaborated by the E&V Task and inserted into an environment where the standard operating procedures include the E&V of support software. 3) The functional requirements for APSE test environments currently being developed by STEP will be supported by a technology which ensures their implementation. Because of these opportunities, a STEP representative is now participating in the E&V task.

NSIA C2 Software Development and Acquisition Study: STEP Team members were recently involved in a NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on software development efforts. Results of this study will feed into the Software T&E Model task.

DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

J. Direct Support to DDT&E

The support to OTEA on the Patriot Follow-on Evaluation III discussed in the previous progress report was continued during this period. A report summarizing STEP findings was delivered to OTEA. This report is not available for public release.
K. Laboratory Support

The equipment to replace our stolen Fortune terminals finally arrived during October. This included one terminal and two keyboards.

In addition, two undergraduate students were commissioned to write a report generator for their senior design project. Senior design projects are classes that undergraduate students at Georgia Tech take in order to get some experience on actual projects. The report generator will be used to aid in the development of quarterly progress reports for government projects.
5. TRAVEL RELATED TO STEP

Date of Trip: 14-16 Oct 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: Hilton Head, SC
Contact: M. Mesecher
Purpose: Brief STEP at NSIA Software Committee Meeting

Date of Trip: 05-08 Nov 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin, W. M. McCracken, J. F. Passafiume
Itinerary: Washington, D.C.
Contact: C. Watt
Purpose: Attend and Participate in ITEA Symposium

Date of Trip: 08 Nov 1984
Individual(s) Traveling: R. M. Harder
Itinerary: Washington, D.C.
Contact: C. Watt
Purpose: Attend ITEA Symposium

Date of Trip: 13 Nov 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: Washington, D.C.
Contact: R. Mathis
Purpose: Brief STEP Progress at STARS Meeting

Date of Trip: 29 Nov 1984
Individual(s) Traveling: R. A. DeMillo, R. J. Martin, W. M. McCracken, J. F. Passafiume
Itinerary: The Pentagon, Washington, D.C.
Contact: D. Greenlee
Purpose: Attend STEP Military Advisory Panel Meeting

Date of Trip: 04 Dec 1984
Individual(s) Traveling: J. F. Passafiume
Itinerary: Washington, D.C.
Contact: H. Stuebing
Purpose: Attend JSSEE Meeting

Date of Trip: 04-07 Dec 1984
Individual(s) Traveling: R. J. Martin
Itinerary: Wright Patterson AFB, OH
Contact: V. Castor
Purpose: Attend E&V Meeting

Date of Trip: 12 Dec 1984
Individual(s) Traveling: W. M. McCracken
Itinerary: Hanscom AFB, MA
Contact: H. Lyness
Purpose: Attend ESD/NSIA Meeting
6. VISITORS

Date of Visit: 24-25 Oct 1984
Visitor(s): STEP Technical Advisory Panel
John Bowen (Hughes Aircraft Company)
Carolyn Gannon (S/W Engineering Consultant)
David Geleperin (S/W Quality Engineering)
Edward Miller, Jr. (S/W Research Assoc.)
Thomas Quindry (DoD PESO - XC)
Raymond Rubey (SofTech)
Ralph San Antonio (Dynamics Research Corp.)
Marilyn Stewart (Logicon)
Peter Wegner (Brown University)
Contact: R. A. DeMillo & STEP Personnel
Purpose: Attend STEP Technical Advisory Panel Meeting
7. PUBLICATIONS, REPORTS, AND BRIEFINGS

Publications
None during this reporting period.

Reports

Briefings
Date of Briefing: 16 Oct 1984
Briefer(s): R. A. DeMillo
Location: Hilton Head, SC
Audience: NSIA Software Committee
Subject of Briefing: STEP Progress and Plans

Date of Briefing: 08 Nov 1984
Briefer(s): R. A. DeMillo
Location: Washington, D.C.
Audience: ITEA Symposium Attendees
Subject of Briefing: Software T&E Guidebook

Date of Briefing: 13 Nov 1984
Briefer(s): R. A. DeMillo
Location: Washington, D.C.
Audience: STARS Meeting Attendees
Subject of Briefing: STEP Progress and Plans (STARS Tasks)

Date of Briefing: 29 Nov 1984
Briefer(s): R. A. DeMillo
Location: The Pentagon, Washington, D.C.
Audience: STEP Military Advisory Panel
Subject of Briefing: Software T&E Guidebook
8. REPORT DISTRIBUTION

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations ...................... 184
Volume 2: Software Test and Evaluation:
    State-of-the-Art Overview ......................... 195
Volume 3: Software Test and Evaluation:
    Current Defense Practices Overview ............ 179
Volume 4: Transcript of STEP Workshop, March 1982 (*) .. 45
Volume 5: Report of Expert Panel on
    Software Test and Evaluation .................. 182
Volume 6: Tactical Computer System
    Applicability Study (**) ...................... 11

(*) Volume 4 distribution is limited to a selected audience.

(**) The distribution of Volume 6 is being handled by its author, Dr. James Leathrum of Clemson University.
Quarterly Progress Report No. 7

THE SOFTWARE TEST AND EVALUATION PROJECT
PHASES III & IV

Supported by

Office of the Director Defense Test and Evaluation
Office of the Secretary of Defense
Washington, D.C.

Naval Air Development Center (NADC)
Contract: F33657-82-G-2083
G17 Project: G36-611

Army Institute for Research in Management
Information and Computer Science (AIRMICS)
Contract: BOA DAAK70-79-D-0067, D.O. No. 0020
G17 Project: G36-623

Report Period: 06 January 1985 - 05 April 1985

31 May 1985

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
THE SOFTWARE TEST AND EVALUATION PROJECT
PHASES III & IV

Quarterly Progress Report Number 7
06 January 1985 - 05 April 1985
31 May 1985

Supported by

Office of the Director Defense Test and Evaluation
Office of the Secretary of Defense
Washington, D. C.

Naval Air Development Center (NADC)
Contract: F33657-82-G-2083
GIT Project: G36-611

Army Institute for Research in Management
Information and Computer Science (AIRMICS)
Contract: BOA DAAK70-79-D-0087, D. O. #0020
GIT Project: G36-623

School of Information and Computer Science
Georgia Institute of Technology
Atlanta, GA 30332
The Software Test and Evaluation Project: Phases III & IV
Quarterly Progress Report Number 7

Table of Contents

1. Introduction ......................................................... 1
2. Organization and Staffing ......................................... 3
3. Current Task Areas ................................................. 4
4. Summary of Progress ............................................... 8
5. Travel Related to STEP ............................................. 14
6. Visitors ............................................................ 16
7. Publications, Reports, and Briefings ............................ 17
8. Report Distribution ................................................. 19
1. INTRODUCTION

This is the Seventh Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.
Project Support

The following contracts provide support for this project.

Title: "Software Test and Evaluation Project, Phases III and IV"
Funding Agency: Naval Air Development Center (NADC)
Contract Number: F33657-82-G-2083
GIT Project Number: G36-611
Principal Investigator: Richard A. DeMillo

Title: "Software Test and Evaluation Project Support"
Funding Agency: Army Institute for Research in Management Information and Computer Science (AIRMICS)
Contract Number: BOA DAAK70-79-D-0087, D. O. #0020
GIT Project Number: G36-623
Principal Investigator: Richard A. DeMillo
2. ORGANIZATION AND STAFFING

Faculty

DeMillo, Richard A. - Professor
Martin, Rhonda J. - Research Scientist II
McCracken, W. Michael - Senior Research Engineer
Passafiume, John F. - Senior Research Engineer

Staff

Richards, Esther E. - Administrative Secretary
Richliew, Ann K. - Clerk/Typist
Seay, Brenda G. - Student Assistant
Wandrick, Greg - Student Assistant

Students

Bilsel, M. Sinasi - Graduate Research Assistant
Harder, Rita M. - Graduate Research Assistant
Offutt, VI, A. Jefferson - Graduate Research Assistant
3. **CURRENT TASK AREAS**

The work to be performed has been organized into the task areas identified below.

**DDT&E Task Areas**

A. Software Test and Evaluation Guidebook

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition life cycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. Software Test and Evaluation Model

Task 8 of the STEP Statement of Work states that "The effects of alternative software life cycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software life cycles. Examples of software life cycles include the traditional waterfall life cycles and the currently advocated evolutionary life cycles.

*Note: The Technology Demonstration Task Area has been deleted due to a lack of funding.*

C. Good Examples

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the software T&E guidebook.
D. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

(*) Ada is a registered trademark of the U.S. Government AJPO
STARS Program Task Areas

E. Technology Management

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USDRe). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

F. Technology Insertion

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USDRe. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

G. APSE Test Environment

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

H. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals, and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

I. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

J. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.
4. SUMMARY OF PROGRESS

A. Software Test and Evaluation Guidebook

The final draft of the Guidebook was prepared and submitted to DDT&E for approval. After consultation with DDT&E and the reviewers mentioned in the previous quarterly report it was determined that the official name of the Guidebook will be "Software Test and Evaluation Manual"; the title of the first volume will be: "Guidelines for the Treatment of Software in Test and Evaluation Master Plans."

During this quarter, DDT&E approved the draft version subject only to a number of minor editorial revisions. The finalized Manual (Volume I) will be issued early in the next quarter.

Work also continued on the structure and contents of Volume II.

B. Software Test and Evaluation Model

Work continued on the software T&E model as discussed in the previous quarterly report.

The primary emphasis during this period was in gathering software development models from a number of current software-intensive programs and relating the status of the software development to major acquisition milestones. It is anticipated that the progress of these and similar programs will continue to be a major source of data against which to validate preliminary versions of the models.

C. Good Examples

Efforts this quarter consisted of data gathering activities on two selected programs. The programs are PAVE PAWS (Air Force ESD) and TACTASS (Navy NAVSEA).

The PAVE PAWS program visits included a meeting with the Air Force Program Manager and his Engineering Chief from MITRE. The program questionnaire, developed for this effort, was reviewed with them and applicable documentation was collected. We also visited the Prime Contractor, Raytheon, reviewed the questionnaire, gathered appropriate documentation, and interviewed the Raytheon Deputy for Software and his chief engineer. We plan on visiting the software subcontractor, IBM, to complete the data gathering portion of this program.
The data gathering activity for TACTASS (AN/SQR-19), consisted of visiting the Navy's program manager and the Prime Contractor (GE). We reviewed the program questionnaire with them and collected appropriate data. There are no more visits planned on TACTASS.

Plans for next period consist of completing the data gathering activities for PAVE PAWS, initiating data gathering on AEGIS (Navy NSWC) and selecting an Army program for data gathering. At the completion of the data gathering phase we will write a summary report of the relevant data supporting the Good Examples of software testing performed on these programs.

D. Risk Model

Progress on the risk model during this reporting period was centered on techniques for defining the total risk associated with the testing effort and on defining a methodology for relating low-level software testing results to overall system objectives.

The primary problem being addressed in both of these activities is the relationship between the evidence provided by a systematic test of the software and the observed behavior of the system. To make the initial analysis simpler, attention was focused on testing for reliability. The system reliability requirements form the basis of the software reliability requirements, and successive stages of software testing are assessed relative to their utility in predicting whether or not overall reliability requirements can be met. The main tools used in this sort of analysis are the methods of Bayesian statistical inference. The results, however do not depend on any particular interpretation of the probability measures.

Additionally, some initial work was carried out to define an external decision model for the risk model. It is expected that such a decision model will eventually form the basis of a computerized tool to calculate relative risks associated with software T&E.
E. Technology Management

During this quarter, the STEP team completed an internal draft of the Management Plan. This draft will be coordinated internally with team members and then informally circulated to the members of the Military Advisory Panel for their comments. These comments will be used to develop the final draft of the Management Plan which will be coordinated with the Office of the Director, CSS to ensure consistency with the directions being planned for the Defense Computer Resources Board (DCRB). Discussions with representatives of the Director, CSS and the STARS Joint Program Office have caused us to rule out the ECSAPP as a feasible alternative organizational structure to perform the management task for software T&E.

F. Technology Insertion

The draft CMS-2 tools baseline report developed during the last reporting period has been reviewed internally and is now being updated and expanded to include more detailed information about the tools. In addition, the team has begun to collect information concerning tools residing within activities of the Army and Air Force. This effort will continue during the next quarter.

G. APSE Test Environment

The comments and additional material described in the previous progress report were integrated and used as a basis for preparing the final draft of the requirements document. It is anticipated that the document will be available early in the next quarter.

Additional technical activities associated with the test environment have centered on the continued investigation of potential design alternatives for an integrated environment.
H. **Coordination with Related Efforts**

The following reflects ongoing efforts with which STEP is currently coordinating.

**Joint Logistics Commanders Initiatives:** The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development/revision of Tri-Service standards for Software Development (DoD-STD-2167) and Software Quality (DoD-STD-SQS). STEP personnel have been involved in the review of these documents and receive updates as they become available.

**JSSEE Team:** This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.

**Measurements Area Coordinating Team (ACT):** This is one of the six area coordinating teams under the STARS program. The team's goal is to develop, apply and evaluate techniques for specifying, predicting, and assessing software quality and life-cycle resource expenditures. STEP personnel coordinate with this team since it represents the STARS area that has goals most similar to those of STEP.

**Evaluation and Validation (E&V) Team:** This Tri-Service Team is tasked with developing the technology necessary to evaluate the suitability and effectiveness of Ada Programming Support Environments (APSEs) and validate conformance to appropriate military standards. Maximum cooperation between the STEP and E&V tasks promises the following benefits: 1) The efforts of the E&V Task will allow the insertion of demonstrated risk reduction technology (in the form of support software which has been thoroughly evaluated and validated) into the acquisition cycle. 2) The testing tool qualification procedures currently being developed by STEP will be elaborated by the E&V Task and inserted into an environment where the standard operating procedures include the E&V of support software. 3) The functional requirements for APSE test environments currently being developed by STEP will be supported by a technology which ensures their implementation. Because of these opportunities, a STEP representative is now participating in the E&V task.
NSIA C2 Software Development and Acquisition Study:
STEP Team members were recently involved in a NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on software development efforts. Results of this study will feed into the Software T&E Model task.

DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

I. Direct Support to DDT&E

Activities during this period were concerned with the review of the software T&E plans for selected programs and initiating STEP support of Dr. Wade's initiative for defining Reliability, Availability, and Maintainability (RAM) criteria for weapon system software.

In reviewing and commenting on software T&E plans, particular emphasis was placed on the application of the criteria contained in the STEP Software Test and Evaluation Manual (see Task A) to these programs. In each case, the relevant organizations were given written and oral commentaries on results of this assessment.

At the direction of USDRE, an initiative to define RAM criteria for software has been organized by the Director Computer Software and Systems (CSS), DUSDRA. In a memorandum from USDRE to CSS dated 21 December, STEP was offered to CSS as a resource in establishing "uniform criteria" for software RAM. Dr. DeMillo has met twice with representatives of CSS and participated in the formulation of a plan for carrying out this task. The plan was delivered to USDRE in a memorandum to CSS dated 8 March. During the next 18 months, STEP personnel -- representing DDT&E -- will be assisting a RAM study panel convened by CSS in formulating RAM criteria.
J. Laboratory Support

In January, Georgia Tech purchased a second Fortune computer to support STEP as well as other activities engaged in by STEP personnel. In addition, an updated Fortune word processor that includes new features and has been optimized to provide improved performance was acquired.

The primary laboratory support problem this quarter concerned disk space. At one point both 20 megabyte disks were completely filled. Strict conservation measures have been implemented in order to keep the Fortune systems up and running.

AT&T recently donated a sizable number of their new 3B2 micro computers to Georgia Tech. One of these has been made available for use by STEP personnel.

Finally, the two undergraduate students who were commissioned to write a report generator as their senior design project continued their work. Senior design projects are classes that undergraduate students at Georgia Tech take in order to get some experience on actual projects. The report generator will be used to aid in the development of quarterly progress reports for government projects.
5. **TRAVEL RELATED TO STEP**

Date of Trip: 15 Jan 1985  
Individual(s) Traveling: R. A. DeMillo, R. J. Martin  
Itinerary: Norfolk, VA  
Contact: E. Sierra  
Purpose: Brief Software T&E Manual to ITEA/Tidewater Chapter

Date of Trip: 22-23 Jan 1985  
Individual(s) Traveling: R. J. Martin  
Itinerary: Washington, D.C.  
Contact: H. Stuebing  
Purpose: Attend JSSEE Meeting

Date of Trip: 31 Jan 1985  
Individual(s) Traveling: R. A. DeMillo, R. J. Martin  
Itinerary: Washington, D.C.  
Contact: D. Greenlee  
Purpose: Brief Software T&E Manual to WIS and AMPE  
Program Personnel, present oral progress report to D. Greenlee, and conduct contract discussions with C. Watt

Date of Trip: 11-12 Feb 1985  
Individual(s) Traveling: W. M. McCracken  
Itinerary: Boston, MA  
Contact: D. Moore (AF/ESD) and P. Hepner (Raytheon)  
Purpose: Gather data for Good Examples Task (FAVEPAWS)

Date of Trip: 11-15 Feb 1985  
Individual(s) Traveling: J. F. Passafiume  
Itinerary: Norfolk, VA  
Contact: J. Cavano  
Purpose: Attend STARS Workshop and initiate coordination with STARS Measurement Area Coordinating Team

Date of Trip: 25 Feb 1985  
Individual(s) Traveling: R. A. DeMillo, W. M. McCracken, J. F. Passafiume  
Itinerary: Washington, D.C.  
Contact: P. Dickenson and BG Salisbury  
Purpose: Attend meeting with OTEA and present STEP briefing to U.S. Army System Software Support Command personnel

Date of Trip: 26 Feb 1985  
Individual(s) Traveling: R. A. DeMillo  
Itinerary: Washington, D.C.  
Contact: D. Greenlee  
Purpose: Present oral progress report
Date of Trip: 26-27 Feb 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Washington, D.C.
Contact: H. Stuebing
Purpose: Attend JSSEE Meeting

Date of Trip: 05-06 Mar 1985
Individual(s) Traveling: J. F. Passafiume
Itinerary: Washington, D.C.
Contact: J. Cavano
Purpose: Attend STARS Measurement Area Coordinating Team Briefing/Meeting

Date of Trip: 05-06 Mar 1985
Individual(s) Traveling: R. J. Martin
Itinerary: San Diego, CA
Contact: V. Castor
Purpose: Attend E&V Meeting

Date of Trip: 07 Mar 1985
Individual(s) Traveling: R. J. Martin
Itinerary: San Diego, CA
Contact: B. Wasgatt
Purpose: Discussions concerning STEP progress and products and applicability to NOSC Integrated Combat Systems Test Facility mission

Date of Trip: 19-20 Mar 1985
Individual(s) Traveling: J. F. Passafiume
Itinerary: Rome, NY
Contact: J. Cavano
Purpose: Attend STARS Measurement Area Coordinating Team Meeting

Date of Trip: 28-29 Mar 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Washington, D.C.
Contact: H. Stuebing
Purpose: Attend JSSEE Meeting

Date of Trip: 03 Apr 1985
Individual(s) Traveling: W. M. McCracken
Itinerary: Washington, D.C.
Contact: D. Horne (NAVSEA)
Purpose: Gather data for Good Examples Task (TACTASS)

Date of Trip: 04 Apr 1985
Individual(s) Traveling: W. M. McCracken
Itinerary: Syracuse, NY
Contact: A. Kerr (GE)
Purpose: Gather data for Good Examples Task (TACTASS)
6. **VISITORS**

   Date of Visit: 04 Mar 1985  
   Visitor(s): James Baca (AFOTEC)  
   Contact: R. A. DeMillo & STEP Personnel  
   Purpose: Discussion plans for STEP participation in AFOTEC Software Risk Assessment Workshop

   Date of Visit: 02 Apr 1985  
   Visitor(s): Capt. Kinneson (MIA)  
   Contact: R. A. DeMillo & STEP Personnel  
   Purpose: Discussions related to future contracting through MIA.
7. **PUBLICATIONS, REPORTS, AND BRIEFINGS**

**Publications**

None during this reporting period.

**Reports**


**Briefings**

Date of Briefing: 15 Jan 1985  
Briefer(s): R. A. DeMillo  
Location: Norfolk, VA  
Audience: ITEA/Tidewater Chapter  
Subject of Briefing: Software T&E Manual

Date of Briefing: 18 Jan 1985  
Briefer(s): R. A. DeMillo  
Location: Atlanta, GA  
Audience: AIRMICS Personnel  
Subject of Briefing: STEP Progress and Plans

Date of Briefing: 31 Jan 1985  
Briefer(s): R. A. DeMillo  
Location: Washington, D.C.  
Audience: WIS and AMPE Program Personnel  
Subject of Briefing: Software T&E Manual
Date of Briefing: 25 Feb 1985
Briefer(s): R. A. DeMillo
Location: Washington, D.C.
Audience: U.S. Army System Software Support Command
Subject of Briefing: STEP Progress and Plans

Date of Briefing: 19 Mar 1985
Briefer(s): J. F. Passafiume
Location: Rome, NY
Audience: STARS Measurement Area Coordinating Team
Subject of Briefing: STEP Progress and Plans
8. **REPORT DISTRIBUTION**

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations ...................... 190

Volume 2: Software Test and Evaluation:
  State-of-the-Art Overview ................................. 203

Volume 3: Software Test and Evaluation:
  Current Defense Practices Overview ............. 184

Volume 4: Transcript of STEP Workshop, March 1982 (*) .. 45

Volume 5: Report of Expert Panel on
  Software Test and Evaluation ......................... 186

Volume 6: Tactical Computer System
  Applicability Study (**) ............................... 11

(*) Volume 4 distribution is limited to a selected audience.

(**) The distribution of Volume 6 is being handled by its author, Dr. James Leathrum of Clemson University.
Table of Contents

1. Introduction .................................................. 1

2. Organization and Staffing .................................. 3

3. Current Task Areas .......................................... 4

4. Summary of Progress ....................................... 8

5. Travel Related to STEP ..................................... 13

6. Visitors ...................................................... 15

7. Publications, Reports, and Briefings ..................... 16

8. Report Distribution .......................................... 17
1. INTRODUCTION

This is the Eighth Quarterly Progress Report prepared covering the Software Test and Evaluation Project (STEP), Phases III & IV.

Project Description

STEP was initiated in 1981 by the Director Defense Test and Evaluation (DDT&E). The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation (T&E) of computer software for mission-critical applications. A number of subsidiary goals have also been established for STEP. Principal subgoals include the stimulation of tool development, the support of policy development, and the identification of research issues and directions in the area of software testing.

Phases I and II of STEP (completed in the Spring of 1983) consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which specific recommendations for improving the state-of-practice in Defense software T&E were developed.

Principal recommendations from the previous STEP phases are intended to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning document (the Test and Evaluation Master Plan or TEMP) and proceeds through the plans and procedures implemented by Project Offices, development organizations, Independent Verification and Validation (IV&V) organizations and independent test organizations. Phases III & IV of STEP, which are currently underway, are designed to define the technology and provide the implementation support for these recommendations.
Project Support

The following contracts provide support for this project.

Title: "Software Test and Evaluation Project, Phases III and IV"
Funding Agency: Naval Air Development Center (NADC)
Contract Number: F33657-82-G-2083
GIT Project Number: G36-611
Principal Investigator: Richard A. DeMillo

Title: "Software Test and Evaluation Project Support"
Funding Agency: Army Institute for Research in Management Information and Computer Science (AIRMICS)
Contract Number: BOA DAAK70-79-D-0087, D. O. #0020
GIT Project Number: G36-623
Principal Investigator: Richard A. DeMillo
2. ORGANIZATION AND STAFFING

Faculty

Balsam, Jeanne H. - Research Scientist II
DeMillo, Richard A. - Professor
Grover, Jeffrey L. - Research Scientist II
Martin, Rhonda J. - Research Scientist II
McCracken, W. Michael - Senior Research Engineer
Passafiume, John F. - Senior Research Engineer

Staff

Barry, Glenn T. - Student Assistant
Hinds, J. Lindsay - Student Assistant
Richards, Esther E. - Administrative Secretary
Richliew, Ann K. - Clerk/Typist

Students

Bilsel, M. Sinasi - Graduate Research Assistant
Flasphler, John C. - Graduate Research Assistant
Harder, Rita M. - Graduate Research Assistant
Offutt, VI, A. Jefferson - Graduate Research Assistant
3. **CURRENT TASK AREAS**

The work to be performed has been organized into the task areas identified below.

**DDT&E Task Areas**

A. **Software Test and Evaluation Guidebook**

Task 9 of the STEP Statement of Work requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition life cycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from the other task areas of STEP Phases III and IV, as appropriate.

B. **Software Test and Evaluation Model**

Task 8 of the STEP Statement of Work states that "The effects of alternative software life cycle models will be examined". This task area is concerned with describing the expected maturity of the software at each decision point for a variety of software life cycles. Examples of software life cycles include the traditional waterfall life cycles and the currently advocated evolutionary life cycles.

Note: The Technology Demonstration Task Area has been deleted due to a lack of funding.

C. **Good Examples**

Task 7(1) of the STEP Statement of Work requires "Gathering objective evidence which indicates the effectiveness of systematic program testing during the development and operational test phases. This evidence should not serve to recommend any particular testing methodology, but should demonstrate the effectiveness of existing technology." The Good Examples Task will examine several systems which are regarded as "successes". Factors identified as contributing to the success of a program will be reported. When such factors may be applied to other programs, they will be incorporated into the software T&E guidebook.
D. Risk Models

Risk modelling is viewed as a formalized decision-making structure in which to resolve software T&E issues. A well-chosen and carefully motivated risk model can serve as a basis for a more quantitative approach to risk reduction for software as well as the structure for a decision support system that can adequately address the unique features of software T&E.

(*) Ada is a registered trademark of the U.S. Government AJPO
**STARS Program Task Areas**

**E. Technology Management**

The goal of this task is to define an organizational structure for managing improvements in the testing process and implementing policy for software T&E. One candidate organization for the ongoing management of T&E activities is the proposed Acquisition Panel which is to be established by the Under-Secretary of Defense for Research and Engineering (USDR&E). The output of this task will be a management plan identifying resource requirements, organizational responsibilities and authority, alternative strategies for implementing DoD and Service policy and guidance, and methods of communicating recommendations concerning software T&E to the acquisition community.

**F. Technology Insertion**

This task is concerned with focusing the responsibility for software T&E-related technology insertion in a single organization. The current candidate organization is the proposed Software Engineering Institute (SEI) which is to be established by USDR&E. One portion of this task will produce a software T&E management and operating plan for the SEI. Another portion will define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion in a warehouse to satisfy test plan requirements. In addition to the identification procedures, a baseline of software testing tools and technology will be provided. Finally, this task will define a procedure for demonstrating and qualifying testing tools for DoD use. Key characteristics to be identified in the qualification procedure will be the support of a quantitative risk model for software T&E, the appropriateness of selected technologies for embedded and mission-critical applications, and documented cost/benefit data.

**G. APSE Test Environment**

This task will offer support to the STARS Program and the Joint Service Software Engineering Environment (JSSEE) Team in the formulation of requirements for candidate APSE test environments (or the identification of existing environments if they should become available soon).
Other Task Areas

H. Coordination with Related Efforts

This task is concerned with coordinating STEP activities with other related efforts. The allows STEP to: 1) gain a large base of support within organizations which will be affected by STEP results, thereby increasing the probability of successful implementation, 2) benefit from the widest possible input, 3) avoid problems encountered when "working in a vacuum", 4) capitalize on other efforts, 5) ensure that other efforts are not in conflict with STEP goals, and 6) obtain increased visibility in the mission-critical computer resource (MCCR) software community.

I. Direct Support to DDT&E

On occasion, STEP personnel are called upon to provide support to DDT&E in ways not specifically detailed in the STEP Phases III & IV Statement of Work. This task has been identified to report on those activities.

J. Laboratory Support

This task encompasses all activities related to providing the necessary computer resources to the STEP Team such that work directly supporting this contract can be accomplished in the most effective manner possible.
4. SUMMARY OF PROGRESS

A. Software Test and Evaluation Guidebook

The Software Test and Evaluation Manual, Volume I, Guidelines for the Treatment of Software in Test and Evaluation Master Plans was completed in May 1985. This manual is aimed at assisting DDT&E personnel when evaluating TEMPs; it will also help with the preparation of TEMPs.

In addition, a new individual joined the STEP team in May and has been assigned to The Software Test and Evaluation Manual, Volume II. This manual is going to be written to assist all organizations that are involved with test and evaluation throughout the acquisition process. The new STEP team member spent time reviewing DoD and Service documents to gather information on the role of T&E in the acquisition process.

B. Software Test and Evaluation Model

This task area was inactive for the duration of the reporting period.

C. Good Examples

Efforts this quarter concentrated on the analysis of the data gathered from the PAVE PAWS (Air Force ESD) and TACTASS (Navy NAVSEA) programs. Each program's data gathering guide, comments from program personnel and associated documentation were analyzed for pertinent examples of successful software testing.

Plans for next period consist of completing the analysis and writing reports that describe the successes on the programs. This information will be condensed for use in the Guidebook task to represent specific examples of successful software testing. In addition, it is planned that the AEGIS (Navy NSWC) data gathering will be conducted.

D. Risk Model

This task area was inactive for the duration of the reporting period.
E. Technology Management

During this quarter, an initial draft of the Management Plan was coordinated with STEP team members. A subsequent draft was informally coordinated with a representative of DDT&E and members of the Military Advisory Panel for their comments. These comments are now being incorporated into a final draft of the Management Plan. It is planned that this draft will be circulated to members of the Military Advisory Panel in anticipation of a meeting to be held in late Summer.

F. Technology Insertion

Writing of the CMS-2 Tools Baseline Report has continued. Testing tool data sheets were finalized for the approved CMS-2 tools. Also during this period, information was received from NADC on their baseline of CMS-2 tools and an updated draft of the STEP CMS-2 report was circulated to the STEP team for comment. The search for Army and Air Force tools continued and those tools that are discovered will be included in the overall baseline report. Finally, work was started on the search methodology writeup.

G. APSE Test Environment

The final draft of the final report for this task was completed during this period and was submitted for editing prior to publication and distribution.

In addition, the report was briefed to the AJPO's Evaluation and Validation Team. Investigation of potential design alternatives for an integrated environment also continued.

H. Coordination with Related Efforts

The following reflects ongoing efforts with which STEP is currently coordinating.

Joint Logistics Commanders Initiatives: The Joint Logistics Commanders are currently sponsoring numerous ongoing efforts, notably the development/revision of Tri-Service standards for Software Development (DoD-STD-2167) and Software Quality (DoD-STD-SQE). STEP personnel have been involved in the review of these documents and receive updates as they become available.
JSSEE Team: This Tri-Service Team is tasked with providing the definition and preliminary design of a Joint Service Software Engineering Environment. Due to the STEP task to develop the functional requirements for an APSE test environment, a STEP representative attends all JSSEE Team meetings.

Measurements Area Coordinating Team (ACT): This is one of the six area coordinating teams under the STARS program. The team's goal is to develop, apply and evaluate techniques for specifying, predicting, and assessing software quality and life-cycle resource expenditures. STEP personnel coordinate with this team since it represents the STARS area that has goals most similar to those of STEP.

Evaluation and Validation (E&V) Team: This Tri-Service Team is tasked with developing the technology necessary to evaluate the suitability and effectiveness of Ada Programming Support Environments (APSEs) and validate conformance to appropriate military standards. Maximum cooperation between the STEP and E&V tasks promises the following benefits:
1) The efforts of the E&V Task will allow the insertion of demonstrated risk reduction technology (in the form of support software which has been thoroughly evaluated and validated) into the acquisition cycle. 2) The testing tool qualification procedures currently being developed by STEP will be elaborated by the E&V Task and inserted into an environment where the standard operating procedures include the E&V of support software. 3) The functional requirements for APSE test environments currently being developed by STEP will be supported by a technology which ensures their implementation. Because of these opportunities, a STEP representative is now participating in the E&V task.

NSIA C2 Software Development and Acquisition Study: STEP Team members were recently involved in a NSIA study for the Air Force Electronic Systems Division on C2 Software Development and Acquisition. More specifically, the study was concerned with the implications of evolutionary acquisition on software development efforts. Results of this study will feed into the Software T&E Model task.
DDT&E Research Initiatives: In addition to STEP, DDT&E is sponsoring research in other areas with the long range goal of developing a decision support system for T&E. STEP personnel participate in DDT&E Program Reviews and Roundtables to maximize the sharing of resources and results between research efforts, and to provide critical evaluations and suggestions for alternative avenues of investigation.

I. Direct Support to DDT&E

Activities during this period were concerned with the review of the software T&E plans for selected programs and participating in the CSS task to define Reliability, Availability, and Maintainability (RAM) criteria for weapon system software.

In reviewing and commenting on software T&E plans, particular emphasis was placed on the application of the criteria contained in the STEP Software Test and Evaluation Manual (see Task A) to these programs. In each case, the relevant organizations were given written and oral commentaries on results of this assessment.

At the direction of USDRE, an initiative to define RAM criteria for software has been organized by the Director Computer Software and Systems (CSS), DUSDRAT. STEP personnel are actively participating in this initiative through membership on the CSS task force which will define RAM criteria. Three meetings were held in the Washington area, and STEP participated in all three. STEP products of these meetings included a compendium of DoD-sanctioned definitions of RAM, a tracking mechanism to enable the association of appropriate definitions with the relevant documents and life cycle phases, and a draft of a RAM management model.

J. Laboratory Support

In May, an undergraduate ICS student who is knowledgeable with respect to Unix systems, particularly the AT&T 3B2s, was hired to fill the Lab Assistant's position. In addition, to support the increasing number of STEP personnel, several new Fortune terminals were ordered. The expected delivery date is late July.
Finally, the two undergraduate students who were commissioned to write a report generator as their senior design project finished their project. Senior design projects are classes that undergraduate students at Georgia Tech take in order to get some experience on actual projects. The report generator will be used to aid in the development of quarterly progress reports for government projects.
5. TRAVEL RELATED TO STEP

Date of Trip: 16 April 1985
Individual(s) Traveling: R. A. DeMillo, W. M. McCracken
Itinerary: Albuquerque, NM
Contact: Lt. Col. Escue
Purpose: IFFN/JTF Program Review

Date of Trip: 24 April 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Washington, D. C.
Contact: H. Stuebing
Purpose: JSSEE Meeting

Date of Trip: 29 April - 02 May 1985
Individual(s) Traveling: R. J. Martin, W. M. McCracken
Itinerary: San Diego, CA
Contact: Lt. Col. Escue
Purpose: STARS Briefing to Industry, NSIA STARS Conference, IFFN/JTF Meetings

Date of Trip: 01 May 1985
Individual(s) Traveling: J. F. Passafiume
Itinerary: Washington, D. C.
Contact: D. Greenlee
Purpose: AMPE TEMP Meeting

Date of Trip: 14-16 May 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Virginia Beach, VA
Contact: D. Green
Purpose: JSSEE OCD Workshop

Date of Trip: 17 May 1985
Individual(s) Traveling: R. J. Martin, W. M. McCracken
Itinerary: Washington, D. C.
Contact: D. Greenlee
Purpose: Contract Discussions

Date of Trip: 03 June 1985
Individual(s) Traveling: R. A. DeMillo, R. J. Martin
Itinerary: Washington, D. C.
Contact: J. Duke
Purpose: RAM Panel Meeting

Date of Trip: 04 June 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Washington, D. C.
Contact: H. Stuebing
Purpose: JSSEE Meeting
Date of Trip: 05 June 1985
Individual(s) Traveling: R. A. DeMillo
Itinerary: Dayton, OH
Contact: V. Castor
Purpose: Brief APSE Test Environment Requirements to E&V Team

Date of Trip: 05-07 June 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Dayton, OH
Contact: V. Castor
Purpose: E&V Meeting

Date of Trip: 10 June 1985
Individual(s) Traveling: R. A. DeMillo, J. L. Grover, R. J. Martin
Itinerary: Washington, D. C.
Contact: J. Duke
Purpose: RAM Panel Meeting

Date of Trip: 11-12 June 1985
Individual(s) Traveling: R. A. DeMillo, J. L. Grover, R. J. Martin, W. M. McCracken
Itinerary: Albuquerque, NM
Contact: J. Baca
Purpose: AFOTEC Software T&E Workshop

Date of Trip: 13 June 1985
Individual(s) Traveling: R. A. DeMillo, J. L. Grover, W. M. McCracken
Itinerary: Albuquerque, NM
Contact: Lt. Col. Crosby
Purpose: IFFN/JTF Program Review

Date of Trip: 14 June 1985
Individual(s) Traveling: R. J. Martin
Itinerary: Huntsville, AL
Contact: Capt. Kinneson
Purpose: Contract Discussions

Date of Trip: 17 June 1985
Individual(s) Traveling: R. A. DeMillo, J. L. Grover, R. J. Martin
Itinerary: Washington, D. C.
Contact: J. Duke
Purpose: RAM Panel Meeting
6. VISITORS

Date of Visit: 29 May 1985
Visitor(s): Capt. Kinneson (MIA)
Contact: R. J. Martin
Purpose: Contract discussions

Date of Visit: 30 May 1985
Visitor(s): Henry Alberts (DSMC)
Contact: R. A. DeMillo and STEP Personnel
Purpose: Discuss STEP participation in ITEA Software T&E Workshop

Date of Visit: 19 June 1985
Visitor(s): Charles Watt
Contact: R. A. DeMillo
Purpose: Contract discussions
7. PUBLICATIONS, REPORTS, AND BRIEFINGS

Publications

None during this reporting period.

Reports


Briefings

Date of Briefing: 02 May 1985
Briefer(s): R. J. Martin
Location: San Diego, CA
Audience: NSIA STARS Conference Attendees
Subject of Briefing: STEP Progress and Plans

Date of Briefing: 05 June 1985
Briefer(s): R. A. DeMillo
Location: Dayton, OH
Audience: APSE E&V Team
Subject of Briefing: APSE Test Environment Requirements
8. REPORT DISTRIBUTION

The number of STEP Phases I and II Final Reports distributed are as follows.

Volume 1: Report and Recommendations ..................... 200

Volume 2: Software Test and Evaluation:
   State-of-the-Art Overview ......................... 213

Volume 3: Software Test and Evaluation:
   Current Defense Practices Overview ............. 194

Volume 4: Transcript of STEP Workshop, March 1982 (*) .. 45

Volume 5: Report of Expert Panel on
   Software Test and Evaluation ..................... 195

Volume 6: Tactical Computer System
   Applicability Study (**) ......................... 11

(* ) Volume 4 distribution is limited to a selected audience.

( ** ) The distribution of Volume 6 is being handled by its
   author, Dr. James Leathrum of Clemson University.
TECHNICAL REPORT

STEP Technical Advisory Panel Meeting Report

By
R. A. DeBoer

Prepared for:
Naval Air Development Center
Warminster, PA 18974

Under
Contract No. F33657-82-C-2583

January 1984

GEORGIA INSTITUTE OF TECHNOLOGY
A Unit of the University System of Georgia
School of Information and Computer Science
Atlanta, Georgia 30332
<table>
<thead>
<tr>
<th>REPORT DOCUMENTATION PAGE</th>
<th>READ INSTRUCTIONS BEFORE COMPLETING FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REPORT NUMBER</td>
<td>2. GOVT ACCESSION NO.</td>
</tr>
<tr>
<td>3. RECIPIENT'S CATALOG NUMBER</td>
<td></td>
</tr>
<tr>
<td>4. TITLE (and Subtitle)</td>
<td>5. TYPE OF REPORT &amp; PERIOD COVERED</td>
</tr>
<tr>
<td>STEP Technical Advisory Panel</td>
<td></td>
</tr>
<tr>
<td>Meeting Report</td>
<td>6. PERFORMING ORG. REPORT NUMBER</td>
</tr>
<tr>
<td></td>
<td>7. AUTHOR(s)</td>
</tr>
<tr>
<td></td>
<td>R.A. DeMillo</td>
</tr>
<tr>
<td>8. CONTRACT OR GRANT NUMBER(s)</td>
<td>F33657-82-G-2083</td>
</tr>
<tr>
<td>9. PERFORMING ORGANIZATION NAME AND ADDRESS</td>
<td>School of Information and Computer Science</td>
</tr>
<tr>
<td></td>
<td>Georgia Institute of Technology</td>
</tr>
<tr>
<td></td>
<td>Atlanta, GA 30332</td>
</tr>
<tr>
<td>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</td>
<td></td>
</tr>
<tr>
<td>11. CONTROLLING OFFICE NAME AND ADDRESS</td>
<td>NADC</td>
</tr>
<tr>
<td></td>
<td>Warminster, PA 18974</td>
</tr>
<tr>
<td>12. REPORT DATE</td>
<td>January, 1984</td>
</tr>
<tr>
<td>13. NUMBER OF PAGES</td>
<td></td>
</tr>
<tr>
<td>14. MONITORING AGENCY NAME &amp; ADDRESS</td>
<td>(if different from Controlling Office)</td>
</tr>
<tr>
<td>15. SECURITY CLASS. (of this report)</td>
<td>Unclassified</td>
</tr>
<tr>
<td>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE</td>
<td></td>
</tr>
<tr>
<td>16. DISTRIBUTION STATEMENT (of this Report)</td>
<td>unlimited</td>
</tr>
<tr>
<td>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</td>
<td>unlimited</td>
</tr>
<tr>
<td>18. SUPPLEMENTARY NOTES</td>
<td></td>
</tr>
<tr>
<td>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</td>
<td></td>
</tr>
<tr>
<td>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</td>
<td></td>
</tr>
</tbody>
</table>
On 08-09 November 1983, STEP held its first meeting of the Technical Advisory Panel at the Georgia Institute of Technology, Atlanta, Ga. The objectives of this meeting were:

1) To report to the panel members on final results of STEP Phases I & II,
2) To describe the tasks being conducted under STEP Phases III & IV, and
3) To determine areas of interest and future involvement of panel members.

This report follows the outline of the agenda by detailing presentations given and listing a few of the relevant discussion points raised. Attachments include the agenda, the attendees list, copies of all slides used, and interest forms received from panel members at the conclusion of the meeting. A sample consulting agreement is also included.

Presentation of STEP Results & Recommendations
(Presented by R. Gagliano)

Following a brief overview of current defense practices and the state-of-the-art in software test and evaluation, each of the recommendations made as a result of the STEP Phases I & II were presented.

Discussion-generated

Questions by the panel analyzed many of the recommendations in detail. The motivation, specific application and intended consequences were all topics clarified during the discussion.

Discussion of Recommendations
(Presented by R. DeMillo)

The recommendations of STEP Phases I/II were concerned with policy changes for software T&E. The changes concentrated on system level test plans and the technology development
needed to support DDT&E's efforts to improve software testing. With modifications to DODD 5000.3, there will be a need to provide coordination and technological help to the services to smooth implementation.

To clarify the intent of the recommendations, a draft of the Test & Evaluation Master Plan (TEMP) checklist was presented. This checklist should be used to evaluate a TEMP in terms of its treatment of software.

**Discussion-generated**

Initially, some members of the panel did not understand how ODDT&E fits in the acquisition cycle and what purposes the TEMP served. After clarification, two opinions emerged:

* The TEMP is extraneous paperwork, and there should be more checks during the development, or
* The TEMP is a good vehicle for its purposes.

Most of the members wished to see a copy of an existing TEMP.

**Overview of STARS Initiative**
(Presented by Col. Vance Mall, OUSD/R&AT)

STEP is now receiving support from the Software Technology for Adaptable, Reliable System (STARS) program. The STARS program is concerned with broad-based improvements in the practice of software engineering. The overview was a historical perspective of the STARS Program, its current status, and its future plans.

**STEP Good Examples**
(Presented by R. Gaqliano)

The Good Examples Task is concerned with demonstrating the effectiveness of existing technology by identifying "successful software systems" and reporting factors which contributed to that success. The overview of this task included motivation, criteria for system selection, areas of investigation and some possible results of reporting good examples to the community. The possibility of hosting a follow-on meeting to the NSIA Software T&E Symposium -- focusing on good examples -- was also discussed.
The discussion centered on the definition of a "successful system" and the best definition seemed to be that of a system which has been completed with no major faults or problems. STEP asked panelists for recommendations and guidance on the selection of examples. They provided some input with promises for more in the future.

**Software Test and Evaluation Guidebook**
(Presented by R. Martin)

The Software Test and Evaluation guidebook is in some senses a culmination of many of STEP's Phase III & IV tasks. Its major goal is to provide, in summary form, information required by the various professionals in the acquisition lifecycle for the successful management of software T&E. In addition to the scope of the guidebook effort, some alternative organizations and styles were presented.

Results of the NADC/Technology Demonstration and Software T&E Model work may also benefit the guidebook as a laboratory framework for carrying out most of the technologically based STEP activities. The relationships of the tasks and ideas for the integration of the results into a guidebook were presented.

**Risk Modelling/Decision-Support-System**
(Presented by R. DeMillo)

Risk modelling is a mathematical way to represent the risks and payoffs inherent in testing. The approach presented borrows heavily from game theory. It looks at issues of risk and utility by attempting to define the costs and arrays of probabilities of events for testing. The empirical information needed to be able to make risk assessments (i.e., cost/benefit and performance) was identified.
**STARS Tasks and Discussion**  
(Presented by J. Passafiume, J. Offutt, & R. DeMillo)

The overall relationship of STARS with STEP was discussed. STARS task areas fall into three categories: management of T&E, technology insertion, and the development of functional requirements for an APSE test environment. The management activities revolve around planning for the hand-off of key T&E functions to STARS and service organizations. Initial efforts toward the development of a test technology baseline will culminate in inserting the technology into military T&E environments. Issues that must be considered when defining the functional requirements for an APSE test environment include: control, modes of test and execution, integration into other lifecycle phases, and human factors. This area is being coordinated with the Joint Service Software Engineering environment (JSEE) Team.

