A STUDY OF COMMERCIAL INDUSTRY BEST PRACTICES IN TEST & EVALUATION WHICH ARE POTENTIALLY APPLICABLE TO DOD DEVELOPMENTAL TEST AND EVALUATION

July 2002

A study performed by
Science Applications International Corporation

For the

Deputy Director, Developmental Test and Evaluation
Office of the Undersecretary of Defense
(Acquisition, Technology and Logistics)
Washington, D.C.

Foreword

The work documented in this report was conducted by Science Applications International Corporation (SAIC) under contract for the Deputy Director, Developmental Test and Evaluation, Office of the Undersecretary of Defense (Acquisition, Technology and Logistics). Publication of this report by SAIC does not constitute endorsement of it's contents by the Department of Defense nor does it constitute an official position of the Department of Defense.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS

The sponsors and authors of this report express their sincere gratitude to the following companies for their generous contributions of time, information and insight. Their achievements in product development validate the best practices in test and evaluation which each of them employ.

Boeing Commercial Airplane Group

Boeing Phantom Works (formerly McDonnell Douglas)

Cummins Engines – Corporate Technical Center

DaimlerChrysler – Corporate Technical Center

Intel Performance Microprocessor Division

Lucent Technologies

Microsoft Office Systems

Motorola

Northrop-Grumman – Integrated Systems Sector

Northrop-Grumman Ship Systems – Ingalls Operations

Xerox

We also wish to thank the following professional test organizations for their generous contributions of time, information and insight. The provision of standards, processes and documentation governing both Government and commercial testing was invaluable.

Underwriters Laboratories

U.S. Army Development Test Command

On Testing –

"If something doesn't have to work,
We can ship it tomorrow"

- Industry test manager

How much is enough?

"Do as little testing as required. But,
<u>DO NOT</u> take required testing out of the schedule"

-Industry test manager

EXECUTIVE SUMMARY

Overview

A highly competitive market-driven economy has forced commercial industry to become more effective and efficient, and this environment has engendered a set of process, organization, management and operations best practices. Some of these best practices may be applicable to the Department of Defense (DoD) Test and Evaluation (T&E) operations. Accordingly, the Deputy Director, Developmental Test and Evaluation (DD, DT&E), Office of the Under Secretary of Defense (Acquisition, Technology and Logistics), is embarked upon a program to determine commercial best practices and has sponsored this study to identify those best practices particularly applicable to DoD T&E operations.

Study Approach

The study team visited commercial companies well known for producing quality products under the assumption that quality T&E is a significant factor contributing to quality. Companies visited are involved in land, sea and air systems and a variety of DoD relevant technologies. The team met with senior corporate managers, engineers and technicians. All visits were unclassified. A critical ground rule was that companies would not be identified in the body of the report with a particular process or practice because of the potential for unintentional release of company proprietary information. The team focused on gathering information in four areas based on the DoD model for T&E. The categories are Philosophy, Policy & Approach; Test Investment; Test Execution; and Test Evaluation.

Selection of Companies

The team chose top companies to determine the best practices that make them successful. The corporate environment of these companies fosters the development, growth and application of best practices. Six particular traits are common across the companies:

- The corporate structures were stable with infrequent turnover of key personnel, leading to consistency of policy, operations, solutions and investment.

- There is top-down focus on time to market (schedule) and prompt introduction of quality products into the marketplace. Best practices are necessary to achieve a reasonable profit in a highly competitive environment.

- There is corporate emphasis on global consistency of operations to include testing. This puts a premium on consistent, configuration controlled standards and processes.

- There is continuing corporate commitment to a program once initiated.

- Senior personnel are knowledgeable of test and understand the value and cost of testing.

- Corporate management has intense interest in a program's schedule, cost and metrics.

Efficiency and Effectiveness

Successful commercial companies retain the competitive edge by efficiently and effectively conducting design, development, T&E and manufacturing. This is reflected and manifested in T&E in three key ways:

- Early Resource Commitment. T&E is consistently part of the early decision, planning and execution processes.