**Panel Discussion**

The final session was reserved for residual discussions and planning for the future involvement of the STEP Technical Advisory Panel.

**Followup**

Subsequent to the meeting, STEP assigned each panelist to a specific task area based on interest and need. Each panelist also received a summary of the meeting, copies of the slides presented, and other pertinent materials.

The next meeting of the STEP Technical Advisory Panel will be held in late March. A firm date and a format for the meeting will be announced.
Preliminary Agenda
STEP Technical Advisory Panel Meeting
November 8-9, 1983
Georgia Institute of Technology

November 8, 1983

8:30 - 9:00 Opening Remarks
9:00 - 10:30 Presentation of STEP Results and Recommendations (Gagliano)
10:30 - 10:45 Break
10:45 - 12:15 Discussion of Recommendations (DeMillo)
12:15 - 1:30 Lunch
1:30 - 2:30 Overview of STARS Initiative (Mall)
2:30 - 3:30 STEP Good Examples (Gagliano)
3:30 - 3:45 Break
3:45 - 4:15 Program Manager's Guidebook (Martin)
4:15 - 5:00 Risk Modelling/Decision Support System (DeMillo)

November 9, 1983

9:00 - 11:00 STARS Task and Discussion
11:00 - 11:15 Break
11:15 - 12:15 Panel Discussion
12:15 - 1:30 Lunch
1:30 - 3:00 Panel Discussion
3:00 Adjourn
LIST OF PARTICIPANTS

Ms. Carolyn Gannon  
General Research Corporation  
5383 Hollister Avenue  
P.O. Box 6770  
Santa Barbara, CA 93111

Ms. Marilyn Stewart  
LOGICON  
1850 N. Lynn St. Suite 900  
Arlington, VA 22209

Mr. John Bowen  
Software Engineering Division  
Hughes Aircraft Company  
Ground Systems Group  
MS 618/P215, P.O. Box 3310  
Fullerton, CA 92634

Mr. Ralph San Antonio  
Dynamics Research Corporation  
7 Lopez Road  
Wilmington, MA 01887

Professor Peter Wegner  
Department of Computer Science  
Brown University  
Box 1910  
Providence, R.I. 02912

Mr. Raymond J. Rubey  
SofTech, Inc.  
3100 Presidential Drive  
Fairborn, OH 45324

Mr. David Gelperin  
Software Quality Engineering  
4825 Olson Memorial Highway, Suite 240  
Golden Valley, MN 55422

Mr. Fred Shockett  
NADC  
Warminster, PA 18974

Lt. Col. Vance Mall  
Ada Joint Program Office  
Room 3D-139 (400A/N Drive)  
The Pentagon  
Washington, D.C. 20301
Professor Richard DeMillo  
Information and Computer Science  
Georgia Institute of Technology  
Atlanta, GA 30332

Professor Ross Gagliano  
Information and Computer Science  
Georgia Institute of Technology  
Atlanta, GA 30332

Professor John Passafiume  
Information and Computer Science  
Georgia Institute of Technology  
Atlanta, GA 30332

Ms. Rhonda J. Martin  
Information and Computer Science  
Georgia Institute of Technology  
Atlanta, GA 30332

Mr. Jeff Offutt  
Information and Computer Science  
Georgia Institute of Technology  
Atlanta, GA 30332

Mr. John Bolino  
Office of the Director Defense  
Room 3D 973  
The Pentagon  
Wash., D.C. 20301

Mr. W. Michael McCracken  
Gen. Electric Company  
Military Electronic Systems Operations  
Court Street Plant  
Building 5, Room 5-P  
P.O. Box 4840  
Syracuse, NY 13221
OSD/DDT&E

SOFTWARE TEST AND EVALUATION PROJECT

GEORGIA INSTITUTE OF TECHNOLOGY
SOFTWARE TEST AND EVALUATION PROJECT (STEP)

- INITIATED BY DDT&E IN 1981

- GOALS

  - DEVELOP NEW DoD POLICY GUIDANCE FOR SOFTWARE T&E
  - STIMULATE TOOL DEVELOPMENT
  - SUPPORT POLICY DEVELOPMENT
  - IDENTIFY RESEARCH DIRECTIONS
OVERVIEW OF PRESENTATION

RATIONALE

STATE-OF-THE-ART

SUMMARY OF FINDINGS

RECOMMENDATIONS
RATIONALE

- SOFTWARE CONTROLS MISSION CRITICAL FUNCTIONS
- IMBALANCES IN DEVELOPMENT OF HARDWARE AND SOFTWARE
- COSTS (TESTING vs. ERRORS)
- INADEQUACIES OF CURRENT GUIDANCE
SOFTWARE CONTROLS
MISSION CRITICAL FUNCTIONS

F/A-18 AIRCRAFT: avionics suite

M-1 TANK: fire control system

AEGIS: distributed processing

TACFIRE: distributed database

MAVERICK: guidance system

WWMCCS: communications systems
IMBALANCES IN DEVELOPMENT OF HARDWARE AND SOFTWARE

1974 DEFENSE SCIENCE TASK FORCE:

"Whereas the hardware development was...monitored, tested and regularly evaluated, the software development was not."

1982 SECRETARY OF DEFENSE:

"...give priority to development of tools and techniques for testing of embedded computers and software."

"Testing of software should...achieve a balanced risk with the hardware..."
COSTS

COSTS DUE TO TESTING

COST OF ERRORS
SOFTWARE LIFECYCLE COST DISTRIBUTION

SOFTWARE OPERATIONAL LIFECYCLE

SOFTWARE DEVELOPMENT LIFECYCLE
INADEQUACIES OF CURRENT GUIDANCE

DoD DIRECTIVE 5000.3

* Quantitative and demonstrable performance objectives and evaluation criteria shall be established...

* Decisions to proceed...based on quantitative demonstration of adequate software performance...
RELATIVE COST
OF ERROR CORRECTION

RELATIVE COST

REQ/SPEC  DESIGN  CODING  INTEGRATE  OPERATION

LIFECYCLE PHASES
DoDD 5000.3 HAS NOT HAD THE DESIRED EFFECT

* "Adequate software performance" is a vague concept.

* Techniques providing "quantitative demonstration" are not used.

* System test objectives may pay little or no attention to software.

* Planning tests to comply with existing guidance is difficult.

* Schedules and budgets squeeze the testing effort.
STEP RECOMMENDATIONS TO DDT&E

NEAR TERM

* Establish chain of software test plans beginning at system level
* Identify test issues for critical software components
* Insert existing technology
* Establish TEMP as major software test planning document

LONG TERM

* Address comprehensive software testing — improve technology
STEP RECOMMENDATIONS TO MILITARY SERVICES

* IMPLEMENT DDT&E RECOMMENDATIONS
* SUPPORT DDT&E RECOMMENDATIONS
* IMPROVE TESTING PRACTICE
  * TECHNOLOGY UPGRADE
  * TEST PLANNING
  * TESTING PROCESS
  * TEST EVALUATION
  * TRI—SERVICE EFFORT
STEP ORGANIZATION

PHASE I
INFORMATION GATHERING

PHASE II
ANALYSIS

PHASE III
FEASIBILITY ASSESSMENT

FEASIBLE
PHASE IVa
IMPLEMENT

NOT FEASIBLE
PHASE IVb
TERMINATE PROJECT
THE STATE-OF-THE-ART
EXISTING TECHNOLOGY: TESTING FOR CORRECTNESS

SPECIFICATION    KNOWN INPUT/OUTPUT

PROGRAM

TEST DATA

MACHINE

COMPARE OUTPUT

IN    OUT
X    Y

OR
## TEST METHODOLOGY

### TEST STRATEGY

Guides overall testing effort

### TESTING TECHNIQUE

Applied within the framework of a test strategy
TEST STRATEGIES

MODULE TESTING

Design of test cases
Coordination of multiple modules

Black Box  White Box
TEST STRATEGIES

INTEGRATION TESTING

INCREMENTAL

NON-INCREMENTSAL

Top-down

Thread Testing
## TESTING TECHNIQUES

### STATIC ANALYSIS

Involves analysis of structural properties but not execution

### DYNAMIC ANALYSIS

Involves execution of program on test data
TESTING TECHNIQUES

STATIC ANALYSIS

Involves analysis of structural properties but not execution

- Requirements analysis
- Specification testing
- Code Inspection
- Walk-throughs
TESTING TECHNIQUES

DYNAMIC ANALYSIS

Involves execution of program on test data

Symbolic Testing
Program Instrumentation
Program Mutation
Input Space Partitioning
Random Testing
Functional Testing
# EVALUATIONS OF TECHNIQUES

<table>
<thead>
<tr>
<th>TECHNIQUE</th>
<th>- subjective</th>
<th>+ effective</th>
<th>SIZE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Static</td>
<td></td>
<td></td>
<td>L</td>
</tr>
<tr>
<td>Symbolic Testing Program Instrument</td>
<td>specification not error oriented</td>
<td>effective debugging time</td>
<td>S L</td>
</tr>
<tr>
<td>Mutation</td>
<td>overtests</td>
<td>effective</td>
<td>M S L</td>
</tr>
<tr>
<td>Partitioning</td>
<td>expensive</td>
<td>reliable</td>
<td></td>
</tr>
<tr>
<td>Random</td>
<td>not error</td>
<td>estimate</td>
<td></td>
</tr>
<tr>
<td>Functional</td>
<td>oriented</td>
<td>effective</td>
<td>L</td>
</tr>
<tr>
<td></td>
<td>isolating</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>functions</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
# TESTING TOOLS

## STATIC ANALYZERS

Analyze structural properties —
no execution

## DYNAMIC ANALYZERS

Execute programs on test data —
analyze program behavior
TESTING TOOLS

STATIC ANALYZERS

Analyze structural properties ——
no execution

Categories
Logic Analyzers
Structural Errors
Front Ends
Coding Style
Interface Consist.

Functions
Data Collection
Error Analysis
Report Gen.
TESTING TOOLS

DYNAMIC ANALYZERS

Execute programs on test data — analyze program behavior

Categories
Symbolic Eval.
Test Data Gen.
Instrumenters
Mutation Tools
Test Drivers

Functions
Test Harness
Execution
Monitor
Eval. Test Data
Gen. Test Data
Gather Statistics
HOW AVAILABLE IS EXISTING TECHNOLOGY?

218 TOOLS CATALOGED ——
42 RESPONDED TO REQUESTS FOR INFORMATION

* MOST TOOLS TAILORED TO PARENT PROJECT
* HIGH COST OF PACKAGING
* NO PUBLISHED COST/BENEFIT STUDIES
* FEW COMPARATIVE EVALUATIONS
* NO PUBLISHED STUDIES ON EFFECTS OF TOOL USAGE
TECHNOLOGY UPGRADE NEEDED

* Investment in test tools
* Techniques for re-testing
* Systematic integration testing
* Systematic real-time testing
* Relate tests to operational characteristics
* Test for performance, etc.
* Quantify test progress and risk
* Demonstrate existing methods
CURRENT DEFENSE PRACTICES
CURRENT PRACTICES

INTERVIEWS

Military Regulations & Standards
Reviews and Inspections
Testing Techniques
Testing Tools
Independent Verif. & Valid.
Risk Assessment

Industry
Applic. Software
Support Software
IV&V Organizations

Military
HQ & Dev. Commds.
Project Offices
OT&E Agencies
PROBLEMS REPORTED BY PROJECT OFFICES

* INSUFFICIENT RESOURCES ALLOCATED TO SOFTWARE DEVELOPMENT AND TESTING

* FEW EFFECTIVE FORMAL PROCEDURES FOR TRACKING PROGRESS DURING EARLY DEVELOPMENT PHASES

* LACK OF EFFECTIVE RISK ASSESSMENT PROCEDURES
SOFTWARE TESTING BY THE CONTRACTORS

* CONTRACTORS RECEIVE LITTLE APPROPRIATE GUIDANCE

* NECESSARY TESTS AND DOCUMENTATION ARE OFTEN NEGLECTED

* EXCESSIVE RELIANCE ON PROGRAMMERS

* TOOLS AND TECHNOLOGY NOT UTILIZED WELL

* CONTRACTORS' PRACTICES NOT AUDITED OR EVALUATED
SOFTWARE IN THE OT&E AGENCIES

* OPERATIONAL TESTERS CONCERNED WITH QUALITY MEASURES OTHER THAN CORRECTNESS

* SOFTWARE IS SINGLED OUT ON AN EXCEPTION—ONLY BASIS

* ERRORS DETECTED DURING OT&E ARE EXPENSIVE TO CORRECT
RECOMMENDATIONS

* ADDRESSED TO DDT&E

* ADDRESSED TO MILITARY SERVICES
MODIFICATION TO DODD 5000.3

IDENTIFY CRITICAL SOFTWARE COMPONENTS

* TEST THROUGHOUT DEVELOPMENT AND INTEGRATION

* TEST RESULTS OBJECTIVE AND TESTS REPEATABLE

* INTERPRET RESULTS IN TERMS OF SYSTEM OBJECTIVES

LEVEL OF TESTING REQUIRED

* MEET SOFTWARE EVALUATION CRITERIA

* BALANCED RISK WITH HARDWARE
TEST AND EVALUATION MASTER PLAN

MISSION/FUNCTION MATRIX

OP. CHARACTERISTICS
TECH. CHARACTERISTICS

DT&E OUTLINE
OT&E OUTLINE

SPECIAL RESOURCE SUMMARY

RELATE CRITICAL SOFTWARE FUNCTIONS TO BE DEMONSTRATED BY TESTING TO MISSION TO BE PERFORMED

REQUIRED SOFTWARE CHARACTERISTICS
SOFTWARE T&E ISSUES

SOFTWARE T&E OUTLINE
TO DATE
PLANNED
OBJECTIVE
SCOPE/EVENTS/SCENARIOS
CRITICAL ITEMS

TEST ARTICLES
TEST TOOLS
INITIATE TOOL DEVELOPMENT PROGRAM

* DEVELOPMENT, PACKAGING, EVALUATION, DISTRIBUTION

* SUPPORT WAREHOUSE, CATALOG OR TEST ENVIRONMENT

* INCLUDE IN TEST PLANS WITHOUT FURTHER ACQUISITION

* DEFINE USAGE CONTEXTS: APPLICATIONS, LANGUAGES, EVAL. CRITERIA

* SERVICE PARTICIPATION — LEVERAGE EXISTING PROGRAMS
DEFINE SOFTWARE T&E MODEL

* CORRELATE WITH ACQUISITION DECISION MILESTONES

* CORRELATE WITH DT&E/OT&E/PAT&E

* CONSIDER SEPARATE SOFTWARE T&E PROGRAM

* GOAL: SUPPORT DT&E/OT&E AND FOCUS ON S/W ISSUES NOT ADDRESSED BY DT&E/OT&E
SET REQUIRED TESTING GOALS FOR TOTAL SOFTWARE SYSTEMS

* DEFINE RISK REDUCTION PROCEDURES FOR
  - SUPPORT SOFTWARE
  - NON-CRITICAL SOFTWARE
  - OTHER DESIGN COMPONENTS

* DEFINE EVALUATION CRITERIA

* GOAL: QUANTITATIVE RISK MODEL FOR TOTAL SOFTWARE SYSTEM

* INCLUDE COMPLETE SOFTWARE ISSUES IN TEST PLANS
IMPLEMENTATION
REQUIRES COORDINATION WITH SERVICES

* IMPLEMENT RECOMMENDATIONS TO DDT&E

* SUPPORT DDT&E IN AREAS NOT DIRECTLY ADDRESSED BY DODD 5000.3

* IMPROVE THE SOFTWARE TESTING PROCESS
STATE-OF-THE-ART IMPROVEMENTS

* **DoD SOFTWARE TECHNOLOGY FOR ADAPTABLE, RELIABLE SYSTEMS (STARS) PROGRAM**

* **INCREASED ATTENTION TO TESTING BY SERVICE RESEARCH OFFICES**

* **SERVICES ASSUME LEAD ROLE IN TEST TOOL PROGRAM**

* **EARLY INSERTION OF TEST TECHNOLOGY IN ADA SUPPORT ENVIRONMENTS**
TEST PLANNING

* WRITTEN TEST PLANS FOR EARLY SOFTWARE DEVELOPMENT PHASES
  o SPECIFY ACCEPTABLE APPROACH
  o EXPLAIN HOW APPROACH SUPPORTS TEMP, HIGHER LEVEL TESTS
  o RIGOROUS ADHERENCE
  o REVIEW FOR DEFICIENCIES
  o REALISTIC ESTIMATE OF SCOPE/EXTENT

* REQUIRE DOCUMENTATION OF UNIT AND MODULE TESTING

* PLANS FOR REGRESSION TESTS, IV&V
TESTING PROCESS

* EMPHASIZE UNIT AND MODULE TESTS TO IDENTIFY ERRORS EARLY
  - SET GOALS FOR TOTAL SOFTWARE SYSTEM
  - RESOLVE MAJOR DEFICIENCIES BEFORE DEDICATED OT&E
  - INCLUDE OPERATION TEST DOCUMENTATION IN TEST DATABASE
  - SYSTEMATIC, APPROPRIATE METHODOLOGIES
  - INTERPRET RESULTS IN TERMS OF SYSTEM OBJECTIVES

* SOFTWARE TESTING TOOLS PROGRAM
  - DEMONSTRATE TOOLS -- SET QUALIFICATION REQUIREMENT
  - INCORPORATE DEMONSTRATED TECHNOLOGY
  - BUILD ON EXISTING PROGRAMS AND ACTIVITIES
TEST EVALUATION

* QUANTITATIVE INDICES OF TEST PROGRESS
  - RELIABLE COST/BENEFIT MEASURES
  - RISK ANALYSIS

* DATA GATHERING, REDUCTION, AND MANAGEMENT
  - AUTOMATED DATA LOGGING
  - ERROR TRACKING
  - METRICS

* DEVELOP EFFECTIVE S/W QA STANDARD

* INTEGRATED DECISION SUPPORT SYSTEM
  - FUNCTIONALLY ORGANIZED TEST INFORMATION
  - DATA FOR PROGRAMMATIC DECISION POINTS
IMPLEMENTING THE RECOMMENDATIONS:

HOW AND WHEN WILL IMPROVEMENTS BE IMPLEMENTED?

* TOOLS AND TECHNOLOGY

* SPECIFICATIONS AND STANDARDS

* POLICY AND GUIDANCE
IMPLEMENTATION STRATEGY

* SELECT SKELETAL SET OF ACTIVITIES ON CRITICAL PATH TO IMPROVEMENTS
  - SMALL NUMBER
  - NEAR TERM PRODUCTS

* GATHER SUPPORT – EXPAND ACTIVITY IN NON–OSD ORGANIZATIONS
  - BUILD ON PHASES I&II
  - SELECTED PROJECT OFFICES
  - TEST ORGANIZATIONS
  - PANEL OF ADVISORS

* IDENTIFY ORGANIZATIONS TO ASSUME ON–GOING RESPONSIBILITY FOR KEY FUNCTIONS
  - BUILD ON EXISTING ACTIVITIES
  - PRECISE MISSIONS/ROLES
  - IDENTIFY LONG–RANGE FUNDING
TOOLS AND TECHNOLOGY
IMPLEMENTATION PLAN

1983: BEGIN INSERTION OF TECHNOLOGY IN SELECTED PROJECTS
1985: PERMANENT PLAN IN PLACE

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>WHO</th>
<th>WHEN</th>
<th>CANDIDATE PRODUCT</th>
<th>ONGOING RESPONSIBILITY</th>
</tr>
</thead>
<tbody>
<tr>
<td>IDENTIFY TECHNOLOGY</td>
<td>STEP</td>
<td>81–83</td>
<td>STEP REPORT</td>
<td>SERVICES SEI</td>
</tr>
<tr>
<td>DEMONSTRATION ON PROJECT</td>
<td>STEP/NADC</td>
<td>83–84</td>
<td>FASP</td>
<td>SERVICES SEI</td>
</tr>
<tr>
<td>INSERT TECHNOLOGY</td>
<td>STEP/NADC</td>
<td>84</td>
<td>S/W T&amp;E GUIDE TOOL MGT. PLAN</td>
<td>SERVICES STARS</td>
</tr>
<tr>
<td></td>
<td>STEP/STARS</td>
<td>84</td>
<td>ADA ENVIRONMT.</td>
<td>S/W ACQ. PANEL</td>
</tr>
<tr>
<td></td>
<td>STEP/STARS</td>
<td>85</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DISTRIBUT/E MARKET</td>
<td>STEP/STARS/</td>
<td>84</td>
<td>S/W T&amp;E GUIDE WAREHOUSE/CATALOG</td>
<td>PROJ. OFFICES DSMC</td>
</tr>
<tr>
<td></td>
<td>SERVICES/</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>STEP PANELS</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
CANDIDATE PROJECT
NADC ASW MODULE (UYK7, CMS2Y)

* CONTINUED SOFTWARE DEVELOPMENT
* DEMONSTRATE:
  EFFECTS OF SYSTEMATIC TEST PLANNING
  TEST DOCUMENTATION AND EVALUATION
  TECHNOLOGY INSERTION ON EXISTING PROJECT
**SPECIFICATIONS AND STANDARDS IMPLEMENTATION PLAN**

1983: PARTICIPATE IN STANDARD FORMULATION

1985: PROVIDE MANUAL FOR TEST STANDARDS TO PROJECT MGRS.

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>WHO</th>
<th>WHEN</th>
<th>PRODUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>MIL-STD-SDS</td>
<td>STEP/JLC</td>
<td>84</td>
<td>TRI-SERVICE STD</td>
</tr>
<tr>
<td>MIL-STD-SQAM</td>
<td>STEP/JLC</td>
<td>---</td>
<td>TRI-SERVICE STD</td>
</tr>
<tr>
<td>IEEE TEST STDs</td>
<td>STEP/IEEE</td>
<td>84</td>
<td>INDUSTRY STD</td>
</tr>
<tr>
<td>MIL-STD-1679</td>
<td>STEP/USN</td>
<td>84</td>
<td>NAVY STD</td>
</tr>
<tr>
<td>PROPAGATION</td>
<td>STEP/NADC</td>
<td>83-85</td>
<td>S/W T&amp;E GUIDE</td>
</tr>
<tr>
<td></td>
<td>MIL. ADV. PANEL</td>
<td></td>
<td>MILESTONE DEFINITION</td>
</tr>
</tbody>
</table>
POLICY AND GUIDANCE
IMPLEMENTATION PLAN

1983: DDT&E REVISION OF DODD 5000.3
1985: ELABORATE POLICY AND PROVIDE DECISION SUPPORT AIDS

<table>
<thead>
<tr>
<th>ACTIVITY</th>
<th>WHO</th>
<th>WHEN</th>
<th>PRODUCT</th>
</tr>
</thead>
<tbody>
<tr>
<td>POLICY RECOMM.</td>
<td>STEP</td>
<td>81–83</td>
<td>STEP REPORT</td>
</tr>
<tr>
<td>POLICY MOD.</td>
<td>DDT&amp;E</td>
<td>83–84</td>
<td>DODD 5000.3</td>
</tr>
<tr>
<td>INTEG. S/W T&amp;E</td>
<td>STEP/BDM/DDT&amp;E</td>
<td>83–85</td>
<td>LIFECYCLE MATRIX</td>
</tr>
<tr>
<td>ELABORATION</td>
<td>DDT&amp;E</td>
<td>84–85</td>
<td>--</td>
</tr>
<tr>
<td>DECISION SUPP.</td>
<td>STEP/AFTEC/BDM/RADC/DDT&amp;E</td>
<td>83–85</td>
<td>RISK MODEL DSS REQUIREMENTS</td>
</tr>
</tbody>
</table>
STEP LEADS TO IMPROVED SOFTWARE T&E

* MEANINGFUL GUIDANCE
  ** CHANGES TO POLICY
  ** NEW STANDARDS
  ** GOOD EXAMPLES
  ** PROGRAM MANAGER'S GUIDE
  ** FOCUSED RESPONSIBILITY

* TECHNOLOGY UPGRADE
  ** TECHNOLOGY BASELINE
  ** INSERTION MECHANISMS
  ** RISK MODELS
  ** NEW TECHNOLOGY
SOFTWARE IN THE MODIFIED TEST AND EVALUATION MASTER PLAN

-- Criteria for evaluation of TEMP

-- Criteria for additional guidance to Project Offices

-- Examples of software characteristics and test issues

-- Examples of software test articles and special support requirements
## PART I - DESCRIPTION

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mission</td>
<td>a. Does the system contain Mission-critical Computer Resources?</td>
</tr>
<tr>
<td>2. System</td>
<td>b. Does software implement critical functions?</td>
</tr>
<tr>
<td></td>
<td>c. Is the system software intensive?</td>
</tr>
</tbody>
</table>
2. System (continued)

a. Key Functions

a. Does Mission/Function Matrix identify primary functional capabilities to be implemented by software?

b. Are the functions:
   -- New?
   -- Modifications of existing capabilities?
   -- Automation of existing capabilities?
   -- Mature?
PART I - DESCRIPTION (CONTINUED)

SECTION

2. System (continued)

   a. Key Functions

   b. Interfaces

   a. Is software important to the interfaces?

   b. Do the interfaces have software implications?
SECTION

2. System (continued)

   a. Key Functions

   b. Interfaces

   c. Unique Characteristics

   a. Does the system use software technology that:
      -- Affects risk?
      -- Has lifecycle impact?
      -- Distinguishes it from other systems?
<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. System</td>
<td></td>
</tr>
<tr>
<td>3. Required Operational Characteristics</td>
<td>a. Are there software characteristics that: -- Are unique to software? -- May be overlooked?</td>
</tr>
<tr>
<td>4. Required Technical Characteristics</td>
<td></td>
</tr>
<tr>
<td>5. Required Software Characteristics</td>
<td></td>
</tr>
</tbody>
</table>
REQUIRED SOFTWARE CHARACTERISTICS

"Software parameters that are primary indicators of conformance to written requirements/specifications and operational suitability/effectiveness."

<table>
<thead>
<tr>
<th>TYPE</th>
<th>PARAMETER</th>
<th>UNIQUE?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>performance</td>
<td>No</td>
</tr>
<tr>
<td>Operational</td>
<td>maintainability</td>
<td>Unique aspects</td>
</tr>
<tr>
<td>Technical</td>
<td>precision/accuracy</td>
<td>Minimal</td>
</tr>
<tr>
<td>Technical</td>
<td>robustness</td>
<td>Yes</td>
</tr>
</tbody>
</table>
PART I - DESCRIPTION (CONTINUED)

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. System</td>
<td></td>
</tr>
<tr>
<td>3. Required Operational Characteristics</td>
<td></td>
</tr>
<tr>
<td>4. Required Technical Characteristics</td>
<td></td>
</tr>
<tr>
<td>5. Required Software Characteristics</td>
<td></td>
</tr>
<tr>
<td></td>
<td>b. Operational Issues</td>
</tr>
<tr>
<td></td>
<td>c. Software Issues</td>
</tr>
<tr>
<td></td>
<td>a. Do the required software characteristics raise unique or easily missed T&amp;E issues?</td>
</tr>
</tbody>
</table>
CRITICAL SOFTWARE T&E ISSUES

"Those aspects of software capability...that must be questioned before a system’s overall worth can be estimated...“

EXAMPLES

<table>
<thead>
<tr>
<th>TYPE</th>
<th>CHARACTERISTIC</th>
<th>ISSUE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>performance</td>
<td>degraded mode</td>
</tr>
<tr>
<td></td>
<td></td>
<td>operation</td>
</tr>
<tr>
<td>Operational</td>
<td>maintainability</td>
<td>adequacy of support</td>
</tr>
<tr>
<td></td>
<td></td>
<td>environment</td>
</tr>
<tr>
<td>Technical</td>
<td>precision/accuracy</td>
<td>algorithm</td>
</tr>
<tr>
<td></td>
<td></td>
<td>correctness</td>
</tr>
<tr>
<td>Technical</td>
<td>robustness</td>
<td>design and architecture</td>
</tr>
</tbody>
</table>
PART II - PROGRAM SUMMARY

SECTION

1. Management

2. Integrated Schedule

QUESTIONS

a. Are key software subsystem demonstrations included on schedule?
T&E OUTLINES

PART III - DT&E Outline

PART IV - OT&E Outline

PART V - Software T&E Outline
SECTION 1. Software T&E to Date

a. Have software characteristics been demonstrated using systematic test methods?

b. Have the planned levels of testing have been achieved? Is the documentation that reports these test results cited?

c. Have testing deficiencies been interpreted in terms of overall system evaluation criteria?
PART V - SOFTWARE T&E OUTLINE (CONTINUED)

SECTION

QUESTIONS

1. Software T&E to date

2. Future Software T&E

   a. Is the test environment (i.e., development, operational, logistics support) appropriate for the characteristics to be demonstrated?
## PART V - SOFTWARE T&E OUTLINE (CONTINUED)

### SECTION

<table>
<thead>
<tr>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>1. Software T&amp;E to date</strong></td>
</tr>
<tr>
<td><strong>2. Future Software T&amp;E (continued)</strong></td>
</tr>
<tr>
<td>a. Software T&amp;E Objectives</td>
</tr>
<tr>
<td>b. Software T&amp;E Events/Scope of Testing/Basic Scenarios</td>
</tr>
<tr>
<td>a. What software characteristics are not adequately addressed in the DT&amp;E and OT&amp;E Outlines?</td>
</tr>
</tbody>
</table>
PART V - SOFTWARE T&E OUTLINE (CONTINUED)

SECTION

QUESTIONS

1. Software T&E to date

2. Future Software T&E (continued)

   a. Software T&E Objectives

   b. Software T&E Events/Scope of Testing/Basic Scenarios

   c. Critical Software T&E Items

   a. Are subsystems needed for adequate software T&E prior to next decision point identified?
## PART VII - SPECIAL RESOURCE SUMMARY

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Test Articles</td>
<td>a. Are critical software components and key subsystems identified?</td>
</tr>
<tr>
<td>SECTION</td>
<td>QUESTIONS</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>1. Test Articles</td>
<td></td>
</tr>
<tr>
<td>2. Special Support Requirements</td>
<td>a. Is there an explanation of how test tools support software test objectives?</td>
</tr>
<tr>
<td></td>
<td>b. Are adequate steps being taken to acquire each tool? Do any of the tools increase risk?</td>
</tr>
</tbody>
</table>
STEP

GOOD EXAMPLES
STEP SOW

Task 7

(1) ...gather objective evidence which indicates the effectiveness of systematic program testing...

...at least two existing systems will be identified for evaluation...

...selected by the degree to which the software has been subjected to systematic testing techniques...
PURPOSE

1. TEMP Checklist
2. Software T&E Guidebook
3. DOD D 5000.3 Recommendations
ARMY SYSTEMS

AAH
ASH
ARTADS
Blackhawk
Bradley Vehicle
Chaparral
Cobra
DIVAD
TACFIRE
Firefinder
FIST Vehicle
Hawk
Hellfire

MICNS
Patriot
Pershing
PLRS
SINCGARS
SEMA
Stinger
RPV
TADS/PVNS
TOW
Roland
Viper
AIR FORCE SYSTEMS

A - 10
ALCM
AWACS
B - 1B
B - 52
C 3
DSCS III
F - 16
GB EO DS System
GPS - NAVSTAR
JTIDS
LANTIRN
Maverick
MX
NORAD Upgrade
OTH
Pavepaws
NAVY SYSTEMS

Aegis
ACDS
Airborne ESM
CG - 17
CGN - 38
LAMPS
MAT3
MK 117
NTDS
P3C
S - 3
Share 7
SIM/STIM
SUBACS
TACC
Trident

JOINT SYSTEMS

AMPE
JT FP
Sat N W
CRITERIA FOR SYSTEMS SELECTION

1. Deployed
2. Software T&E Success
3. Information Access
4. Generation/Release of New Information
5. Mix of Systems
   a. Three Services
   b. Wide spectrum of systems
   c. Different types of Software
FINAL SET OF SYSTEMS

NAVY

Aegis
LAMPS

AIR FORCE

AWACS
Pavepaws

ARMY

Firefinder/REMBASS
(AN/TPQ-36, AN/TPQ-37)

Missileminder
(An/TSQ-73)
LEVELS OF ORGANIZATIONS

Headquarters

Development Commands

Proponent

System Program Office

Testing Organization(s)

Prime Contractor(s)

Sub-contractor(s) - Software

V & V Group
DOCUMENT TRACE

MENS - Mission Element Needs Statement
SON - Statement of Need
ROC - Required Operational Capability
RFP - Request for Proposal
SOW - Statement of Work
Proposal

TEMP - Test and Evaluation Master Plan

Plans
Development
Life-cycle Management
Test

Procedures
Reports
Reviews
Audits
Discussion Subjects

Military Regulations & Standards
Organization & Management
Plans & Procedures
Techniques & Tools
Verification & Validation
Audits & Reports
Reviews & Inspections
Risk Assessment
Quality Assurance
Simulators
Environments
QUESTIONS

What was the nature of the environments?

How was the software quality accomplished?

What was done to balance the Hardware and Software risks?

To what other unique features do you attribute the Software T & E success?
NATURE OF THE ENVIRONMENTS

1. What was the Software T & E environment for:
   Development Tests?
   Operational Tests?

2. What special effort was required in transitioning to the support environment?

3. Were there other special environments; i.e., maintenance?

4. What are the long-term aspects of the tools in the various environments?
SOFTWARE QUALITY

1. What test procedures were used?

2. How were they enforced?

3. How were reviews and audits conducted?

4. How were reviews and audits documented?
BALANCING HW AND SW RISKS

1. Were automated tools used in:
   Development?
   Testing?

2. Were simulators used?

3. Was regression testing utilized?

4. Was V & V independent?
OTHER UNIQUE FEATURES

1. Were there attempts to insert new SW technology?

2. Were there unusual features of the system?

3. Were there special relationships between the government agencies and the prime?

4. Was there scheduled sequencing of SW delivery, testing and acceptance?
PRELIMINARY OBSERVATIONS

TECHNICAL

Prior Simulator Work
Automated Support Facilities

MANAGEMENT

Software Responsibility
Early Government Involvement
## LOCATIONS

<table>
<thead>
<tr>
<th>System</th>
<th>DevComd</th>
<th>ProgOfc</th>
<th>Prime</th>
<th>Sub-SW</th>
<th>IV &amp; V</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aegis</td>
<td>NAVA MAT/NSSC</td>
<td>NSWC</td>
<td>RCA</td>
<td>CSC</td>
<td>RCA</td>
</tr>
<tr>
<td>LAMPS</td>
<td>NASC</td>
<td>NADC</td>
<td>IBM</td>
<td>IBM/R'th'nNADC</td>
<td></td>
</tr>
<tr>
<td>TPQ-36</td>
<td>ERADCOM</td>
<td>CSTA</td>
<td>Hughes</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TSQ-73</td>
<td>MICOM</td>
<td>Redstone</td>
<td>Sperry</td>
<td></td>
<td></td>
</tr>
<tr>
<td>AWACS</td>
<td>AFSC</td>
<td>ESD</td>
<td>Boeing</td>
<td></td>
<td></td>
</tr>
<tr>
<td>P'Paws</td>
<td>AFSC</td>
<td>ESD</td>
<td>Raytheon</td>
<td>ESD</td>
<td></td>
</tr>
</tbody>
</table>
# SCHEDULE OF VISITS

<table>
<thead>
<tr>
<th>Location</th>
<th>Agency</th>
<th>Tentative Dates</th>
</tr>
</thead>
<tbody>
<tr>
<td>Huntsville, AL</td>
<td>MICOM</td>
<td>Nov - Dec '83</td>
</tr>
<tr>
<td></td>
<td>TSQ-73 PMO</td>
<td></td>
</tr>
<tr>
<td>Washington</td>
<td>NAVMAT (2)</td>
<td>Dec '83 - Jan '84</td>
</tr>
<tr>
<td></td>
<td>ERADCOM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>AFSC (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>NSWWC</td>
<td></td>
</tr>
<tr>
<td></td>
<td>IBM</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CSC</td>
<td></td>
</tr>
<tr>
<td>Hanscom AFB</td>
<td>ESD (2)</td>
<td>Jan '84</td>
</tr>
<tr>
<td></td>
<td>Raytheon (2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sperry</td>
<td></td>
</tr>
<tr>
<td>Owego, NY</td>
<td>IBM</td>
<td>Jan '84</td>
</tr>
<tr>
<td>Ft. Monmouth</td>
<td>TPQ-36 PMO</td>
<td>Feb '84</td>
</tr>
<tr>
<td>Mooresstown</td>
<td>RCA</td>
<td>Feb '84</td>
</tr>
<tr>
<td>Fullerton, CA</td>
<td>Hughes</td>
<td>Feb '84</td>
</tr>
<tr>
<td>Seattle, WA</td>
<td>Boeing</td>
<td>Feb '84</td>
</tr>
</tbody>
</table>
Program Manager's

Guidebook

R. J. Martin
08 Nov 83
What ???

When ???

Why ???

How ???
Program Manager's Guidebook

- Checklists

- Supporting Information
Program Manager's Guidebook

SLW T&E model
Technology Baseline
TEMP Checklist

Risk Modelling

Good Examples

NADG Technology Demonstrations
Existing Documentation
Plans for NADC Technology Demonstration

1. Gather information on FASP

2. Select project

3. Select tools

4. Apply STEP recommendations
   - Identify critical components
   - Develop test plans, etc.
   - Conduct tests
   - Evaluate & document results

5. Track errors found per level of test ... report to NADC.
Software T&E Model

DT&E / OT&E / PAT&E

RS.

SIW T&E
RISK MODELS

OR:

HOW MUCH TESTING IS ENOUGH?

PART IV
3 LIFECYCLES

SOFTWARE

APPLICATION SYSTEM

DECISION

CRITICAL FACTOR IN DECISIONS - RISK

TESTING - REDUCE DECISION RISK
ISSUES

GENERALIZED COST

LIKELIHOOD OF ERROR

PROGRAM ERRORS VS. FUNCTIONAL ERRORS
TESTING — RISK REDUCING ACTIVITY

\[ S_1 = \text{ERROR X IS OBSERVED} \]

\[ S_2 = \text{ERROR X IS NOT OBSERVED} \]

\[ u_{ij} = "\text{desirability" of event (i,j)} \]
CHOOSE TEST STRATEGY FROM AMONG $A_1, A_2, \ldots, A_n$

EVENT $S_i \in \{S_1, S_2, \ldots, S_m\}$ IS OBSERVED WHEN SYSTEM IS OPERATED

$\text{NATURE}$
$S_1 \quad S_2 \quad S_3 \quad \ldots \quad S_m$

$(A_i, S_j) = \text{OUTCOME}$
Goal - RANK TESTER'S CHOICES WRT

- \( u_{ij} \)'s (resp. \( r_{ij} \)'s)
- Optimality criteria

\( S_1 = \text{CRITICAL SYSTEM ERROR OCCURS} \)

\( S_2 = \overline{S_1} \)

\( A_2 = \text{high cost test designed to insure that} \ S_1 \ \text{does not occur} \)

\( A_1 = \text{low cost test} \)

Does tester choose \( A_1 \) or \( A_2 \)?
Policy #1: High cost test should be applied to reduce risk of $S_1$.

\[ U_{12} \quad U_{21} \]

\[ U_{22} \quad (1) \]

\[ U_{11} \quad (0) \]

\[
\begin{pmatrix}
S_1 \\ S_2
\end{pmatrix}
\begin{pmatrix}
0 & 2 \\
2 & 1
\end{pmatrix}
\]

\[
\begin{pmatrix}
S_1 \\ S_2
\end{pmatrix}
\begin{pmatrix}
2 & 0 \\
0 & 1
\end{pmatrix}
\]
Policy #2: Apply Least Cost Test and Justify Incremental Costs

\[ U_{12} \]

\( U_{11} \rightarrow U_{22} (2) \)

\( U_{21} (0) \)

\[
\begin{bmatrix}
A_1 \\
A_2 \\
\end{bmatrix}
\begin{bmatrix}
S_1 \\
S_2 \\
\end{bmatrix}
\]

\[
\begin{bmatrix}
0 \\
2 \\
2 \\
1 \\
\end{bmatrix}
\]
Policy #3: Another Interpretation of Lowest-Cost Strategy

\[
\begin{array}{cc}
\text{A}_1 & \text{A}_2 \\
\begin{pmatrix}
\text{u}_{11} & \text{u}_{12} (\infty) \\
\text{u}_{21} & \text{u}_{22} (1)
\end{pmatrix}
\end{array}
\]

\[
\begin{array}{cc}
\text{S}_1 & \text{S}_2 \\
\text{A}_1 & 0 & 10^6 \\
\text{A}_2 & \text{10}^{-6} & \text{10}^{-5}
\end{array}
\]

Always best if Nature is conservative at 10^{10}.
Cost of Failure:

\[ \text{COST}(S_j) = \text{"LOST FUNCTION" COST} + \text{CORRECTION COST} + \ldots \]

\[ \frac{(1 - \frac{\text{COST}_{i}}{\text{COST}_j})}{\text{COST}(A_i)} \]

\[ S \]
Utility should also measure "surprise"

\[
\begin{bmatrix}
S & \bar{S} \\
\begin{bmatrix}
u_{RS} & u_{R\bar{S}} \\
u_{PS} & u_{P\bar{S}} \\
u_{CS} & u_{C\bar{S}}
\end{bmatrix}
\end{bmatrix}
\]

Random Path
Correctness Proof

\[\text{Cost} \downarrow\]

\(\oplus\) passes Occurs Surprise Utility
ACTORS AFFECTING DEFINITION OF TEST CASES

- number of variables
- number of domains
- structural complexity

DEFINE SEQUENCE OF TEST CASES

T₀, T₁, T₂, ..., Tᵢ, Tᵢ₊₁, ...

sᵢ = number of factor variations
to obtain Tᵢ from Tᵢ₋₁

s* = number of possible variations

REFERENCE FACTOR:

Δₗ ≡ sᵢ / s*

Δₗ is a measure of cost
Utility of Test $T_j$

$$U_j \triangleq \Delta_j \cdot \Sigma_i \cos(r)$$

$r$ is a remaining error

**Conjecture (Ruby)**

$$\Delta_j \triangleq \Sigma_i \text{Preco}[T_j \text{ detects error of}]$$
Utility of event \((T_i, S_j)\) = 

\[ u_{ij} = \text{COST}(T_i, S_j) \times \text{Prob}\{S_j|\Theta \text{ passes } T_i\} \]

**But**

\[ \text{Prob}\{S_j|\Theta \text{ passes } T_i\} \text{ may not be:} \]

1. Known
2. Easily determined
3. Meaningful

**APPROACH:**

Define: (i) \(P_{ij}\) such that

\[ P_{ij} \leq P_{kj} \iff \text{Prob}\{S_j|\Theta \text{ passes } T_i\} \leq \text{Prob}\{S_j|\Theta \text{ passes } T_k\} \]

(2) Test selection criteria that respects ordering
What strategies capable of detecting errors that can cause failure $S_1$?
Selecting Optimal Test $\hat{T}$

Problems with Max-min (min-max risk)

1) Small advantage can outweigh large ones
2) May select $\hat{T} = T_3$ from $\{T_1, T_2, T_3, T_4\}$
   
   $\hat{T} = T_2$ from $\{T_1, T_2, T_3\}$

Alternatives:

1) $\forall i \neq j$ choose $T$ that minimize risk
   $\in \{T_i, T_j\}$

2) $\hat{T} =$ winner of tournament $(\pm)$
   (intransitivities $\Rightarrow \hat{T}$ is not unique)
ESSIMISM-OPTIMISM INDEX
max-min looks at worst consequence

\[
\begin{array}{c}
S_1, S_2, \ldots, S_n \\
T_i \left| u_{11}, u_{12}, \ldots, u_{in} \right.
\end{array}
\]

\[s = \min \{x : \alpha = 0 \text{ is min-max} \}
\]
\[\sigma \text{ can be chosen empirically}
\]
\[\text{Randomisation over optimal } T^* \text{ not nec. optimal}
\]
\[\text{Violates "complete ignorance"}
\]
Choose $T_i$ to maximize $\frac{1}{n} \sum s_j$.

**Objection:** Different partitioning of nature $\Rightarrow$ different solutions
Axiomatics

1. Completely order test strategies
2. Isolate optimal set

\[ \Delta = \{ T_1, \ldots, T_m \} \]
\[ \Delta = \{ S_1, \ldots, S_n \} \]
\[ J^* = \text{optimal set} \]

\[ T_i \text{ is ADMISSIBLE IF there is no } j \text{ s.t. } T_j \geq T_i \]
Independence of Irrelevant Alternatives

if \( T_j \succ T \) for some \( T_j \in S \) then each \( T \in S^\ast \) is also optimal for \( S \cup \{T_j\} \)
SUBOPTIMALITY

Let $P_j = \text{Prob} \{ S_j \} \text{ observed}$

$p_j = \text{assumed distribution.}$

\[
\text{val}_{p_i} = P_1u_{i1} + P_2u_{i2} + \cdots
\]

\[
\text{val}_{p_j} = P_1u_{j1} + P_2u_{j2} + \cdots
\]

Characterize $P, p$ such that for $T_i$ chosen according to $p$

\[
c \cdot \text{val}_{p_i} \geq \text{val}_{p_j}
\]

where $c \geq 1$ depends only on $P, p$
To QUANTIFY ELEMENTS OF RISK
THE DoD SOFTWARE TECHNOLOGY FOR ADAPTABLE, RELIABLE SYSTEMS (STARS) PROGRAM

- THE GOALS OF STEP AND STARS OVERLAP IN THE AREA OF TECHNOLOGY IMPROVEMENT

STARS: Improve the state of practice in the acquisition, management, development and support of computer software...

STEP: Improve the policy, regulations, and technology for the test and evaluation of computer software...

- STEP WILL RELY ON STARS TO IMPLEMENT LONG-RANGE IMPROVEMENTS IN SOFTWARE T&E
  - On-going organizations and support
  - Broadly-based technology upgrades
  - Education and training
  - Modern software engineering environments
PARALLELS BETWEEN STEP AND STARS IMPLEMENTATION STRATEGIES

- **STARS:** EXPLOIT EXISTING TECHNOLOGY BASE, BUILD ON EXISTING DoD EFFORTS, AND COORDINATE TALENTS AND EXPERTISE OF MANY DoD ORGANIZATIONS

- **STEP:** IDENTIFY AND DEMONSTRATE EXISTING TEST TECHNOLOGY, EXPAND ACTIVITY IN NON-OSD ORGANIZATIONS, IDENTIFY ORGANIZATIONS TO ASSUME ON-GOING RESPONSIBILITY FOR KEY FUNCTIONS
STARS ORGANIZATIONS
KEY TO STEP IMPLEMENTATION EFFORTS

0 THE SOFTWARE ENGINEERING INSTITUTE

-- Demonstration and qualification of new test technology

-- Identification and productizing of promising tools and methodologies

-- Distribution and "marketing" of qualified tools to Service warehouses

-- Education and Research

0 THE ACQUISITION PANEL

-- Oversight and management of software T&E technology

-- Plan implementation of policy and regulations
STARS Tasks

T&E MANAGEMENT PLAN FOR ACQUISITION PANEL -

Define organizational structure for managing improvements in the testing process and implementing policy for software T&E.

T&E MANAGEMENT AND OPERATING PLANS FOR SEI -

Define the role of SEI in supporting test tool development and in encouraging the integration of state-of-the-art test technologies into modern programming support environments.

TEST TECHNOLOGY BASELINE AND IDENTIFICATION PROCEDURES FOR SEI -

Define an initial procedure for identifying test tools, environments, and methodologies for Service qualification and inclusion in a warehouse to satisfy test plan requirements.

PROVIDE DEMONSTRATION AND QUALIFICATION PROCEDURES FOR SEI -

Define a procedure for demonstrating and qualifying tools for DoD use.

PROVIDE FUNCTIONAL REQUIREMENTS FOR APSE TEST ENVIRONMENTS -

Offer support to STARS, SEI planners and the Joint Service Team for Ada Environment definition in the formulation of requirements for candidate test environments.
TOOL CLASSIFICATIONS

* Static Analysis Tools

* Dynamic Analysis Tools

* Test Support Tools
STATIC ANALYSIS TOOLS

* Code Audit/Consistency
* Cross Reference
* Interface Analysis
* I/O Specification Analysis
* Data Flow Analysis
* Type Analysis
* Units Analysis
DYNAMIC ANALYSIS TOOLS

* Symbolic Evaluators

* Test Data Generator

* Program Instrumentor

* Program Mutation Analyzer
TEST SUPPORT TOOLS

* Automatic Test Drivers

* Comparators
GOALS

Short Range

* Facility for Automated Software Production (FASP)

* Tool Warehouse

* Find and Catalog Available Tools
GOALS

Long Range

* Develop Criteria and Standards for Evaluating and Designing Tools

* Improve Test Tool Technology

* Develop/Define Tools for Ada Test Environment
The Issue of Control

In an operating systems, errors are abnormal conditions.

In test environments, errors are standard conditions.
Modes of Execution

Kernel
Full Test
Data Generation
Data Gathering/Execution Monitoring
No Test
The Execution Environments

If the test environment is: development support operational support operational

The target can be: development support operational operational
Human Factors Issues

Display
Data Reduction
Interference and Orthogonality
Gaming
There are lots of other issues:

Virtual Machine State
Plug Compatible Tools
Internal Representations
Please respond with a short (one paragraph) answer to each question.

Your view, what is the role of the Advisory Panel?
To Advise and comment on tasks performed by STEP.

What specific contributions do you think that you can make to STEP?
To review and comment on STEP papers.
My strength is writing and editing - put me on the distribution list for reports.

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?
Not a specialist in any of these areas.

Which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?
Good Examples
Risk Analysis

Are you interested in submitting a “good example” for consideration by the panel? Also, is the topic suitable for the ITEA good example meeting?
Probably not.

I would like to participate in writing a paper on the role of testing in the software life cycle, that could serve as a basis for input to STARS and the standards.
In your view, what is the role of the Advisory Panel?

- Review STEP activities (plans)
- Suggest STEP implementation strategies
- Solicit user input and comments

What specific contributions do you think that you can make to STEP?

- Developing a process for transferring existing T&E technology
to new projects

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

Technology Insertion

In which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?

- STARS tasks
  - T&E Management and Operating Plans for SEI
  - T&E Technology transfer and Identification Procedures for SEI

Are you interested in submitting a "good example" for consideration by the panel? If so, is the topic suitable for the ITEA good example meeting?

- Provide Functional Requirements for ASE Test Environments
Please respond with a short (one paragraph) answer to each question.

In your view, what is the role of the Advisory Panel?

Provide selected expertise (state-of-the-art and state-of-the-practice) as necessary to support STEP team in the Implementation phase.

What specific contributions do you think that you can make to STEP? For large projects:

- I can provide analysis of test phase effectiveness,
- error class distribution of errors detected during test phases, and error severity distribution by test phases.
- I can also provide insight to the usefulness of error seeding and cumulative software reliability models. I am currently evaluating software testing tools.

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

Technology insertion

In which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?

1. STARS tasks (Test Technology Baseline and Tool Qualification Procedures)
2. risk models

Are you interested in submitting a “good example” for consideration by the panel? If so, is the topic suitable for the ITEA good example meeting?

Possibly. I have in mind PLRS-II which was developed for the U.S. Army by Hughes-Fullerton. I'll have to update my error database with O&M results in order to determine whether PLRS-II is really a good example. Also need to see if both Hughes and the Army will release sufficient information.
Name: Carolyn Cannon

Please respond with a short (one paragraph) answer to each question.

In your view, what is the role of the Advisory Panel?

Provide a sounding board and direction to STEP staff at various stages of STEP task implementation. Also provide expertise in selected technical and policy-making areas.

What specific contributions do you think that you can make to STEP?

1. Obtain, summarize, and evaluate the Software Test Handbook from RADC.
2. If possible, obtain preliminary documentation on the Software Acquisition Manager's Guidebook - summarize and evaluate.

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

Technology insertion.

Which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting? Yes.

T&E guidebook.

Are you interested in submitting a "good example" for consideration by the panel? No.

So, is the topic suitable for the ITEA good example meeting?
Name: Ralph San Antonio

Please respond with a short (one paragraph) answer to each question.

In your view, what is the role of the Advisory Panel?

Evaluate findings, recommendations and products of STEP, in light of background and experience of individual members. Comment on acceptability and usability of STEP products in context of present work environment.

What specific contributions do you think that you can make to STEP?

Evaluate STEP products for acceptability by software acquisition managers. Assess consistency with other software initiatives (SEI, SPARK) and products such as MIL STD 881. Provide expertise in area of tool and measurement.

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

Background and expertise in all these areas.

In which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?

T&E model, good example, risk models, STARS tasks in order.

Are you interested in submitting a "good example" for consideration by the panel? If so, is the topic suitable for the ITEA good example meeting?

Interested, but need more details before I can say that I am able to. What are the restrictions — the good examples for a "successful" program?
In your view, what is the role of the Advisory Panel?

To serve as a source of information to the STEP project and to review STEP objectives & findings.

What specific contributions do you think that you can make to STEP?

- Thorough understanding of the policy & most DoD & service SW policy
- Extensive IV & V experience
- Can review policy & standards recommended by STEP; also can review guidebook(s).

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

Policy/procedures

In which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?

T&E guidebook

Are you interested in submitting a "good example" for consideration by the panel? If so, is the topic suitable for the ITEA good example meeting?

Possibly - a good IV & V plan.
Name: David Gelperin

Please respond with a short (one paragraph) answer to each question.

In your view, what is the role of the Advisory Panel?

1) To review & comment on the STEP program before & during execution
2) To advise particular projects within the program
3) To participate in particular projects

What specific contributions do you think that you can make to STEP?

1) review overall program
2) participate in demonstration project

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

- policy/procedures
- standards/regulations

In which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?

- develop unit tests for demo
- work on development of test "methodology"

Are you interested in submitting a "good example" for consideration by the panel? If so, is the topic suitable for the ITEA good example meeting?

- No
Name: Raymond J. Rubey

Please respond with a short (one paragraph) answer to each question.

In your view, what is the role of the Advisory Panel?

Question and evaluate STEP plans and results. Report on the expected reception of the results and recommendations by the "users" of STEP.

What specific contributions do you think that you can make to STEP?

Factor into the results the practical experience in embedded S/W development and test.

Which of the areas of the implementation plan (policy/procedures, technology insertion, standards/regulations) most suits your background and interests?

All

In which of the task areas (e.g., good examples, T&E guidebook, STARS tasks, risk models) are you interested in working? Are you willing to make a short (10-20 minute) presentation at the next panel meeting?

T&E Guidebook: no particular preference
Risk models: I am willing to make a presentation

Are you interested in submitting a "good example" for consideration by the panel? If so, is the topic suitable for the ITEA good example meeting?

* F-16 Avionics (WPAFB ASD)
* Yes
Subject: Consulting Agreement No.

Dear:

This letter is to serve as an agreement between the Georgia Institute of Technology (GIT) and for the furnishing of consulting services to GIT as noted herein. This agreement is to begin effective , and terminate on

The services covered by this agreement are as follows:

Compensation will be
, up to a maximum of

GIT's maximum liability hereunder is

Invoices may be submitted on a monthly basis, for work performed during the preceding month, to:

Georgia Institute of Technology, Engineering Experiment Station, Atlanta, Georgia 30332 upon the completion of the rendering of such services.

It is agreed that you will hold GIT harmless from any claims, demands, or actions arising out of the services performed by you under this Agreement. It is also agreed that you will not use the name of the Georgia Institute of Technology or any of its divisions or units, in any advertising or publicity matter, without the prior written consent of the Georgia Institute of Technology. The same shall apply to the Georgia Tech Research Institute. It is understood that this Agreement may be terminated at any time, either for the convenience of GIT or for the failure to perform properly hereunder.

All writings produced by you specifically under this Agreement shall be the property of GIT. Additionally, it is understood that by your acknowledgement below, you represent that there is no direct or indirect interest which you have which would conflict in any manner with the performance by you under this Agreement.
Your services hereunder are provided as an independent contractor to GI7. Consequently, you are responsible for all applicable federal and state regulations relating to income tax, social security, worker's compensation and unemployment insurance.

If the terms of this agreement are acceptable, please so indicate as noted below and return one copy of this Agreement to the undersigned.

Sincerely,

Michael Drew

Richard P. Dobb
OFFICE OF CONTRACT ADMINISTRATION

Receipt Acknowledged and Accepted

By ___________________________

Date __________________________

Soc. Sec. _______________________
STEP Technical Advisory Panel Meeting Report

By
R. A. DeMillo

Prepared for
Naval Air Development Center
Warminster, PA 18974

Under
Contract No. F33657-82-G-2083

April 1984

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
<table>
<thead>
<tr>
<th><strong>4. TITLE (and Subtitle)</strong></th>
<th>STEP Technical Advisory Panel Meeting Report</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>7. AUTHOR(s)</strong></td>
<td>R.A. DeMillo</td>
</tr>
<tr>
<td><strong>8. CONTRACT OR GRANT NUMBER(s)</strong></td>
<td>F33657-82-G-2083</td>
</tr>
<tr>
<td><strong>9. PERFORMING ORGANIZATION NAME AND ADDRESS</strong></td>
<td>School of Information and Computer Science, Georgia Institute of Technology, Atlanta, GA 30332</td>
</tr>
<tr>
<td><strong>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</strong></td>
<td>NADC, Warminster, PA 18974</td>
</tr>
<tr>
<td><strong>11. CONTROLLING OFFICE NAME AND ADDRESS</strong></td>
<td>NADC, Warminster, PA 18974</td>
</tr>
<tr>
<td><strong>12. REPORT DATE</strong></td>
<td>April, 1984</td>
</tr>
<tr>
<td><strong>13. NUMBER OF PAGES</strong></td>
<td></td>
</tr>
<tr>
<td><strong>15. SECURITY CLASS. (of this report)</strong></td>
<td>Unclassified</td>
</tr>
<tr>
<td><strong>16. DISTRIBUTION STATEMENT (of this Report)</strong></td>
<td>unlimited</td>
</tr>
<tr>
<td><strong>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</strong></td>
<td>unlimited</td>
</tr>
<tr>
<td><strong>19. KEY WORDS</strong></td>
<td></td>
</tr>
<tr>
<td><strong>20. ABSTRACT</strong></td>
<td></td>
</tr>
</tbody>
</table>
TABLE OF CONTENTS

Introduction .................................................. 2
Presentation of STEP Status Update ......................... 2

Overviews by Task Leaders .................................
  Guidebook/Good Examples ................................... 3
  Risk Models ................................................... 3
  STARS Initiative .............................................. 3

Attachment I, Agenda ......................................... 4

Attachment II, List of Participants ......................... 5

Attachment III, Task area Reports .........................
  Software Test and Evaluation Guidebook Report ............ 7
  Good Examples Report ....................................... 9
  Risk Models Task Group Report ............................ 9
  STARS Tasks - Task Group Report ......................... 11

Attachment IV, Presentation Material .....................
  Guidebook/Good Examples Plans ........................... 16
  Risk Models for Software Testing ........................ 18
  STARS Update ............................................... 26
  STARS Tasks ................................................ 32
On April 04-05, 1984, STEP held its second meeting of the Technical Advisory Panel at the Georgia Institute of Technology, Atlanta, Georgia. The objectives of this meeting were:

1) To review the overall status of the STEP project.

2) To report to the panel members progress in the four major task areas:
   a) Software Test and Evaluation Guidebook
   b) STEP Good Examples
   c) Risk Modelling/Decision Support Systems
   d) STARS Tasks

3) To engage in detailed discussions and solicit review comments from the panel members on the above topic areas.

The meeting consisted of the status presentation followed by an overview presentation of each topic area. The group then divided into separate task areas to enter into detailed discussion. At the conclusion on Thursday April 05, the task area discussion leaders presented reports of their results to the complete group.

This report contains a summary of the overall STEP status and each area's discussions/conclusions. Attachments include the agenda, the attendees list, the reports of each group, and copies of all slides used.

Presentation of STEP Status Update
(Presented by R. A. DeMillo)

Professor DeMillo presented a summary of STEP activities that have occurred since the last Technical Advisory Panel meeting. The presentation served as a review of STEP and an introduction to the topics to be discussed in detail during the remainder of the meeting. Items discussed that would not have any detailed discussion included the planning currently being initiated for the technology demonstration. The A7-E program at China Lake has been suggested as a vehicle to perform a technology demonstration. It will be an important part of STEP in that it would allow for the validation of the STEP recommendations prior to their formal implementation.
Overviews by Task Leaders

The Task Leaders presented an overview of their plans for the remainder of the meeting for the individual task areas.

Guidebook/Good Examples
(Presented by W. M. McCracken)

The Guidebook and Good Examples task areas were combined because of their interrelated themes and the common interest of the panel members. Since the last meeting the guidebook activity concentrated on the development of the checklist that is a part of the January 1984 report on Policy Recommendations for Software T&E. The advisory panel team was asked to review the checklist from their experiential perspective and produce a critique to be used to finalize the checklist. The team was also asked to review the list of good examples and to offer their ideas, including criteria for selection.

Risk Models
(Presented by R. A. Gagliano)

The group was presented with the work to date on risk models. The activity has concentrated on the application of game theoretic techniques to evaluating the cost versus completeness issue. Basically the model assures an opportunity for choice and allows the tester to select his test strategies versus available resources and assess the risk of his choices. The advisory panel team was charged with reviewing and critiquing the model.

STARS Initiative
(Presented by J. F. Passafiume)

The group was presented with the current status of STARS and its relation to STEP. The presentation covered the areas of STARS activities to be addressed by the task group: "Test Technology Baseline and Identification Procedures for SEI"; and "Provide Demonstration and Qualification Procedures for SEI".
Attachment I

AGENDA
STEP Technical Advisory Panel Meeting
April 4-5, 1984
Georgia Institute of Technology

<table>
<thead>
<tr>
<th>April 4 1984</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>8:30 - 9:00</td>
<td>Welcome and Opening Remarks</td>
</tr>
<tr>
<td>9:00 - 10:30</td>
<td>Presentations by STEP Task Leaders</td>
</tr>
<tr>
<td>10:30 - 10:45</td>
<td>Break</td>
</tr>
<tr>
<td>10:45 - 12:15</td>
<td>Group Discussions in Task Groups</td>
</tr>
<tr>
<td>12:15 - 1:30</td>
<td>Lunch</td>
</tr>
<tr>
<td>1:30 - 5:00</td>
<td>Group Discussions (Continued)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>April 5, 1984</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>9:00 - 10:45</td>
<td>Group Discussions (Continued)</td>
</tr>
<tr>
<td>10:45 - 11:00</td>
<td>Break</td>
</tr>
<tr>
<td>11:00 - 12:15</td>
<td>Group Reports</td>
</tr>
<tr>
<td>12:15 - 1:30</td>
<td>Lunch</td>
</tr>
<tr>
<td>1:30 - 2:30</td>
<td>Group Reports</td>
</tr>
<tr>
<td>2:30 - 3:00</td>
<td>Closing Remarks (DeMillo)</td>
</tr>
<tr>
<td>3:00</td>
<td>Adjourn</td>
</tr>
</tbody>
</table>
## Attachment II

### LIST OF PARTICIPANTS

<table>
<thead>
<tr>
<th>NAME</th>
<th>TASK AREA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Sinasi Bilsel</td>
<td>Risk Modelling/Decision Support Systems</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td>(404) 894-3180</td>
</tr>
<tr>
<td>Richard A. DeMillo</td>
<td>Risk Modelling/Decision Support Systems</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td>(404) 894-3180</td>
</tr>
<tr>
<td>Ross A. Cagliano</td>
<td>Risk Modelling/Decision Support Systems</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td>(404) 894-5130</td>
</tr>
<tr>
<td>Richard J. Lipton</td>
<td>Risk Modelling/Decision Support Systems</td>
</tr>
<tr>
<td>Dept. EE and Computer Science Princeton University Princeton, NJ 08540</td>
<td></td>
</tr>
<tr>
<td>Carolyn Gannon</td>
<td>Software Test and Evaluation Guidebook/Good Examples</td>
</tr>
<tr>
<td>General Research Corporation</td>
<td>5383 Hollister Avenue</td>
</tr>
<tr>
<td>P.O. Box 6770</td>
<td>Santa Barbara, CA 93111</td>
</tr>
<tr>
<td>W. Michael McCracken</td>
<td>Software Test and Evaluation Guidebook/Good Examples</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td>(404) 894-3180</td>
</tr>
<tr>
<td>Edward F. Miller, Jr.</td>
<td>Software Test and Evaluation Guidebook/Good Examples</td>
</tr>
<tr>
<td>Technical Director</td>
<td></td>
</tr>
<tr>
<td>Software Research Associates</td>
<td>P.O. Box 2432</td>
</tr>
<tr>
<td>San Francisco, CA 94126</td>
<td></td>
</tr>
<tr>
<td>Capt. W. P. Nelson</td>
<td>Software Test and Evaluation Guidebook/Good Examples</td>
</tr>
<tr>
<td>6575 School Squadron</td>
<td></td>
</tr>
<tr>
<td>Brooks AFB</td>
<td></td>
</tr>
<tr>
<td>Texas 78235</td>
<td></td>
</tr>
<tr>
<td>John Bowen</td>
<td>STARS</td>
</tr>
<tr>
<td>Software Engineering Division Hughes Aircraft Company Ground Systems Group</td>
<td></td>
</tr>
<tr>
<td>MS 618/P215, P.O. Box 3310</td>
<td></td>
</tr>
<tr>
<td>Fullerton, CA 92634</td>
<td></td>
</tr>
</tbody>
</table>
Matija Exel
Georgia Institute of Technology
(404) 894-3180

Jeff Offutt
Georgia Institute of Technology
(404) 894-3180

John F. Passafiume
Georgia Institute of Technology
(404) 894-3982

Fred Sayward
ITT
1000 Oronoque Lane
Stratford, Conn 06497

Fred Shocket
NADC
Westminster, PA 18974
Attachment III

TASK AREA REPORTS

SOFTWARE TEST AND EVALUATION GUIDEBOOK REPORT

Team Members:  C. Gannon
               M. McCracken
               E. Miller
               W. Nelson

The team reviewed the current checklist. The following pages are the recommendations of the team. The team first formulated what they felt were the eight (8) critical questions to be answered to allow the assessment of adequate and suitable testing. This list was then compared against the current checklist to determine how well it answered the eight questions (item). The team also reviewed the checklist for format and ease of use (items 1-4).

COMMENTS ON THE CURRENT CHECKLIST

1. Follow TEMP structure more closely, including the separation of DT&E and OT&E.

   Example:
   
   System
   a. Key Functions
      (questions)

2. Ouestions should be formulated so that the answers clearly reflect a strength or weakness of the TEMP.

   Example:
   
   "If the system implements critical functions, are they clearly identified in the TEMP?"

   rather than

   "Does the software implement critical functions?"

3. Consider using the same format (or model) as the DSARC Guidebook for Embedded Computer Resources.

4. Note: The TEMP is used in almost all systems (even less-than-major defense systems) by the services, as well as OSD.
5. The relationship between the eight items (see attached list) and the TEMP is as follows:

<table>
<thead>
<tr>
<th>Question</th>
<th>TEMP Cross Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Part II#1 (not currently covered)</td>
</tr>
<tr>
<td>2</td>
<td>Part VI (needs additional coverage)</td>
</tr>
<tr>
<td>3</td>
<td>Part III#2c (accountability procedures are not covered)</td>
</tr>
<tr>
<td>4</td>
<td>Part III#2c (needs to be more global. Also question achievability)</td>
</tr>
<tr>
<td>5</td>
<td>Part III#1 and Part IV (add preservation and machine readability of data)</td>
</tr>
<tr>
<td>6</td>
<td>Part II#1 (add agencies that are not directly involved in the tests)</td>
</tr>
<tr>
<td>7</td>
<td>Is a new item. Belongs in Section B2. Addresses the implementation of the TEMP.</td>
</tr>
<tr>
<td>8</td>
<td>Is a new item. Belongs in Part #1 and B3, between Milestones 2 and 3. Emphasizes continued update.</td>
</tr>
</tbody>
</table>

At each "level" (e.g., PAT&E, OT&E, DT&E) of test for a given program, the following information is necessary to judge the suitability and adequacy of the planned testing:

1. Has the appropriate agency been assigned the responsibility and does it have the capability of performing the test?

2. Will all necessary resources (time, funding, equipment, personnel) be available?

3. What accountability procedures and evaluation criteria are planned?

4. Is the methodology (general approach, tool base, techniques) suitable and achievable?

5. Will test results, including error data, be properly recorded and preserved (preferably in machine readable form)?

6. Will test information be disseminated to all appropriate test, development, supporting and using agencies?

7. Are the provisions of the TEMP properly supported by related program documents (e.g., the CRISP, MOA, ICSP, SEMP, etc.) and the contract?

8. Are there provisions for updating the TEMP itself as the program evolves?
GOOD EXAMPLES REPORT

Due to lack of time the Good Examples Area was not thoroughly reviewed and no specific recommendations were made.

RISK MODELS TASK GROUP REPORT

Task Group:  S. Bilsel
R. DeMillo
R. Gagliano
R. Lipton

Overview

The concept of risk modeling was discussed from two perspectives: the game-theoretic modeling approach that was presented at the November 1983 Technical Advisory Panel Meeting and in contrast to other standard methodologies, e.g., cost benefit analysis. The previous STEP efforts had sought to distinguish the application of risk modeling in software T&E from those in other domains.

Group Discussions

R. Lipton presided over the group discussions which focused mainly on the approach using a two-person game. Several questions were generated as indicated below:

1. What is unique about software T&E in this model?
2. Are the strategies strictly adaptive?
3. Are there means of partitioning the outcomes (nature)?
4. How are these terms differentiated: error, failure, undesirable event or behavior, etc.?
5. Are the utility functions unique?
6. How are they generated?
7. Are there other software parameters, such as performance, that may be of interest?
8. Is the use of Ruby's Conjecture proper?

9. What are the ramifications of the shape of the matrix?

10. How is an "else clause;" (i.e., one does not have to "play the game,"") interpreted in terms of SW?

Lipton summarized his experiences with such models as they have been applied to various other settings: investments, product development and oil field exploration. In the latter, a strong case was made for an analogy that modified the usual oil field exploration strategies such that: the event of no oil (errors) discovered implies another oil company (or many, signifying a great number of SW users) can buy the right to explore for oil at, say, one dollar.

There was not total agreement among the members of the task group on the following issues:

a. Ruby's Conjecture - its interpretation and applicability

b. The impact of the shape of the matrix and/or the sizes of \( m \) and \( n \) (the number of outcomes versus the number of strategies, respectively)

c. How to develop and utilize so-called regret functions, etc.

Agreement was reached on the following issues:

I. The game-theoretic risk modeling approach seems promising for software test and evaluation.

II. Software test and evaluation may, in fact, be a new and fertile area for risk assessment.

III. Much work needs to be done to construct the utility functions; an obvious way is to relate to costs.

IV. A symposium would be in order in the near future for Risk Assessment in Software Test and Evaluation.

V. Other examples would be extremely beneficial and useful.

Group Report

As task leader, Lipton presented via overhead projector many of the ideas indicated above (see attached). In particular, he developed the oil field exploration analogy and described the two substantive meta-issues: choosing to "play" and changing the matrix.
The STARS Tasks that we intended to address were as follows:

1. T&E Management Plan for Acquisition Panel.
   Define organizational structure for managing improvements in the testing process and implementing policy for software T&E.

2. T&E Management and Operating Plans for SEI.
   Define the role of SEI in supporting test tool development and in encouraging the integration of state-of-the-art test technologies into modern programming support environments.

3. Test Technology Baseline and Identification Procedures for SEI.
   Define an initial procedure for identifying test tools, environments, and methodologies for service qualification and inclusion into a warehouse to satisfy test plan requirements.

4. Provide Demonstration and Qualification Procedures for SEI.
   Provide a procedure for demonstrating and qualifying tools for DoD use.

5. Provide functional Requirements for APSE Test Environments.
   Offer support to STARS, SEI planners, and the Joint Service Team for Ada Environment definition in the formulation of requirements for candidate test environments.

After an initial discussion, John Passafiume established a priority order for addressing the tasks. Tasks were prioritized as follows: 3 - 4 - 2 - 1 - 5. The rationale for the priority ordering was the schedule for deliverables. The exception was the task for the Acquisition Panel. Here this task was placed fourth on the list due to the uncertainty with respect to the formation of the Acquisition Panel.

The Test Subenvironment task was placed last on the priority list as we are not quite ready to discuss the strawman that had been prepared.
by Dr. DeMillo and internally circulated within the STEP team. (More work is needed before this document can be released for outside review). Within the available time we were only able to address the first and second priority tasks and had just a brief discussion with no reportable conclusions with respect to the third. This and the remaining tasks should be addressed in detail at the next panel meeting. The following is related to the first two priority tasks:

Priority 1 - Test Technology Baseline and Identification Procedures for SEI.

Issues Considered:

* Lack of cooperation from tool development community.
* Influence of language, software size, application.
* No common definition for test tool (emulator, tool set, etc.).
* Application to software life cycle.
* What is the SEI role, pre-Ada? Will SEI support non-Ada tools?

Identification Procedure:

1. Inventory what sets of tools and methodologies exist and are available.
   a. If they are not used on DoD projects, then a trial test should be done before feeding into the evaluation task (SEI). The aim here is to reduce the risk of something that hasn't really been found useful before.
   b. If they are being used on DoD projects, then proceed to evaluation task (SEI).
   c. If the tool does not exist:
      case a: Manual procedure used on a DoD project.
      (1) Construct requirement document for tool.
      (2) Build or have built for SEI.
      case b: Tool or methodology not being used or widely recognized (feasibility study).

Additional suggestions:

* STEP should provide SEI with inventory of available tools.
* STEP should provide SEI with lessons learned from previous STEP activities.
* Recommended cover letter and survey questionnaire to include a section (optional) for detail on a particular tool or methodology should be provided to SEI.
Priority 2 - Provide Demonstration and Qualification Procedures for SEI.

Issues Considered:

* Derivation of testing goals from DoD standards (recommended changed).
* Determination of responsive tool types to goals.
* Who will take life cycle maintenance responsibility for tool?
* What are GFE Tools? Who maintains them?

Evaluation Procedure:

1. Classify the considered tool by functionality and lifecycle applicability.

Preconditions are:
   a. Define lifecycle and lifecycle applicability (use Software Development Standard as source).
   b. Define ideal and required testing environment DoD software (ala APSE).
   c. Define functionality classification.

2. Establish evaluation criteria for each tool type. The following is a list of criteria that we feel is important:

   a. Resource utilization (# of cycles, etc.). Need also to consider people intensity as cycles could be cheaper than people. Consider a manual walk-thru and the people cost involved versus a large number cruncher using cycles during background processing. The cost of the walk-thru could be considerably higher than the cost of using the large machine.
   b. Usability
   c. Portability
   d. Functional performance
   e. Documentation quality
   f. Installation restrictions
      (1) Hardware
      (2) Software
   g. Maintainability
   h. Extensibility
   i. Interoperability (within the Test Subenvironment or with other software)

3. Generate benchmark software. (Some "generic software may already be available.") It should have the following characteristics (not an exhaustive list):
a. Stress the program.
b. General performance.
c. Representative of the three Services.
d. The relationship between benchmark software and criteria to evaluate must be established.

(We considered the role of SEI and the amount of involvement that they would be able to have. It was not clear if SEI would be able to do this. We discussed the issue of SEI maintaining the library and tailoring the tools or benchmark software when needed for a particular program.)

Benchmarking procedure:

a. Define error criteria.
b. Seed errors in benchmark if applicable.
c. Use some error free benchmarks.

4. Generate evaluation procedure, using 2 and 3, above.

5. Run evaluation.

NOTE: We cannot automatically assume that SEI will be equipped and staffed to be able to do all of this work. A fall-back position is to ask the vendor to provide some of the information and to ask the service components to help in a particular area, e.g., language control facility provided by the US Air Force.

6. Record results.

7. For acceptable tools, distribute results and add to the inventory. (Acceptability criteria must be defined).
Attachment IV

PRESENTATION MATERIAL
GUIDEBOOK/GOOD EXAMPLES PLANS

* GUIDEBOOK

  * REVIEW TEMP CHECKLIST (APPENDIX A - TECH. REPORT)
  * NEED TO ADD SUPPORT SOFTWARE
  * GENERATE IDEAS FOR DOWN-STREAM CHECKLISTS
  * HOW DO I WRITE A TEMP TO RESPOND TO 5000.3? (SERVICES)
  * HOW DO I RESPOND TO A TEMP? (PROGRAM OFFICE)
  * GENERATE IDEAS FOR THE GUIDEBOOK

* GOOD EXAMPLES

  * REVIEW CANDIDATES / SELECTION CRITERIA
  * GENERATE LIST OF KEY QUESTIONS
  * PANEL'S GOOD EXAMPLES / CRITERIA
CHECKLIST

* SYSTEM FUNCTIONS, INTERFACES, AND UNIQUE CHARACTERISTICS
  * DETERMINE SOFTWARE CONTRIBUTION
  * DETERMINE DEGREE OF RISK RELATED TO SOFTWARE

* REQUIRED CHARACTERISTICS AND CRITICAL T & E ISSUES
  * RED LIGHT ITEMS

* PROGRAM SUMMARY
  * IS THE TEST PROGRAM BEING MANAGED

* T & E OUTLINE
  * TO DATE
    * HAS THE WORK BEEN DONE?
  * FUTURE
    * IS THE WORK TO BE DONE?

* SPECIAL RESOURCE SUMMARY
  * LIST OF KEY / CRITICAL SOFTWARE
  * LIST OF NECESSARY RESOURCES
Risk Models for Software Testing
\[ u_{ij} = \text{utility of } T_i \text{ and } S_j \]

Criterion
- maximin
- maximax
- minimax risk
- Hurwicz
- Insufficient reason
How is software different?

- Test Plans are often adaptive

- Many possible states $S_1, ..., S_m$
Where do we drill oil wells?
Where do we drill oil wells?
Where do we drill oil wells?

when do we adopt and conclude no oil?

what is the risk?
Where do we drill oil wells?

When do we adopt and conclude no oil?

What is the risk?

- Meta Issues
  - cannot choose do not play
  - can change the matrix
STARS UPDATE
GOALS: IMPROVE PRODUCTIVITY WHILE ACHIEVING GREATER SYSTEM RELIABILITY AND ADAPTABILITY (IN THE FACE OF INCREASINGLY DEMANDING REQUIREMENTS) THROUGH SOFTWARE DEVELOPMENT AND IN-SERVICE SUPPORT PROCESSES THAT ARE MORE RESPONSIVE, PREDICTABLE, AND COST-EFFECTIVE.
STARS ADDRESSES PEOPLE, TOOLS AND METHODS, AND THE NEED TO IMPROVE THE STATE OF PRACTICE, AND NOT JUST THE RESEARCH STATE OF THE ART.
CURRENT PLAN

* PROVIDES RAPID BUILD-UP IN EFFORT AND CHANGES IN PRACTICE

* INSTITUTIONALIZE THE PROCESS OF CONTINUING IMPROVEMENT

* TRANSFER EFFORTS TO DOD COMPONENTS
STARS FEATURES FOR RAPID CHANGE

* SPECIAL LEADERSHIP

* COORDINATION

* FUNDING

* LIMITED LIFE THRU FY90
STARS ELEMENTS AND FUNDING

* AUTOMATED SOFTWARE ENGINEERING ENVIRONMENT ($113M)

* MEASUREMENT ($27M)

* ACQUISITION AND PROJECT MANAGEMENT ($29M)

* APPLICATION-SPECIFIC EFFORTS ($37M)

* METHODOLOGY ($45M)

* HUMAN RESOURCES AND ENGINEERING ($42M)
STARS TASKS
T&E MANAGEMENT PLAN FOR ACQUISITION PANEL

DEFINE ORGANIZATIONAL STRUCTURE FOR MANAGING IMPROVEMENTS IN THE TESTING PROCESS AND IMPLEMENTING POLICY FOR SOFTWARE T&E
I&E MANAGEMENT AND OPERATING PLANS FOR SEI

DEFINE THE ROLE OF SEI IN SUPPORTING TEST TOOL DEVELOPMENT AND IN ENCOURAGING THE INTEGRATION OF STATE-OF-THE-ART TEST TECHNOLOGIES INTO MODERN PROGRAMMING SUPPORT ENVIRONMENTS
TEST TECHNOLOGY BASELINE AND IDENTIFICATION PROCEDURES FOR SEI

DEFINE AN INITIAL PROCEDURE FOR IDENTIFYING TEST TOOLS, ENVIRONMENTS, AND METHODOLOGIES FOR SERVICE QUALIFICATION AND INCLUSION INTO A WHAREHOUSE TO SATISFY TEST PLAN REQUIREMENTS
PROVIDE DEMONSTRATION AND QUALIFICATION PROCEDURES FOR SEI

DEFINE A PROCEDURE FOR DEMONSTRATING AND QUALIFYING TOOLS FOR DOD USE
PROVIDE FUNCTIONAL REQUIREMENTS FOR APSE TEST ENVIRONMENTS

OFFER SUPPORT TO STARS, SEI PLANNERS, AND THE JOINT SERVICE TEAM FOR ADA ENVIRONMENT DEFINITION IN THE FORMULATION OF REQUIREMENTS FOR CANDIDATE TEST ENVIRONMENTS
**REPORT DOCUMENTATION PAGE**

<table>
<thead>
<tr>
<th>1. REPORT NUMBER</th>
<th>2. GOVT ACCESSION NO.</th>
<th>3. RECIPIENT'S CATALOG NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>4. TITLE (and Subtitle)</th>
<th>5. TYPE OF REPORT &amp; PERIOD COVERED</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEP Military Advisory Panel Meeting Report</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>6. PERFORMING ORG. REPORT NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td>F33657-82-G-2083</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>7. AUTHOR(s)</th>
<th>8. CONTRACT OR GRANT NUMBER(s)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F33657-82-G-2083</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>9. PERFORMING ORGANIZATION NAME AND ADDRESS</th>
</tr>
</thead>
<tbody>
<tr>
<td>School of Information and Computer Science</td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
</tr>
<tr>
<td>Atlanta, GA 30332</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>11. CONTROLLING OFFICE NAME AND ADDRESS</th>
<th>12. REPORT DATE</th>
</tr>
</thead>
<tbody>
<tr>
<td>NADC</td>
<td>June, 1984</td>
</tr>
<tr>
<td>Warminster, PA 18974</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>13. NUMBER OF PAGES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>14. MONITORING AGENCY NAME &amp; ADDRESS (if different from Controlling Office)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>15. SECURITY CLASS. (of this report)</th>
<th>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE</th>
</tr>
</thead>
<tbody>
<tr>
<td>unclassified</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>16. DISTRIBUTION STATEMENT (of this Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>unlimited</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</th>
</tr>
</thead>
<tbody>
<tr>
<td>unlimited</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>18. SUPPLEMENTARY NOTES</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
</tbody>
</table>
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>2</td>
</tr>
<tr>
<td>Attachment I, Agenda</td>
<td>3</td>
</tr>
<tr>
<td>Attachment II, List of Participants</td>
<td>4</td>
</tr>
<tr>
<td>Attachment III, T&amp;E Management Plan</td>
<td>5</td>
</tr>
<tr>
<td>Attachment IV, Good Examples.</td>
<td>21</td>
</tr>
<tr>
<td>Attachment V, CMS.2 Testing Tools Baseline</td>
<td>30</td>
</tr>
</tbody>
</table>
The Military Advisory Panel of the Software Test and Evaluation Project (STEP) met on 18 May, 1984. The purpose of the panel is to give STEP the advantage of the experience and perspective of the military mission critical computer resource personnel in reviewing progress to date and planned activities. The panel is composed of representatives of the Service's Development Commands and representatives of the Service's Operational Test and Evaluation Agencies. A list of attendees is attached.

The meeting was convened by Don Greenlee, Deputy Director for Defense Test and Evaluation.

Dr. DeMillo (the project director) presented the agenda for the day and introduced STEP to the panel. The agenda is attached. Following Dr. DeMillo's presentation, Mr. McCracken gave a detailed review of STEP, its progress to date, and the plans for STEP for the next two years. Three of the task areas of STEP (Good Examples, T & E Management Plan, and Baseline) were then described and discussed with the panel. The presentation material and notes of the three areas are attached.
ATTACHMENT I

Military Advisory Panel

First Meeting: 18 May 1984

Agenda

0830 - 0845  Welcoming Remarks  
D. R. Greenlee  
DDT&E, Strategic, Naval and C3 Systems

0845 - 0900  Introduction to Software Test and Evaluation  
Project and Charter of Advisory Panel  
R. A. DeMillo  
Georgia Tech, STEP

0900 - 1000  STEP Progress and Plans  
W. M. McCracken  
Georgia Tech, STEP

1000 - 1015  Break

1015 - 1045  T&E Management Plan for Acquisition Panel  
J. F. Passafiume  
Georgia Tech, STEP

1045 - 1115  Good Examples of Software T&E  
R. A. Gagliano  
Georgia Tech, STEP

1115 - 1145  Testing Tools: Technology Baseline  
R. A. DeMillo

1145 - 1215  Open Discussion  
Group

1215 - 1230  Closing Remarks and Summary of Action Items  
R. A. DeMillo

1230  Adjourn
## ATTACHMENT II

### Participants

<table>
<thead>
<tr>
<th>Name</th>
<th>Organization</th>
<th>Phone</th>
</tr>
</thead>
<tbody>
<tr>
<td>Don Greenlee</td>
<td>OUSDRE (T&amp;E)</td>
<td>(202) 695-7175</td>
</tr>
<tr>
<td>LCDR Dennis E. Ray</td>
<td>COMOPTEVFOR Norfolk VA</td>
<td>(804) 444-5026</td>
</tr>
<tr>
<td>Stephen French</td>
<td>USA QTEA</td>
<td>(202) 756-2164</td>
</tr>
<tr>
<td>Jim Baca</td>
<td>AFOTEC</td>
<td>(505) 844-9421</td>
</tr>
<tr>
<td>CAPT Dave Boslaugh</td>
<td>NAVMAT O8Y</td>
<td>(202) 692-3966</td>
</tr>
<tr>
<td>COL Harold Archibald</td>
<td>HQ DARCOM</td>
<td>(202) 274-9310/11/14</td>
</tr>
<tr>
<td>Owen McOmber</td>
<td>HQ NAVMAT (MAT O8Y)</td>
<td>(202) 692-3966</td>
</tr>
<tr>
<td>John Passafiume</td>
<td>Georgia Tech</td>
<td>(404) 894-3982</td>
</tr>
<tr>
<td>Pete Fonash</td>
<td>HQ DARCOM</td>
<td>(202) 274-9318</td>
</tr>
<tr>
<td>Richard DeMillo</td>
<td>GA Tech - STEP</td>
<td>(404) 894-3180</td>
</tr>
<tr>
<td>John Bolino</td>
<td>ODOT&amp;E</td>
<td>(202) 697-3898</td>
</tr>
<tr>
<td>Michael McCracken</td>
<td>GA Tech - STEP</td>
<td>(404) 894-3192</td>
</tr>
<tr>
<td>Ross Gagliano</td>
<td>GA Tech - STEP</td>
<td>(404) 894-5130</td>
</tr>
<tr>
<td>COL Ken Nidiffer</td>
<td>Hq AFSC</td>
<td>(301) 981-5731</td>
</tr>
<tr>
<td>R. J. Martin</td>
<td>GA Tech - STEP</td>
<td>(404) 894-3738</td>
</tr>
</tbody>
</table>
ATTACHMENT III

T&E MANAGEMENT PLAN
John Passafiume led a discussion with the Panel Members using the preliminary draft of the management plan as a strawman. He noted that the STARS Program Office has developed a Draft Charter for the Proposed Acquisition Panel and that the Charter calls for a standing subgroup to address Test and Evaluation issues. The panel members indicated that the draft charter had not been provided to them for review. We expressed some concern since the panel members would more than likely be the service representatives on the proposed Acquisition Panel. We agreed that STEP would approach the STARS Program Office to attempt to obtain a copy of the draft charter. Although the Panel Members did not fully agree with the list of issues and proposed solutions, a lively and profitable discussion provided us with further insights into the problem of institutionalizing STEP. Mr. Greenlee expressed the opinion that we seemed to be limiting ourselves to organizational solutions and expressed the opinion that some of the proposed solutions, although worthy, would be difficult to realize due to the political realities. We all agreed that training and education was a key ingredient and decided to ask a representative from DSMC to attend our next meeting. One of the panelists pointed out that DSMC training would only reach part of the target audience so that we ought to also be concerned with a review of training at the Service level. The military advisory panel members were asked to review the strawman issues and solutions and be prepared to provide a "service position" at the next meeting of the panel.
STRATEGY

- PROBLEM STATEMENT
- LIST OF ISSUES
- PROPOSED SOLUTIONS
- SOLICIT SERVICE POSITIONS
The Problem

Who will be the permanent advocate for SW T&E within DOD?
ISSUES

- REVIEW DODD 5000.3
- REVIEW IMPLEMENTING REGS
- INTERFACE WITH SEI?
- GFE ENVIRONMENTS
- DOD ADVOCACY FOR SW T&E
- INDUSTRY ADVOCACY FOR SW T&E
- GOVT VS IEEE/OTHER INDUSTRY STANDARDS FOR SW T&E
ISSUES (CONTINUED)

- RELATIONSHIP BETWEEN DDT&E AND DOT&E W/RESP TO SW T&E

- OTHER OSD AGENCIES/ACTIVITIES WITH T&E RESPONSIBILITY
**SOLUTIONS**

- Form JLC Subcommittee on SW Testing

- Acquisition Panel provide emphasis on T&E

- Deputy to DDT&E for SW T&E

- Deputy to DOT&E for SW T&E

- Industry Organization(s) establish standing SW T&E panels

- SEI will have responsibility for oversight of SW T&E
SOLUTIONS (CONTINUED)

- DSMC CURRICULUM REVIEW REQUIRED

- SERVICES & OTHER DOD COMPONENTS APPOINT FOCAL POINT FOR SWT&E
MANGEMENT PLAN FOR ACQUISITION PANEL

Objective:

To develop a T&E management plan for the Acquisition Panel, or some other existing or proposed tri-service organization. (See Appendix A for background).

Basic Strategy:

Utilize the Military Advisory Panel by presenting them with a statement of the problem, a List of issues, and proposed alternative means of providing for a management oversight over software testing in the future. The members of the panel will be requested to provide a Service position on the issues and the proposed solutions.

Statement of The Problem:

The STEP Phases I and II uncovered several deficiencies in the conduct of SW Test and Evaluation. These deficiencies were addressed by the STEP Phases I and II recommendations. The recommendations addressed changes needed in regulations and guidance, programmatic changes, state-of-the art improvements, test planning, the testing process, test evaluation, and the need for tri-service standards. The current STEP phase is addressing some of these recommendations (e.g., The T&E Model, T&E Guidebook, APSE Test Environment), but the momentum created by STEP must be continued after STEP finishes. There is a need for a permanent agency or activity such as the proposed Acquisition Panel to be tasked to continue to address the issues highlighted by the STEP recommendations. In addition, there are other issues that also need to be addressed which were not included in the above recommendations.
Issues Related to SW T&E:

1. Need to review DoDD 5000.3 and recommend changes to DDT&E.

2. Need to review implementing regulations and insure changes are made.

3. Interface with SEI - how to set this up?

4. Implications of GFE environment requirements.


8. Relationship between DOT&E and DDT&E with respect to SW T&E.

9. Other OSD agencies/activities with T&E responsibility.
Proposed Solutions:

1. Formation of a JLC Sub-committee on SW Testing.

2. Acquisition Panel will provide emphasis on T&E.
   * Formation of a working group for SW T&E.

3. DDT&E will have a Deputy for SW T&E.

4. DOT&E will have a Deputy for SW T&E.

5. Industry Organization will establish standing SW T&E Panels.

6. SEI will have responsibility for oversight of SW T&E.

7. DSMC Curriculum Review required.

8. Services and other DoD Components appoint focal point for SW T&E.
<table>
<thead>
<tr>
<th><strong>ISSUES</strong></th>
<th><strong>SOLUTIONS</strong></th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Need to review DODD 5000.3 &amp; recommend changes to DOTEE</td>
<td>1. Form JLC Subcommittee on SW Testing</td>
</tr>
<tr>
<td>2. Need to review implementing regs &amp; insure changes are made</td>
<td>2. Acquisition Panel provides emphasis on SW T&amp;E</td>
</tr>
<tr>
<td>3. DOTEE will have a deputy for SW T&amp;E</td>
<td>3. DOTEE will have a deputy for SW T&amp;E</td>
</tr>
<tr>
<td>4. DOTEE will have a deputy for SW T&amp;E</td>
<td>4. DOTEE will have a deputy for SW T&amp;E</td>
</tr>
<tr>
<td>5. Industry Organization will establish standing SW T&amp;E Panels</td>
<td>5. Industry organization will establish standing SW T&amp;E Panels</td>
</tr>
<tr>
<td>6. SEI will have responsibility for oversight of SW T&amp;E</td>
<td>6. SEI will have responsibility for oversight of SW T&amp;E</td>
</tr>
<tr>
<td>7. DSMC curriculum review required</td>
<td>7. DSMC Curriculum review required</td>
</tr>
<tr>
<td>8. Services/Other Components appoint focal point for SW T&amp;E</td>
<td>8. Services/Other Components appoint focal point for SW T&amp;E</td>
</tr>
</tbody>
</table>
APPENDIX A - Acquisition Panel Background

The Acquisition Panel will be established by the Undersecretary of Defense for Research and Engineering (USDR&E) and report to the part of OSD responsible for overall management of mission-critical computer resources within the DoD such as the Defense Computer Resources Board. The panel will work with and be funded by the STARS program office and will be responsible for managing all STARS acquisition-related activities. It will represent DoD elements significantly impacted by or having significant impact on defense software acquisition. The panel will recommend appropriate acquisition policies, contract incentive mechanisms, and related guidelines to encourage contractor participation in defense software efforts; encourage use of modern software practices that decrease life-cycle costs; and encourage developments of reusable software components. Other objectives include identifying means to strengthen the review process of the defense system acquisition review council; recommending changes to the Defense Acquisition Regulation so that it supports the goals and objectives of the software initiative; and recommending improvements to defense software acquisition procedures, processes and related directives. The panel will forward recommendations for specific improvements to the appropriate organization after coordinating with the OSD.

The panel will include representatives from each service's embedded computer systems software acquisition organization; a STARS project office member; a member from the OUSDR&E responsible for command, control, communications and intelligence (C31); and a member from the ODDT&E; a technology specialist approved by OUSDR&E (Research and Advanced Technology) and a productivity specialist approved by the OUSDR&E (acquisition management).

The panel must communicate with representatives of government, academia, and user activities that are affected by software acquisition. Communication will be achieved through working groups, which will review acquisition procedures and practices and recommend improvements. Some working groups will include representatives from the DSMC, private organizations, and industry associations. Specialized working groups will be established to address concerns and requirements of the software end user and those involved in the acquisition and of mission-critical computer resources.

The panel will maintain liaison with the Joint Logistic Commanders, the Defense Acquisition Regulation Council, the Federal Acquisition Regulation Council, and other appropriate offices within the OSD. The service representatives on the panel also will be liaison with the Joint Logistics Commanders/Joint Policy Coordinating Group for Computer Resources Management.
The Panel will provide a representative to participate in DAR/FAR council and subcommittee deliberations in matters affecting defense software acquisitions. Of particular interest is DoD rights in data and computer software acquisition. Proposed changes to the DAR will be coordinated with the acquisition panel.

The second task in the pre-award area is to assess the impact of new technology on systems acquisition. Emerging software technologies will be analyzed to determine the rewards and risks for acquisition programs considering the use of these technologies. New technology will be assessed in terms of shorter development period, increased programmer productivity, improved software quality and reliability, lower lifecycle cost, enhanced software supportability and adaptability, startup/front-end cost, and support availability.

Emerging technologies that are candidates for this research include:

- Ada/Ada Environments
- VHSIC and near VHSIC Products
- Artificial intelligence and associated languages
- New computer architectures
- The Navy's software cost reduction program
- Program development language and associated support
- Requirements tracking tools such as SREM and PSL/PSA

The result of this effort will be a continually updated compilation of rewards and risks, which will be provided to government and industry for use in future acquisition/development strategies and system design trade-off studies.
Figure 2. Components of the software acquisition panel.
APPENDIX B - Bibliography

Department of Defense, Software Technology for Adaptable, Reliable Systems (STARS) Program Strategy, 1 April 1983, Department of Defense, Washington, DC


STEP Phases I and II Final Report

Attachment IV

GOOD EXAMPLES
STEP

GOOD EXAMPLES
STEP SOW

Task 7

(1) ...gather objective evidence which indicates the effectiveness of systematic program testing...

...at least two existing systems will be identified for evaluation...

...selected by the degree to which the software has been subjected to systematic testing techniques...
GUIDEBOOK

VALIDATE
STEP
Recommendations

GOOD

EXAMPLES

THE MODEL

OTHER
Recommendations

RESEARCH GAPS
CRITERIA FOR SYSTEMS SELECTION

1. Deployed
2. Software T&E Success
3. Information Access
4. Generation/Release of New Information
5. Mix of Systems
   a. Three Services
   b. Wide spectrum of systems
   c. Different types of Software
FINAL SET OF SYSTEMS

NAVY

Aegis
LMAPS

AIR FORCE

AWACS
Pavepaws

ARMY

Firefinder/REMBASS
(AN/TPQ-36, AN/TPQ-37)

Missileminder
(AN/TSQ-73)
"Good Examples" of

Indications of early SW test planning
Special resources for SW test tools
Independent Verification & Validation
Systematic Unit Testing
Consistent Regression Testing
Test data-bases
Automated support for risk assessment
Evidence in Documents

TEMP

Dev/Test Plans

STDs/Regs

Procedures

Test Results

Contracts

Management

Plans, Structure, Interfaces

Schedule

Methodology

Technology/Tools
Military Advisory Panel

- Information Sharing
- Solicitation of Assistance
- Imminent Contacts
- PMO Focus
- Access to Personnel/Documents
Attachment V

CMS.2 Testing Tools Baseline

Goal: Define the software T&E technology (e.g., tools) baseline for CMS.2 based systems. The technology baseline will consist of those automated T&E methodologies that are mature, demonstrated, and can be inserted directly into practice on current planned Navy programs. As in previous STEP efforts, a software testing tool is defined to be support software that directly implements or aids in the implementation of a systematic software test methodology. Thus, for example, special-purpose simulators are not considered to be part of the baseline.

Progress to Date: So far, only the following two technologies have been identified as candidates for inclusion in the baseline.

Facility for Automated Software Production (NADC): A total software development and support facility which includes a number of software T&E tools.

Systematic T&E (FCDSSA-San Diego): This tool is of unknown functionality because promised documentation has not been sent.

Navy Facilities Contacted: The following sites were contacted in an attempt to define the baseline:

FCDSSA-Dam Neck
FCDSSA-San Diego
MAT 08Y
NADC
NAVAIR
NAVELEX
NAVSEA
NOSC
NSWC
NTDS
NUSC
NWC
ONR
Point Heueneme
Issues: The major unresolved issues involve the completeness and accuracy of the CMS.2 baseline and the procedures for distributing baseline status reports to Navy/OSD organizations.

1. Are there tools currently in use at the Navy sites listed above that have not been identified to STEP?

2. Are there other Navy sites that have tools that should be included in the baseline?

3. Are there tools (other than proprietary tools that are not available externally) in contractors' software environment that can be included in the baseline?

4. Are the two tools listed above appropriate for inclusion in the baseline?

5. What should be Navy reaction to the baseline?

6. How can STEP distribute baseline status reports to maximize the chances of developing new T&E technology?

General Baseline Construction: The techniques used to construct the CMS.2 baseline are currently being applied to Jovial, Fortran, Ada and other DoD languages. The utilization of the CMS.2 baseline within DDT&E, DOT&E, DUSD(R&A)/STARS/SEI, and the Navy will be a starting point for the more general effort.

Georgia Tech Contact

Mr. Jeff Offutt
The Software Test and Evaluation Project
School of Information and Computer Science
Georgia Institute of Technology
Atlanta, Georgia 30332
Phone: 404-894-3180
Net: rad.gatech @ csnet-relay
STEP Technical Advisory Panel Meeting Report

by
J. F. Passafiume

Prepared for
Naval Air Development Center
Warminster, PA 18974

Under
Contract No. F33657-82-G-2083

January 1985

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
<table>
<thead>
<tr>
<th>REPORT NUMBER</th>
<th>GOVT ACCESSION NO.</th>
<th>RECIPIENT'S CATALOG NUMBER</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TITLE (and Subtitle)</th>
<th>STEP Technical Advisory Panel Meeting Report</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>AUTHOR(s)</th>
<th>J.F. Passafiume</th>
</tr>
</thead>
</table>

| PERFORMING ORGANIZATION NAME AND ADDRESS | Georgia Institute of Technology  
School of Information and Computer Science  
Atlanta, Georgia 30332 |
|------------------------------------------|-------------------------------------------------|

<table>
<thead>
<tr>
<th>CONTRACT OR GRANT NUMBER(s)</th>
<th>F33657-82-G-2083</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>PERFORMING ORGANIZATION REPORT NUMBER</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>REPORT DATE</th>
<th>January 1985</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>NUMBER OF PAGES</th>
<th>1</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>DISTRIBUTION STATEMENT (of this Report)</th>
<th>unlimited</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</th>
<th>unlimited</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>SUPPLEMENTARY NOTES</th>
<th></th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>KEY WORDS</th>
<th>(Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>ABSTRACT</th>
<th>(Continue on reverse side if necessary and identify by block number)</th>
</tr>
</thead>
</table>
TABLE OF CONTENTS

Introduction ........................................... 1

Overviews by Task Leaders ............................. 2

  Software T&E Guidebook ............................... 2
  Software T&E Model .................................. 2
  Software Test Environment ............................ 2

Attachment I, Agenda .................................. 3

Attachment II, List of Participants .................... 4

Attachment III, Task Area Discussions ................. 5

  Software T&E Guidebook ............................... 6
  Software T&E Model .................................. 8
  Software Test Environment ............................ 9

Attachment IV, Presentation Material .................. 16

  Software T&E Guidebook ............................... 17
  Software T&E Model .................................. 23
On October 24-25, 1984, the Software Test and Evaluation Project (STEP) held the third meeting of its Technical Advisory Panel (TAP) at the Georgia Institute of Technology, Atlanta, Georgia. The objectives of the meeting were:

1) To report to Panel members on progress in three major task areas:
   a) Software Test and Evaluation (T&E) Guidebook
   b) Software T&E Model
   c) Software Test Environment

2) To engage in detailed discussions and solicit review comments from panel members on the above topic areas.

The meeting was divided into three sessions devoted to the above topic areas. After a presentation on each area, discussion leaders solicited inputs/comments from the group concerning refinements of the ideas presented and approaches for future work. At the conclusion of each segment, the designated discussion leader presented a wrap-up summarizing the main issues discussed.

This report contains a summary of each of the area's discussion/conclusions. Attachments include the agenda, the attendees list, the presentations made to each group and any conclusions reached, as well as copies of all presentation material.
OVERVIEWS

Software T&E Guidebook Session  
(Presented by R. J. Martin)

Ms. Martin gave a brief review of the motivation for this task, an outline of the environment within which the guidebooks will exist, and a short discussion of the current view of the ultimate contents of these guidebooks. The principal purpose of this segment was to determine the role of the various organizations with respect to test documents, as well as the form and general content of the set of guidebooks.

Software T&E Model Session  
(Presented by W. M. McCracken)

Mr. McCracken initiated discussion by presenting the model of the total system acquisition lifecycle and how the software lifecycle related to it. After the introduction the presentation centered on the various alternative software life cycles that could occur due to different program implementation strategies. These strategies include: off-the-shelf hardware, off-the-shelf software, and evolutionary acquisition.

Software Test Environment Session  
(Presented by R. A. DeMillo)

Dr. DeMillo presented an outline of a STEP draft report entitled "Requirements for a Test Subenvironment of an Advanced Software Engineering Environment" to the panel members. After the introduction of this report, the group was divided into two sections with the purpose of evaluating the requirements. One group was charged with dealing with the requirements from the perspective of persons with management functions while the second group represented the perspective of the technical user.
Wednesday, October 24

0830-0900  Coffee and Welcoming Remarks
0900-1030  Software Test and Evaluation Guidebook
            Discussion Leader:  Ronnie Martin
1030-1045  Break
1045-1200  Continuation of Guidebook Discussion
1200-1330  Lunch
            Guthridge Lounge
            Faculty Alumni House (Hosted by STEP)
1330-1500  Software T&E Model
            Discussion Leader:  Mike McCracken
1500-1515  Break
1515-1700  Continuation of Software T&E Model Discussion
1700-1730  Wrap-up
            Discussion Leader:  Rich DeMillo
1900       Dinner
            Pittypat's Porch

Thursday, October 25

0845-0900  Reconvene
0900-1030  Software Test Environment Discussion
            Discussion Leaders:  Rich DeMillo, Ronnie Martin
            Mike McCracken, John Passafiume
1030-1045  Break
1045-1215  Continuation of Software Test Environment Discussion
1215-1230  Wrap-up and Adjourn
Attachment II

LIST OF PARTICIPANTS

Bilsel, Sinasi
Georgia Institute of Technology
(404) 894-6266

Miller, Edward F., Jr.
Technical Director
Software Research Associates
P.O. Box 2432
San Francisco, CA 94126

Bowen, John
Software Engineering Division
Hughes Aircraft Company
Ground Systems Group
MS 618/P215, P.O. Box 3310
Fullerton, CA 92634

Offutt, Jeff
Georgia Institute of Technology
(404) 894-6266

DeMillo, Richard A.
Georgia Institute of Technology
(404) 894-3180

Passafiume, John F., Col.(Ret.)
Georgia Institute of Technology
(404) 894-3982

Gannon, Carolyn
4701 Patrick Henry Dr., #301
Santa Clara, CA 95050

Quindry, Thomas L.
DoD PESO - XC
Defense Logistics Agency
Cameron Station
Alexandria, VA 22304

Gelperin, David
Software Quality Engineering
4825 Olson Memorial Highway
Suite 240
Golden Valley, MN 55422

Rubey, Raymond J.
SofTech
3100 Presidential Drive
Fairborn, OH 45324

Harder, Rita
Georgia Institute of Technology
(404) 894-6266

San Antonio, Ralph
Dynamics Research Corporation
7 Lopez Road
Wilmington, MA 01887

Martin, Rhonda J.
Georgia Institute of Technology
(404) 894-3837

Stewart, Marilyn
LOGICON
1850 N. Lynn St., Suite 900
Arlington, VA 22209

McCracken, W. Michael
Georgia Institute of Technology
(404) 894-6172

Wegner, Peter, Ph.D.
Department of Computer Science
Brown University
Box 1910
Providence, R.I. 02912
Attachment III

TASK AREA DISCUSSIONS
Session 1: Software Test and Evaluation Guidebook

Discussion Leader: R. J. Martin

This session began with a brief review of the motivation for the task, an outline of the environment within which the guidebooks will exist, a brief discussion of the current view of the ultimate contents of the guidebooks, and a short presentation of progress to date in this area. Each of the panel members was mailed a copy of the "Software Test and Evaluation Guidebook, Part I: The Test and Evaluation Master Plan" prior to the meeting. Therefore, in general, the panel members were familiar with the approach being taken with respect to this task. The background remarks were followed by a brief introduction of the session topics and expectations. Handouts were distributed with the goal of focusing the efforts of the session and providing panel members with a common basis for purposes of the discussion (see Attachment IV).

The principal topic of the session discussion was that of the Role of the Various Organizations with respect to Test Documents. The questions asked and summaries of the ensuing discussions follow.

1. Is it possible to consider a generic form of each document type as the basis for the generation of a guidebook which would be useful when applied to all such test documents?

Since the panel was not aware of the details of the existing range of test documents, they were not comfortable with the idea of allowing generic documents to represent the specifics for purposes of guidebook generation. It was suggested that, prior to deciding this issue, the STEP team investigate the specific documents which would be represented by each generic document type.

2. Are the descriptions of the generic documents appropriate?

Unable to comment at this time. See discussion above.

3. Can the matrix be collapsed (i.e., do some organizations have similar responsibilities with respect to testing)?

The panel recommended that the rows of the matrix be reduced as indicated in the figure below (see Question 5). The additional columns represent the panel's wariness of generic document types.

4. Is it possible to assign primary responsibility for generation of the generic document types?

Unable to comment at this time. See discussion above.
5. **What should the elements of the matrix be?**

See figure below. Time only allowed discussion of the row representing the Program Manager.

<table>
<thead>
<tr>
<th></th>
<th>TEMP</th>
<th>Test Reqmt's</th>
<th>Test Plan</th>
<th>Test Des/Pro</th>
<th>Test Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSD/DDT&amp;E</td>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prog Mgr</td>
<td>G</td>
<td>?</td>
<td>A</td>
<td>A</td>
<td>?</td>
</tr>
<tr>
<td>OT&amp;E'ers</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key:     G = generate     R = review     I = input     A = approve
Session 2: Software T&E Model

Discussion Leader: W. M. McCracken

Purpose:

The T&E Model session of the TAP was concerned with determining the relationship of different software life cycles to the overall system acquisition life cycle.

Basis of Discussion:

It has been conjectured that different software life cycles may in fact occur in a major system acquisition as a function of the activities associated with the acquisition. The TAP was presented with initial models of the different types of software life cycles that could occur and the rationale for the models. The panel was asked to review the models and comment based on their experience and understanding of the system acquisition process and the software life cycle.

Discussion:

A presentation was made to first review the major system acquisition process and its associated life cycle. After the review, a presentation was made of the standard software life cycle and its relationship to the acquisition life cycle. Discussion by the TAP at this point centered on understanding the information presented and relating it to the members' experiences.

The discussion was then turned to examining other software life cycles that could occur when a "nonstandard" acquisition process was used. The types of nonstandard processes included: the use of off-the-shelf hardware, the use of off-the-shelf software and the use of evolutionary acquisition. The presentation then discussed the different models of the software life cycles and the ramifications of the different acquisition processes on the software life cycles.

Several questions were raised during the discussion. The questions can be grouped into a single statement related to whether or not the nonstandard acquisition processes in fact impact the software life cycle. No general consensus was reached.

Conclusions:

The TAP did not offer any specific conclusions for this topic. They did suggest many improvements of the models and areas for further investigation by the STEP team. Further suggestions were centered around the question grouping raised in the discussion session: "Does the software life cycle actually get impacted when different acquisition processes are occurring in a major system acquisition?"
Introduction

An outline of the STEP draft report "Requirements for a Test Subenvironment of an Advanced Software Engineering Environment" was presented to the TAP. The major topics covered in this outline are listed below.

1. Concept of a Software Engineering Environment (SEE)
   a. Integrated Tools, Methodologies, Databases
   b. Lifecycle Based (as opposed to code-test)
   c. Evolution toward Automated Environment

2. Concept of a Test Subenvironment (TSE)
   a. Support Test Community Within SEE
   b. Minimize Impact on Non-testers
   c. Minimize Architectural Constraints on overall SEE

3. User Capabilities
   a. The Test Community (Note: The codes T, M, B denote functions that are primarily Technical, Managerial, or Both)

<table>
<thead>
<tr>
<th>Contractor</th>
<th>Government</th>
</tr>
</thead>
<tbody>
<tr>
<td>T Software Engineer</td>
<td>X</td>
</tr>
<tr>
<td>T Tester</td>
<td>X</td>
</tr>
<tr>
<td>T IV&amp;V'er</td>
<td>X</td>
</tr>
<tr>
<td>B QA'er</td>
<td>X</td>
</tr>
<tr>
<td>B CM'er</td>
<td>X</td>
</tr>
<tr>
<td>M Functional Mgr.</td>
<td>X</td>
</tr>
<tr>
<td>M Project Mgr.</td>
<td>X</td>
</tr>
<tr>
<td>T Operational Tester</td>
<td>X</td>
</tr>
<tr>
<td>M DDT&amp;E</td>
<td>X</td>
</tr>
<tr>
<td>B Maintainers</td>
<td>X</td>
</tr>
<tr>
<td>T End Users</td>
<td>X</td>
</tr>
<tr>
<td>T Measurers</td>
<td>X</td>
</tr>
</tbody>
</table>
b. Capability Attributes

** cost
** feasibility
** technical approach
** criticality
** development cost
** application dependency
** user demands

Analysis of Capabilities

The TAP was divided into two groups. One group was to review and evaluate the requirements that related primarily to users who executed a management function. The other group dealt primarily with technical users. The groups were charged with developing capabilities lists that supported the common user scenarios for each of the categories of users.

Management Functions

The analysis of user requirements for each category of management-oriented user yielded the following:

OT&E
- realistic environment
- tie SEE to real environment
- reproduce operational errors in SEE
- instrumentation of target

DDT&E
- too high level
- summary of other information
- test requirements matrix...

Maintainers
- change analysis, regression testing, all of above... histories

End Users
- diagnostics, simulators, data collectors

Measurers
- measurement tools... data capturing, monitoring and reporting

Software Engineer
- requirements, design, correct way to ensure tests by others
- error location

Tester
- stimulate and observe behavior of software to identify errors or confirm correct behavior and report results

IV&V'ers
- capability to independently see that product satisfies user needs
QA'ers
- product complies with contractual obligations
- proper standards set and followed

CM'er
- identify products and assess changes
- control and manage multiple baselines
- status accounting

Functional Manager
- statistics on people and organizational performance
- job assignment scheduling aids

Project Manager
- demonstrate meeting contract agreements and right level of testing for risk involved

The capabilities for Project Managers were studied in detail, and related to Life cycle Phases as follows:

<table>
<thead>
<tr>
<th>Capabilities to be Supported</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>7</th>
<th>8</th>
</tr>
</thead>
<tbody>
<tr>
<td>Life Cycle Phase</td>
<td>Req. Analysis</td>
<td>Top Level Design</td>
<td>Detailed Design</td>
<td>Code/Unit Test</td>
<td>Integration Test</td>
<td>HW/SW Test</td>
<td>System Test</td>
<td>Maintenance</td>
</tr>
<tr>
<td>------------------------------</td>
<td>---------------</td>
<td>------------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-----------------</td>
<td>-------------</td>
<td>-------------</td>
<td>-------------</td>
</tr>
<tr>
<td>Req. Analysis</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Top Level Design</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Detailed Design</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Code/Unit Test</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Integration Test</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HW/SW Test</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>System Test</td>
<td></td>
<td></td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Maintenance</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The capabilities are coded as follows:

1. Capability to abstract management statistics from traceability matrix

2. Capability to examine the requirements selectively by characteristics, e.g., completeness measures, required classes; assess impact of changes

3. Capability for automated documentation throughout the test and development chain

4. Capability for baseline statistics

5. Test requirements matrix that maps test methods, tools, and test levels to software and system test requirements

6. Capability to abstract development status statistics

7. Capability to incorporate the error statistics into the risk model

8. Capability to obtain performance statistics
Additionally, the group derived the following set of General Requirements categorized by management function.

Costing Capabilities
- Cost Breakdown by activity
- Regression Costing
- Built-in test model that addresses cost (e.g., risk model)

Scheduling Capabilities
- Spreadsheet
- PERT, CPM capabilities
- "What if..." analysis
- Capability to compare past and present systems
- Capability to assess status of hardware

Tracking Capabilities
- Work Breakdown Structure
- Technical Progress Tracing
- Word Processing Capabilities
- Requirements/Design methodologies that result in a database
- Capability to display the contract compliance requirements
- Capability to assess organizational or individual performance
- System usage characteristics
- File of action items, problem reports, discrepancy reports, including resolution/status indicators, from all sources (e.g., IV&V, QA)

Technical Functions

After discussing the typical usage scenarios for the technical functions, this group derived a matrix of capabilities categorized by Cost, Feasibility, Technical Approach (T/A), Criticality, Size of Application Program to which capability can be economically applied (Appl.), and the intended users of the capability (a bar above a user code indicates "all but"). The matrix is given below.

Key:  
\[ \begin{array}{cccccc}
\text{H} & \text{M} & \text{L} & \text{H} & \text{M} & \text{L} \\
\text{high} & \text{med} & \text{low} & \text{high} & \text{med} & \text{low} \\
1 & 2 & 3 & 4 & 5 & 6 \\
\text{Soft.Eng.} & \text{Tester} & \text{IV&V'er} & \text{QA'er} & \text{CM'er} & \text{OT&E'er} \\
4 & 5 & 6 & \text{QA'er} & \text{CM'er} & \text{OT&E'er} \\
7 & 8 & 9 & \text{Maintainer} & \text{End Users} & \text{Measurers} \\
\end{array} \]
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1) Provide realistic test inputs and corresponding valid outputs for functional tests.</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>2) Provide boundary/limit data for stress tests and corresponding error response expected</td>
<td>H</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>5</td>
<td>9</td>
</tr>
<tr>
<td>3) Maintain reqmt., design, code, test traceability data base</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>4) Maintain problem report data base (keyed to ?)</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>all</td>
<td></td>
</tr>
<tr>
<td>5) Provide test coverage analysis and reporting</td>
<td>M/L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>6) Maintain test history</td>
<td>H/M</td>
<td>M</td>
<td>H</td>
<td>L</td>
<td></td>
<td></td>
</tr>
<tr>
<td>7) Provide endurance (i.e. long duration) tests</td>
<td>H</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>8) Provide regression test tools -- preserve data identify change impact compare results auto rerun</td>
<td>H</td>
<td>M/H</td>
<td>H</td>
<td>M</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>9) Evaluate test for redundancy, completeness, validity</td>
<td>?</td>
<td>L</td>
<td>M</td>
<td>?</td>
<td>5</td>
<td>8</td>
</tr>
<tr>
<td>10) Standards audit (code evaluation)</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>11) Units consistency analysis</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>12) Data flow anomaly detection</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>8</td>
<td>5</td>
</tr>
<tr>
<td>13) Interface consistency analysis (req., design, code)</td>
<td>L/M</td>
<td>H</td>
<td>H</td>
<td>M/L</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>14) Requirements consistency analysis</td>
<td>M</td>
<td>M</td>
<td>Reg.</td>
<td>H</td>
<td>L</td>
<td>8</td>
</tr>
</tbody>
</table>

13
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15) Design metric evaluation</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td>1,4,9</td>
</tr>
<tr>
<td>16) Quality metric evaluation</td>
<td>L</td>
<td>M</td>
<td>RADC</td>
<td>M</td>
<td>L</td>
</tr>
<tr>
<td>17) Provide stubs</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>1,2,3,7</td>
</tr>
<tr>
<td>18) Provide drivers</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>1,2,3,7</td>
</tr>
<tr>
<td>19) &quot;Display&quot; intermediate results</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>M</td>
<td>9, 5</td>
</tr>
<tr>
<td>20) Simulate &quot;typical&quot; users (see table)</td>
<td>H</td>
<td>M/L</td>
<td>M</td>
<td>H</td>
<td>1,2,3,7</td>
</tr>
<tr>
<td>21) Verify requirements</td>
<td>?</td>
<td>L</td>
<td>(e.g. H)</td>
<td>L</td>
<td>1,2,3,4,8 prototyping</td>
</tr>
<tr>
<td>22) Verify design</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>23) Symbolic assertion checker</td>
<td>M/L</td>
<td>L</td>
<td>M</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>24) Enable symbolic execution to generate assertions</td>
<td>L</td>
<td>H</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>25) Measure timing between arbitrary points</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>26) Generate cause-effect graphs and associated test cases</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>27) Mutation analysis</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>L</td>
<td></td>
</tr>
<tr>
<td>28) Automatic program proofs</td>
<td>H</td>
<td>?</td>
<td>M</td>
<td>?</td>
<td></td>
</tr>
<tr>
<td>29) Reformat, structure, re-draw, program listing for readability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>30) Simulate/emulate target environments</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>5</td>
</tr>
<tr>
<td>31) Connect SEE to target to provide &quot;visibility&quot; into actual execution</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>5</td>
</tr>
<tr>
<td>32) Descriptive symbol table lists, set/used tables</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>8</td>
</tr>
<tr>
<td>------------</td>
<td>------</td>
<td>-------</td>
<td>-----</td>
<td>-------</td>
<td>-------</td>
</tr>
<tr>
<td>33) Invocation trees</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>8</td>
</tr>
<tr>
<td>34) Reproduce &quot;operational&quot; case test &quot;cases&quot; to duplicate problems in SEE</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>35) Reduce test data and show in graphic form</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>36) Provide error models and predict future</td>
<td>L</td>
<td>M</td>
<td>H</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>37) Provide instrumentation for operational testing</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>38) Simulate target computer H/W and other faults.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>39) Provide closed-loop simulation case for full system capacity</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>40) Simulate requirements</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>41) Simulate design</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>42) Selectively turn on diagnostic outputs and monitor results</td>
<td>L</td>
<td>H</td>
<td>H</td>
<td>M/L</td>
<td>5</td>
</tr>
<tr>
<td>43) Expert system for test data selection</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>44) Error diagnosing expert system</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>H</td>
<td></td>
</tr>
<tr>
<td>45) Expert system for prediction</td>
<td>L</td>
<td>M</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>46) Maintain CM records, versions, document ties, etc.</td>
<td>M</td>
<td>H</td>
<td>H</td>
<td>L</td>
<td>all</td>
</tr>
<tr>
<td>47) Estimate change cost and retest time, effort</td>
<td>L</td>
<td>M</td>
<td>M/L</td>
<td>M</td>
<td>4,7,5</td>
</tr>
<tr>
<td>48) Accuracy study processor</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>M</td>
<td></td>
</tr>
<tr>
<td>49) Usage frequency analysis</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>5</td>
</tr>
<tr>
<td>50) Complexity metric</td>
<td>L</td>
<td>H</td>
<td>M</td>
<td>L</td>
<td>1,2,4,6,9</td>
</tr>
</tbody>
</table>
Attachment IV

PRESENTATION MATERIAL
Software Test and Evaluation Guidebook

Background

One of the current STEP task areas requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition life cycle the information required for the successful management of software T&E. The guidebook will be based upon the STEP Phases I and II recommendations and will incorporate results from other task areas of STEP Phases III and IV, as appropriate.

During the acquisition life cycle, a variety of organizations become involved with the software. The usual form which this involvement takes is that of reviewing documentation. The documents reviewed serve different purposes and vary in the amount of detail presented concerning the software and software T&E. In addition, each organization has specific interests and objectives which must be satisfied through the review process. Documents specifically concerned with testing range from the Test and Evaluation Master Plan (TEMP) through the lower level test plans and reports.

The Software Test and Evaluation Guidebook is intended to support the various organizations during this review process. Checklists are being developed to help ensure that information required by the STEP Phases I and II recommendations is included, as appropriate, in each document type. The checklists will include references to supporting materials and "good examples", whenever possible, to further aid the reviewing organizations as well as those involved in the development of documents.

Efforts to date on the Software Test and Evaluation Guidebook have concentrated on the development of "Software Test and Evaluation Guidebook, Part I: The Test and Evaluation Master Plan" for use by DDT&E.
Role of Various Organizations wrt Test Documents

<table>
<thead>
<tr>
<th>TEMP</th>
<th>Test Reqmt's</th>
<th>Test Plan</th>
<th>Test Des/Pro</th>
<th>Test Report</th>
</tr>
</thead>
<tbody>
<tr>
<td>OSD/DDT&amp;E</td>
<td>E</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HQ's</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dev Cmd</td>
<td>R</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prod Div</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Prog Mgr</td>
<td>G</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>DT&amp;E'ers</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>OT&amp;E'ers</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>PDSS'ers</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Users</td>
<td>I</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Key: G = generate  R = review  I = input  E = evaluate

Descriptions of Organizations

OSD/DDT&E: Office of the Secretary of Defense/Director, Defense Test and Evaluation

HQ's: Headquarters of the Army, Navy, Air Force

Dev Cmd: Development Command of the Army (Army Materiel Command), Navy (Naval Material Command), and Air Force (Air Force Systems Command)

Prod Div: Product Division. Examples include the Army's Communication and Electronics Command, the Navy's Sea Systems Command, and the Air Force's Aeronautical Systems Division.

Prog Mgr: Program Manager.

DT&E'ers: Development Test and Evaluation Agency. For example, the Army's Test and Evaluation Command.
Descriptions of Organizations (cont.)

OT&E'ers: Operational Test and Evaluation Agency of the Army (Operational Test and Evaluation Agency), Navy (Operational Test and Evaluation Force), and Air Force (Air Force Operational Test and Evaluation Center).

PDSS'ers: Post Deployment Software Support Organization.

Users: The ultimate Using Command or an appropriate representative.

Contents of test documents

TEMP:

The Test and Evaluation Master Plan includes a brief need statement and a description of the system which emphasizes how that need is satisfied. Further information includes required technical and operational characteristics and their associated critical T&E issues. The management structure of the program development and the test program are summarized and the integrated schedule is presented. Outlines describe all DT&E and OT&E conducted to date or planned for the future including objectives, events/scope of testing/basic scenarios and critical items needed to accomplish the test plan. Production Acceptance T&E is also described. Finally, the special resource summary details test articles needs and special support requirements.

Test Requirements:

The Test Requirements Document describes the overall philosophy of testing and assigns responsibilities for the test program. Qualification methods and levels of test to be employed are also addressed. Formal quality conformance requirements are detailed including data collection and recording requirements. Finally, any special qualification requirements such as tools, facilities, test formulas, or acceptance tolerance limits are included. Another item which is often included is a qualification cross reference table which relates the system requirements to the test requirements.
Contents of test documents (cont.)

Test Plan:
The Test Plan Document presents the limitations and assumptions of the testing and specific testing requirements. Management information such as responsibilities, schedules, and resource requirements are included. Tests are identified and related to the requirements to be tested and described in terms of test levels, test classes, and test methods. In addition, the purpose, scope, and objectives of each test are described. Test data generation and validation, and test process validation procedures are described. Finally, data reduction and analysis requirements and test reporting requirements are summarized.

Test Description/Procedure:
The Test Description/Procedure Document includes the limitations of the test, the test schedule, and required resources. Required pre-test procedures for the equipment or software are described. Specific test cases are presented in extreme detail which includes initialization procedures, inputs, actions required by test operators or equipment, intermediate test results, actions to be performed in the event of a program stop or indicated error, evaluation criteria, and expected results. Finally, data reduction/data analysis procedures and the methodology to be applied for correlating test results with test requirements are presented.

Test Report:
The Test Report Document describes any deviations from the test procedures which occurred during testing and their effects on the interpretation of test results. Test results are presented in summary form as well as in a detailed form along with a detailed test history. Results of the test evaluation are reported and recommendations for improvement are described.
Questions

Is it possible to consider a generic form of each document type (e.g. Test Plan) as the basis for the generation of a guidebook which would be useful when applied to all test plans?

Are the descriptions of the generic documents appropriate?

Can the matrix be collapsed (i.e., do each of the organizations have unique responsibilities with respect to testing)?

Is it possible to assign primary responsibility for generation of the generic document types?

What should the elements of the matrix be?
Topics of the TEMP Guidebook

1. Software's contribution to system capabilities and risk
2. Required operational and technical software characteristics
3. Critical software T&E issues
4. Management of the software program and the T&E program
5. Software milestones on the integrated schedule
6. Software T&E in terms of the system's DT&E and OT&E
7. Software test articles and special support requirements.

Questions

Are these topics appropriate for the guidebooks which remain to be produced?

Could the TEMP guidebook serve as a baseline to be expanded and tailored for the remaining guidebooks?

What mapping exists between the TEMP Guidebook topics and the topics to be addressed in the other guidebooks?

How should the original topics be modified for the new guidebooks?

IS THIS A GOOD FRAMEWORK FOR THE DEVELOPMENT OF FUTURE GUIDEBOOKS?
Software Test and Evaluation Model
**MILESTONE DECISIONS & PHASES OF ACTIVITY**
(per DoDD 5000.1, 29 Mar 82)

**Phase 0**
- Concept Exploration
- Mission Need Decision
  - Approval of JMSNS & Authorization to Proceed
  - Decision by Sec Def

**Phase I**
- Demonstration & Validation
- Milestone I Decision
  - Concept Selection & Authorization to Proceed based on SCP & TEMP
  - Recommendation by DSARC...
  - Decision by Sec Def

**Phase II**
- Full-Scale Development
- Milestone II Decision
  - Go-Ahead & Approval to Proceed based on DCP/IPS & TEMP
  - Recommendation by DSARC...
  - Decision by Sec Def

**Phase III**
- Production & Deployment
- Milestone III Decision
  - Authorization to Proceed
  - Decision may be delegated to DoD Component

**JMSNS = Justification Major System New Starts**
**SCP = System Concept Paper**
**DCP/IPS = Decision Coordinating Paper/Integrated Program Summary**
<table>
<thead>
<tr>
<th>CONCEPT EXPLORATION</th>
<th>DEMONSTRATION &amp; VALIDATION</th>
<th>FULL-SCALE DEVELOPMENT</th>
<th>PRODUCTION &amp; DEPLOYMENT</th>
</tr>
</thead>
<tbody>
<tr>
<td>PHASE 0</td>
<td>PHASE 1</td>
<td>PHASE 2</td>
<td>PHASE 3</td>
</tr>
<tr>
<td>IDENTITY AND VALIDATE NEED. EXPLORE ALTERNATE CONCEPTS</td>
<td>VALIDATE SYSTEM REQUIREMENTS</td>
<td>DESIGN, DEVELOP TEST &amp; EVALUATE</td>
<td>PRODUCTION &amp; DEPLOYMENT</td>
</tr>
</tbody>
</table>

INITIAL SYSTEM SPEC

FINAL SYSTEM SPEC, H/W AND S/W CI SPECs, INITIAL COMP RESOURCE LIFE/CYCLE MGT PLAN

MILESTONE DECISIONS AND PHASES OF ACTIVITY
CONCEPT EXPLORATION

PHASE 0

DEMONSTRATION & VALIDATION

PHASE 1

FULL-SCALE DEVELOPMENT

PHASE 2

PRODUCTION & DEPLOYMENT

PHASE 3

S/W LIFECYCLE

STANDARD SOFTWARE LIFE CYCLE
STEP Military Advisory Panel Meeting Report

Prepared for
Naval Air Development Center
Warminster, PA 18974

Under
Contract No. F33657-82-G-2003

February 1985

GEORGIA INSTITUTE OF TECHNOLOGY
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
**Report Documentation Page**

<table>
<thead>
<tr>
<th>Report Number</th>
<th>Govt Accession No.</th>
<th>Recipient's Catalog Number</th>
</tr>
</thead>
</table>

<table>
<thead>
<tr>
<th>Title</th>
<th>Type of Report &amp; Period Covered</th>
<th>Performing Org. Report Number</th>
</tr>
</thead>
<tbody>
<tr>
<td>STEP Military Advisory Panel Meeting Report</td>
<td></td>
<td>F33657-82-G-2083</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Performing Organization Name and Address</th>
<th>Program Element, Project, Task Area &amp; Work Unit Numbers</th>
</tr>
</thead>
<tbody>
<tr>
<td>School of Information and Computer Science Georgia Institute of Technology Atlanta, GA 30332</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Controlling Office Name and Address</th>
<th>Report Date</th>
<th>Number of Pages</th>
</tr>
</thead>
<tbody>
<tr>
<td>NADC Warminster, PA 18974</td>
<td>February, 1985</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution Statement (of this Report)</th>
<th>Security Class. (of this report)</th>
<th>Declassification Downgrading Schedule</th>
</tr>
</thead>
<tbody>
<tr>
<td>unlimited</td>
<td>unclassified</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Distribution Statement (of the abstract entered in Block 20, if different from Report)</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>unlimited</td>
<td></td>
</tr>
</tbody>
</table>

**Key Words**

**Abstract** (Continue on reverse side if necessary and identify by block number)

<table>
<thead>
<tr>
<th>19. KEY WORDS</th>
<th>20. ABSTRACT</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Continue on reverse side if necessary and identify by block number)</td>
<td>(Continue on reverse side if necessary and identify by block number)</td>
</tr>
</tbody>
</table>

EDITION OF 1 NOV 65 IS OBSOLETE
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>Discussion</td>
<td>2</td>
</tr>
<tr>
<td>Attachment I, Agenda</td>
<td>3</td>
</tr>
<tr>
<td>Attachment II, List of \ Participants</td>
<td>4</td>
</tr>
<tr>
<td>Attachment III, Presentation Material</td>
<td>5</td>
</tr>
</tbody>
</table>
STEP Military Advisory Panel
Meeting Report

Introduction.

On November 29, 1984, the Software Test and Evaluation Project (STEP) held the second meeting of the Military Advisory Panel (MAP) at the Pentagon in Washington, D.C. The objectives of the meeting were:

1) To present an overview of the Software Test and Evaluation Guidebook.

2) To critique the Guidebook and discuss the method of implementation and the form and content of any follow-on documents in the series.

Background.

One of the current STEP task areas requires that "Methods for implementing new DoD guidance in software T&E" be identified. The Software Test and Evaluation Guidebook is intended to be one such method. Its major goal is to provide the various professionals in the acquisition life cycle the information required for the successful management of software T&E. The Guidebook is based upon the STEP Phases I and II recommendations and will incorporate results from other task areas of STEP Phases III and IV, as appropriate.

During the acquisition life cycle, a variety of organizations become involved with the software. The usual form which this involvement takes is that of reviewing documentation. The documents reviewed serve different purposes and vary in the amount of detail presented concerning the software and software T&E. In addition, each organization has specific interests and objectives which must be satisfied through the review process. Documents specifically concerned with testing range from the Test and Evaluation Master Plan (TEMP) through the lower level test plans and reports.

The Software Test and Evaluation Guidebook is intended to support the various organizations during this review process. Checklists are being developed to help ensure that information required by the STEP Phases I and II recommendations is included, as appropriate, in each document type. The checklists will include references to supporting materials and "good examples", whenever possible, to further aid the reviewing organizations as well as those involved in the development of documents.

Efforts to date on the Software Test and Evaluation Guidebook have concentrated on the development of "Software Test and Evaluation Guidebook, Part I: The Test and Evaluation Master Plan" for use by DDT&E. The purpose of this meeting, as stated above, was to obtain comments from the MAP on this document.
Discussion.

The meeting was convened in room 3D973 of the Pentagon at 1000 hours. Mr. Greenlee opened the meeting with an introduction and presented a rationale for the Guidebook from the perspective of the Director, Defense Test and Evaluation (DDT&E). Dr. DeMillo then followed with a detailed presentation of the various portions of the Guidebook and the rationale for each of the individual component entries. Ms. Martin led discussions concerning the implementation of the Guidebook and the form of any follow-on documents. A good deal of the discussion that followed was concerned with how the Guidebook would actually be used by various members of the T&E Community.

One of the major issues was the handling of any follow-on documents. It was the consensus that one book should be developed that would be organized into parts that would be related to various users and activities.

A second issue was the level of detail. Several attendees felt that the Guidebook went into considerably more detail than one would normally expect to find in a TEMP. It was generally agreed that as a minimum, the guidance should give a reviewer the information that enables a decision on the worth of the TEMP, and give the TEMP author a feeling for the level of detail that must be provided.

The third main issue concerned the method of implementation. It was agreed that the Guidebook should be promulgated to the Services with a letter from Dr. Wade; this method appearing to have the most opportunity of forcing compliance from the Services.

The Panel made the following recommendations:

1) The Guidebook should be a two-part document; the first essentially as the one under review but restructured for ease of use, the second a treatment of the issues that various organizations should be concerned with as an acquisition proceeds.

2) The Guidebook should be promulgated by USDR&E.
Attachment I

AGENDA

1000-1030 Overview of Guidebook
   ** purpose
   ** structure
   ** users

1030-1100 Critique of Guidebook

1100-1200 Implementation and Follow-on
   ** users of follow-on guides
   ** traceability to TEMP, DoD-STD-SDS, common life cycles and major test planning and reporting documents
   ** promulgation strategies
   ** maintenance strategies (e.g., identifying an advocate or transitioning to STARS)

1200 Adjourn
LIST OF PARTICIPANTS

Mr. Donald Greenlee - ODDTE
Mr. John Bolino - ODOTE
Col. Robert Christopher - DSMC
LCol. Richard Stanley - AMC/AJPO
Capt. David Boslaugh - NAVMAT
Cdr. Steve Modlin - COMOPTEVFOR
Major Ed Stevens - AFSC
Mr. James Baca - AFOTEC
Dr. Richard A. DeMillo - Georgia Institute of Technology
Ms. R. J. Martin - Georgia Institute of Technology
Mr. W. Michael McCracken - Georgia Institute of Technology
Col(Ret) John F. Passafiume - Georgia Institute of Technology
Attachment III

PRESENTATION MATERIAL
THE SOFTWARE TEST AND EVALUATION PROJECT

PROGRESS AND PLANS
1984–1985

DIRECTOR, DEFENSE TEST AND EVALUATION
OFFICE OF THE SECRETARY OF DEFENSE

THE SOFTWARE TEST AND EVALUATION PROJECT
GEORGIA INSTITUTE OF TECHNOLOGY
STEP INITIATED BY DDT&E IN 1981

GOAL IS TO IMPROVE THE PRACTICE OF SOFTWARE TEST AND EVALUATION

* DEVELOP POLICY AND STANDARDS
* INSERT EXISTING TECHNOLOGY
* COORDINATE WITH RELATED EFFORTS
RATIONALE FOR STEP

SOFTWARE T&E IS A CRITICAL TECHNOLOGY

* MISSION CRITICAL SOFTWARE
* HARDWARE AND SOFTWARE IMBALANCES
* COSTS: TESTS VS. ERRORS
* INADEQUACIES OF CURRENT GUIDANCE
STUDY RESULTS: 20 RECOMMENDATIONS

MODIFY EXISTING POLICY AND PROVIDE NEAR AND LONG TERM IMPLEMENTATION SUPPORT

* Emphasis on test planning: chain beginning at system level
* Concentrate resources on critical software components
* Insert available technology into practice

THE SOFTWARE TEST AND EVALUATION PROJECT
GEORGIA INSTITUTE OF TECHNOLOGY
### Policy Implementation Through Test and Evaluation Master Plan

**Major Planning Document — Submitted and Evaluated Before Milestones**

<table>
<thead>
<tr>
<th>Mission/Function Matrix</th>
<th>Relate functions to be demonstrated to mission to be performed</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational Characteristics</td>
<td>Primary indicators of conformance to written specifications or operational requirements</td>
</tr>
<tr>
<td>Technical Characteristics</td>
<td>Aspects of system capability questioned before estimating overall worth</td>
</tr>
<tr>
<td>Operational Issues</td>
<td>To date, planned, objective, scope</td>
</tr>
<tr>
<td>Technical Issues</td>
<td>Test articles, test tools</td>
</tr>
</tbody>
</table>

**The Software Test and Evaluation Project**

**Georgia Institute of Technology**
SOFTWARE IN THE MODIFIED
TEST AND EVALUATION MASTER PLAN

--- Criteria for evaluation of TEMP

--- Criteria for additional guidance
to Project Offices

--- Examples of software characteristics and test issues

--- Examples of software test articles and special support requirements

GEORGIA INSTITUTE OF TECHNOLOGY
THE SOFTWARE TEST AND EVALUATION PROJECT
<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mission</td>
<td></td>
</tr>
<tr>
<td>2. System</td>
<td>a. Does the system contain Mission-critical Computer Resources?</td>
</tr>
<tr>
<td></td>
<td>b. Does software implement critical functions?</td>
</tr>
<tr>
<td></td>
<td>c. Is the system software intensive?</td>
</tr>
</tbody>
</table>
### PART I - DESCRIPTION (CONTINUED)

#### SECTION 2. System (continued)

<table>
<thead>
<tr>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>a. Key Functions</td>
</tr>
<tr>
<td>a. Does Mission/Function Matrix identify primary functional capabilities to be implemented by software?</td>
</tr>
<tr>
<td>b. Are the functions:</td>
</tr>
<tr>
<td>--New?</td>
</tr>
<tr>
<td>--Modifications of existing capabilities?</td>
</tr>
<tr>
<td>--Automation of existing capabilities?</td>
</tr>
<tr>
<td>--Mature?</td>
</tr>
</tbody>
</table>
PART I – DESCRIPTION (CONTINUED)

SECTION

2. System (continued)

a. Key Functions

b. Interfaces

QUESTIONS

a. Is software important to the interfaces?

b. Do the interfaces have software implications?
<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. System (continued)</td>
<td>a. Key Functions</td>
</tr>
<tr>
<td></td>
<td>b. Interfaces</td>
</tr>
<tr>
<td></td>
<td>c. Unique Characteristics</td>
</tr>
<tr>
<td></td>
<td>a. Does the system use software technology that:</td>
</tr>
<tr>
<td></td>
<td>--Affects risk?</td>
</tr>
<tr>
<td></td>
<td>--Has lifecycle impact?</td>
</tr>
<tr>
<td></td>
<td>--Distinguishes it from other systems?</td>
</tr>
<tr>
<td>SECTION</td>
<td>QUESTIONS</td>
</tr>
<tr>
<td>-------------------------------</td>
<td>---------------------------------------------------------------------------</td>
</tr>
<tr>
<td>2. System</td>
<td></td>
</tr>
<tr>
<td>3. Required Operational</td>
<td>a. Are there software characteristics that:</td>
</tr>
<tr>
<td>Characteristics</td>
<td>-- Are unique to software?</td>
</tr>
<tr>
<td>4. Required Technical</td>
<td>-- May be overlooked?</td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
</tr>
<tr>
<td>5. Required Software</td>
<td></td>
</tr>
<tr>
<td>Characteristics</td>
<td></td>
</tr>
</tbody>
</table>

GEORGIA INSTITUTE OF TECHNOLOGY
THE SOFTWARE TEST AND EVALUATION PROJECT
REQUIRED SOFTWARE CHARACTERISTICS

"Software parameters that are primary indicators of conformance to written requirements/specifications and operational suitability/effectiveness."

EXAMPLES

<table>
<thead>
<tr>
<th>TYPE</th>
<th>PARAMETER</th>
<th>UNIQUE?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>performance</td>
<td>No</td>
</tr>
<tr>
<td>Operational</td>
<td>maintainability</td>
<td>Unique aspects</td>
</tr>
<tr>
<td>Technical</td>
<td>precision/accuracy</td>
<td>Minimal</td>
</tr>
<tr>
<td>Technical</td>
<td>robustness</td>
<td>Yes</td>
</tr>
<tr>
<td>SECTION</td>
<td>QUESTIONS</td>
<td></td>
</tr>
<tr>
<td>--------------------------------------</td>
<td>---------------------------------------------------------------------------</td>
<td></td>
</tr>
<tr>
<td>2. System</td>
<td></td>
<td></td>
</tr>
<tr>
<td>3. Required Operational Characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Required Technical Characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>5. Required Software Characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>6. Critical T&amp;E Issues</td>
<td>a. Do the required software characteristics raise unique or easily missed T&amp;E issues?</td>
<td></td>
</tr>
<tr>
<td>a. Technical Issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>b. Operational Issues</td>
<td></td>
<td></td>
</tr>
<tr>
<td>c. Software Issues</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
"Those aspects of software capability...that must be questioned before a system's overall worth can be estimated..."

### EXAMPLES

<table>
<thead>
<tr>
<th>TYPE</th>
<th>CHARACTERISTIC</th>
<th>ISSUE?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>performance</td>
<td>degraded mode operation</td>
</tr>
<tr>
<td>Operational</td>
<td>maintainability</td>
<td>Adequacy of support env.</td>
</tr>
<tr>
<td>Technical</td>
<td>precision/accuracy</td>
<td>algorithm success</td>
</tr>
<tr>
<td>Technical</td>
<td>robustness</td>
<td>design and architecture</td>
</tr>
</tbody>
</table>

GEORGIA INSTITUTE OF TECHNOLOGY
THE SOFTWARE TEST AND EVALUATION PROJECT
PART II – PROGRAM SUMMARY

SECTION                        QUESTIONS

1. Management

2. Integrated Schedule        a. Are key software subsystem demonstration included on schedule?
T&E OUTLINES

PART III — DT&E Outline

PART IV — OT&E Outline

PART V — Software T&E Outline
## PART V - SOFTWARE T&E OUTLINE

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
</table>
| 1. Software T&E to Date | a. Have software characteristics been demonstrated using systematic test methods?  
  b. Have the planned levels of testing been achieved? Is the documentation cited?  
  c. Have testing deficiencies been interpreted in terms of overall system evaluation criteria? |

GEORGIA INSTITUTE OF TECHNOLOGY  
THE SOFTWARE TEST AND EVALUATION PROJECT
<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Software T&amp;E to Date</td>
<td></td>
</tr>
<tr>
<td>2. Future Software T&amp;E</td>
<td>a. Is the test environment (i.e., development, operational, logistics support) appropriate for the characteristics to be demonstrated?</td>
</tr>
<tr>
<td>SECTION</td>
<td>QUESTIONS</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>1. Software T&amp;E to Date</td>
<td></td>
</tr>
<tr>
<td>2. Future Software T&amp;E (continued)</td>
<td>a. What software characteristics are not adequately addressed in the DT&amp;E and OT&amp;E Outlines?</td>
</tr>
<tr>
<td>a. Software T&amp;E Objectives</td>
<td>a.</td>
</tr>
<tr>
<td>b. Software T&amp;E Events/Scope of Testing/Basic Scenarios</td>
<td></td>
</tr>
</tbody>
</table>
## PART V - SOFTWARE T&E OUTLINE (CONT)

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Software T&amp;E to Date</td>
<td></td>
</tr>
<tr>
<td>2. Future Software T&amp;E (continued)</td>
<td>a. Are subsystems needed for adequate software T&amp;E prior to next decision point identified?</td>
</tr>
<tr>
<td>c. Critical Software T&amp;E Items</td>
<td></td>
</tr>
<tr>
<td>SECTION</td>
<td>QUESTIONS</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>1. Test Articles</td>
<td>a. Are critical software components and key subsystems identified?</td>
</tr>
</tbody>
</table>
### PART VII – SPECIAL RESOURCE SUMMARY

#### SECTION

<table>
<thead>
<tr>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>1. Test Articles</td>
</tr>
<tr>
<td>2. Special Support Requirements</td>
</tr>
<tr>
<td>a. Is there an explanation of how test tools support software test objectives?</td>
</tr>
<tr>
<td>b. Are adequate steps being taken to acquire each tool? Do any of the tools increase risk?</td>
</tr>
</tbody>
</table>

GEORGIA INSTITUTE OF TECHNOLOGY

THE SOFTWARE TEST AND EVALUATION PROJECT
TECHNICAL REPORT NO. GIT-ICS 84/11

SUPPLEMENT TO THE PROCEEDINGS OF THE INTERNATIONAL TEST AND EVALUATION ASSOCIATION 1984 SYMPOSIUM (STEP SESSION)

Prepared for
Naval Air Development Center
Warminster, PA 18974

Under
Contract No. F33657-82-G-2083

November 8, 1984

GEORGIA INSTITUTE OF TECHNOLOGY
A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA
SCHOOL OF INFORMATION AND COMPUTER SCIENCE
ATLANTA, GEORGIA 30332
SUPPLEMENT TO THE PROCEEDINGS

OF THE

INTERNATIONAL TEST AND EVALUATION ASSOCIATION

1984 SYMPOSIUM (STEP SESSION)

THE SHOREHAM HOTEL
WASHINGTON, D.C.

08 NOVEMBER, 1984
<table>
<thead>
<tr>
<th>REPORT DOCUMENTATION PAGE</th>
<th>READ INSTRUCTIONS BEFORE COMPLETING FORM</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. REPORT NUMBER</td>
<td>GIT-ICS 85/05</td>
</tr>
<tr>
<td>2. GOVT ACCESSION NO.</td>
<td></td>
</tr>
<tr>
<td>3. RECIPIENT'S CATALOG NUMBER</td>
<td></td>
</tr>
<tr>
<td>4. TITLE (and Subtitle)</td>
<td>Supplement to the Proceedings of the INTERNATIONAL TEST AND EVALUATION ASSOCIATION 1984 Symposium (STEP Session)</td>
</tr>
<tr>
<td>5. TYPE OF REPORT &amp; PERIOD COVERED</td>
<td></td>
</tr>
<tr>
<td>6. PERFORMING ORG. REPORT NUMBER</td>
<td></td>
</tr>
<tr>
<td>7. AUTHOR(s)</td>
<td></td>
</tr>
<tr>
<td>8. CONTRACT OR GRANT NUMBER(s)</td>
<td>F33657-82-G-2083</td>
</tr>
<tr>
<td>9. PERFORMING ORGANIZATION NAME AND ADDRESS</td>
<td>Georgia Institute of Technology School of Information and Computer Science Atlanta, Georgia 30332</td>
</tr>
<tr>
<td>10. PROGRAM ELEMENT, PROJECT, TASK AREA &amp; WORK UNIT NUMBERS</td>
<td></td>
</tr>
<tr>
<td>11. CONTROLLING OFFICE NAME AND ADDRESS</td>
<td>NADC Warminster, PA 18974</td>
</tr>
<tr>
<td>12. REPORT DATE</td>
<td>April 1985</td>
</tr>
<tr>
<td>13. NUMBER OF PAGES</td>
<td></td>
</tr>
<tr>
<td>14. MONITORING AGENCY NAME &amp; ADDRESS (if different from Controlling Office)</td>
<td></td>
</tr>
<tr>
<td>15. SECURITY CLASS. (of this report)</td>
<td></td>
</tr>
<tr>
<td>15a. DECLASSIFICATION/DOWNGRADING SCHEDULE</td>
<td></td>
</tr>
<tr>
<td>16. DISTRIBUTION STATEMENT (of this Report)</td>
<td>unlimited</td>
</tr>
<tr>
<td>17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)</td>
<td>unlimited</td>
</tr>
<tr>
<td>18. SUPPLEMENTARY NOTES</td>
<td></td>
</tr>
<tr>
<td>19. KEY WORDS (Continue on reverse side if necessary and identify by block number)</td>
<td></td>
</tr>
<tr>
<td>20. ABSTRACT (Continue on reverse side if necessary and identify by block number)</td>
<td></td>
</tr>
</tbody>
</table>
FOREWORD

This supplement contains the presentations and panel session of the Software Test and Evaluation Project (STEP) Session of the International Test and Evaluation Association Symposium on November 8, 1984.

The motivation for this session of the ITEA symposium is to showcase some good examples of software testing in the DoD and to update the test community on some current activities of the DoD to improve the testing of software. The session is organized around three technical areas; Software T&E Policy and Guidance; Software Test Management; and Software Test Tools. The Software T&E Policy and Guidance section is concerned with presenting policy level changes to DoDD 5000.3 that will improve the visibility and accountability of software in major systems acquisitions. Included in this section is a presentation of a guidebook that will be issued to aid in implementing DoD policy. The Software Test Management section presents two perspectives of the management of software testing during development test and operational test. The Software Test Tools section is used to describe the effectiveness of using software test tools on two DoD programs. In addition, it describes the to be issued DoD-Std-SDS and SOS, the new Joint Logistics Commanders Standards for software development and software quality assurance.

A panel session to discuss in an open forum the current experiences of the panel and the audience related to testing software completes this session.

W. Michael McCracken
Session Chairman
Software Test and Evaluation Project
Georgia Institute of Technology
# TABLE OF CONTENTS

## SOFTWARE T&E POLICY AND GUIDANCE

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Pending Revisions to DoDD 5000.3</td>
<td>1</td>
</tr>
<tr>
<td>Donald R. Greenlee</td>
<td></td>
</tr>
<tr>
<td>Office of the Director Defense Test and Evaluation</td>
<td></td>
</tr>
<tr>
<td>Software Test and Evaluation Guidebook</td>
<td>3</td>
</tr>
<tr>
<td>Richard A. DeMillo</td>
<td></td>
</tr>
<tr>
<td>School of Information and Computer Science</td>
<td></td>
</tr>
<tr>
<td>Georgia Institute of Technology</td>
<td></td>
</tr>
</tbody>
</table>

## SOFTWARE TEST MANAGEMENT

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>TECOM Policy, Procedures and Responsibilities for T&amp;E of Software Embedded in Battlefield Automated Systems</td>
<td>16</td>
</tr>
<tr>
<td>Francis Bartosik</td>
<td></td>
</tr>
<tr>
<td>U.S. Army Test and Evaluation Command</td>
<td></td>
</tr>
<tr>
<td>AFOTEC Software Test Managers Handbook</td>
<td>27</td>
</tr>
<tr>
<td>Major Frederick Foster</td>
<td></td>
</tr>
<tr>
<td>U.S. Air Force Operational Test and Evaluation Command</td>
<td></td>
</tr>
</tbody>
</table>

## SOFTWARE TEST TOOLS

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Effects of Testing Tools on Aegis Lifetime Support</td>
<td>43</td>
</tr>
<tr>
<td>P. Graham O'Neil</td>
<td></td>
</tr>
<tr>
<td>Sanders Associates</td>
<td></td>
</tr>
<tr>
<td>A-7E Test Program</td>
<td>60</td>
</tr>
<tr>
<td>Paul Clements</td>
<td></td>
</tr>
<tr>
<td>Naval Research Labs</td>
<td></td>
</tr>
<tr>
<td>DoD-STD-SDS and SQS</td>
<td>74</td>
</tr>
<tr>
<td>LCDR Michael T. Gehl</td>
<td></td>
</tr>
<tr>
<td>U.S. Navy Electronics Command</td>
<td></td>
</tr>
</tbody>
</table>

## PANEL SESSION

<table>
<thead>
<tr>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Transcript</td>
<td>89</td>
</tr>
<tr>
<td>Donald R. Greenlee</td>
<td></td>
</tr>
<tr>
<td>Richard A. DeMillo</td>
<td></td>
</tr>
<tr>
<td>Francis Bartosik</td>
<td></td>
</tr>
<tr>
<td>Major Frederick Foster</td>
<td></td>
</tr>
<tr>
<td>Paul Clements</td>
<td></td>
</tr>
<tr>
<td>LCDR Michael T. Gehl</td>
<td></td>
</tr>
</tbody>
</table>
Pending Revisions to DoDD 5000.3

Donald R. Greenlee

Pending revisions to DoDD 5000.3 are presented to inform the test community of these changes. The changes establish improved guidelines for the testing of software that is a part of the mission critical computer resources of major weapons systems. The material presented is for information only, as the changes have not been coordinated in the Services.

The changes to DoDD 5000.3 that are related to software testing include the following:

Test and Evaluation of Computer Software: additional language has been added to the draft Directive that clarifies the intent of this section. In summary, this language is intended to require that software components that implement critical functions of a system be identified and that those components be tested throughout the development/integration portion of the software lifecycle. This section is planned to require that the results of those tests be objective, repeatable, available to subsequent test groups and interpretable in terms of the overall system objectives. The level of testing of these components should be sufficient to achieve a balanced risk with the hardware on which they are implemented.

In Part I, Section 3 of the TEMP, there will be a description of the critical software issues that must be addressed by testing. The description will contain an outline of the software program and the T&E management responsibilities of the participating T&E organizations. It will show the integrated time sequencing of critical software T&E events. It will also state the corresponding test-verifiable objectives and the corresponding goals, specifications and thresholds for each critical software issue. It will contain a summary of the software T&E already conducted as well as the planned software T&E. Finally, it will contain the identification of critical software components and subsystems required for testing in each phase and the test tools required, including how they support the software test objectives.

Mr. Greenlee's presentation did not use viewgraphs nor was there any material for handouts.
Donald R. Greenlee

Mr. Greenlee is currently acting as Deputy Director, Defense Test and Evaluation for Strategic, Naval and Communications, Command, Control and Intelligence Systems. Prior to joining the Office of the Secretary of Defense, he was Associate Head of the Data Processing Systems Department in the MITRE Corporation's National Command and Control Systems Division. He held earlier positions with the Johns Hopkins University Applied Physics Laboratory and the Apollo Support and Defense Systems Departments of the General Electric Company.

Mr. Greenlee attended the Universities of Florida and Maryland, Syracuse and Ohio State Universities and MIT; he holds a B.S. in Mathematics and Master's degrees in Mathematics and in Engineering. He is a member of the Tau Beta Pi and Alpha Pi Mu national engineering honor societies, a U.S. Soccer Federation-licensed coach and referee, and a master ultramarathoner, for which he trains in the corridors of the Pentagon.
The focus of this presentation is a guidebook that is being developed by the Software Test and Evaluation Project (STEP) for the Director, Defense Test and Evaluation. The Software Test and Evaluation Project was initiated in 1981 by the Director Defense Test and Evaluation. The primary objective of STEP is to develop new DoD guidance and policy for the test and evaluation of computer software for mission-critical applications.

The Software Test and Evaluation Guidebook is a two volume reference set that provides checklists and guidance to DoD components in the area of software test and evaluation for major Defense system acquisition. Volume I is devoted to the evaluation of the treatment of software in Test and Evaluation Master Plans. Volume II (not presented) addresses the structuring, planning, conduct, and evaluation of software tests throughout the acquisition process.
Richard A. DeMillo

Richard A. DeMillo has an extensive background in computer science research and computing practice. Dr. DeMillo holds the Ph.D. degree in Information and Computer Science from the Georgia Institute of Technology (1972). He is the author or co-author of over 50 reports, books and articles reporting basic research results in theoretical computer science, computer security, and software engineering.

Prior to receiving his Ph.D., Dr. DeMillo held programming positions and research assistantships. From 1972 to 1976 he was an Assistant Professor of Electrical Engineering and Computer Science at the University of Wisconsin-Milwaukee. In 1976, he returned to Georgia Tech and was promoted to the rank of Professor in 1981.

In the area of software engineering, Dr. DeMillo has concentrated on software reliability and the emerging area of software metrics. He is a co-developer of the program mutation approach to software testing and is the author of the book *Program Mutation: an approach to software testing*. Dr. DeMillo is currently Principal Investigator for the DoD Software Test and Evaluation Project (STEP). The goal of STEP is to revise the policy and procedures in the Department of Defense for the testing of software in major DoD systems. In this capacity, Dr. DeMillo reports to the Office of the Secretary of Defense.

Dr. DeMillo is on the editorial board of several professional journals and has served as chairman of a number of national meetings. He has been a Distinguished Visitor of the IEEE Computer Society and is a frequent speaker at professional and technical gatherings.
THE SOFTWARE TEST AND EVALUATION PROJECT

PROGRESS AND PLANS
1984–1985

DIRECTOR, DEFENSE TEST AND EVALUATION
OFFICE OF THE SECRETARY OF DEFENSE

STEP INITIATED BY DDT&E IN 1981

GOAL IS TO IMPROVE THE PRACTICE
OF SOFTWARE TEST AND EVALUATION

* DEVELOP POLICY AND STANDARDS
* INSERT EXISTING TECHNOLOGY
* COORDINATE WITH RELATED EFFORTS
RATIONALE FOR STEP

SOFTWARE T&E IS A CRITICAL TECHNOLOGY

* MISSION CRITICAL SOFTWARE
* HARDWARE AND SOFTWARE IMBALANCES
* COSTS: TESTS VS. ERRORS
* INADEQUACIES OF CURRENT GUIDANCE

STUDY RESULTS: 20 RECOMMENDATIONS

MODIFY EXISTING POLICY AND PROVIDE NEAR AND LONG TERM IMPLEMENTATION SUPPORT

* Emphasis on test planning: chain beginning at system level
* Concentrate resources on critical software components
* Insert available technology into practice
SOFTWARE IN THE MODIFIED TEST AND EVALUATION MASTER PLAN

--- Criteria for evaluation of TEMP

--- Criteria for additional guidance to Project Offices

--- Examples of software characteristics and test issues

--- Examples of software test articles and special support requirements

GEORGIA INSTITUTE OF TECHNOLOGY
THE SOFTWARE TEST AND EVALUATION PROJECT

POLICY IMPLEMENTATION THROUGH TEST AND EVALUATION MASTER PLAN

MAJOR PLANNING DOCUMENT -- SUBMITTED AND EVALUATED BEFORE MILESTONES

MISSION/FUNCTION MATRIX
OPERATIONAL CHARACTERISTICS
TECHNICAL CHARACTERISTICS
OPERATIONAL ISSUES
TECHNICAL ISSUES
T&E OUTLINES
SPECIAL RESOURCE SUMMARY

Relate functions to be demonstrated to mission to be performed
Primary indicators of conformance to written specifications or operational requirements
Aspects of system capability questioned before estimating overall worth
To date, planned, objective, scope
Test articles, test tools

THE SOFTWARE TEST AND EVALUATION PROJECT
GEORGIA INSTITUTE OF TECHNOLOGY
## PART I - DESCRIPTION

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Mission</td>
<td>a. Does the system contain Mission-critical Computer Resources?</td>
</tr>
<tr>
<td></td>
<td>b. Does software implement critical functions?</td>
</tr>
<tr>
<td></td>
<td>c. Is the system software intensive?</td>
</tr>
<tr>
<td>2. System</td>
<td>a. Does the system contain Mission-critical Computer Resources?</td>
</tr>
<tr>
<td></td>
<td>b. Does software implement critical functions?</td>
</tr>
<tr>
<td></td>
<td>c. Is the system software intensive?</td>
</tr>
<tr>
<td></td>
<td>a. Does Mission/Function Matrix identify primary functional capabilities to be implemented by software?</td>
</tr>
<tr>
<td></td>
<td>b. Are the functions:</td>
</tr>
<tr>
<td></td>
<td>-- New?</td>
</tr>
<tr>
<td></td>
<td>-- Modifications of existing capabilities?</td>
</tr>
<tr>
<td></td>
<td>-- Automation of existing capabilities?</td>
</tr>
<tr>
<td></td>
<td>-- Mature?</td>
</tr>
</tbody>
</table>
### PART I — DESCRIPTION (CONTINUED)

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. System (continued)</td>
<td></td>
</tr>
<tr>
<td>a. Key Functions</td>
<td>a. Is software important to the interfaces?</td>
</tr>
<tr>
<td>b. Interfaces</td>
<td>b. Do the interfaces have software implications?</td>
</tr>
<tr>
<td>c. Unique Characteristics</td>
<td></td>
</tr>
</tbody>
</table>
### PART I — DESCRIPTION (CONTINUED)

#### SECTION

2. System
3. Required Operational Characteristics
4. Required Technical Characteristics
5. Required Software Characteristics

#### QUESTIONS

- Are there software characteristics that:
  - Are unique to software?
  - May be overlooked?

### REQUIRED SOFTWARE CHARACTERISTICS

"Software parameters that are primary indicators of conformance to written requirements/specifications and operational suitability/effectiveness."

#### EXAMPLES

<table>
<thead>
<tr>
<th>TYPE</th>
<th>PARAMETER</th>
<th>UNIQUE?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>performance</td>
<td>No</td>
</tr>
<tr>
<td>Operational</td>
<td>maintainability</td>
<td>Unique aspects</td>
</tr>
<tr>
<td>Technical</td>
<td>precision/accuracy</td>
<td>Minimal</td>
</tr>
<tr>
<td>Technical</td>
<td>robustness</td>
<td>Yes</td>
</tr>
</tbody>
</table>

GEORGIA INSTITUTE OF TECHNOLOGY
THE SOFTWARE TEST AND EVALUATION PROJECT
## PART I — DESCRIPTION (CONTINUED)

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. System</td>
<td></td>
</tr>
<tr>
<td>3. Required Operational Characteristics</td>
<td></td>
</tr>
<tr>
<td>4. Required Technical Characteristics</td>
<td></td>
</tr>
<tr>
<td>5. Required Software Characteristics</td>
<td></td>
</tr>
<tr>
<td>6. Critical T&amp;E Issues</td>
<td>a. Do the required software characteristics raise unique or easily missed T&amp;E issues?</td>
</tr>
<tr>
<td>a. Technical Issues</td>
<td></td>
</tr>
<tr>
<td>b. Operational Issues</td>
<td></td>
</tr>
<tr>
<td>c. Software Issues</td>
<td></td>
</tr>
</tbody>
</table>

### GEORGIA INSTITUTE OF TECHNOLOGY

THE SOFTWARE TEST AND EVALUATION PROJECT

## REQUIRED SOFTWARE CHARACTERISTICS

"Those aspects of software capability...that must be questioned before a system's overall worth can be estimated..."

### EXAMPLES

<table>
<thead>
<tr>
<th>TYPE</th>
<th>CHARACTERISTIC</th>
<th>ISSUE?</th>
</tr>
</thead>
<tbody>
<tr>
<td>Operational</td>
<td>performance</td>
<td>degraded mode operation</td>
</tr>
<tr>
<td>Operational</td>
<td>maintainability</td>
<td>Adequacy of support env.</td>
</tr>
<tr>
<td>Technical</td>
<td>precision/accuracy</td>
<td>algorithm success</td>
</tr>
<tr>
<td>Technical</td>
<td>robustness</td>
<td>design and architecture</td>
</tr>
</tbody>
</table>
PART II — PROGRAM SUMMARY

SECTION | QUESTIONS
---|---
1. Management | 
2. Integrated Schedule | a. Are key software subsystem demonstration included on schedule?

T&E OUTLINES

PART III — DT&E Outline

PART IV — OT&E Outline

PART V — Software T&E Outline
### PART V — SOFTWARE T&E OUTLINE

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
</table>
| 1. Software T&E to Date | a. Have software characteristics been demonstrated using systematic test methods?  
  b. Have the planned levels of testing been achieved? Is the documentation cited?  
  c. Have testing deficiencies been interpreted in terms of overall system evaluation criteria? |

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>2. Future Software T&amp;E</td>
<td>a. Is the test environment (i.e., development, operational, logistics support) appropriate for the characteristics to be demonstrated?</td>
</tr>
<tr>
<td>SECTION</td>
<td>QUESTIONS</td>
</tr>
<tr>
<td>---------</td>
<td>-----------</td>
</tr>
<tr>
<td>1. Software T&amp;E to Date</td>
<td></td>
</tr>
<tr>
<td>2. Future Software T&amp;E (continued)</td>
<td>a. What software characteristics are not adequately addressed in the DT&amp;E and OT&amp;E Outlines?</td>
</tr>
<tr>
<td>a. Software T&amp;E Objectives</td>
<td>b. Software T&amp;E Events/Scope of Testing/Basic Scenarios</td>
</tr>
<tr>
<td>c. Critical Software T&amp;E Items</td>
<td>a. Are subsystems needed for adequate software T&amp;E prior to next decision point identified?</td>
</tr>
</tbody>
</table>
### PART VII - SPECIAL RESOURCE SUMMARY

<table>
<thead>
<tr>
<th>SECTION</th>
<th>QUESTIONS</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Test Articles</td>
<td>a. Are critical software components and key subsystems identified?</td>
</tr>
<tr>
<td>2. Special Support</td>
<td>a. Is there an explanation of how test tools support software test objectives?</td>
</tr>
<tr>
<td>Requirements</td>
<td>b. Are adequate steps being taken to acquire each tool? Do any of the tools increase risk?</td>
</tr>
</tbody>
</table>
TECOM Policy, Procedures and Responsibilities for
T&E of Software Embedded in Battlefield Automated Systems

Francis Bartosik

The U.S. Army Test and Evaluation Command is responsible for test and evaluation of assigned Army and other service material systems. TECOM is headquartered at Aberdeen Proving Ground, MD. TECOM contains nine subordinate installations/field operating agencies. TECOM relies on repeatable, highly instrumented, system-level testing to demonstrate system capabilities. To aid in the testing of software portions of systems, TECOM has developed a policy for that activity. This policy is used to reduce the art of testing to more of a standardized, scientific process. The intentions behind implementing this policy are to: (1) establish consistency among HQ staff elements and field test agencies in the approach to software testing; (2) to formalize and discipline the software test process and the scheduling of the events comprising that process to better support the total TECOM mission of test and evaluation; (3) to focus attention on the mission-critical embedded computer resources in the battlefield automated systems being developed.
Francis Bartosik

Frank Bartosik is a native of Newell, PA, a small town thirty miles south of Pittsburgh. He received his BS degree in secondary education with a major in mathematics from California State University of Pennsylvania, and has done graduate work at the University of Delaware.

He began his career with the government in July 1967 as a mathematician at the Combat Systems Test activity at Aberdeen Proving Ground, MD. After three years of active duty in the U.S. Navy in the Pacific, Mr. Bartosik returned to Aberdeen's combat systems test activity in April 1972. His duties at that time included computer programming, and were to broaden in scope in the next several years to encompass activities in planning for the modernization of Aberdeen's data acquisition, reduction, analysis, and processing capabilities and planning, programming and budgeting for the test activity's ADP Center.

In February of 1978, he was selected for the position of operations research analyst in the software evaluation group in the simulation and software division of the Headquarters, U.S. Army Test and Evaluation Command at Aberdeen. His duties comprised the testing and the evaluation of the software-related aspects of battlefield automated systems. On September 1 of this year, he was selected as the senior analyst of the software evaluation group. On November 5, he was designated as the Chief, Long Range Development Group, HQ, Test and Evaluation Command, responsible for the development and publication of long range planning for the command.

Mr. Bartosik is married to the former Cynthia Ann Williams of Hellertown, PA. They live in Newark, Delaware with their seven year old son, Joey, and their four year old daughter, Katie.
TECOM SOFTWARE TESTING POLICY

- BACKGROUND
- POLICY
  WHY
  WHAT
  HOW
- EXPECTATIONS
- PERSONAL OBSERVATIONS

[Diagram of organizational structure with DOD, HQDA, AMC, CECOM, TECOM, MICOM, DPG, EPG, WSMR, YPG, TTC, JPG, APG, ADTA, CRTC]
TECOM SOFTWARE TESTING POLICY

- **TECOM TESTING**

  - CONTROLLED
  - REPEATABLE
  - INSTRUMENTED
TECOM SOFTWARE TESTING POLICY

WHY IS IT NECESSARY?

- INCONSISTENCY IN FIELD AGENCIES' SOFTWARE TESTING APPROACH
- INVOLVEMENT BY TESTER TOO LATE FOR ALL BUT "BLACK BOX" TESTING
- ADEQUATE TESTING PRECLUDED BY (AMONG OTHERS) INSUFFICIENT KNOWLEDGE OF SYSTEM TO BE TESTED

TECOM SOFTWARE TESTING POLICY

WHAT IS THE INTENTION?

- ESTABLISH CONSISTENT APPROACH TO SOFTWARE TESTING THROUGHOUT TECOM
- FORMALIZE SOFTWARE T&E PROCESS IN SUPPORT OF TECOM MISSION
- EMPHASIZE MISSION-CRITICAL EMBEDDED COMPUTER RESOURCES
TECOM SOFTWARE TESTING POLICY

**Basis**

- Practical Experience (School of Hard Knocks)
- Methodology Studies 1974-76
- ARTADS Software Acquisition Study 1975
- DARCOM Study 1976
- TECOM Technical Report SY-2-77 1977
- Test Operating Procedure 1-1-056 1977
- JLC
  - Electronics Reliability Software Working Group 1975
  - Ad Hoc Group on T&E Planning 1978
  - Coordinating Group on Computer Resource Mgmt 1978
  - Monterey III 1979/81
  - Orlando I 1983

**Basis (Cont'd)**

- Conferences and Symposia
  - Software Quality Mgmt Conference 1971
  - Second Army Software Symposium 1978
  - DARCOM Tactical Computer Conference 1978
  - DOD/NSIA Symposium on Software T&E 1983
  - Multi-Service DT&E Commanders 1980-84

- Other Activities
  - BAISEMP 1978-81
  - ACCSSEIMP 1982-84
  - Defense Management Journal March 1979
  - STEP 1983—
TECOM SOFTWARE TESTING POLICY

PROVISIONS

- Embedded computer resources addressed in context of system T&E
- Software assessed against specifications and standards
- Impact of software performance on system performance assessed
- Software quality assessed in terms of usability, correctness, interoperability, maintainability
- Capability to test and evaluate battlefield automated systems updated through feedback to command methodology and instrumentation programs

WHAT ARE THE RESPONSIBILITIES?

- TECOM HQ
  • As evaluator
  • As test manager
- Field test agencies
TECOM SOFTWARE TESTING POLICY

RESPONSIBILITIES

- TECOM HQ, AS EVALUATOR, WILL:
  - DEVELOP SOFTWARE PLANNING DIRECTIVE IN TIME FOR FIELD AGENCY EARLY INVOLVEMENT
  - DEVELOP INDEPENDENT EVALUATION PLANS (IEP)
  - DEVELOP TEST DESIGN PLANS (TDP)
  - ASSURE ADEQUACY OF DETAILED TEST PLANS (DTP)
  - Develop Independent Evaluation Reports (IER)

RESPONSIBILITIES (Cont’d)

- TECOM HQ, AS TEST MANAGER, WILL:
  - DEVELOP SOFTWARE PLANNING DIRECTIVE
  - ASSURE AVAILABILITY OF TEST TOOLS/INSTRUMENTATION
  - ASSURE ADEQUACY OF DTP
  - APPROVE TEST REPORT
  - OBTAIN AND FORWARD DATA TO INDEPENDENT EVALUATOR
SOFTWARE TESTING OBJECTIVES

1. PARTICIPATE IN SPEC DEVELOPMENT
2. ASSURE ALGORITHMIC CORRECTNESS
3. DETERMINE DOCUMENTATION ADEQUACY
4. MONITOR IV&V ACTIVITIES
5. VALIDATE SIMULATIONS
6. SYSTEMATICALLY DETECT AND ANALYZE SOFTWARE FAILURES
7. DETERMINE RESOURCE CONSTRAINTS AND EXCESSES
8. ASSESS SOFTWARE ASPECTS OF SYSTEM SPEC COMPLIANCE
9. DETERMINE RETEST REQUIREMENTS
10. MEASURE OPERATING SYSTEM FUNCTION
11. RELATE SYSTEM AND COMPUTER TIME LINES
12. ENSURE A CONSISTENT SOFTWARE INCIDENT REPORTING SYSTEM
13. SUPPORT SOFTWARE PORTION OF SYSTEM EVALUATION
TECOM SOFTWARE TESTING POLICY

WHAT ARE THE MECHANICS?

- SOFTWARE PLANNING DIRECTIVE
- EARLY PARTICIPATION IN SYSTEM DEVELOPMENT
- INCREASED RELIANCE ON SINGLE INTEGRATED DEVELOPMENT TEST CONCEPT (SIDTC)

- MATERIEL DEVELOPER
- DEVELOPING CONTRACTOR(S)
- IV&V AGENT
- SYSTEM DEVELOPMENT TESTER
- COMBAT DEVELOPER
- OPERATIONAL TESTER
- DT AND OT EVALUATORS
- LIFE CYCLE SOFTWARE SUPPORT CENTER

WHAT ARE OUR EXPECTATIONS?

- UNIFIED COMMAND-WIDE APPROACH TO SOFTWARE TESTING
- NO SURPRISES TO PM'S
- AUDIT TRAIL FOR SOFTWARE TESTING (MONEY, MANPOWER, TOOLS)
- EARLY INVOLVEMENT FOR TESTER AND EVALUATOR
- INCREASE TESTER'S KNOWLEDGE OF SYSTEM
- FACILITATE TIMELY IDENTIFICATION OF NECESSARY TEST TOOLS
- INCREASE APPROPRIATENESS/QUALITY OF TESTS
- INCREASE QUALITY OF TEST DATA/TEST REPORTS
- INCREASE QUALITY OF EVALUATION
The Air Force Operational Test and Evaluation Center (AFOTEC) is the USAF independent test agency. It is responsible for testing new and modified weapons systems under realistic, operational conditions. The Software Evaluation Division (LG5) is tasked with evaluating the operational effectiveness and suitability of mission critical computer resources (MCCR) embedded within those systems. This presentation describes the requirement for operational test and evaluation (OT&E) of MCCR software and AFOTEC's role in managing the OT&E program.

AFOTEC/LG5 was organized to meet the intent of DoD 5000.3 which requires the development of software performance objectives, operational testing of software, and participation by OT&E agencies during planning and development activities. At the Center, software test managers plan for OT&E of the MCCR software. They estimate test requirements and write the software portion of the OT&E test plan. They devise an organizational structure for software testing and insure the test team is adequately staffed. The Deputy Test Director for Software Evaluation (DSE) is a key AFOTEC position on the field test team. The DSE is supported by software evaluators from the eventual using, supporting and training organizations, to collect, evaluate, and analyze software data during test. The Software Test Manager works closely with the test team during and after test to prepare the interim and final reports.

AFOTEC/LG5 has published guidelines, AFOTECP 800-2 Volumes 1 and 2, describing the activities of the software test manager and deputy for software evaluation. Additionally, Volumes 3 through 5 of the same series support the evaluation of source listings and documentation, the operator-machine interface, and the software support resources. These guidelines are "living" documents, responsive to a strong lessons learned program, and provide a crossfeed of information between space/strategic, tactical/avionic, and C3I software OT&E programs.

The Operational Test Center is an active participant in software test and evaluation. AFOTEC/LG5 currently uses standardized assessment tools for software maintainability, operator-machine interface, and software support resources. Future methodologies under development are aimed at evaluating computer system security, performance, and software risk analysis.
Major Frederick J. Foster

Major Foster is a branch chief with the Software Evaluation Division, Air Force Operational Test and Evaluation Center. Since 1970 he has participated in designing and testing management information and mission critical computer systems for state government and military applications. He has a Bachelor of Science in Systems Engineering and a Master of Business Administration. Major Foster is a senior pilot and a graduate of the Air Command and Staff College.
SOFTWARE EVALUATION DIVISION

SOFTWARE TEST AND EVALUATION

→ BACKGROUND

- AFOTEC ORGANIZATION FOR SOFTWARE QT&E
- AFOTEC APPROACH TO SOFTWARE QT&E
- FUTURE EFFORTS IN SOFTWARE QT&E
BACKGROUND

THE 1974 DEFENSE SCIENCE TASK FORCE ON TEST AND EVALUATION

THE TASK FORCE FOUND THAT:

"WHEREAS THE HARDWARE DEVELOPMENT WAS FOR THE MOST PART SCHEDULED, MONITORED, TESTED, AND REGULARLY EVALUATED, THE SOFTWARE DEVELOPMENT WAS NOT."

SOFTWARE DEVELOPMENT SEQUENCE

SOFTWARE ERROR OCCURRENCE AND DISCOVERY
AFOTEC OT&E DIRECTION

DOD 5000.3 (26 DEC 79)

- PERFORMANCE OBJECTIVES SHALL BE ESTABLISHED FOR SOFTWARE DURING EACH SYSTEM ACQUISITION PHASE

- SOFTWARE SHALL UNDERGO OPERATIONAL TESTING... UTILIZING TYPICAL OPERATOR PERSONNEL

- OT&E AGENCIES SHALL PARTICIPATE IN SOFTWARE PLANNING AND DEVELOPMENT TO ENSURE CONSIDERATION (OF THE) OPERATIONAL ENVIRONMENT AND EARLY DEVELOPMENT OF OPERATIONAL TEST OBJECTIVES

AFR 80-14 (12 SEP 80)

- OPERATIONAL TESTING OF SOFTWARE MUST EXAMINE ITS FUNCTIONAL PERFORMANCE, THE INTERFACE BETWEEN OPERATOR AND MACHINE, AND ITS MAINTAINABILITY.

SOFTWARE TEST AND EVALUATION

- BACKGROUND

- AFOTEC ORGANIZATION FOR SOFTWARE OT&E

- AFOTEC APPROACH TO SOFTWARE OT&E

- FUTURE EFFORTS IN SOFTWARE OT&E
MISSION STATEMENT

RESPONSIBLE FOR PROVIDING SOFTWARE EXPERTISE FOR PLANNING, EVALUATING, AND REPORTING THE OPERATIONAL EFFECTIVENESS AND SUITABILITY OF AIR FORCE SYSTEMS WITH EMBEDDED COMPUTERS.
SOFTWARE TEST AND EVALUATION

• BACKGROUND

• AFOTEC ORGANIZATION FOR SOFTWARE OT&E

• AFOTEC APPROACH TO SOFTWARE OT&E

• FUTURE EFFORTS IN SOFTWARE OT&E
SOFTWARE EVALUATION

- TEST PREPARATION
  - EARLY PLANNING WITH IMPLEMENTING, USING, SUPPORTING AGENCIES
  - PREPARE OBJECTIVES, MEASURES, METHODOLOGY
  - DESIGN REVIEWS, CRWG, TPWG

- TEST CONDUCT
  - IN-PLANT TESTING
  - ON-SITE TESTING

- EVALUATION
  - TEST DATA ANALYSIS
  - TEST DATA EVALUATION
  - REPORT PREPARATION

S/W OT&E GUIDELINES

- SOFTWARE TEST MANAGER'S GUIDE
- GUIDE FOR THE TEST TEAM
  DEPUTY FOR SOFTWARE EVALUATION
- SOFTWARE MAINTAINABILITY EVALUATOR'S GUIDE
- SOFTWARE OPERATOR-MACHINE INTERFACE EVALUATOR'S GUIDE
- SOFTWARE SUPPORT RESOURCES EVALUATION GUIDE
SYSTEM QT&E

WEAPON SYSTEM

OPERATIONAL EFFECTIVENESS

SOFTWARE EVALUATION

OPERATIONAL SUITABILITY

- DOES THE SOFTWARE DEGRADE SYSTEM PERFORMANCE?
  - TEST DESIGN
  - SYSTEM FAILURES

- IS THE SOFTWARE EASY TO USE?
  - OPERATOR-MACHINE INTERFACE SOFTWARE

- IS THE SOFTWARE EASY TO CHANGE?

SOFTWARE OPERATOR-MACHINE INTERFACE EVALUATION

- USES STANDARD STRUCTURED QUESTIONNAIRE
- USES SYSTEM OPERATORS AS EVALUATORS
- DETERMINES PRESENCE/ABSENCE OF DESIRABLE ATTRIBUTES
- RESULTS ARE QUANTITATIVE VALUES
SAMPLE QUESTION

Q: 6. OPERATOR INPUT ERRORS ARE DETECTED, AND THE CAUSE OF THE ERROR IS DISPLAYED TO THE OPERATOR.

TEST FACTOR OR CHARACTERISTIC: ASSURABILITY

RESPONSE (CHOOSE ONE) POINTS

COMPLETELY AGREE 6
STRONGLY AGREE 5
GENERALLY AGREE 4
GENERALLY DISAGREE 3
STRONGLY DISAGREE 2
COMPLETELY DISAGREE 1

ELEMENTS OF SOFTWARE SUPPORTABILITY

SOFTWARE SUPPORTABILITY

MAINTAINABILITY
DESIGN CONSIDERATIONS
- MODULARITY
- DESCRIPTIVENESS
- CONSISTENCY
- SIMPLICITY
- EXPANDABILITY
- INSTRUMENTATION

SOFTWARE SUPPORT RESOURCES
- PERSONNEL
- SUPPORT SYSTEMS
- FACILITIES
SOFTWARE MAINTAINABILITY EVALUATION METHOD

APPROACH

- Evaluate documentation and source listings
- Use standard questionnaires for all software
- Use evaluators from eventual support agency

SOFTWARE MAINTAINABILITY EVALUATION METHOD

- Questions are organized by desirable characteristics:
  - Modularity
  - Descriptiveness
  - Consistency
  - Simplicity
  - Expandability
  - Instrumentation
- 10 to 15 questions per characteristic
SAMPLE QUESTION

Q: S-65 _THE NUMBER OF EXECUTABLE STATEMENTS IN THIS MODULE IS MANAGEABLE._

TEST FACTOR OR CHARACTERISTIC: SIMPLICITY

RESPONSE (CHOOSE ONE):

- COMPLETELY AGREE
- STRONGLY AGREE
- GENERALLY AGREE
- GENERALLY DISAGREE
- STRONGLY DISAGREE
- COMPLETELY DISAGREE

POINTS:

- 6
- 5
- 4
- 3
- 2
- 1

ELEMENTS OF SOFTWARE SUPPORTABILITY

SOFTWARE SUPPORTABILITY

MAINTAINABILITY

- MODULARITY
- DESCRIPTIVENESS
- CONSISTENCY
- SIMPLICITY
- EXPANDABILITY
- INSTRUMENTATION

SOFTWARE SUPPORT RESOURCES

- PERSONNEL
- SUPPORT SYSTEMS
- FACILITIES
SOFTWARE SUPPORT RESOURCES EVALUATION METHOD

APPROACH

- EVALUATE EXISTING OR PLANS FOR SUPPORT FACILITY
- DETERMINE IF NECESSARY SUPPORT ASSETS ARE/WILL BE ALLOCATED
- USE A TAILORED QUESTIONNAIRE FOR EACH SYSTEM
- USE EVALUATORS FROM EVENTUAL SUPPORT AGENCY

SOFTWARE SUPPORT RESOURCES EVALUATION METHOD QUESTIONNAIRE HIERARCHY.
SOFTWARE TEST AND EVALUATION

OTHER EVALUATION EFFORTS

- IV&V
- COMPUTER TIMING AND SIZING
- SOFTWARE MATURITY

SOFTWARE TEST AND EVALUATION

- BACKGROUND
- AFOTEC ORGANIZATION FOR SOFTWARE OT&E
- AFOTEC APPROACH TO SOFTWARE OT&E
  ➔ FUTURE EFFORTS IN SOFTWARE OT&E
ACTIVE PARTICIPATION WITH SOFTWARE COMMUNITY

- RADG
- STEP
- ADA
  - ADA/JUG
  - AFWAL E&V
- STARS
- IEEE
- MC'S'

CURRENT INITIATIVES

- AFOTEC DEVELOPING METHODOLOGIES
  - SOFTWARE ERROR TRACKING
  - SUPPORTABILITY RISK ASSESSMENT
  - COMPUTER SYSTEM SECURITY
  - DIAGNOSTICS

41
SUMMARY

- SOFTWARE EVALUATION TOOLS EXIST
  - MAINTAINABILITY
  - COMPUTER SUPPORT RESOURCES
  - OPERATOR-MACHINE INTERFACE SOFTWARE
Effect of Testing Tools on Aegis Lifetime Support

P. Graham O'Neil

This report is on the use of tools in large scale system testing. This will include an overview of the system, the environment and the schedule. Attention will be given to areas in which improvements could be made. Cases where results can be compared to theory will be presented and assessments of tool utilization and productivity will be presented. The focus of the testing effort was in the integration phase and the goals were to provide a stable, highly reliable, well understood set of computer programs to control the system.
P. Graham O'Neil

P. Graham O'Neil received the BS degree with majors in mathematics, psychology and English from the University of Richmond. After graduation, he joined the Naval Surface Weapons Center at Dahlgren, VA. During this 17 years, he was responsible for generating six degree of freedom simulations for aircraft, missiles, and projectiles. For seven years, he was the Combat System Engineer for the AEGIS Combat System. He is presently Senior Principal System Engineer at Sanders Associates, Nashua, NH. His current interests are software tools, knowledge base system development and large scale system engineering.
EFFECT OF TESTING TOOLS ON AEGIS LIFETIME SUPPORT

OUTLINE

I. INTRODUCTION

II. CHARACTERISTICS OF EACH PHASE

III. NET RESULTS FOR TWO SHIPS

IV. SHORTFALL EXAMINATION

V. ASSESSMENT OF EFFORTS

VI. SUMMARY
<table>
<thead>
<tr>
<th>ELEMENT</th>
<th>SIZE (K WORDS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ADS</td>
<td>60</td>
</tr>
<tr>
<td>ACTS</td>
<td>60</td>
</tr>
<tr>
<td>CND</td>
<td>205</td>
</tr>
<tr>
<td>PCS</td>
<td>40</td>
</tr>
<tr>
<td>ORTS</td>
<td>64</td>
</tr>
<tr>
<td>SPY</td>
<td>200</td>
</tr>
<tr>
<td>NCS</td>
<td>195</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TOOLS</th>
<th>SIZE (K WORDS)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISS</td>
<td>350</td>
</tr>
<tr>
<td>ISA</td>
<td>260+</td>
</tr>
<tr>
<td>ADAR</td>
<td>200</td>
</tr>
<tr>
<td>PMTA</td>
<td>30</td>
</tr>
<tr>
<td>IVG</td>
<td>180</td>
</tr>
</tbody>
</table>

TACTICAL PROGRAM AND SUPPORTING TOOL SIZE

OUTLINE

I. INTRODUCTION

II. CHARACTERISTICS OF EACH PHASE

III. NET RESULTS FOR TWO SHIPS

IV. SHORTFALL EXAMINATION

V. ASSESSMENT OF EFFORTS

VI. SUMMARY
REQUIREMENTS REVIEW CHARACTERISTICS

1. Used PSL/PSA to construct as-specified and as-built data bases.

2. Data bases reflect logical and physical structure.

RESULTS

1. Uncovered omissions in requirements.

2. Was used to determine testing requirements of the simulator system.

3. Supported testing efforts.

4. Being used for lifetime support.

DESIGN CHARACTERISTICS

1. Top-down design of the tactical programs.

2. Simulator family used skeleton design.

RESULTS

1. Design integrity of system.

2. Ease of analysis to a desired level.

3. Simulator family provided easily transportable/reusable software.

4. Simulator family provided common control from script processor.
CONSTRUCTION CHARACTERISTICS

1. SIMULATOR USAGE COMBINED WITH REAL WORLD STIMULUS.
2. TESTING SUPPORT SYSTEM USED IN CONJUNCTION WITH SIMULATOR.
3. STRESS TESTING FOCUSED ON SATURATION.
4. DATA DICTIONARIES.
5. COUNTER SOLUTION PROGRAMS NOT AVAILABLE.

RESULTS

1. SIMULATION COMPLEMENTED REAL WORLD FOR STRESS TESTING.
2. TESTING SUPPORT SYSTEM USED FOR ARCHIVAL OF TEST SETS, RESULTS AND REPORTS.
3. TESTING SUPPORT SYSTEM USED FOR TEST DATA GENERATION AND EXPECTED RESULTS.
4. HIGH LEVEL OF AUTOMATION, BUT MORE WAS POSSIBLE.

INTEGRATION CHARACTERISTICS

1. REAL EQUIPMENT.
2. AUTOMATED SCHEDULING AND COMPLETION PROGRAMS FOR TRACKING PROGRESS.
3. ACCESSIBLE PROBLEM REPORT DATA BASE.

RESULTS

1. SOME SURPRISES AND SHORTFALLS WITH REAL EQUIPMENT.
2. SIMULATOR USED AS WORK AROUND UNTIL CORRECTIONS WERE SUCCESSFUL.
3. SOME INTUITIVE TESTERS HIGHLY EFFICIENT.
4. GROWING PAINS.
OPERATION AND MAINTENANCE CHARACTERISTICS

1. FOCUS ON REGRESSION TESTING AND PROGRAM RELIABILITY.
2. MULTIPLE USERS/SITES AND BASELINES.
3. AUTOMATIC DATA COLLECTION SYSTEM FOR SOURCE CODE.

RESULTS

1. UNINTENTIONAL ERROR SEEDING AND PROGRAM MUTATIONS.
2. NEW TOOLS TO ANALYZE RELIABILITY.

OUTLINE

I. INTRODUCTION
II. CHARACTERISTICS OF EACH PHASE
III. NET RESULTS FOR TWO SHIPS
IV. SHORTFALL EXAMINATION
V. ASSESSMENT OF EFFORTS
VI. SUMMARY
OUTLINE

I. INTRODUCTION

II. CHARACTERISTICS OF EACH PHASE

III. NET RESULTS FOR TWO SHIPS

IV. SHORTFALL EXAMINATION

V. ASSESSMENT OF EFFORTS

VI. SUMMARY

INTEGRATION SHORTFALLS - FIRE CONTROL

PROBLEM CONTRIBUTORS

1) INCOMPLETE AND INCONSISTENT REQUIREMENTS/DESIGN.

2) LATE INTRODUCTION OF REAL EQUIPMENT.

3) DECEPTIVE SIMULATORS.

4) LACK OF EXCEPTION TESTING.

5) MASSIVE AMOUNTS OF UNORGANIZED DATA.

RESOLUTION CONTRIBUTORS

1) FUNCTIONAL DESIGN INTENT CAPTURED.

2) STATE DIAGRAM KEYED EXCEPTION TESTING.

3) SHORTFALL WAS NO SURPRISE.

4) DEDICATED PEOPLE.
INTEGRATION SHORTFALLS - READINESS DATA BASE

PROBLEM CONTRIBUTORS

1) BLURRED SUBSYSTEM BOUNDARIES.

2) LATE INTRODUCTION OF COMPLETE EQUIPMENT SUITE.

3) SIMULATOR LACK.

4) COMPUTER IS TIME, I/O, AND CORE BOUND.

5) TASK UNDERESTIMATED.

OUTLINE

I. INTRODUCTION

II. CHARACTERISTICS OF EACH PHASE

III. Net RESULTS FOR TWO SHIPS

IV. SHORTFALL EXAMINATION

V. ASSESSMENT OF EFFORTS

VI. SUMMARY
AEGIS INTEGRATED DATA BASE SYSTEM (AIDS)

1. ANALYZED REQUIREMENTS/SPECIFICATIONS BY EVALUATING TRACEABILITY, COMPLETENESS, CONSISTENCY AND DATA USAGE.

2. DETECTED TEST ANOMALIES BY:
   - ANALYZING FUNCTIONAL PATHS.
   - IDENTIFYING MISSING THRESHOLDS OR TOLERANCES.

3. DESIGNED TEST CONDITIONS TO REVEAL PROBLEMS BY:
   - DETERMINING AFFECTED PROCESSES/PROCEDURES.
   - IDENTIFYING DATA USED OR MODIFIED.

AIDS PROBLEM ANALYSIS

ISOLATES CAUSE BY IDENTIFYING
   FUNCTIONAL PATH.
   TRIGGERED EVENTS.
   DATA SET/USED.
   INTERACTION OF DATA AND PROCESS.

AIDS IN RESOLUTION BY
   EVALUATING PROBLEM REPORT BY POSSIBLE CAUSE.
   DETERMINING EFFECTS OF POSSIBLE SOLUTIONS.
   IDENTIFIES RE-TEST REQUIREMENTS.
AEGIS EXPERT SYSTEM

TEST IMPLEMENTATION BREAKOUT
RELIABILITY EFFORT FOCUS

1. USEFUL TO MANAGE THE DELIVERY/DEVELOPMENT OF QUALITY SOFTWARE.

2. IDENTIFY MODULES WITH INTRINSIC WEAKNESS.

3. IMPROVE MTBM/CE.

4. IMPROVE PROGRAMMER/TESTER PRODUCTIVITY.

5. DUAL STUDY OF PROBLEM REPORT LOG AND CODE CHANGES.
### FIELD

<table>
<thead>
<tr>
<th></th>
<th>SELECT</th>
<th>SORT</th>
<th>PRINT</th>
</tr>
</thead>
<tbody>
<tr>
<td>CPPR Number</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Log Date</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Originating Site</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Element</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Problem Type</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Priority</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Baseline</td>
<td>X</td>
<td>X</td>
<td>X</td>
</tr>
<tr>
<td>Level</td>
<td></td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Technical Evaluation</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Code Version</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>SCCB Status</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Element Code Status</td>
<td>X</td>
<td></td>
<td>X</td>
</tr>
<tr>
<td>Affected Function</td>
<td></td>
<td></td>
<td>X</td>
</tr>
</tbody>
</table>

**MINIMUM DATA COLLECTION CAPABILITIES FOR RELIABILITY ASSESSMENT**
### CPFR Error Count Data

**CG-47 Fire Control System**

<table>
<thead>
<tr>
<th>DATE</th>
<th>PRIORITY</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>HIGH</td>
<td>MEDIUM</td>
</tr>
<tr>
<td>APRIL 1982</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>MAY 1982</td>
<td>8</td>
<td>4</td>
</tr>
<tr>
<td>JUNE 1982</td>
<td>5</td>
<td>3</td>
</tr>
<tr>
<td>JULY 1982</td>
<td>14</td>
<td>5</td>
</tr>
<tr>
<td>AUGUST 1982</td>
<td>8</td>
<td>1</td>
</tr>
<tr>
<td>SEPTEMBER 1982</td>
<td>11</td>
<td>1</td>
</tr>
<tr>
<td>OCTOBER 1982</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>NOVEMBER 1982</td>
<td>4</td>
<td>7</td>
</tr>
<tr>
<td>DECEMBER 1982</td>
<td>6</td>
<td>1</td>
</tr>
<tr>
<td>JANUARY 1983</td>
<td>2</td>
<td>0</td>
</tr>
<tr>
<td>FEBRUARY 1983</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>MARCH 1983</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

![Schneidewind Model](image)

**Schneidewind Model**

The graph shows a decreasing trend in error counts over time, with the highest counts in July and the lowest in March. The trend line suggests a downward trend in errors as the intervals increase.
## Expected Number of Errors in Future Testing Intervals

<table>
<thead>
<tr>
<th>Future Interval</th>
<th>Generalized Poisson</th>
<th>Nonhomogeneous Poisson Process</th>
<th>Brooks &amp; Motley</th>
<th>Schneidewind</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>4.92</td>
<td>5.51</td>
<td>5.67</td>
<td>5.51</td>
</tr>
<tr>
<td>2</td>
<td>4.26</td>
<td>5.11</td>
<td>5.18</td>
<td>5.11</td>
</tr>
<tr>
<td>3</td>
<td>4.04</td>
<td>4.74</td>
<td>4.74</td>
<td>4.74</td>
</tr>
<tr>
<td>4</td>
<td>3.91</td>
<td>4.40</td>
<td>4.33</td>
<td>4.40</td>
</tr>
<tr>
<td>5</td>
<td>3.81</td>
<td>4.08</td>
<td>3.95</td>
<td>4.08</td>
</tr>
<tr>
<td>6</td>
<td>3.73</td>
<td>3.79</td>
<td>3.62</td>
<td>3.79</td>
</tr>
</tbody>
</table>

## Reliability Results

I. Problem modules could be predicted by several metrics.

II. Little correlation between modules with numerous problems reported and those with numerous problems fixed.

III. Reliability predictors appear accurate enough to use for decisions.

IV. Statistics are required as compensation for evaluation variances.

V. Using these statistics as a basis for taking action to improve programmer/tester performance has not been accomplished.
OUTLINE

I. INTRODUCTION

II. CHARACTERISTICS OF EACH PHASE

III. NET RESULTS FOR TWO SHIPS

IV. SHORTFALL EXAMINATION

V. ASSESSMENT OF EFFORTS

VI. SUMMARY

SUMMARY

1. Problem reports from each ship are half the previous numbers.
2. Analysis of requirements and design provides a good indicator of functions/modules with significant problems.
3. Extensive counter solution data is not an absolute requirement.
4. In an extensive system, consistent error reporting is unlikely to be achieved.
5. Definite needs exist for automated tools.
6. Usage of these tools must be planned and encouraged to have effect.
7. Where sufficient level of simulation fidelity cannot be achieved, other action is required.
8. Error seeds in programs were discovered in short time spans.
A-7E Test Program

Paul Clements

The paper will describe the Naval Research Laboratory's Software Cost Reduction (SCR) project, a project to investigate the application of advanced software engineering techniques to real-time embedded software systems with tight requirements and resource constraints. Under the direction of Dr. David Parnas, SCR is providing a model of software development, including model documents, specifications, code, and procedures, for other software projects to emulate. The example application chosen is the Navy's A-7E aircraft.

SCR testing philosophy is that design and specification of software is as important to testing as is the execution of after-the-fact test cases. SCR testing embodies the following areas:

- specification in the software requirement of required exception-handling;
- specification of undesired events in all module interface design documents;
- black box module testing, based on the module interface specifications;
- clear box module testing based on module implementations;
- subset testing; and
- hand-over to standard Navy Test procedures.

Finally, the paper will describe promising indications of drastically reduced integration and testing time achieved as a result of implementing a small subset of the system.
Paul Clements

Paul Clements is the project manager for the Naval Research Laboratory's Software Cost Reduction (SCR) project. Under the guidance of Dr. David Parnas, SCR is investigating the application of advanced software engineering techniques to real-time embedded software systems with tight requirements and resource constraints. Mr. Clements graduated from the University of North Carolina at Chapel Hill in 1977 with a BS degree in Mathematical Sciences, and in 1980 with an MS degree in Computer Science.
PROBLEM

- NAVY SOFTWARE IS EXPENSIVE TO MAINTAIN
  - Changes Required but Risky
  - Difficult to Understand
  - Poorly Documented
  - Poorly Structured
  - Difficult to Validate
  - Difficult to Train New People

PROBLEM

- NAVY SOFTWARE IS UNRELIABLE
  - Delivered software contains errors
  - Changes often introduce errors
  - No good theory for software QA
MAJOR PROBLEM CAUSES

- Many Arbitrary Details
  - Syntax & semantics of programming languages
  - Procedures needed to communicate with devices; e.g., radar
  - Format of input data; e.g., message format
  - Communications protocols
- Rapidly changing requirements
- Long lifetime of systems
- Lack of good models
- Lack of appropriate formalisms

SOFTWARE COST REDUCTION

TEST & EVALUATION

SCR OBJECTIVE: REBUILD A-7E OPERATIONAL FLIGHT PROGRAM (OFP) USING RECENT SOFTWARE ENGINEERING TECHNOLOGY

- PRODUCE MODEL DOCUMENTS
- PRODUCE A WORKING OFP
- FLY THE PLANE
- COMPARE OFPs
USEFUL PRINCIPLES

- Separation of Concerns
  - A principle for isolating independent problems
  - Worker should only have to think about one thing at a time
  - Examples: reading input data, detecting events, allocating resources (scheduling)

USEFUL PRINCIPLES

- Abstraction
  - A principle for exploiting commonalities
  - Abstract means conceived apart from special cases
  - Many-to-one mapping
  - Examples: differential equations, graphs, data types
COROLLARIES

- Process Structuring (Cooperating Sequential Processes)
  - Separation of concerns to permit scheduling independent of function
- Undesired Events
  - Separation of concerns to permit appropriate error handling

REQUIREMENTS SPECIFICATION - COMPLETENESS

- Computer Characteristics
- Input and Output Interfaces
- System States and History
- Output Values
- Timing
- Accuracy
- Likely Changes
  - Exception-handling
  - Subsets
  - Other likely changes
- Dictionary
SOFTWARE DESIGN

- The (Information-Hiding) Modular Structure

- Separate Concerns of:
  - Hardware details (Hardware-Hiding Module)
  - Requirements (Behavior-Hiding Module)
  - Software design decisions (Software Decision Module)

FIRST-LEVEL MODULES

- Hardware-Hiding Module
- Behavior-Hiding Module
- Software-Decision-Hiding Module
SECOND-LEVEL MODULES

- Hardware-Hiding Module
  - Extended Computer Module
  - Device Interface Module

- Behavior-Hiding Module
  - Function Driver Module
  - Shared Services Module

- Software-Decision-Hiding Module
  - Physical Models Module
  - Data Banker Module
  - Application Data Types Module
  - System Generation Module
  - Software Utility Module

THIRD-LEVEL MODULES

Extended Computer Module
  - Data Type Module
  - Input/Output Module
  - Parallelism Control Module
  - Sequence Control Module
  - State Module
  - Test Module
  - Timer Module
  - Virtual Memory Module
  - Interrupt Handler Module

Device Interface Module
  - Air Data Computer
  - Angle of Attack Sensor
  - Audible Signal Device
  - Computer Fail Device
  - Doppler Radar Set
  - Flight Information Displays
  - Forward Looking Radar
  - Head-Up Display
  - Inertial Measurement Set
  - Panel
  - Projected Map Display Set
  - Radar Altimeter
  - Shipboard Inertial Navigation
  - Slow Control
  - Switch Bank
  - TACAN
  - Visual Indicators
  - Waypoint Information System
  - Weapon Characteristics
  - Weapon Release System
  - Weight on Gear

Function Driver Module
  - Air Data Computer Functions
  - Audible Signal Functions
  - Computer Fail Signal Functions
  - Doppler Radar Functions
  - Flight Information Display Functions
  - Forward Looking Radar Functions
  - Head-Up Display Functions
  - Inertial Measurement Set Functions
  - Panel Functions
  - Projected Map Display Set Functions
  - SINS Functions
  - Visual Indicator Functions
  - Weapon Release Functions

Shared Services Module
  - Mode Determination Module
  - Stage Director Module
  - Shared Subroutine Module
  - System Value Module
  - Panel I/O Support Module
  - Input Display Format
  - Configuration

Application Data Type Module
  - Numeric Data Type Module
  - State Transition Event Module

Physical Models Module
  - Earth Characteristics Module
  - Aircraft Motion Module
  - Spatial Relations Module
  - Human Factors Module
  - Weapon Behavior Module
  - Target Behavior Module
  - Filter Behavior Module

Data Banker Module
  (one submodule per producing submodule)

System Generation Module
  - System Generation Parameter Module
  - Module Version Selection Module
  - Subset Selection Module
  - Support Software Module

Software Utility Module
SOFTWARE DESIGN

- Module Interface Specification
  - Apply abstraction to produce abstract interface
  - Users see (use) only externally visible functions
  - Redundancy for error checking (review)
  - Formalism for precision
  - UE specification to highlight significant error handling decisions

EXAMPLE OF ABSTRACT INTERFACE SPECIFICATION

- Chosen from Device Interface Module

DI.DRS
DOPPLER RADAR SET

DI.DRS.1. INTRODUCTION
The Doppler Radar Set (DRS) is a sensor that measures aircraft ground speed and drift angle during flight.

DI.DRS.2. INTERFACE OVERVIEW

DI.DRS.2.1 ACCESS PROGRAM TABLE

<table>
<thead>
<tr>
<th>Program Name</th>
<th>Parm type</th>
<th>Parm info</th>
<th>Undesired events</th>
</tr>
</thead>
<tbody>
<tr>
<td>+G_DRS_MODE+</td>
<td>pl:drs_mode:0</td>
<td>+DRS mode+</td>
<td>None</td>
</tr>
</tbody>
</table>

| +G_DRS_DRIFT_ANGLE+   | pl:angle:0  | !=drift angle DRS+ | Wrong DRS model  |
| +G_DRS_GROUND_SPEED+  | pl:speed;0   | !=ground speed DRS+ |                  |
| +G_DRS_RELIABILITY+   | pl:boolean;0 | !=DRS reliable+   |                  |
| +G_DRS_TEST_RESULTS+  | pl:boolean;0 | !=DRS test result+ |                  |
| +START_DRS+           | none        |                      |                  |
| +STOP_DRS+            | none        |                      |                  |

Program Effects

+START_DRS+
Changes the mode from OFF to one of the other modes.
The mode entered depends on conditions beyond the control of the software.

+STOP_DRS+
Sets the mode to OFF.

DI.DRS.2.2. EVENTS SIGNALLED

ST(=DRS mode changed+)
ST/OFT=FT(!=DRS reliable+)

DI.DRS.2.3 MODES OF THE MODULE
The modes of this module are OFF, OPERATE, MEMORY, STANDBY, and TEST. They are mutually exclusive. The following table defines what program calls are legal in what mode; L=legal, NL=not legal.
SCR TEST & EVALUATION

DEVELOPMENT OVERVIEW

* DESIGN THE MODULAR STRUCTURE
* SPECIFY THE MODULE INTERFACES
* DESIGN SUBSETS
  * CODE SUBSET FUNCTIONS
  * TEST (PARTIAL MODULES)
  * TEST SUBSET ON TC-2 SIMULATOR
* DELIVER COMPLETE OFP FOR FLIGHT TEST
* COMPARE

APPLICATION OF PRINCIPLES

DEFINE IMPLEMENTATION-INDEPENDENT REQUIREMENTS

ORGANIZE SYSTEM INTO MODULES

DEFINE MODULE INTERFACES  DESIGN PROCESS STRUCTURE

WRITE PSEUDO-CODE  DEFINE USES

WRITE IMPLEMENTATION CODE

TEST ON FLIGHT SIMULATOR
SCR TEST & EVALUATION

MODULE TEST PHILOSOPHY

* BLACK BOX TESTS

- USE ONLY INTERFACE SPECIFICATIONS IN CONSTRUCTING TESTS

- TEST INTERFACE FUNCTIONS OVER INPUT DOMAIN

- CAUSE FUNCTIONS TO PRODUCE RESULTS OVER OUTPUT RANGE

- (PSEUDO) CONTINUOUS DOMAIN INPUT
  
  = USE BOUNDARY VALUES & TYPICAL VALUES

- DISCRETE INPUT DOMAIN
  
  = USE EVERY INPUT VALUE, IF POSSIBLE, OTHERWISE SPECIAL VALUES & AT LEAST ONE NON-SPECIAL VALUE

- (PSEUDO) CONTINUOUS OUTPUT
  
  = PRODUCE BOUNDARY VALUES & TYPICAL VALUES

- DISCRETE OUTPUT
  
  = PRODUCE EVERY VALUE, IF POSSIBLE, OTHERWISE SPECIAL VALUES & AT LEAST ONE NON-SPECIAL VALUE
SCR TEST & EVALUATION

MODULE TEST PHILOSOPHY

* CLEAR BOX TESTS

- CODE READING

  = ORGANIZE PROGRAM INTO EQUIVALENCE CLASSES OF STATES

  2 STATES ARE IN THE SAME CLASS IFF ONE CAN PROVE THAT IF THE PROGRAM WORKS CORRECTLY WHEN STARTED IN ONE STATE, IT WILL WORK CORRECTLY WHEN STARTED IN THE OTHER STATE

- SELECT ONE TEST CASE FROM EACH EQUIVALENCE CLASS

SCR TEST & EVALUATION

SUBSET TESTING

* SUBSET CAN BE VIEWED AS A MODULE; APPLY MODULE TEST PRINCIPLES

- IF SUBMODULES OF THE SUBSET ARE ALL EXTERNALLY-VISIBLE, SUBSET TESTS ARE THE SAME AS MODULE TESTS (SUBSET INTERFACE IS THE UNION OF THE SUBMODULE INTERFACES)

- IF THERE ARE HIDDEN SUBMODULES, THEN THE SUBSET INTERFACE IS DIFFERENT THAN THE UNION OF THE SUBMODULE INTERFACES; SUBSET TESTS MAY BE DIFFERENT FROM MODULE TESTS
WHAT WAS TESTED AT NWC IN OCTOBER

- PRELIMINARY DEMONSTRATION SUBSET
  - EXERCISE EVERY 2nd LEVEL MODULE IN THE OFP RUNNING ON THE TC-2
  - TAKE INPUTS AND MOVE SYMBOLS ON HUD

- SUPPORT TOOL SUBSET
  - TRANSLATOR GENERATOR
  - EC TRANSLATOR

INTEGRATION AT NWC

- INTEGRATED IN 3 NIGHTS
- TOTAL OF 9 ERRORS
- NO HAND-PATCHES - ALL WERE CORRECTED IN SOURCE

NO ERROR CROSSED A MODULE BOUNDARY
CODING AND TESTING EFFORT

- CLASSICAL: 1 WORK HOUR OF TESTING FOR EVERY 1/2 WORK HOUR OF CODING
  
  \textit{(Wolverton (1974), Daly (1977), Boehm (1981))}

- IN PDS WE FOUND: 10 WORK MINUTES OF TESTING FOR EVERY 1/2 WORK HOUR OF CODING

<table>
<thead>
<tr>
<th>CLASSICAL</th>
<th>CODING</th>
<th>TESTING</th>
</tr>
</thead>
<tbody>
<tr>
<td>SCR</td>
<td>CODING</td>
<td>TESTING</td>
</tr>
</tbody>
</table>

In 1977, the Joint Logistics Commanders chartered a Joint Policy Coordinating Group on Computer Resource Management (CRM). The mission of the CRM was to coordinate and ensure consistency in the preparation of new and revised regulations and standards, to provide recommendations on critical resource areas, and to provide a focal point for coordinating standardization programs. The CRM subsequently chartered a subgroup for Computer Software Management (CSM) to review policies, procedures, regulations, and standards relating to computer software and forward specific recommendations to the CRM on critical areas related to software acquisition and management, including software development, quality, testing, and post-development support.

The CSM has structured their activities into three projects; the software development project, the software quality project, and the post-development software support (PDSS) project. A software development cycle model was developed and is the fundamental framework for the CSM projects. The CSM has hosted three joint industry/government workshops. The first two, held in Monterey, California in 1979 and 1981, dealt with issues concerned with software development and software quality. The third workshop, held in Orlando, Florida, in 1983, dealt exclusively with PDSS issues. The results of the workshops are being incorporated into the products of the CSM projects.

The software development project consists of a Joint Regulation on the Management of Computer Resources in Defense Systems, a tri-service coordinated DoD Standard on Software Development (DoD-STD-SDS), a collection of 25 Data Item Descriptions, proposed changes to several existing Military Standards, and a guidebook for program managers on implementing the new standard. The software quality project consists of a Joint Regulation on the Software Quality Program, a DoD Standard on Software Quality (DoD-STD-SQS), and a guidebook for program managers. The post development software support project has only recently been initiated, and currently consists of an action plan which, when approved, will start several projects in the area of PDSS.
Lieutenant Commander Gehl is currently the Assistant Division Director for the Computer Resources Division at the Naval Electronic Systems Command (NAVELEX). He is the primary interface with the Embedded Computer Program Office at the Naval Material Command (NAVMAT) for matters relating to the Navy Standard Computer Program, and for NAVMAT ECR policies.

LCDR Gehl received a Master of Science degree in Computer Science from the Naval Postgraduate School in Monterey, California, prior to reporting to NAVELEX in September of 1982. Prior to his postgraduate education, he served on USS Henderson (DD 785), USS Jouett (CG 29), USS Lynde McCormick (DDG-8), and COMDESRON TWO SEVEN. He graduated from Iowa State University in 1971 with a BS degree in Electronic Engineering.
JOINT LOGISTIC COMMANDER'S SOFTWARE STANDARDIZATION PROGRAM

JOINT LOGISTICS COMMANDERS' (JLC) ORGANIZATION

DARCOM  NAVMAT  AFSC  AFLC  USMC

JOINT SECRETARIAT

JOINT POLICY COORDINATING GROUP ON COMPUTER RESOURCE MANAGEMENT (CRM) ESTABLISHED DEC 77

COMPUTER SOFTWARE MANAGEMENT (CSM) GROUP ESTABLISHED APR 78
MAJOR EVENTS

1977 — CRM FORMED
1978 — CSM FORMED
1979 — MONTEREY I
1981 — MONTEREY II
1982 — DRAFT POLICY, STANDARDS, DIDs
1983 — DRAFT QUALITY POLICY, STANDARDS
1983 — ORLANDO I

JOINT POLICY COORDINATING GROUP (JPCG) ON
COMPUTER RESOURCE MANAGEMENT (CRM)

- U.S. ARMY MATERIAL COMMAND (AMC) COL H. ARCHIBALD
- NAVAL MATERIAL COMMAND (NMC) CAPT D. BOSLAUGH (CHAIRPERSON)
- AIR FORCE LOGISTIC COMMAND (AFLC) LT COL J. HARRINGTON
- AIR FORCE SYSTEMS COMMAND (AFSC) COL K. NIDIFFER
- U.S. MARINE CORPS (USMC) MAJ K. PTACK
CRM MISSION

- To coordinate and insure consistency in the preparation of new and revised regulations and standards
- To provide recommendations on critical resource areas
- To provide a focal point for coordinating standardization programs

COMPUTER SOFTWARE MANAGEMENT (CSM) SUBGROUP

- AMC  C. Oglesby
- NMC  LCDR M. Gehl
- AFLC  O. Kvenvold
- AFSC  CAPT L. Cooper (Chairperson)
CSM SUBGROUP MISSION

TO REVIEW POLICIES, PROCEDURES, REGULATIONS, AND STANDARDS
RELATING TO COMPUTER SOFTWARE AND FORWARD SPECIFIC
RECOMMENDATIONS TO THE 1PCG-CRM ON CRITICAL AREAS RELATED
TO SOFTWARE ACQUISITION AND MANAGEMENT, INCLUDING SOFTWARE
DEVELOPMENT, QUALITY, TESTING, AND POST-DEVELOPMENT SUPPORT.

MONTEREY I
AREAS OF CONCERN

- SOFTWARE ACQUISITION POLICY
- SOFTWARE ACQUISITION AND DEVELOPMENT STANDARDS
- SOFTWARE DOCUMENTATION STANDARDS
- SOFTWARE QUALITY ASSURANCE STANDARDS
- SOFTWARE ACCEPTANCE CRITERIA
MONTEREY II AREAS OF CONCERN

- DRAFT DID REVIEW
- HARDWARE/SOFTWARE/FIRMWARE CONFIGURATION ITEM SELECTION CRITERIA
- STANDARDIZATION AND ACCREDITATION OF COMPUTER ARCHITECTURES
- SOFTWARE COST ESTIMATING
- SOFTWARE REUSEABILITY

ORLANDO I AREAS OF CONCERN

- INDUSTRY/GOVERNMENT WORKFORCE MIX
- IV & V BY SUPPORT PERSONNEL
- COST OF OWNERSHIP
- SOFTWARE SUPPORT ENVIRONMENT
- CHANGE IMPLEMENTATION
- CONFIGURATION MANAGEMENT
SOFTWARE DEVELOPMENT STANDARDIZATION PROJECT

DATA ITEM DESCRIPTIONS
DOD-STD-SDS
JOINT REGULATION MANAGEMENT OF COMPUTER RESOURCES IN DEFENSE SYSTEMS
IMPLEMENTATION PLANS
TRAINING COURSE
GUIDEBOOKS

JOINT SERVICE SOFTWARE STANDARDS: PROGRAM OBJECTIVE

- PRODUCE A COMPLETE, CONSISTENT SET OF SOFTWARE ACQUISITION AND DEVELOPMENT STANDARDS WHICH
  - ESTABLISH A WELL-DEFINED AND EASILY UNDERSTOOD SOFTWARE ACQUISITION AND DEVELOPMENT PROCESS
  - PROVIDE ADEQUATE VISIBILITY AND CONTROL MECHANISMS THROUGHOUT ALL PHASES OF ACQUISITION
  - REDUCE CONFUSION AND ELIMINATE CONFLICTS IN EXISTING SOFTWARE STANDARDS
CSM SUBGROUP APPROACH

- DEFINE SOFTWARE DEVELOPMENT FRAMEWORK
- ESTABLISH JOINT POLICY
- DEVELOP TRI-SERVICE DEVELOPMENT/DOCUMENTATION STANDARDS
- DEVELOP SOFTWARE QUALITY ASSESSMENT/MEASUREMENT STANDARDS
- DEVELOP TAILORING GUIDEBOOKS AND TRAINING COURSES
- COORDINATE/IMPLEMENT ON NEW SYSTEMS

SOFTWARE DOCUMENTATION STANDARDS

DIDS DEVELOPED FOR
- MANAGEMENT PLANS
- ENGINEERING SPECIFICATIONS/DOCUMENTATION
- TEST PLANS/PROCEDURES/REPORTS
- SUPPORT DOCUMENTATION

WILL PROVIDE COMMON DOCUMENTATION SET
SOFTWARE QUALITY STANDARDIZATION PROJECT

OBJECTIVE

- TO DEVELOP AND IMPLEMENT A SYSTEMATIC AND WELL-UNDERSTOOD TRI-SERVICE SOFTWARE QUALITY PROGRAM THAT IS FULLY INTEGRATED WITH THE SOFTWARE DEVELOPMENT PROCESS

- TO PROVIDE GUIDANCE AND TRAINING TO APPROPRIATE PERSONNEL IN THE IMPLEMENTATION OF THE SOFTWARE QUALITY PROGRAM
SOFTWARE QUALITY STANDARDIZATION PROJECT

PROGRAM THRUST

- END-USE ORIENTED
- NOT ORGANIZATIONALLY LIMITED
- PLACES EMPHASIS ON PROCESS AS WELL AS PRODUCT
- PLACES EMPHASIS ON PROJECT FRONT END
- TOTAL PROGRAM
  - SOFTWARE QUALITY PROGRAM EMBEDDED WITHIN MANAGEMENT OF SOFTWARE PROJECT
  - PLACES RESPONSIBILITY FOR PRODUCT QUALITY ON PROGRAM MANAGER
- SOFTWARE QUALITY PROGRAM ≠ SQA

SOFTWARE QUALITY STANDARDIZATION PROJECT

SOFTWARE QUALITY PROGRAM

- A SOFTWARE QUALITY PROGRAM SHALL BE IMPLEMENTED AND INCLUDE ACTIVITIES NECESSARY TO ACHIEVE A DEFINITION OF QUALITY REQUIREMENTS, THE METHODS TO BUILD QUALITY INTO THE PRODUCT AND MAINTAIN THAT LEVEL OF QUALITY, AND THE DETERMINATION OF THE QUALITY LEVEL ATTAINED IN THE PRODUCT
## SOFTWARE QUALITY STANDARDIZATION PROJECT

### SOFTWARE QUALITY REQUIREMENTS (STANDARDS)

<table>
<thead>
<tr>
<th>SDS</th>
<th>SQAM</th>
<th>REQUIREMENTS</th>
<th>STANDARD</th>
<th>FAR</th>
<th>MIL-Q-9858A</th>
<th>MIL-S-52779A</th>
<th>DOD-STD-1679A</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.4</td>
<td>4.8</td>
<td>SUBCONTRACTOR CONTROLS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>4.8</td>
<td>- ESTABLISH</td>
<td></td>
<td>5.2</td>
<td>3.3</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>4.4</td>
<td>4.8</td>
<td>- EVALUATE</td>
<td></td>
<td>5.1</td>
<td>3.3</td>
<td>4.6</td>
<td></td>
</tr>
<tr>
<td>5.7.1.4</td>
<td>-</td>
<td>CORRECTIVE ACTION SYSTEM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>5.7.1.4</td>
<td>-</td>
<td>- ESTABLISH</td>
<td></td>
<td>3.5</td>
<td>3.2.9</td>
<td>5.8.4(p)</td>
<td></td>
</tr>
<tr>
<td>5.7.1.4</td>
<td>-</td>
<td>- IMPLEMENT</td>
<td></td>
<td>3.5</td>
<td>3.2.9</td>
<td>5.8.4(p)</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>4.5</td>
<td>- EVALUATE/ASSESS</td>
<td></td>
<td>6.5</td>
<td>3.2.9</td>
<td>5.9.10</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>5.7.1</td>
<td>PERFORM ACCEPTANCE INSPECTION</td>
<td>46.5</td>
<td>5.8.1</td>
<td>6.3</td>
<td>5.10 (p)</td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>5.7.1</td>
<td>- LIBRARY CONTROLS</td>
<td>46.202-1</td>
<td></td>
<td>3.2.5</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>- ESTABLISH</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>- PERFORM</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-</td>
<td>5.8.1</td>
<td>- EVALUATE/ASSESS</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

## SOFTWARE QUALITY STANDARDIZATION PROJECT

### EXPECTED RESULTS

- **ESTABLISHES A FULLY INTEGRATED SYSTEMATIC AND WELL-DEFINED SOFTWARE QUALITY PROGRAM**
- **INSTITUTIONALIZES THE TRAINING PROGRAMS AND GUIDANCE ESSENTIAL TO IMPLEMENTATION**
- **PROVIDES COMMONALITY AMONG SERVICES**
  - FEWER REGULATIONS AND STANDARDS
  - FEWER DATA ITEM DESCRIPTIONS
- **REDUCES COST TO CONTRACTORS AND GOVERNMENT**
  - IMPROVES CONTRACTOR PRODUCTIVITY
  - ALLOWS CONTRACTORS TO STANDARDIZE AND AUTOMATE PROCESSES
- **PROVIDES MORE ACCURATE VISIBILITY INTO SOFTWARE DEVELOPMENT STATUS**
JOINT POLICY
MANAGEMENT OF COMPUTER RESOURCES IN DEFENSE SYSTEMS

TEST PLANNING

"TEST & EVALUATION PLANNING SHALL BEGIN AS EARLY AS POSSIBLE AND SHALL CONTINUE THROUGHOUT THE SYSTEM ACQUISITION LIFE CYCLE"*

"QUANTITATIVE & DEMONSTRABLE PERFORMANCE OBJECTIVES AND EVALUATION CRITERIA SHALL BE ESTABLISHED..."*

"TEST PLANNING SHALL INCLUDE BOTH FORMAL & INFORMAL TESTING"*

TEST PLANNING/CONDUCT

GOVERNMENT TEST PLANNING
- TEST PLANNING WORKING GROUP (TPWG)
- TEST & EVALUATION MASTER PLAN (TEMP)
- COMPUTER RESOURCES LIFE CYCLE MANAGEMENT PLAN (CRLCMP)

CONTRACTOR TEST PLANNING
- SOFTWARE TEST PLAN (STP)
- SOFTWARE TEST DESCRIPTION (STD)
- SOFTWARE TEST PROCEDURES (STPR)
TEST READINESS REVIEW

A TEST READINESS REVIEW SHALL BE CONDUCTED FOR EACH CSCI TO DETERMINE THAT THE SOFTWARE TEST PROCEDURES ARE COMPLETE AND TO ENSURE THAT THE CONTRACTOR IS READY FOR FORMAL TESTING.

LEVELS OF TESTING

FORMAL TESTING
- A TEST WHICH IS CONDUCTED IN ACCORDANCE WITH TEST PLANS AND PROCEDURES APPROVED BY THE PROCURING AGENCY AND WITNESSED BY AN AUTHORIZED PROCURING AGENCY REPRESENTATIVE.
- CSCI TESTING
- MAY INCLUDE SOME UNIT AND/OR CSC TESTING
- SYSTEM TESTING
- DOCUMENTED IN SOFTWARE TEST REPORT

INFORMAL TESTING
- ANY TEST WHICH DOES NOT MEET ALL THE REQUIREMENTS OF A FORMAL TEST
- UNIT TESTING
- CSC INTEGRATION TESTING
- DOCUMENTED IN SOFTWARE DEVELOPMENT FOLDERS
DeMillo: Mike read a list of issues this morning and I would like to get to some of them during this panel discussion. To follow the format that Mr. Watt laid out during the Panel discussion yesterday, we should lead off with discussions from the audience. I know that Dr. Phil Dickinson had a comment about an hour ago that would be appropriate for the Panel. The issue was whether or not MIL-STD-SDS or the methodologies that are growing up around it should handle concerns on the right hand side of the life cycle. The issue that Phil was talking about, in particular, was that he sees in the operational test community the need for hooks in the software and, more generally, design for testability issues in the software.

Gehl: I think that we agree with the requirement that the entire system life cycle needs to be looked at, back end as well as the front end. All of the life cycle has to be looked at when you are defining the requirements and developing the system. We think that is fundamental. I don't see that as an immediate goal in any effort that I know of that is ongoing. I talked to you already about the proposed revisions to SDS. Even though we don't have it on the street yet, there is a proposed revision already started and we do plan on looking at several issues that we just can't fix in the time frame for getting the first version out. One of those issues is what we have been calling the system-engineering concept: for example, the question of how the software that we are developing fits into the entire system. That is where most of these issues of testing and post development software support come in. Those issues come in when you realize that it is not just software, not just a black box, but a system.
DeMillo: Let me see if I can carry Phil's comments a little farther. It seems to me that from the testing community's point of view that we are spending a lot of time validating Ada compilers, talking about standard instruction set architectures, and doing a lot of very careful funneling of the design process along lines that we think are going to help the system. Yet there is one area - the area of making sure that the software that gets put into operational systems can be instrumented, tested and evaluated in a rational way - that is not being addressed.

Gehl: We would certainly like to see that. But you have to understand that we four members of the CSM cannot solve all the problems. We can't say here's how to fix the problems for testing, configuration management, and quality. We need input from you, the test community, so that when it is appropriate, you can say we think that you can incorporate this particular technology into your document. We would welcome input about how to incorporate that either into SDS or into MIL-STD-499 "System Engineering Requirements" or, in general, wherever you think it should best be placed. I think that is what needs to be done. We try to do too much and now we need to get the rest of the community involved.

DeMillo: Any other followups on that? Yes, Ronnie?

Martin: I'd like to hear responses from the other members of the Panel on that.

Clements: Shall I start? As I understand the question, we are talking about pieces of software such that you can tell things about the operation of the software at runtime to help you test and debug. The undesirable events are taken into account or designed by flags that we can raise when they occur. Now, it turns out that we cannot support that kind of software in our application in the production version. There is not enough room. A lot of systems have that problem, there is just not enough room to have those kinds of frills. I am sorry that is the only word that comes to my mind and I don't mean to trivialize the matter by calling it frills, but they are extras. So what we are going to do is to build our system, with the undesired event handling code present. We are going to implement in subsets so we have the room and then piece by piece we will take that out as we are convinced that those undesired events can no longer occur. We handle the problem up front by putting it in our software requirements document. If you want hooks in the software that is delivered to you, you had better require it. That is how SCR handles that issue.
Dickinson: It has implications at the build level, at the OT level, and at the fielded level so that when it does hiccup, as Rich says, you can tell what is going on. It's probably also useful for training because you can start getting human interactions. You can tell what's going on and sort out whether a human screwed it up or the system is really doing something strange.

Foster: At AFOTEC we try and do it by getting testability put in early in the various groups that we work with. We try to encourage them to build testability into the system. We did review the SDS document and we expect to see testability put into it as part of the development standards. It is something that you can't just stick on the end, it has got to be built in from the start.

DeMillo: Don, do you have anything to add?

Greenlee: I certainly go along with all the comments relating to early-on hooks, flags and other support functions for testability but I am going to pick up on something that Phil Dickinson said which is the need to discriminate between different types of errors. We see numerous examples of systems coming up for a major milestone review in which growth parameters such as reliability and maintainability are less than expected but the developer is unable to distinguish whether those are hardware or software errors. A specific example is a communication system in which the MTBF is less than the minimum acceptable value and some of the errors which were presented during the course of the review were software errors. These were promised to be eliminated in subsequent versions of the software when "it matured". Later on incidentally, it turned out that an appreciable number of those were actually hardware errors. There is a tendency to ascribe problems to the software. So in addition to moving up earlier in the process and not leaving everything to testing, it is important to think about the problem of trying to discriminate between hardware and software and the user errors.

DeMillo: Yes, we have a question from the audience.
Bartosik: TECOM has commented on the SDS document and I must admit that I was not one of the guys who contributed to that effort. Phil Dickinson and I have been involved in the Army C3I Studies over the past year. One issue has been the software hooks and hardware monitor points on systems. These should be designed in from the very beginning from the system inception and, as Paul said, made a requirement. We at TECOM certainly support that. We have been fighting for that ourselves for a long time but we haven't had any more success with it than the operational testers have had so far.

Gehl: I think I would like to bring up a problem with that; it's the same one that has been mentioned: the size of the production version. Usually, what I think of as OT&E is the stuff done by OPTEVFOR. It is not the OT&E that's done as part of the development but it is the evaluation that is done after the system is ready to go and we are ready to go for a production decision and the system is essentially what is going to be put in the fleet. In the Navy, we have a lot of programs that were constrained by size. I don't know what my bottom line is, but I think it would be hard to come up with a standard way of putting hooks in the software except for the kind of hook that says, "do it earlier, maybe at OT1 or OT2, rather than OT3". Possibly that could get into the standards, but even SDS and MIL-STD 1679 don't talk about OT1 and OT2 or the different types of operational testing. But I think we do need to distinguish between the phases when that kind of thing should occur.

DeMillo: I think that the kind of test environments that I was talking about and Phil was talking about is when, during an operational test, you have an integrated system, you are driving it around on the desert, and instrumenting all kinds of things that happen in the system, except the software. That is fine when the software works the way it should, but when the system hiccups you really don't have any visibility into the system software. The problem that I see is that you can't really get the visibility that you want unless you have the software architected to allow it.

Gehl: One way of doing that, which in the hardware corresponds to built-in test, is to require or specify built-in software tests so that you can isolate things down to that level.
capacity standpoint, the physical size of the memory may not be such a limiting factor. The thing that will have to be carefully engineered will be the timing requirements on the execution of the software. I can see where having the routines in the software that will pump out the data for testing might slow the execution down to an unacceptable level. The time line will have to be looked after very carefully to see that the speed that is needed to get bits of information from one point of the architecture to another isn't impeded to the point where the mission is jeopardized.

DeMillo: Isn't it a mistake to argue about systems and technologies that are 20 years old in light of what we can do today? It is very unlikely that you will find 16 bit limitations in technologies where you are packing highly dense circuits.

Dickinson: We have to keep in mind that you may be grandfathering existing systems but I agree with the general principal. In a newer system you should do it better.

Audience: I have an alternative for testers that might be useful. Under the old way of doing things, when you charged for errors during a test, you wouldn't charge errors to software unless you could prove the software was at fault. An alternative is that the software developers would be responsible for errors until you can prove that it wasn't the software that was at fault.

DeMillo: That comment leads into one of the issues that Mike McCracken mentioned earlier this morning. Specifically, I would like to ask if the grass is always greener. If you knock around software conferences long enough, we see a lot of concern over the status of software. Your suggestion indicates to me that the software is more suspect than the hardware and I would like to spend just a few minutes getting the feeling of the panel as to whether or not things are really as bad as we have been led to believe.

Greenlee: Yes.

DeMillo: Thank you.

Bartosik: Yes.

Audience: (Laughter)
Audience: I would like to disagree with the entire panel. Now that you have all answered it seems that the best answer to this question is it would be nice to have the capacity to judge these errors. I am a member of the test community, and we require in the case of mechanics, for example, that there be access ports to mechanical equipment so it can be observed during tests. In the case of electronics you have test measurement and diagnostic equipment that is required. Now what is it about software that makes it an exception to this? Is software some kind of magic so that we can't require this? It seems to me that the test and evaluation community position should be that these things should be a requirement.

Clements: If you have an absolute weight limitation on your vehicle would you rather throw away the test equipment or the weapon?

Audience: That strikes me as the same kind of objection we encounter in the hardware world when someone says, "it's too hard to put the ports or the hand holds in that place", but it is important to realize that you can require that sort of thing.

Clements: I am glad to hear that. I think it is important to realize that those requirements can be placed on software too. It is also important to realize that we have to have enough resources and capabilities so that we are able to do that. Sometimes the world is not like that.

Bartosik: If I can just add something. The theme of this ITEA conference is the impact of high technology on test and evaluation. It seems to me that with VLSI, VHSIC, and Crays being readily available, the smaller physical size of processes, lower power requirements, lower heat output, and the multi-millions of bits of storage you can put on boards these days, that I don't think the capacity of memory might be as much of a limiting factor as it would have been five years ago. I think that technology is going to permit us to do the kinds of things, at least in software, that we have been talking about. Understand that I am not an electronics engineer, so that I don't understand all the intricacies of what might happen to a system when you attach a probe at one point or another but I do think that from a software standpoint and from a memory
DeMillo: Let me tell you what motivated the question to begin with. I was at a Design Automation Conference a couple of years ago on a Panel that had to do with testing. And it turns out that I was the only software guy on the Panel. And the question was put to me, "How do you software guys do it? How come your testing of reliability is so much better than ours." When I picked myself up off the floor and thought about that and I wondered if you tend to look at those other technologies through slightly rose colored glasses. I think there are examples of systems where the software gets a lot of blame laid on it. The software gets a lot of blame because you are dealing with case analysis, because you don't really know or have the technology to know what is going on inside the system. In those cases, certainly the software ends up looking quite bad. Now if there is no controversy over that, I will go on to something else, but I offer that as a maybe unconventional view of the world.

Audience: I think maybe there is some truth to that and the purpose of my comment was to point out how difficult it is to really check.

Foster: I'd like to throw something out here. Just as a thought. We have seen that recently as hardware becomes more redundant where the software incorporates some fault tolerance that these problems can be pushed off onto the software. The failure really was in the hardware and the software is supposed to make a switch and it doesn't, then it is now a software problem.

O'Neil: One of the things that bothers me is the level of complexity involved in some of this. I am thinking particularly in the case of radar. Instead of using a computer to keep track of a file of several targets, we are using some mechanical device to do that. Don't you have a lot of problems with reliability and maintainability doing that sort of thing? You end up putting your data in a slightly different format and the problem becomes more difficult and that is what is responsible for a lot of maintainability problems.
DeMillo: I think there is a fundamental difference at work here. And it has to do with the manufacturing process. The manufacturing process for software is purely a design process. The platforms are manufactured. They are built on assembly lines and are subject to reliability measures that you get from engineering physical materials. We all talk about the cost of software but if you take any reasonably large program, put in hundreds and sometimes even thousands of platforms, the software represents a tiny portion of the capitalization on that system. I don't know where that leads but...

French: We hear a lot of talk around the Government that when you go to a contractor to buy something, you get exactly what you pay for or you get exactly what you ask for. So, training government contract people about how to ask software contractors for good products is an important concern. In particular, how do you ask for maintainable software, and the state-of-the-art in software design. Presumably if we can solve that problem, we can avoid some of the difficulties.

DeMillo: I think we heard about that already. One of the purposes of SDS for instance, is to aid that problem.

Gehl: Well, SDS merely carries on what the RADC group is doing in their tech reports and what MIL-STD-1679 is doing in terms of their attempts to get good design principals in the software in the first place. The training of program managers is being addressed by several isolated efforts right now. There is no consolidated effort that I know of. We were asked by the tri-service test commanders about a year ago to come up with a training program to implement SDS. We are going to be faced with a lot of problems; for instance, how do we reach everyone that needs it. The contractors need it, the government program managers need it, the DCAS people need it to monitor it. The magnitude of the task is bigger than all resources can accomplish. We can develop a training course but we are going to need lots of help from the service training communities to get that implemented.

French: Do you think that would be one of the most effective ways to improve the state of the situation?

Gehl: Yes.
Gehl: I think that as a first rule the OT&E guy shouldn't have much to do because the DT&E guy should have done it for you for software. By the time the software comes through the design process which is also the construction process, it has been produced. By the time you give it to the operational tester, it should have completed all of its tests except for possibly system integration and stuff like that. The first rule for OT&E is to make sure that the DT&E guys did their job.

Clements: Precise, unambiguous and complete specifications of the requirements. I am next to last so I get to look very safe and wise. I have had a lot of time to think about it but that is my bottom line.

Dickinson: Precise, unambiguous and complete, but does it fit the needs?

Clements: So you would have the operational tester review the requirements to make sure that user needs are properly represented.

Dickinson: Does that include the verbage on the screen that talks to the sailor?

Clements: We have some pretty clear ideas about how you write a specification like that. When you say verbage I cringe. We try to use as little English as possible.

Audience: (Laughter)

Clements: We have some forms. Our software requirements are specified using tables and the tables have such properties that you can do completeness checking and consistency and all those things. We found that a document like that which is divided up into parts and each part is divided into parts lends itself very well to test case generation. Once the system is thrown over the fence, you can generate test cases for it from the tables to tell you what the system ought to do and you can find out whether or not it really does.

Greenlee: You guys said exactly what I was going to say, But the fifth principle is to decide how much testing is enough. Just like in the case of hardware testing, it relates to a very good point that you raised before and
Greenlee: I don't have an answer to that question but I just thought of an answer to another question. Actually, it relates because specifying what you want software to do is related to testing. Going back to the genesis of STEP, it was not uncommon to find when the software testing was elaborated for a decision point review, that the PM had done the best he could, literally. If, for some reason, that were held in question then the PM had the right to turn around to us and say, "well, what did you want me to do"? This is in terms of evaluating the software testing, validating or whatever you want to do. It seems to be a cultural phenomena: people are comfortable with sonar gear, jet engines, or many other things which are more physically tangible. Perhaps, career wise we have grown up with more familiarities with the hardware. Maybe this is something that will change with time but it is just basically a mysterious area of endeavor. That is why we are trying to improve the state-of-the-art in T&E guidelines. I think that will ultimately play back in terms of specifying software and specifying the properties of what you want it to do.

Audience: I would like to ask those panel members that failed to respond what is the very first thing that you should keep in mind as far as testing software goes?

Bartosik: I would say that one of the best things that came out of the Army study on C3I was the suggestion that we get the operational tester into the contractor's plant just as soon as they start coming up with something that looks like the system so that that tester can sit down and interface with that prototype - I am going to use the word loosely. The tester should see the system prototype at a very early stage, when the thing is first getting off the ground. And if you want to transfer that again into good requirements definition, feel free to do that but that is the first thing that came to my mind.

Foster: From my standpoint, I think it is important to plan for test. You have to make sure that you have the resources and that everything is there together. That you have the hardware to run the testing on, that you have the right kind of people and that they are ready to go into test. Don't frustrate yourself by sitting there with partial pieces of hardware and software while you are getting no data whatsoever. So be ready for the test and plan for the test.
Bartosik: Do you mean the developers or the operational tester?

Grover: I mean the contractors.

Bartosik: Into the test plans?

Grover: Yes, I mean to offer insight into those areas that should be tested and help DoD personnel in making decisions about what should go into a test.

Gehl: I personally don't see anything wrong with that. I would guess that if I were an operational tester, I would like having that kind of input to help me design my test. That's a cheaper way than having me sit in the plant everyday watching the code being built. I would like that kind of input. I don't know that it should be required. At least definitely not on every contract.

Clements: It would be nice to run some experiments. I personally have no problem with that approach. I would like to see an experiment where you have a test plan with that input and try to discover the kinds of errors that you uncover with the developer's input that you wouldn't have found out otherwise.

Foster: We see quite a bit of that in the hand off between DT and OT. What happens is that we get that information indirectly by seeing what the DT development tester discovers. We take that data and if it is usable in the operational environment, that is not just a laboratory environment, you could both have some direct application in the operational environment when that good operational test data. In other words, you don't run that test but we make inferences from it.

Bartosik: I think if you have a real team concept such as the one being tossed around in the Army these days, it is possible to closely coordinate everyone from the outset. Now I understand that there are a lot of vague terms here like "close coordination". Just what does that mean? For example, how many people are willing to go to all these meetings and reviews? How much documentation is everybody going to read? How much are they going to understand? But those problems aside, the team concept leads to good test plans. Those players that are now involved to that extent are going to know a lot more about that system than they currently do. From that standpoint, I think it will happen as a matter of course. Now to what degree you will want it to happen even more, is another question.
that is how much are you willing to invest in testing. Just like we really don't have an answer to how much testing is really enough for the hardware, I don't think we have even that degree of certainty for the software. That is one reason that we are interested in trying to develop techniques for risk assessment models which will tell you how much testing we have to do. I think the passion for built-in testing is probably over emphasized in most of the systems. I think that we are learning now that in a lot of cases it is not worth the effort and one fools oneself because of a false sense of security. How much do you want to pay for that extra increment of testing? I would like to have Georgia Tech look at that, and I know AFOTEC is looking at risk assessment. I think it is fine for us to be whipped into a lather in righteous indignation over the problems of software testing, but we also have to recognize there will never be enough resources on any given program to do as much testing of the software as we want. So I think that principal five ought to be started early on and decide how much testing is commensurate with the overall objective of the program.

DeMillo: Since I am dead last, I get to agree with everything. I think the only thing I would add to what has been said is the term "goals and thresholds". As much as possible, what the software should do should be quantified. That is a requirements issue. I really strongly disagree with Dan McDonald's idea yesterday of munging around with the software to see what happens. Software has the peculiar property that for every test that you design, there is a little twittle to the software that will make it pass that test. It's like conducting an experiment in a chemistry lab. You should say what you are going to do before you do it so that you can tell whether or not you have done it. Are there any questions from the audience?

Grover: Do you feel designers and implementers should have some role designing tests and test plans? The hardware gives some idea of what engineering principles are involved in putting the system together. With the software it is much more difficult to do unless you have more detailed knowledge about the architecture and design of the software. In general as you delve deeper into the systems the black box approach seems to be less acceptable.
Clements: I hear you attributing that to the size of the team. The small size of the team from beginning to end.

Passafiume: That sounds like the Marine Corps.

McCracken: Yes, you are probably doing the same thing on the A7 project. I think that is probably one of the reasons your success is so great there. That is a situation where you don't have all those communications problems to worry out.

Audience: Isn't there a more general principal at work here like the less government help, the better.

McCracken: No, absolutely not. The point I was trying to make is that these six guys were technically motivated. They questioned things, they wanted to understand the system, they wanted to make sure it was done right. They weren't motivated by filling out forms. They were technically interested. I think that, in fact, they were more involved in the program than their U.S. counterparts would have been. When you deal with our own Government contracting, thousands descend and thousands leave and you have a certain amount of time to submit written comments and then thousands descend and thousands leave. By contrast, we established a personal rapport with the engineers.

Bartosik: What was their penalty if they failed?

McCracken: I don't think they looked at it from that perspective. They wanted the system to succeed.

Bartosik: That wasn't my question. My question was, what was their penalty if they didn't succeed?

McCracken: I don't know. They were charged with delivering a radar system. If they hadn't succeeded, I am sure they would have suffered. But the point is they didn't and they knew what they were going to do.

Gehl: I have a couple comments on that. I don't think it is feasible for us to do. For one thing, the way we are structured in the Navy with seven program managers, each of which is staffed with 30 people who have responsibility for over 125 projects, we don't have the resources to be able to do that on every program. Secondly, the DoD acquisition process requires you to
Dickinson: As Rich DeMillo knows, that happened to some extent this past summer on Patriot. We had Raytheon and a number of contractors supporting our follow-on evaluation and supporting the program office. Even Georgia Tech.

DeMillo: We are running up against quitting time here.

Martin: Don Greenlee hasn't responded to that last question yet.

Greenlee: OT&E is supposed to be independent but that doesn't mean innocent or ignorant. What you say is, in fact, common place. The operational tester is obligated to get as much mileage as possible out of all prior testing information whether it comes from the developer or whether it comes from the contractor or other source. He doesn't have to believe it, he doesn't have to build a test around it but I think he is obligated to assess and consider and evaluate it for all for what it might bring to his own team. In fact, it is absolute T&E policies to promote concurrency to the degree possible without sacrificing the integrity of the operational tester.

McCracken: I was wondering if the panel would like to comment on a recent experience I had during my previous life as a contractor. The question is whether or not there is anything to be learned from buying systems with a very small dedicated team. I had that experience with a radar contract for a foreign country. The customer placed in our plant a small group of engineers who were responsible for the total program, i.e., systems engineering, design, software, testing, logistics and even operational testing when the system was fielded. The point is that with a small dedicated team, the communications problem was non-existent.

Bartosik: I would like to know how to manage that.

McCracken: Pardon me?

Bartosik: I'd like to know how they do business over there. You paint quite a contrast for us.

McCracken: Yes, I agree. When I experienced it there were six people who were very well trained, and their responsibility was to get this radar system into production as soon as possible. This group was concerned about the system requirements, the software requirements, logistics, training, development testing, operational testing. I think it was important that they had personal responsibility for the system from beginning to end and they got an outstanding product.
justify your money and your people every year over a seven or eight year period. That means that you can't behave in the way that you describe which, as I see it, is essentially like the commercial contractor.

McCracken: John Passafiume made a comment "It wasn't the Marine Corps, was it?" I think that's significant. It turns out the Marine Corps buys systems in a very similar fashion.

DeMillo: I would like to thank the Panel for their energy and help this afternoon. I would like to thank the audience for their endurance and enthusiasm. Thank you.
In anticipation of the need to continue the momentum created by the Software Test and Evaluation Project (STEP) which was initiated in 1981 by the Director Defense Test and Evaluation to improve the practice of test and evaluation of computer software in mission-critical applications, the Software Technology for Adaptable, Reliable Systems (STARS) Joint Program Office tasked STEP with defining an organizational structure for managing future improvements in the testing process. Options analyzed include: (1) Formation of a Software Test and Evaluation Working Group within the Embedded Computer Systems Acquisition...
Policy Panel, (2) Services and Other DoD Components Appoint a Focal Point for Software Test and Evaluation, (3) Formation of a Joint Logistics Commanders' Subgroup for Software Testing, (4) the Software Engineering Institute Given Oversight Responsibility for Software Test and Evaluation, and (5) Formation of a Software Test and Evaluation Working Group under the Defense Computer Resources Board. Of the options considered, only the Defense Computer Resources Board was found to have the necessary authority and organizational positioning to provide sufficient emphasis to managing improvements in software test and evaluation. It was recommended that a permanent Software Test and Evaluation Working Group be established under the Defense Computer Resources Board's Standing Committee for Test and Evaluation. Included in this report is a Management Plan for the Software Test and Evaluation Working Group that defines membership and resource requirements, organizational authority and responsibilities, and methods for communicating recommendations concerning software T&E to the acquisition community.
### Contents

**Executive Summary**

1. Introduction ................................................. 1

2. STEP Phases I and II Recommendations .................... 3

3. Analysis of Candidate Organizational Structures .......... 7

4. Recommended Organizational Structure ..................... 9


6. Conclusions .................................................. 15

References ....................................................... 16
Executive Summary

The Software Test and Evaluation Project (STEP) was initiated in 1981 by the Director Defense Test and Evaluation to improve the practice of test and evaluation of computer software in mission-critical applications. Initial STEP activities resulted in a series of recommendations to establish a chain of test planning, documentation, and evaluation criteria that starts at the most general planning documents and proceeds through the plans and procedures implemented by Project Offices, development organizations, independent verification and validation organizations, and independent test organizations. Current STEP efforts are designed to define the technology and provide implementation support for the recommendations.

In anticipation of the need to continue the momentum created by STEP, the Software Technology for Adaptable, Reliable Systems (STARS) Joint Program Office tasked STEP with defining an organizational structure for managing future improvements in the testing process. Options analyzed include: (1) Formation of a Software Test and Evaluation Working Group within the Embedded Computer Systems Acquisition Policy Panel, (2) Services and Other DoD Components Appoint a Focal Point for Software Test and Evaluation, (3) Formation of a Joint Logistics Commanders' Subgroup for Software Testing, (4) the Software Engineering Institute Given Oversight Responsibility for Software Test and Evaluation, and (5) Formation of a Software Test and Evaluation Working Group under the Defense Computer Resources Board. Of the options considered, only the Defense Computer Resources Board was found to have the necessary authority and organizational positioning to provide sufficient emphasis to managing improvements in software test and evaluation. It was further recommended that a permanent Software Test and Evaluation Working Group be established under the Defense Computer Resources Board's Standing Committee for Test and Evaluation.

Also included in this report is a Management Plan for the Software Test and Evaluation Working Group that defines membership and resource requirements, organizational authority and responsibilities, and methods for communicating recommendations concerning software T&E to the acquisition community.
1. Introduction

This report defines an organizational structure for managing improvements in the software test and evaluation (T&E) process, and provides a management plan for ongoing activities. It is provided in response to tasking by the Software Technology for Adaptable Reliable Systems (STARS) Joint Program Office (SJPO).

In an extensive study of current practices, the Software T&E Project (STEP) uncovered several deficiencies in the conduct of software T&E. In response, STEP provided 28 recommendations for improvement to the Director Defense Test and Evaluation (DDT&E) in July 1983. A five volume report detailed the need for policy and guidance modifications, programmatic changes, and technological improvements [1,2,3,4,5].

Implementation of the 1983 STEP recommendations is underway. For instance, a revision to DoD Directive (DoDD) 5000.3 "Test and Evaluation" is currently being staffed and will be promulgated shortly [6]. This revised directive is intended to include several of the modifications recommended by STEP concerning policy for software T&E at the level of the Secretary of Defense. DoD-STD-2167, the Defense System Software Development Standard, also incorporates many STEP recommendations.

In addition, current STEP efforts are designed to define the technology and provide implementation support for the recommendations. For example, Volume I of the Software T&E Manual, "Guidelines for the Treatment of Software in Test and Evaluation Master Plans" (TEMPs), supports the review of TEMPs for systems that contain mission-critical software components, are software intensive, or present software testing issues that significantly affect risk [7]. Volume II of the Software T&E Manual, "Guidelines for Software Test and Evaluation within the DoD Acquisition Process", has a major goal of providing the various professionals in the acquisition life cycle with the information required for the successful management of software T&E [8]. The "Software Testing Tools Identification, Demonstration, and Qualification Procedures" allow the location and selection of test tools, environments, and methodologies to satisfy test plan requirements based on key characteristics such as support of quantitative risk models for software T&E, appropriateness for mission-critical applications, and documented cost/benefit data [9]. And, the "Requirements for a Test and Evaluation Subenvironment of an Advanced Software Engineering Environment" offers support to the STARS Program and the STARS Software Engineering Environment (STARS-SEE) Team in the formulation of test support requirements [10]. The ultimate goal of these and other STEP activities is to improve the practice of software testing throughout the DoD.
The momentum created by STEP must be continued and embedded into the fabric of the system acquisition process. The most effective approach to institutionalizing the necessary improvements is to assign the responsibility for the implementation of the STEP recommendations, and the investigation of software T&E issues that arise in the future, to a permanent organization. This report summarizes the 1983 STEP recommendations, analyzes several candidate organizational structures, recommends one for the oversight of improvements in software T&E, and provides a management plan for its use.
2. **STEP Recommendations**

Phases I and II of STEP, completed in the Spring of 1983, consisted of extensive information-gathering and analysis efforts. The chief goal of these phases was to assess the feasibility of modifying DoD policy and guidance for software T&E. In support of the feasibility assessment, a broad overview of the state-of-the-art and the current state of Defense practices in software T&E was constructed [2,3]. Input was sought from DoD components, industrial representatives, selected experts and consultants, and specially convened workshop and symposium participants. In addition, extensive surveys of both the software T&E literature and vendors of automated software T&E tools were prepared. These sources provided a consistent view of software T&E needs and capabilities which formed the basis from which 28 specific recommendations for improving the state of practice in Defense software T&E were developed [1]. The recommendations addressed planning for the software testing process in the context of a major system acquisition and outlined needs for research and development of software test related capabilities. This section summarizes the STEP Phases I and II recommendations.

**Test Planning Recommendations**

The TEMP should be established as the major planning document for software testing, thereby ensuring early incorporation of software test issues into the overall test program. In addition, project offices should set goals for the testing of the total software system, including those components not specified in the TEMP. The nature and extent of the testing required for these components should be sufficient to achieve a balanced risk with mission-critical components.

A chain of test plans and evaluation criteria that begins at the level of system test objectives and proceeds through the detailed testing of software components within development organizations should be established. Program offices should encourage and support the development of written test plans for tests to be conducted during early phases of software development. These plans should:

* contain a specification of what constitutes an acceptable approach to testing,
* explain how the approach adopted supports objectives of the higher level test,
* include provisions for regression testing for all software, specification, and requirements modifications and updates,
integrate the involvement of IV&V contractors (when required) into the overall testing effort,

be adhered to rigorously by Program Managers,

be critically reviewed for deficiencies, and

reflect a realistic, worst-case estimate of the scope and extent of the required testing effort.

Existing technology should be inserted into the T&E process using software that represents the highest decision risk as the focus of the software test plan. Project offices should ensure that unit and module tests exercise critical functions with a systematic test methodology. In selecting a test methodology, primary considerations should be:

* the appropriateness of the methodology,

* known cost/benefit ratios,

* established error detection capabilities of the methodology, and

* the extent to which test results are interpretable in terms of software test objectives set forth in the TEMP.

The relationship between tests performed and the errors to be discovered must be explicit in the test methodology. This relationship should be a principal consideration in determining the appropriateness of the test.

Project offices should require documentation of unit and module tests. Documentation requirements should include:

* resource requirements,

* simulation requirements for inputs,

* analysis requirements for outputs,

* test case cross references to system requirements, and

* sufficient supporting information to allow the reconstruction and repetition of tests.
Numerous organizations are involved in software certification, approval, standards enforcement, and T&E activities for each program. Respective organizational responsibilities and technical roles should be well defined in order to effectively utilize resources, minimize duplication of effort, and raise confidence that test procedures and results of other organizations' efforts are reliable and to be trusted for internal use. In addition, a software test data base should be established for each program and should contain documentation and results from development, operational, and certification testing, as well as other related test information.

Areas Requiring Further Investigation

Research funding should be continued at an accelerated pace for software T&E methodologies and the tools to support these methodologies. Research should also concentrate on establishing usage contexts for the methodologies and the experimental determination of error detection capabilities.

An ongoing program of software testing tools development, packaging, evaluation, distribution, and support should be initiated to provide a warehouse, catalog, or test environment of approved testing tools which can be referenced in the TEMP without acquisition or further approval. Included in this effort should be a program to identify and qualify tools for early use in the development cycle. The qualification requirements should specify usage contexts for specific tools and provide comparative analyses of costs and effectiveness. Provisions should be made for generalizing and improving tools which implement state-of-the-art test techniques and strategies. In addition, contract funding patterns should be evaluated to allow special purpose tools developed in support of the contract deliverables to become deliverable items under the same contract.

APSE and SEE development plans should be modified and expanded to include substantial provisions for test support environments. Test support tools should be made available in the first generation of APSE's and SEE's that are used to develop software-intensive systems.

Quantitative indices of software testing progress should be developed using cost and risk as essential factors. These should include:

* reliable cost/benefit measures for testing software,
* cost/effectiveness measures of testing tools,
* cost/benefit aspects of IV&V, and
* quantitative risk analysis techniques for software errors.

Efforts to provide improved data gathering, reduction, and measurement capabilities should be expanded. Automated data logging and data base systems should be developed to track and record errors on software-intensive systems. The relationship between measurable characteristics of software products and the processes used to produce the software should be validated. Those metrics which are reliable predictors of software quality should be applied to enhance the evaluation process. In addition, the development of testing techniques that take into account quality measurements other than correctness should be encouraged and supported.

Software evaluation criteria should be defined for:
* support software,
* software that is required for system operation but does not directly implement mission-critical functions, and
* other software design components.

These criteria should form the basis for a quantitative risk model of the software T&E process to be used in the evaluation of the overall software testing effort.

An integrated decision support system for software T&E that combines functionally organized test information and evaluations with data that is required for major programmatic decision points should be developed.
3. Analysis of Candidate Organizational Structures

The following organizational structures/solutions were proposed and evaluated in terms of their capabilities to initiate and manage software T&E improvements. Emphasis was placed on organizations that could function within the existing framework of DoD software development and acquisition.

**Formation of a Software T&E Working Group within the Embedded Computer Systems Acquisition Policy Panel (ECSAPP)**

The ECSAPP was recommended by the STARS Program to be established within USDRE(R&AT) to serve as a focal point for technical assistance in the definition of strategies and plans related to the acquisition of MCCR. The ECSAPP was also the original target for the software T&E management focus; however, representatives of the Director, Computer Systems and Software (CSS) and SJPO have discouraged this alternative. The ECSAPP is not positioned organizationally to make a major impact on software T&E practices.

**Services and Other DoD Components Appoint a Focal Point for Software T&E**

This alternative is really no different from the status quo. It is easy to implement since the Services could assign this responsibility to an existing position such as the Computer Resource Manager at the Service and component headquarters organizations. However, although it might focus Service activities, there would be no central focal point at the DoD level and there is no guarantee that the Services would always work in concert toward the advancement of software T&E.

**Formation of a Joint Logistics Commanders' (JLC) Subgroup for Software Testing**

The JLC are and have been involved in several efforts to promote improvements to the state of practice in software development and acquisition. The JLC Joint Policy Coordinating Group for Computer Resource Management published a Joint Regulation on Management of Computer Resources in Defense Systems; has been instrumental in the development of DoD-STD-2167, the Defense System Software Development Standard (also known as SDS); and is currently supporting the development and coordination of the proposed DoD-STD-2168, Software Quality Evaluation Standard (SQE). Since the JLC organizational structure is already defined and operating in these and other areas, it may be possible to add a new subgroup to address software T&E or to add this mission to that of an existing subgroup.
The Software Engineering Institute (SEI) Given Oversight Responsibility for Software T&E

The SEI is charged with the overall mission of accelerating the development of modern software engineering techniques and methods, and promulgating their use throughout the mission-critical computer systems community. The main focus of the effort is: (1) the development of a resource of professional talent and technology; (2) the transformation of this potential into usable and effective methods, techniques and supporting systems; and (3) the maintenance of a showcase of methods, techniques and supporting systems to share with the mission-critical systems community to achieve standards of excellence in practice. The SEI is chartered as a Federally Funded Research and Development Center [11]. Therefore, it is not positioned to provide direction to the Services and DoD Agencies or to perform in an oversight role. While it is evident that the SEI should play a major role in the area of software testing, it should not take the DoD lead. A more appropriate role for the SEI is that of a consultant providing support to DoD activities. In addition, the SEI could serve as a clearinghouse for the latest software testing technology.

Formation of a Software T&E Working Group under the Defense Computer Resources Board (DCRB)

The DCRB is established by DoDD 5000.29 "Development, Acquisition, and Management of Mission-Critical Computer Resources (MCCR)", which emphasizes mission-critical computer systems/software and charges the heads of DoD components with designating a Senior Official for MCCR and providing the resources necessary for implementation of the directive [12]. All Services and OSD Agencies have Flag level members on the DCRB and have appointed Senior Officials for MCCR at the Under Secretary level. The DCRB functions at a level where it can focus on the critical aspects of software T&E described above and has the authority to oversee the implementation of recommendations for improvement. No significant disadvantages of this alternative have been identified.
4. Recommended Organizational Structure

The DCRB is an executive committee created to provide a high level and uniform approach to handling MCCR. Its charter was approved by the Deputy Under Secretary of Defense for Research and Advanced Technology (DUSDRE(R&AT)) on 10 August 1984 [13]. The DCRB has been tasked to provide advice and guidance over all MCCR matters, specifically including but not limited to functional areas previously handled by the Management Steering Committee for Embedded Computer Resources. The DCRB has also been charged with the establishment and tasking of subordinate committees to address specific areas of MCCR acquisition and utilization.

Of the organizations proposed above, only the DCRB has the necessary authority and organizational positioning to provide sufficient emphasis to the tasks at hand. Although the formation of a JLC Subgroup for Software Testing was also a good alternative, official OSD sponsorship is preferred. Additional attractive attributes of the DCRB are that it is currently active, has a Secretariat to provide administrative support, and has established funding channels to support existing activities. The remainder of this discussion is concerned with the establishment of a Software T&E Working Group under the auspices of the DCRB.

The organization of the DCRB provides for Standing Committees that address functional areas with recurring issues. Committees may establish working groups to address specific problems. It is recommended that a permanent Software T&E Working Group be established. Two of the Standing Committees appear to be candidates to house this working group: (1) the Standing Committee for Embedded Computer Systems and (2) the Standing Committee for T&E. The Embedded Computer Systems Standing Committee is planned to address complex problems associated with embedded systems including battlefield requirements, unique applications, and the degree of commonality needed for both hardware and software. The Test and Evaluation Standing Committee is planned to evaluate existing and proposed T&E processes and to promote the use of standard hardware and software capabilities which reduce the number of latent defects.

Although many critical software T&E issues are associated with the development and deployment of Embedded Computer Systems, other issues exist that are not specifically related to embedded systems. A Software T&E Working Group established under the Standing Committee for Embedded Computer Systems would be forced to neglect a significant portion of the software testing community's problems and needs. It is, therefore, recommended that a permanent Software T&E Working Group be established under the Standing Committee for T&E*.

*Another option would be for several Standing Committees to jointly establish the Software T&E Working Group.
Just as software is not acquired, developed, or operated in isolation, likewise, the Software T&E Working Group must not function in isolation. Coordination with other groups, both internal and external to the DCRB, is essential. As a minimum, interfaces must be established immediately with the Standing Committee for Embedded Computer Systems as well as with the Standing Committees for Standards, Policy and Plans, Quality Assurance, Environments and Languages, and Life Cycle Management.

The Management Plan for the Software T&E Working Group defines membership and resource requirements, organizational authority and responsibilities, and methods for communicating recommendations concerning software T&E to the acquisition community.

Software T&E Working Group Membership

The Software T&E Working Group should contain representatives from each of the following sponsoring organizations (see Figure 1):

* the Office of the Director, Operational T&E (DOT&E),
* the Office of the Assistant Secretary of Defense for Acquisition and Logistics (ASD(A&L)),
* the Office of the Assistant Secretary of Defense for Command, Control, Communications, and Intelligence (ASD(C3I)),
* the Office of the Deputy Under Secretary of Defense for Test and Evaluation (DUSDRE(T&E)),
* the Office of the Deputy Under Secretary of Defense for Reliability and Maintainability (DUSDRE(R&M)),
* the Secretaries of the Military Departments,
* the Military Departments (both development test and operational test organizations),
* the Defense Communications Agency (DCA),
* the National Security Agency (NSA),
* other organizations and agencies as approved by the Senior Official for MCCR, and
* other non-DoD agencies as approved by their cognizant officials and the Senior Official for MCCR.

The primary responsibility of a sponsoring organization is to supply a representative to the Software T&E Working Group at an appropriate level of effort. The representatives should have sufficient expertise in the areas of software testing and acquisition to allow for maximum contribution to, and participation in, working group discussions. In return, the representatives have the opportunity to present their organizations' software T&E issues directly to the working group for consideration.
Figure 1.
Resource Requirements

In addition to the administrative support supplied by the DCRB Secretariat, it will be necessary to provide either Government or contractual support for the preparation and coordination of meeting minutes, working group plans, and reports. Additional assistance for the procurement process may also be required.

Software T&E Working Group Authority and Responsibilities

The DCRB is established by DoDD 5000.29, "Development, Acquisition, and Management of Mission-Critical Computer Resources" to assist and advise DUSDRE(R&AT) [12]. This document recommends the formation of a permanent Software T&E Working Group within the Standing Committee for T&E.

The Software T&E Working Group is responsible for:

* identifying new issues (including those supplied by sponsors) to supplement those addressed by the 1983 STEP recommendations,
* determining and prioritizing appropriate actions,
* making recommendations to the T&E Standing Committee,
* preparing and executing plans of action as directed by the T&E Standing Committee,
* updating plans of action as necessitated by changes in technology, threat, and environment,
* tracking status of related activities,
* keeping records of Software T&E Working Group activities, and
* publishing plans and reports.

In addition, members of the Software T&E Working Group are responsible for coordinating within their home organizations to ensure that technological advances are merged with current practices, and that the appropriate level of support and attention to software T&E is maintained.
Methods for Communicating Recommendations to the Acquisition Community

There are four communication paths that software T&E recommendations can take: (1) recommendations to other Standing Committees of the DCRB via the Standing Committee for T&E or the chair of the DCRB, (2) recommendations to DUSDRE(R&AT) via the DCRB and the Standing Committee for T&E, (3) recommendations to a specific sponsoring organization via its representative, or (4) recommendations broadcast to all sponsoring organizations. In the case of option four, attempts should be made to coordinate independent activities concerning similar recommendations.

Regularly scheduled status briefings should be presented to sponsoring organizations and to other key individuals. In order to assure information interchange with a wider audience, conferences should also be held to address specific topics. Finally, to maximize cooperation, Software T&E Working Group reports should be distributed throughout the mission-critical software community.
6. Conclusions

STEP Phases I and II activities and subsequent follow-on efforts have revealed that current software T&E practices suffer from several critical deficiencies. In order to correct these deficiencies, it is recommended that a permanent DoD organization be established to analyze issues, determine solutions, and provide direction and support for the implementation of those solutions. Evaluation of the alternative organizational structures has indicated that the DCRB is the most appropriate entity to manage improvements in software T&E. A permanent Software T&E Working Group sponsored by organizations that have system/software acquisition and testing responsibility will ensure that the needs of mission-critical software testing are addressed.
References


May 17, 1984

Ms. Pat Heitmuller
OCA/SSD
Campus

Dear Ms. Heitmuller:

Please find enclosed one copy of a draft report entitled "Requirements for a Test and Evaluation Subenvironment of an Advanced Software Engineering Environment". This report is due under the deliverable schedule of the STEP Contract (F33657-82-G-2083).

Sincerely,

Richard A. DeMillo
Professor

RAD/rp

Enclosure(1)

Copies of draft sent to: DDT&E
                      STARS
                      NADC
REQUIREMENTS FOR A TEST AND EVALUATION
SUBENVIRONMENT OF AN ADVANCED SOFTWARE ENGINEERING
ENVIRONMENT

Richard A. DeMillo
The Software Test and Evaluation Project
School of Information and Computer Science
Georgia Institute of Technology
Atlanta, Georgia 30332
# Contents

Executive Summary ................................................. 1

Introduction ......................................................... 1

General Requirements ..............................................

  User Needs ......................................................

  The Concept of a Subenvironment ..............................

  Resource Classifications .....................................

  Modes of Operation ...........................................

  Host-Target Selectability .....................................

Specific Requirements ............................................

  Tool Building Services ......................................

  Test Building Services ......................................

  Test Description Services ...................................

  Test Preparation and Support Services .....................

  Manual and Executable Specifications ......................

  Control of Test-Related Processes .........................

  Other Passive Resources and Decision Support Services

  Human Factors and User Interfaces .........................

Concluding Remarks ................................................

Bibliography ........................................................
Introduction

This report constitutes the recommendations of the Software Test and Evaluation Project (STEP) concerning the requirements for software testing in current and planned Ada (registered trademark) software engineering environments. STEP is tasked by USD (USDR&E/DDT&E and DUSDRA&T/STARS) to analyse the policy and technology underlying the current practice of software T&E and to make recommendations for improving that practice. Among the specific concerns of STEP is the rapid insertion of available software testing technologies into Ada environments. The recommended requirements which follow have been derived after extensive study of DoD practices and needs in the area of software testing and are intended to form an essential implementation phase of policy, regulation and standards modifications being considered by DoD components. The background leading to these recommendations has been described in a multi-volume report issued by STEP [1].
These requirements have been formulated in coordination with a number of related efforts. The Joint Service Software Engineering Environment (JSSEE) effort is one such activity. The goal of JSSEE is to carry out the preliminary design of an advanced software engineering environment which will generalize current Ada programming support environments by addressing the total software lifecycle. STEP has participated in a liaison role on the JSSEE definition team since October 1983. The software T&E requirements contained in this report are intended to be an additional set of JSSEE capabilities and should be viewed as implemented in an early JSSEE. Another related effort is the Joint Logistics Commanders' MIL-STD for Defense System Software Development (now nearing implementation). STEP participated in the development of MIL-STD-SDS, particularly in the area of software testing. The requirements which follow have been derived from a study of MIL-STD-SDS and are supportive of the lifecycle provisions of the new standard. Additional coordination has taken place with a variety of DoD, industrial and academic groups through committee memberships, consultancies to STEP and military advisory groups.

The capabilities discussed below are intended to form a software Test Sub-Environment (TSE), that is, a capability subset of a SEE. At the same time, the capabilities of the TSE comprise a unique set of functions that can be used to compose tests and support evaluations for all facets of software testing. It should be possible to customize TSE capabilities to reflect the unique requirements of: (1) the current software lifecycle phase, (2) project-specific technology and management, (3) application-specific technology, (4) applicable DoD and Service policy, and (5) applicable contracts, regulations, and standards.

The existence of a TSE will require the design of interfaces and subsystem boundaries. The TSE will require convenient access to that portion of the basic environment (e.g., virtual operating system interface and I/O services) needed to create, execute and analyze processes. Similarly, the TSE will be accessed by other users and processes. The manner in which the TSE interfaces with other components of the SEE is a matter of design. However, since software testing is often a heavy consumer of resources, these interfaces (environment-at-large, TSE, operating system, user) must be designed to optimize demands on those resources subject to certain constraints. These constraints may dictate the placement of subsystem boundaries -- as may be required, for instance, to avoid contamination between test-related and non-test-related processes.
Figure 1 illustrates the major categories of TSE resources (passive, active, and support), and the major components of the SEE with which interfaces must be constructed. Figure 1 also shows the manner in which processes may be run in relation to the TSE. For example, a "user" may edit and debug a program ("Process A" in the diagram) while under the watchful eye of a passive "kernel" test capability. On the other hand, an expert "tester" may use the kernel capability to compose a sophisticated active test resource which has total control of a software component being-tested ("Process B"). In still another scenario, a user process ("Process A") may create another process ("Process C", in the diagram) that in turn accesses a support resource of the TSE (e.g., an editor, tool builder, or interpreter) for purposes that are unrelated to testing. Finally, a class of "privileged" user may gain access to the TSE to perform executive activities such as system maintenance.

The following paragraphs expand on each of the major functions, interfaces, and architectural constraints of the TSE. For the most part, architectural constraints are contained in a "General Requirements" section, while a "Specific Requirements" section enumerates the technical characteristics and functional components that will be needed to plan, carry out, document and report software tests.
........Insert Figure 1. (functional organization) here ..
General Requirements

Since the SEE is to support the entire lifecycle, it must also support the test and evaluation portion of that lifecycle. If this is to be accomplished, an understanding of the special needs of the acquisition process and the role of test and evaluation in that process is important.

DoD policy for major system acquisition resides in the 5000 series directives and instructions. The acquisition process consists of four phases: Concept Exploration (Phase 0); Demonstration and Validation (Phase I); Full-Scale Development (Phase II); and Production and Deployment (Phase III). Prior to the beginning of each phase, the system under development must pass the milestone review which is conducted by the Defense System Acquisition Review Council (DSARC). The decision of the DSARC is based, in part, on the test and evaluation assessment as presented by the Directors, Defense Test and Evaluation and Operational Test and Evaluation. DoDD 5000.3, "Test and Evaluation", requires the preparation and approval of a Test and Evaluation Master Plan (TEMP) for each major system prior to each decision milestone (beginning with Milestone I). The information in this document provides the basis upon which test and evaluation assessments are developed. The TEMP includes a program management summary, an outline of the system's development test and evaluation, and an outline of the system's operational test and evaluation. Therefore, the sources of information for the TEMP may include the program office, the development test and evaluation agency, the independent test and evaluation agency, the development contractor, and the IV&V contractor, if one exists. Additional constraints are placed on these organizations, and the test and evaluation activities they perform, by the Service regulations which implement DoD policy and the Military Standards which are applied to the efforts of the defense contractors. The result is a set of user needs and requirements which must be satisfied by the SEE.
User Needs

The software test and evaluation capabilities of advanced software engineering environments should be available to testers and test-related managers and engineers at all stages of the software lifecycle and should support the usage scenarios which are most likely to occur in practice.

To be effective in implementing meaningful policy revisions for software testing, the software engineering environments should support the needs of testers, decision-makers and managers in a coherent way. Among the professionals who will interact with such test capabilities are the following:

1. Acquisition decision-makers: this category includes decision-makers in OSD who will rely on the output of the various stages of software T&E to provide the information that is needed by their decision support systems.

2. Project managers: project management includes both DoD and contractor project managers who need project information and analyses.

3. Programmers: the activities of programmers contribute a great deal to the effectiveness of the overall T&E effort and should be supported and documented by their environment.

4. Testers: the category of testers includes military and commercial-sector professionals who have the responsibility to conduct planned, documented, and audited testing in accordance with contractual or internal corporate requirements. Included in this category are engineers responsible for dedicated development testing and operational (system) testing.

5. IV&V and Lifecycle Software Engineers: this category of engineer includes all those who must rely on the existence of standard environment features and capabilities to review, audit, or evaluate the development practices or build upon those practices during support, maintenance or evolutionary phases of the software lifecycle.

....expand with usage scenarios for 1-5 above....
The Concept of a Subenvironment

Test and evaluation capabilities should be delivered in a sub-environment (the Test Subenvironment -- TSE) of the overall software engineering environment (SEE). The TSE should constitute a closed, self-contained system with clearly defined interfaces to the rest of the SEE.

The requirement for a subenvironment does not preempt SEE designs in which the TSE is highly integrated with the rest of the environment. It is a useful abstraction, however, when the following requirements are taken into account:

1. The TSE should consist of components that are highly integrated with each other and are together easily integrated with the SEE.

2. For protection of non-test processes, efficiency of operation, economies of scale in duplicating operating system functions, and provision of secure audit trails, TSE processes should be logically isolated from the rest of the SEE.

3. Since the system level test requirements that drive the more detailed stages of program testing originate before acquisition Milestone I and may continue well into the operational life of the system, there may be a set of competing stability requirements that must be satisfied by the TSE but are not needed in the rest of the SEE.

The exact nature of the TSE-SEE interface depends to a large extent on the functions and services that can be provided to the TSE by the SEE, balancing the (sometimes competing) requirements for autonomy cited above. Obviously, whatever is not provided by the SEE must be included in the TSE.

Like the rest of the environment, the components devoted to testing both rely on and deliver services. In this respect, that portion of the environment which is used primarily for testing should inherit its design characteristics from the environment as a whole. There are a number of reasons for adopting such an attitude. A test environment that is disconnected from the rest of development environment is:

1. duplicative of resources,
2. jarring to users,
3. difficult to support and maintain, and
4. expensive to develop and interface.

Even so, it is essential to adopt a view of the test environment as a closed, self-contained system with clearly defined interfaces to the rest of the environment -- in short, the testing services should be delivered by a subenvironment.
The concept of a test subenvironment (TSE) leads to a number of design issues that must be resolved early. The most important of these issues is the degree to which a tester is physically isolated from the rest of the environment. Figure 2 illustrates four distinct design approaches to the problem of isolation. At one extreme is the approach taken by the designers of mutation systems[]. In mutation systems, the user invokes the test tool and is thereafter insulated from virtually all other operating system features and functions. Intermediate approaches are those of Berkeley Unix and the Xerox Cedar environments. In Unix a degree of isolation is possible by invoking one of two command language interfaces or shells[]. Each shell offers independent access to common system resources, but invocation of shell capabilities must be carried out by explicit reference to the desired shell. Cedar, on the other hand provides a single interface in which a degree of separation is enforced by system architecture; the workstations on which Cedar programmers construct software are connected by an Ethernet to each other and to a variety of "servers" which provide common resources to all other processes. At another extreme are those environments -- most operating systems, for example -- in which the system enforces no subsystem boundaries beyond those constructed by user processes. Even in these cases, however, subsystems can be usefully defined. The Georgia Tech swt subsystem (a simulation of the Unix Bourne shell) of the commercial Primos operating system is an example[].

In the remainder of this section we outline the extent to which the test environment is required to be a self-contained subsystem, concentrating on four important aspect of the TSE: the integration of tools and services, the requirements for isolation, the external constraints and how they translate into stability requirements, and the interface between test services and the remaining services of the environment.

Integration. When automated testing aids are applied, they tend to be applied only sporadically. A major finding of the data-gathering and analysis phases of STEP[] was that automated testing tools are not used in practice. This finding has been supported by independent studies (see, e.g., []). Three factors contribute to this state. First, the tools that are available, tend to be stepchildren and therefore tailored to a parent project but useful for few other tasks. Second, the existing tools are difficult to compose. Even though test requirements may call for a continuous and coordinated testing activity, there is little support for smooth transitions between methodologies or tools. Third, most software development environments are based on a model of design and implementation (edit, compile, run) that leaves little room for the layering of reviews, reports and evaluations that are required to develop complex systems for variegated users under often rigid budgetary constraints.
The common thread underlying these factors is a lack of integration. Test tools and technologies are not integrated with the rest of the software engineering environment (when the parent project moves on, the ad-hoc test tool dies), with each other (the work involved in the pairwise interfacing of a number of independent -- unintegrated -- tools is quadratic in the number of tools when it should be linear, i.e., it should not cost much more to use a number of tools in composition than it does to use them independently) or with the activities that are most commonly associated with software development (the most heavily used tools -- editors and compilers, for example -- do not contain the "hooks" that allow easy and natural access to test facilities).

The exact nature of the integration required in the TSE spills over into design and is best left to the architects of the SEE. However, designers might be guided by the degree of integration achieved in the Interlisp environment [1]:

Interlisp is not merely a collection of independent programming tools, but an integrated system. By integration, we mean that there need not be any explicit context switch between tasks or programming tools, in switching, for example, from debugging to editing to interrogating MasterScope about the program...the various facilities can use each other in important ways...the integration of facilities increases their power.

A requirement for integration points strongly to a subsystem view of test resources. For example, user (non-test) processes that call upon test services should be able to access those services through a common interface and thus avoid the overhead in planning an interface with each tool or technology that is needed. Plainly, a process that interacts with the test sub-environment can expect a single interface; contrast this with a process for which each interaction requires a new interface. By the same token, integration of the test tools with each other also points toward a subsystem: test tools rely on system resources that are mainly disjoint from non-test-related resources. Burdening the SEE at large with supplying services that are needed only to compose and define test resources flies in the face of economic good sense.

Isolation. The TSE should possess the following three characteristics:

(1) integrity and reliability
(2) security
(3) efficiency.
The integrity and reliability of an environment are the extent to which processes and functions are immune to malicious or unpredicted failures. Unlike most SEE processes, TSE processes will fail at an extremely high rate. Furthermore, the failures that are induced during tests may involve external data sources, system-wide storage, or run-time (target) executive functions. The design of the TSE must be sufficiently robust to insure that (1) data and programs that are not otherwise involved in a fault do not become contaminated by it, and (2) overall TSE and SEE services are not interrupted even in the presence of severe faults.

The TSE must also deliver a class of secure services. These capabilities will usually be dictated by contractual or management requirements. Furthermore, audit trails and control over access to certain SEE/TSE objects are necessary to insure a minimal amount of visibility into the software development process. An example of a TSE-secure service that has been frequently suggested is the automatic logging of errors. If error logging is required, then it is probably also a requirement that the reporting of errors not be biased and that the contents of the error database not be entrusted to users with "vested" interests in skewing the error data. Bias in the reporting of errors can arise in a number of ways -- for example, statistical sampling procedures can be altered or influenced. Insuring that error data is not contaminated and is not released to unauthorized users is essentially a problem of access control.

Efficiency in TSE operation has an impact in a number of realms. First, it affects human users. Inefficient tools and cumbersome sequences of operations not only discourage voluntary use, they may encourage the use of schemes which purposely avoid the most useful test methodologies. Second, inefficiency may frustrate certain kinds of testing.

For example, a program under test may generate certain "fatal" run-time errors that are normally detected and handled by agents of the operating system. When a process has a relatively low expectation of failure, this exception-handling can be a fairly involved procedure without affecting overall system performance. In error-prone processes, however, there is frequently a trade-off between performance and reliability. A primitive example of such a trade-off is the use of diagnostic monitors in tactical weapons systems; in the midst of battle, diagnostic routines may consume most system resources, reducing overall system effectiveness below a minimum threshold. In these cases, monitors are either disabled or ignored on the strength of the argument that the availability of a faulty process is more desirable than the absence of the process altogether.
In some cases, efficiency constraints may actually rule out certain tests that should be performed. For example, most dynamic analyzers execute an internal form of the source program in an interpretive mode. The reasoning is that turning control of the entire test process over to the operating system -- when a "divide-by-zero" error is generated, for instance -- is too cumbersome, requiring a user to frequently enter the operating system and re-inverse test tools. On the other hand, an interpreter that is part of the tool can simply choose its own exception handling options without ever consulting the operating system. This makes most serious performance testing impossible since interpreters significantly degrade software performance.

In each of these instances, the most natural design tends to isolate the TSE in some manner. In addressing reliability issues, for example, the requirement for fault tolerance translates into a redundancy requirement (which, in turn gives rise to a subsystem like Co-Pilot in Mesa []) or into other isolating mechanisms (e.g., recovery blocks []) that insure availability of services. By the same token, security requirements are most naturally interpreted as requirements for logical isolation [].

Stability of the Environment. The Mesa environment [] defines the concept of a client process. A client is a purchaser of services from an environment resource. Possible clients of test services and the services they might want have been described in the usage scenarios above. A striking feature of these clients is the degree to which they are constrained and interconnected. Beginning with acquisition policy and following a chain to the most technical of design and development activities, testing is a planned, documented, reported and evaluated collection of procedures and protocols.

One effect of these constraints is on the stability of the TSE -- for example, the release protocols for the TSE updates and revisions may differ from the release structure of the SEE at a whole. Such differences are especially apparent when a TSE component or function specifically supports a policy-level report that is tied to the acquisition process instead of the software lifecycle or the SEE lifecycle. As a simple example, consider the "Management" section of the Program Summary (i.e., Part II) of a TEMP (cf. []) that specifies "arrangements between participants for test data sharing." If the TEMP reflects the T&E process, then agreements in force for sharing test data must also be respected by the TSE; in particular a function that implements such an agreement must be common to all participants in the agreement and cannot be altered by an arbitrary revision of the TSE.
Stability problems have been solved in commercial environments [] by defining release protocols that carefully promote new or revised capabilities according to the extent to which they may be relied upon (to reflect higher-order schedules, for instance). In such protocols, tools and functions that are not to be relied upon at all may be distributed, but will not be involved in any "critical path" decisions or activities. As a capability becomes more stable, it is elevated along a hierarchy (the depth and structure of which varies) until it is incorporated into the existing system release or version. Any such protocols in force for the SEE should take into account the need for an independent TSE protocol.

This set of externally imposed characteristics leads us to view the test resources as a self-contained subsystem.

**Resource Classifications**

TSE capabilities should be classified according to services provided in support of MIL-STD-SDS as well as according to system-dependent capabilities.

The TSE is simultaneously a system resource and a tool to implement policy and regulatory constraints on software development. In one taxonomy (Figure 3), TSE resources are categorized by reference to the software lifecycle, priorities for inclusion (e.g., contractually required, technically desirable, useful but not generally available) and the type (e.g., data gathering, planning and reporting, testing, support) of service offered. In the second taxonomy (Figure 4), resources are classified hierarchically by reference to system capabilities (e.g., system services, interface services, and services reserved to the TSE).
\[
\text{Figure 3. Lifecycle Classification of TSE Functions}
\]
Figure 4. System Classification of Functions
(To be filled in later)
Modes of Operation

The TSE should support at least four modes of operation: a kernel mode (a basic set of test capabilities that are delivered to all users), an active test mode (which may require expertise on the part of the user), a mode to allow access to certain TSE resources by non-test processes, and a privileged mode in which no test activities are active.

Each of these modes supports a specific set of client processes. In kernel mode, audit trails are insured since all users are given an appropriate T&E environment in which to work (although the form that the kernel takes may vary between categories of users). Since the kernel mode is delivered to all users, it should place minimal demands (e.g., knowledge of test technology) on the user. In the active test mode, the user is assumed to be a knowledgable engineer who has requested a test resource and is therefore willing to suffer performance penalties or interface as an expert with an advanced T&E tool. The no-test mode is reserved for the delivery of services that do not involve testing. A programmer (who is already in kernel mode) may desire the use of a TSE resource -- a syntax-directed data editor, for example -- for reasons unrelated to testing. The no-test mode will allow the delivery of such services without placing the user in an active test mode. The privileged mode bypasses audit and data gathering facilities and should be reserved for computer system management.

... detailed discussion of modes of operation...
Host-Target Selectability

The TSE should support the eight most likely configurations of host machines, target machines, and target environment. The host may be connected to, or disconnected from, the target machine. The target machine may be emulated or not. The target environment may be simulated or not. Target machines and environments that are not simulated on the host may consist of operational hardware or the result of external simulators/emulators.

These configurations have been formulated to meet the host/target approach to software design supported by Ada environments and are the minimal set of configurations that will be encountered in development/operational testing scenarios. Support of these configurations in Ada Program Support Environments is explicitly required.

The eight configurations, as enumerated in [], are the following (cf. Figure 5):

C0: Host emulated target: the host computer simulates both the target machine and the target environment.

C1: Host emulated target machine: the host computer emulates the target machine and interacts with an external environment.

C2: Host connected target machine: the host computer interacts with an external target machine while simulating the target environment.

C3: Host connected target: the host computer interacts with an external target machine and external target environment.

C4: Host disconnected target machine: the host computer simulates the target environment but does not substantially interact with the external target machine.

C5: Host disconnected target: the target machine and environment are external to the host, and the host interacts with neither of them.

C6: Simultaneous host-target emulated target environment: the host and target computers are the same but they simulate the target environment.

C7: Simultaneous host-target external target environment: the host and target computers are the same; they interact with an external target environment.
....Insert Figure 5 (host-target configurations) here.....
The host-target selection provisions of the SEE-at-large may not be adequate for the TSE. Unless special provisions are made in the design of the SEE it may not be possible to instrument or monitor target processes, capture target environment scenarios in a data definition language or properly contain faulty processes. Therefore, host-target selectability in a form that is suitable for testing is an explicit requirement of the TSE.

...The rest of this section will discuss the implications of host-target selectability on the design of a TSE...
Specific Requirements

...intro to technical requirements...

Tool Building Services

Tool-building services should be provided and should be structured to promote (1) sharing of common software T&E functions among tools, (2) plug compatibility among tools designed for different SEE's, and (3) ease of implementation of specialized testing tools. These services should include, as a minimum the front-end (e.g., parsing, internal form manipulation, and execution) facilities common in static and dynamic analyzers.

The tool-building services of the TSE should encapsulate the commonality of function seen in many testing tools. The actual design of the tool building facility may be incorporated into the SEE at large (as in the Tajo facility in the Mesa environment) or may comprise a relatively independent set of tools and features that each solve a specific application problem (as in the YACC-LEX-AWK tool set in Unix (registered trademark)).

The type of commonality that should be embodied in the tool building facility can be illustrated by the "front-end" processing stages of a number of dynamic analyzers, which parse source language statements into an internal form, maintain the source code for display and manipulation, attach test-specific information to the internal form thus making the form an executable object, and finally attach semantics to the internal form by providing an execution facility (e.g., an interpreter). The process of tool-building should not require the duplication of these functions. Many test methodologies (e.g., program mutation, various partitioning schemes, weak mutation) can be implemented as specialized test tools that "plug into" the language-dependent front end processors.

Demanding human interfaces or file handling capabilities are also likely to present common problems to tool designers and should also be accommodated by the tool-building services.
Test-Building Services

Specialized test-building capabilities should be user generated from a relatively small number of primitive test generation functions that include (1) path and domain selection, (2) symbolic aids, (3) static aids, and (4) coverage analyzers.

These capabilities are the basis of most dynamic and static test data generation schemes and should not only be supported by the TSE but should be integrated to the extent that a skilled tool-builder can compose these primitives into more extensive test generators. It is not required that these primitives be uniform in the degree of automation supported for test generation.

For example, the path and domain selection primitives may take the form of interactive aids that generate "path expressions" when the user wants to isolate a path or domain, while the symbolic aids may vary from verification condition generators to optimization algorithms that solve systems of equations produced by the path selection tools. The coverage analyzers should be sufficiently general to gauge both component coverage (e.g., statement coverage analysis) and error/fault coverage (e.g., mutation coverage analysis).

Test Description Services

The TSE should contain capabilities for communication between dedicated software test generation services and system test generation services. As a minimum, the TSE should support a data definition language that allows the translation of system test data into host-readable form for analysis by TSE tools.

This requirement insures that there is a facility available for sharing of test results and data between some of the life-cycle stages. For example, it should be possible to achieve a unit test threshold or goal by combining dedicated unit test results and results of a system test in configuration C3.
Test Preparation and Support Services

The clerical burden of test preparation should be eased with services that provide: (1) structured data editors, (2) spreadsheet data auditors, (3) file comparators, and (4) data and file services.

Some of these capabilities (e.g., file services) will undoubtedly be provided by the SEE and are present in the TSE by virtue of its interface. Other capabilities (e.g., structured data editors) will most likely be present in the TSE for technical reasons but will provide a useful resource that will be exploited by processes in no-test mode. At the other extreme are test-specialized services (e.g., a spreadsheet auditor that updates by formula all data fields dependent on a field that changes) that will be used outside the TSE only infrequently.

Manual and Executable Specifications

Automated and manual aids should be provided for determining input-output and correctness behavior of software. As a minimum, these services should include data reduction interfaces that allow user determination of correctness and an executable specification language that allows the uniform expression of input-output correctness conditions.

This addresses the so-called "oracle" problem for software testing. While no single solution to the problem is known, existing test methodologies rely on an amalgam of manual and automated specification techniques that must be supported by the TSE.
Control of Test-Related Processes

The TSE should have sufficient control over test related processes to insure the following: (1) that other SEE processes are protected from faulty processes, (2) that initiation of test activities can be enforced and managed, (3) that privileged processes can exempt themselves from enforcement, (4) that virtual (configured) machine states can be saved and manipulated.

The effect of these requirements is to place a burden on the TSE to either access or duplicate operating system functions. For example, storing machines states (e.g., to allow reproduction of error conditions in real-time processes) or branching to arbitrary machine states (e.g., to avoid the expense of restarting faulty processes during repetitive or regression testing) are typically capabilities that are reserved to operating systems. On the other hand, for Ada-based SEE's there may be language dependencies (e.g., the Ada exception-handling capabilities) that can be exploited to reduce the extent to which the TSE duplicates or makes use of operating system functions.

Other Passive Resources and Decision Support Services

... e.g., data gathering, error logging, metrics, reports.

Human Factors and User Interfaces

... gaming, data reduction and presentation, bit-mapped displays, mice, windowing, availability of source code.
Concluding Remarks

The TSE will certainly inherit many other capabilities from the larger SEE. For example, the SEE is expected to incorporate extensive interactive aids including advanced screen management functions that will allow for the windowing, graphics and data display services that should be present in the TSE. By the same token, standardization on a "goal TSE" should follow the same course taken by the SEE at large to insure not only commonality of environments but also interoperability with related systems.

No known technical barriers exist to the implementation of these recommendations. In fact the prototypes which offer sample designs in furtherance of these requirements exist in academic, government and industrial research laboratories throughout the country. It is the recommendation of STEP that the first designs of a TSE attempt to exploit the already considerable research investment in software test technology that has been made by the nationwide community of software engineers.
The Software Test and Evaluation Project was asked by the Director Defense Test and Evaluation (OSD/USDR&E) to assist in reviewing the Identify Friend, Foe, Neutral Software Test Program. The review was conducted from March 1985 through October 1985. This report contains the results and recommendations of the review.
Table of Contents

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0</td>
<td>1</td>
</tr>
<tr>
<td>1.0</td>
<td>Introduction</td>
</tr>
<tr>
<td>2.0</td>
<td>2</td>
</tr>
<tr>
<td>2.0</td>
<td>Content of the Report</td>
</tr>
<tr>
<td>3.0</td>
<td>2</td>
</tr>
<tr>
<td>3.0</td>
<td>Letter to Col. Crosby Documenting the Results of the 13 June Meeting at Kirtland Air Force Base</td>
</tr>
<tr>
<td>3.1</td>
<td>4</td>
</tr>
<tr>
<td>3.1</td>
<td>Attachment I to 28 June Letter, Minutes of 13 June Meeting</td>
</tr>
<tr>
<td>3.2</td>
<td>6</td>
</tr>
<tr>
<td>3.2</td>
<td>Attachment II to 28 June Letter, Written Comments on the Statement of Work and the V&amp;V Plan for the Follow-on IV&amp;V Contract</td>
</tr>
<tr>
<td>4.0</td>
<td>8</td>
</tr>
<tr>
<td>4.0</td>
<td>Review of the Technical Evaluation Plan</td>
</tr>
<tr>
<td>5.0</td>
<td>14</td>
</tr>
<tr>
<td>5.0</td>
<td>Review of the Technical Proposals for the Follow-on IV&amp;V Contract</td>
</tr>
<tr>
<td>6.0</td>
<td>19</td>
</tr>
<tr>
<td>6.0</td>
<td>Letter to Col. Archino Stating Recommendations for Enhancing the Overall Software Testing Activities on IFFN</td>
</tr>
<tr>
<td>Attachment I</td>
<td>21</td>
</tr>
<tr>
<td>Attachment I</td>
<td>Chronological Listing of Meetings Attended</td>
</tr>
</tbody>
</table>
1.0 Introduction

The Identify Friend, Foe, Neutral (IFFN) Program Review Task was initiated by the Joint Test Director of IFFN, Colonel David T. Archino, in concert with the Director Defense Test and Evaluation, Mr. Charles K. Watt, on 6 March 1985. The program was perceived by Col. Archino and Mr. Watt as having some problems in the software. The issue was initially raised due to a six month schedule slip in the software and the fact that one of the subcontractors' software was not able to be integrated into the system. Col. Archino and Mr. Watt asked the Software Test and Evaluation Project (STEP) to review the software, its development, test and management, and make recommendations to Col. Archino.

2.0 Content of the Report

A list of mutually agreed to review tasks for STEP to perform on IFFN was developed at a meeting on 13 June, 1985. The results of each review task were documented in a memorandum to the IFFN Joint Test Force (JTF) documenting the results of each review. The content of this report is a compendium of those memoranda.

This report has been organized in chronological order to facilitate a historical perspective of the activities we undertook in reviewing IFFN. Attachment I is a listing of meetings attended as a part of the review activity.
3.0 Letter to Col. Crosby Documenting the Results of the 13 June Meeting at Kirtland Air Force Base
June 28, 1985

Lt. Col. Crosby
IFFN-JTF
Kirtland Air Force Base
Albuquerque, NM 87117-5000

Dear Lt. Col. Crosby:

Enclosed herein are two attachments that document the phone conversations we have had over the past two weeks.

The first attachment is the minutes of the meeting on June 13, 1985 in Albuquerque. Included in those minutes were our recommendations for the testing of IFFN.

The second is the written comments on the IV&V SOW and V&V Plan. The written comments include those items I transmitted to you over the phone on June 21, 1985 per the meeting minutes referenced above. The additional items included in the written comments are editorial in nature and are included for your disposition.

Sincerely,

W. Michael McCracken
The Software Test and Evaluation Project

WMM/esr
cc: with attachment 1 Col. Moore
      with attachment 1 Col. Archino
MEMORANDUM

TO: Lt. Colonel Crosby

CC: Colonel Archino
    Colonel Moore

FROM: W M McCracken

DATE: 6/17/85

SUBJECT: IFFN Meeting at Albuquerque, 13 June 1985

The day started with presentations by Logicon on their proposed test approach (to be initiated with Stage 2 testing) and by SAIC on their proposed Technical Evaluation Plan.

The discussions between the JTF and Georgia Tech centered on three areas of interest related to the IFFN program. The three areas were: short term plans, long term plans and quick response items.

1. **Short Term Plans:** Most of the day's discussions centered on the short term plans required between now and 1 August 1985. The short term plans are concentrated on the tasks necessary to complete the review and issuance of the bid package for the IV&V contract. We have been requested to review the SOW for the IV&V bid package by 21 June 1985. At the request of the JTF, we will also support the proposal evaluation during the month of August. In addition, we have been requested to review the SAIC Technical Evaluation Plan by August 1. Finally, we have been requested to review the IFFN JTF Test Support Analysis document on a time as available basis with a suspense date of August 1.

2. **Long Term Plans:** The discussions associated with the long term plans were centered on the possible technical risk associated with the Scenario Planner and the Real Time System. Agreement was made that the Scenario Planner and the Real Time System would probably support the Stage 1 Tests. The question raised by Georgia Tech was, "are the Scenario Planner and Real Time System capable of supporting the remaining tests?" Since it was recognized by both parties that this question was based on subjective observation, it was agreed that Georgia Tech would formulate their ideas on what should be reviewed in these areas.
3. Quick Response Items: These discussions were centered on the Stage 1 test activities that would be occurring prior to the new IV&V contract and the implementation of the new test approach by Logicon. The discussions were concerned with risk aversion. That is, what should the IFFN team be doing that would minimize the risk of downstream problems during Stage 1 testing or as a result of Stage 1 testing.

Georgia Tech suggested that the IV&V contractor be focused on determining the functional capabilities being tested during Stage 1. Since Stage 1 testing is not using a structured approach, the evaluation has to be more subjective. The IV&V contractor should therefore be using techniques that support that type of evaluation. This should include the evaluation of the adequacy of test cases in testing requirements.
3.2 Attachment II to 28 June Letter, Written Comments on the Statement of Work and the V&V Plan for the Follow-on IV&V Contract

MEMORANDUM

TO: Lt. Col. Crosby
FROM: W M McCracken
DATE: 28 June 1985

SUBJECT: Recommended Changes to IV&V SOW for IFFN

The following changes are recommended to be made to the IFFN IV&V SOW and IV&V plan.

1. Changes to the SOW.


1.2 Paragraph 4.1, Change second sentence to read: "The contractor shall establish and maintain an on-site (Albuquerque, NM) technical team, shall designate a team leader to act on the program manager's behalf when appropriate and shall perform all work on-site unless approved by the JTF."

1.3 Paragraph 4.3.1 a, Correct the paragraph to reflect the proposed new test approach to be used by the prime contractor.

2. Changes to the V&V Plan.

2.1 Paragraph 1.2, Correct the overview to reflect the proposed new test approach to be used by the prime contractor.

2.2 Paragraph 2.5, If the prime contractor is generating traceability matrices (as presented at the June 13 meeting), do you want the IV&V contractor to also produce them or do you want him to just review them? This paragraph doesn't explicitly require them to be produced nor could I find a reference to their production by the IV&V contractor elsewhere.

2.3 Paragraph 4.2.1, Change the third sentence to read: "documentation performs according to", to "documentation conforms with".

2.4 Paragraph 4.2.1 j, Delete this sentence. The compliance with DID's should be performed by QA and should be auditable.
2.5 Paragraph 4.2.2, Delete this sentence. The compliance with DID's should be performed by QA and should be auditable.

2.6 Paragraph 4.2.2.1, The second sentence implies a contract requirement on the prime contractor that allows for test documentation to be delivered at the end of testing. If this is not the case, recommend reviewing this documentation at CDR. If this is a contract requirement, recommend early look for technical adequacy at CDR.

2.7 Paragraph 5.2.3.2, second paragraph. Insert the following sentences at the beginning of the second paragraph: "The IV&V contractor is responsible to the JTF to determine the adequacy of testing by the prime contractor to insure that the overall requirements of the system are met. This determination consists of: analyzing the adequacy of the test cases to determine that the requirements under test have been verified; analyzing the completeness of the tests; assessing the impact and risk of any inadequate or incomplete testing; and proposing additions/changes to the planned testing to improve the adequacy and/or completeness of the tests."
4.0 Comments on the Review of the Technical Evaluation Plan

MEMORANDUM

TO:        Lt. Col. Crosby
FROM:      W M McCracken
DATE:      16 June 1985

Enclosed herein are the comments on the above referenced plan. The review consisted of evaluating Sections 1, 3, 4, 5, 8, Appendices B-D and Attachment 1.

The criteria of review was as follows:

1. Does the plan address the issues of the IFFN test program identified to date?

2. Does the TE Plan clearly define the proposed tests and differentiate between those tests used for further evaluation of the system (beyond that planned by the development contractor) and those deemed necessary to overcome deficiencies of the development contractors planned testing.

3. Does the TE Plan include the information necessary to allow evaluation of the proposed TE testing?

4. Does the TE Plan meet the requirements of the IV&V SOW and IV&V plan for IFFN?

5. Does the TE Plan clearly identify all resources necessary to plan, conduct and report the proposed tests?

Note: The actual review comments are grouped according to the four levels of Independent Tests identified in the TE Plan. In all cases, specific comments are referenced by the appropriate section identifier.
General Comment - Nowhere in this plan is there a resource summary beyond that of equipment and conduct personnel. Is the contractor furnishing this data to the JTF for evaluation? If not, there are to many caveats and unknowns in this plan to be able to determine if this is within the scope of the current contract.

Section 1.
Recommend deleting second paragraph of 1.1, the JTF doesn't need to be sold on the merits of IV&V.

1.3.2 - Regression Tests. These are not independent tests. Reference comments on Regression Tests below.

Section 3.
General Comment on Section 3.
This section does not contain enough detail for review of the adequacy of the planned testing. The detailed test areas of Attachment 1 and Appendix B-D contain adequate details for evaluation (reference specific comments below). The additional areas beyond those are general statements of tests and are neither quantified nor are they justified, e.g., items a-k of 3.2.1.3.2 and 3.2.1.3.3, 3.2.1.3.4. and in some cases 3.2.3 and 3.2.4

ETS Suitability Tests
3.2.1.1 - Description

Last paragraph. Is the evaluation of the ETS facility a requirement of IV&V or should it be performed by the Operations Directorate?

3.2.1.3.1 - Evaluation of Computer System Operations

This test has the potential to be a very valuable test for the IFFN JTF. The operation of a system as complicated as the IFFN requires the type of evaluation proposed in this section. This section should be more specific in its definition of the items to be tested and the criteria for evaluation.

Suggest clearer definition of the specific items to be looked at (instead of "will include") and as good a definition as possible of the evaluation criteria. (It is recognized that most of this evaluation is subjective and at best is based on opinions derived from experience with systems similar to IFFN).
3.2.1.3.2 - A-Spec Conformance Testing

The first paragraph describes the use of Attachment 1 for the tracking of A-Spec requirements and the completeness of the testing of those requirements. The statement that Attachment 1 will be used to determine "candidates" for independent test is unacceptable. The attachment or some other table should list the requirements that will not be tested and which requirements should be subjected to independent testing by the IV&V contractor.

The items a-k represent items that will be evaluated beyond those specific items in Attachment 1. Was there a basis for the determination that these additional items are of criticality to the success of IFFN?

The following are specific comments on items a-k.

a. - This item is confusing because it refers directly to A-Spec requirements to be evaluated. Are there additional requirements being evaluated here? Or are there concerns with the planned or conducted tests for this area and these tests are to reduce those concerns?

b. - What is different about this item from FCA/PCA?

c. - How is this item different from Human Factors Evaluation, ref. 3.2.1.3.3?

d. - Is this item based on an identified critical item of IFFN? How does this differ from Attachment 1?

e. - Is the stress portion of this test different from that of 3.2.3?

f. - How is this item different from Attachment 1?

g. - It is assumed that this test is a result of the concerns with the Pre-test software. What is the criteria of success in this test?

h. and i. - These two items are subjective and difficult to evaluate. The intent of these tests is understood but some observable and measurable criteria should be applied.

j. - Is this item based on a concern of the JTF?

k. - The validation of this item has the potential for consuming large resources and these resources should be planned for. This comment is made because of the potential problems of non-repeatable tests and the analysis required to determine why the test was not repeatable.
Technical review of Attachment 1 was not performed due to lack of understanding of the tests performed to date on the IFFN. One comment to be made is the matrix should be marked to delineate which requirements are applicable to each Stage, since it is assumed that every requirement in the A-Spec is not only applicable to Stage 1.

3.2.1.2.2 - Human Factors Evaluation

This test might improved if all of the operator positions were manned by JTF personnel (ref. 5.0). Also resource requirements could be reduced.

3.2.1.3.4 - Environmental Factors Evaluation

Are there governing requirements for these tests (e.g. A-Spec and/or Mil Stds? If so they should be somehow referenced. If not the tests become very subjective.

Regression Tests

General Comment - The test cases referred to in Appendices B-D are the development contractors test cases. If the IV&V contractor is suggesting that these test cases be run by them they are not performing independent testing of IFFN. This comment is not made to criticize the management of JTF, if they have requested the IV&V contractor to run these test cases. It is noted that this is not independent testing because independent testing requires the development of independent test plans, procedures and reports as well as the actual conduct of the tests.

Reliability and Stress Tests

General Comments - The intention of these tests is well founded. The lack of detail of the tests provokes some concern. Stress testing and reliability testing can consume large amounts of resources. For the JTF to properly evaluate the quality of these tests they should be detailed. E. G., page 34, first paragraph "Error conditions can consist of ....". The error conditions of interest should be defined. The definitions should be derived based on testing to date and criticality to JTF. The reference to Paragraph 1.3.3 doesn't clarify this issue. Item a. is repeated in this paragraph. Item b is questioned because it doesn't relate to any requirements of the system. If the IV&V contractor is trying to discover items for system improvement, this should be directed by the JTF.
3.2.3.3 - Requirements/Functions To Be Tested

The reliability test section should be detailed further to state the exact error conditions. The second paragraph of 3.2.3.1, Description, says "Error conditions can consist of...". Whereas, this section explicitly states these as the error conditions to be tested. Each actual error condition should at least have example errors to be tested. Are these errors based on IFFN requirements or experience with IFFN?

The stress test section is more explicitly detailed and justified.

The description starting on page 38 seems to more properly belong with the reliability portion of this section. If the JTF agrees to this type of activity, more detail needs to be furnished. Specifically, what are the empirical results that have been previously measured with this technique and what are the expected results on IFFN. Also, if the empirical technique is referenced it should have been applied on an actual development and not under laboratory conditions to be valid. The JTF, if they agree with the use of this technique, should be given exact cost/resource information, so that they may make the cost/benefit trade of its use. That is, is the cost of using this technique less than the cost of the impact of the undetected errors in the system without using it.

Timing and Response Tests

3.2.4.3 Requirements/Function To Be Tested

The evaluation of I/O should be capable of being measured. It may not be an important factor on this system and thus the resources should not be expended to verify this requirement.

Section 8.

Was the establishment of these priorities agreed to by JTF? The first priority of regression testing minimizes the amount of independent testing being performed by the IV&V contractor. If the regression testing is performed by the development contractor, the remaining resources and items in this plan will result in a reasonable evaluation of the ability of the IFFN to meet its overall system requirements. The three remaining items and possibly some of the requirements listed as not planned to be tested by the development contractor are a reasonable attack at testing the system beyond that of a development contractor and thus will afford the JTF with an independent evaluation of the system.
The schedule as shown does not contain enough detail to reasonably evaluate the planned testing. The schedule should contain enough detail to evaluate each test type for completeness and accuracy in estimated time. The proposed testing if taken as a whole is likely to consume significantly more schedule than shown. The priorities shown in the Appendices are not adequate to allow evaluation of the schedule. Since the schedule does not show the time for each portion, the priorities don't mean anything relative to each other.
5.0  Review of the Technical Proposals for the Follow-on IV&V Contract

MEMORANDUM

To:    Major Paleck
cc:    Col. Crosby
From:  W.M. McCracken
        J.L. Grover
Date:  22 August 1985

This memorandum is in response to your request for assistance in review of the technical proposals for the IFFN, follow-on IV&V effort.

The review of the offerors proposals was based on applying the review criteria, Attachment I, to the Technical Proposal of each offer on a stand alone basis. We reviewed the other aspects of the proposals (Management and Appendices) as background or overview information only. The technical management task (4.1) was referenced as background data. Attachments II-IV contain the comments on each proposal on a strength/weakness basis against each item of the review criteria.
Attachment I

Review Criteria

The review of the offeror's proposals was based on criteria developed from previous recommendations made to the JTF by Georgia Tech.

1) Is the IV&V contractor able to effectively perform the evaluation of IFFN within a constrained budget and schedule.

2) Is the IV&V contractor capable of focusing his effort on the key and critical functions of IFFN while maintaining a balanced evaluation of the total system.

3) Is the IV&V contractor capable of determining and assessing the risks of IFFN, making recommendations for solutions and then determining that the risks have been minimized.

4) Is the IV&V contractor able to report to IFFN management the information necessary to allow them to make programmatic decisions, while simultaneously being able to evaluate and assess the underlying technical aspects used to develop these reports.

5) Is the IV&V contractor capable of being responsive to the JTF's changing requirements for IFFN.
1) The proposal alluded to the constraints of budget in several sections, but did not offer recommendations on how to solve the issue.

Page II-41-1., 5th sentence from bottom of page, does address the issue relative to test requirements matrices.

Several references are made (e.g. Page II, 4.1.1., 4th sentence from bottom) to the benefit of SAIC's current work being able to avoid a new contractor's having to repeat the work as a start up task.

Bottom Line: SAIC feels its current approach if continued will satisfy this item.

2) The proposal alluded to this item in II-8.2.1.1, page II-52 4th sentence and on, in talking about coordinating criticality of functions with JTF.

Bottom Line: Generally SAIC did not address this item throughout the technical section. The item was addressed at the SAT level.

3) Proposal generally addressed this item, but there was no specific risk assessment activity noted in the system context. The identification of problems and solutions was addressed. The Model Requirement Analysis Section (II-5) did talk about the importance and difficulty of evaluating the models and the risk to the program.

Bottom Line: SAIC did present techniques for problem identification (and examples of previous work) and solution but did not describe assessment in the context of program risk.

4) SAIC did not address this item (not required by SOW or CDRL).

5) The proposal mentioned this is an issue that needs to be addressed and they felt they would be able to address it. The proposal did give examples of how they have adapted to change. E.g. change from original test approach by Logicon to SIT, (II-8-1, page II-46) and Validation Tasks, Responsiveness (II-7, a.4, page II-39).

Bottom Line: SAIC proposes to continue its current methods based on their experience on the current contract.
1) The proposal specifically addressed the item. (Reference, page 2-8, Determination of Priorities and 1st paragraph of page 2-8).

2) The proposal specifically addressed this item in all sections of the proposal.

3) The proposal alluded to all parts of this item. It specifically addressed identification of risk on page 2-18, 2.2.2, 3rd sentence page 2-19, and page 2-52, 2.7.2.1, 5th bullet. It did not specifically address recommendations or determination of risk reduction.

4) This item was not addressed in the proposal (not required by SOW or CDRL).

5) This item is referred to in the overview, page 2-8 but is not specifically addressed in the remainder of the proposal.

Bottom Line: Syscon appears to have addressed the first three issues of the review criteria. The concern is whether or not Syscon is technically able to implement its approach within the current schedule.
Attachment IV

-TITAN-

1) This item is addressed in 2.2, pages 2-20, second paragraph, by the use of RETRACC, page 2-20, paragraph 2.2.1. And in general with a proposed approach of reducing the verification effort (sometimes required to be scoped by the JTP).

2) This item is alluded to in the proposal but not specifically addressed until 2.9, Tech Eval, page 2-63.

3) This item is partially addressed on page 2-6 (issue identification and workarounds), but is not substantiated in the remainder of the proposal.

4) This item is addressed on page 2-72, 4th paragraph.

5) This item was not addressed in the proposal.

Bottom Line: This proposal opted to minimize costs by reducing the verification tasks. It did not address what the positive and negative ramifications of that approach to verification would be. It did not address what validation should accomplish other than in a general manner.
6.0 Letter to Col. Archino Stating Recommendations for Enhancing the Overall Software Testing Activities on IFFN.
September 23, 1985

Col. David T. Archino  
Joint Test Director  
IFFN/JTF  
Kirtland Air Force Base  
Albuquerque, NM 87117-5000

Dear Col. Archino,

As a result of our review of the software testing of IFFN, we feel that one of the areas that can be improved is the effectiveness of the IV&V contractor. The current recompetition is a major step to improving the quality of the IV&V effort, but the IV&V contract should be increased in dollar value as well. The current IV&V contract value, as we understand it, is approximately 5% of the total software development. The ratios of IV&V to development effort are typically 10% to 100% (less than a 10% effort is generally not able to generate any significant added value to the extant review capabilities of a program). The range of the ratio is determined by the complexity of the software and the criticality of the application. IFFN is a system of medium complexity, but the application criticality is low (not man or nuclear rated nor critical to the direct defense of the country). Since IFFN is a system that is important to the overall improvement of the defense of the country, we feel you should at least increase the IV&V budget to 10% of the development effort. This would allow for a more effective review and independent test of the system for the follow on stages of development.

Another area we feel could be easily improved is the active participation by the JTF in the engineering support of the system development. Currently, a large portion of your staff is devoted to the operations of IFFN and the result is the software engineering function is minimized. If you can increase the staffing of the operations side of the program, some of your software trained people would be able to directly participate in the development and test of IFFN.

Please let me know if I may be of further help to you in the success of IFFN.

Sincerely,

Richard A. DeMillo  
Project Director  
The Software Test and Evaluation Project

A UNIT OF THE UNIVERSITY SYSTEM OF GEORGIA  
AN EQUAL EDUCATION AND EMPLOYMENT OPPORTUNITY INSTITUTION
# Attachment I

## Chronological Listing of Meetings Attended

<table>
<thead>
<tr>
<th>Date</th>
<th>Meeting</th>
<th>Location</th>
</tr>
</thead>
<tbody>
<tr>
<td>4/15/85</td>
<td>Program Management Review</td>
<td>Kirtland Air Force Base</td>
</tr>
<tr>
<td>4/28/85</td>
<td>Status Briefings</td>
<td>Logicon, San Diego SAIC, La Jolla</td>
</tr>
<tr>
<td>6/10/85</td>
<td>Status Briefings</td>
<td>Kirtland Air Force Base</td>
</tr>
<tr>
<td>8/21/85</td>
<td>Review of IV&amp;V Proposals</td>
<td>Kirtland Air Force Base</td>
</tr>
</tbody>
</table>