- Essentiality of T&E. Corporate management recognizes the value and cost of T&E. Commercial companies are more likely to increase T&E in both number of cycles and stringency to assure quality.

- Ensuring Capability. Corporate management takes the responsibility to fund a robust test capability.

Success Engenders Best Practices and Vice Versa.

Successful commercial companies recognize that an environment that breeds new ideas and innovation is essential in maintaining a competitive edge. Their T&E environments are no exception.

Conclusions

A number of commercial best practices are applicable to DoD T&E. The following list of best practices is not presented in any priority order but, rather, in the order these practices are described in the report. However, those practices highlighted in bold print are considered as providing the highest value for the effort involved.

Best Practices

Philosophy, Policy, Approach

- **Recognize that testing is a way to identify and solve problems early in the process in order to control time, cost and schedule late in the process.**
- Recognize that best practices generate success and vice versa.
- Stabilize corporate leadership and test staff and commit to T&E as a key enabler.
- Focus on quality of product and process to drive the efficiency and effectiveness of T&E.
- **Develop consistent processes to ensure consistent products. Understand the value and cost of T&E.**
- Implement efficient and effective test processes in order to compete. Keys:
    - Ensure T&E is consistently part of the decision, planning and execution process.
    - **Early commitment by all stakeholders on required T&E resources.**
    - Certification of T&E processes and organizations (~ISO 9000).
    - Ensuring capital capability.
- **Increase T&E to assure product quality rather than reduce it to save T&E cost.**
- Use metrics and quality control processes to understand how well the test process is operating.

Test Investment

- **Ensure early determination of the investment costs to acquire new capability for program support.**
- Require analytically sound ROI analysis for test investments.
- **Ensure cohesive (year-to-year) investment plans.**
- **Charge cost of test investment to program.**

Test Execution

- **Involve testers and evaluators very early:**
    - **Ensures testers know test requirements.**
    - **Ensures developers know requirements for test.**
- **Capture test costs at program initiation.**
- Emphasize concurrent and integrated T&E.
- Institute formal quality check processes.
- Use System Integration Laboratories and embedded instrumentation.
- Give proper consideration to the use of external test capability in test planning.
- Ensure testers control test planning, equipment, facilities, instrumentation and test resources.
- Continue to increase the use of modeling and simulation to expand the test process.
- Do not generally support the outsourcing of testing and evaluation.
- Frequently use the Six Sigma (6s) or similar quality processes.

Test Execution continued

- Automate data collection and archiving.
- Benchmark in-house and within industry.
- **Use measurements and metrics.**
- Initiate programs to seek ten-fold reductions in the number of software tests required.
- **Integrate Master Test Plans and test execution with program resources and milestones.**
- **Charge full cost of testing to the program.**
- **Establish measures of effectiveness.**
- Quantify risk for management decision when considering reduced testing.
- **Train the in-house test workforce in test engineering disciplines.**
- Emphasize multi-use T&E platforms.

Test Evaluation

- Continue to increase the use of modeling and simulation to expand the evaluation context based on verified test data.
- Correlate faults and solutions in a closed loop process to ensure problems are resolved
- Use Physics of Failure as a tool to predict and analyze system performance and shortfalls.

Test Philosophy/Process/Evaluation
(A combined category)

- Establish corporate internal web based sites for exchange of ideas, benchmarks, data, applications and processes. Address:
    - Data collection retrieval/archiving.
    - Modeling and Simulation.
    - Test and Evaluation methods.

Recommendations:

- Implement or reinforce the bolded best practices in DoD as soon as possible.

- Development implementation or reinforcement strategies for the remainder of the best practices using DoD stakeholder teams.

- Present the results of this study to the DoD acquisition and T&E communities.

SECTION I - STUDY FRAMEWORK

Best Practices in Test and Evaluation

Following World War II, the command and staff model that successfully prosecuted the Allied victory was embraced by the United States' industrial sector as an efficient and coordinated means to expand and capture global markets. This was, in effect, a transfer of "Best Practices." The model worked as U.S. industries became a dominant global force. However, as the world's economies recovered, intense foreign and domestic competition pushed United States industry into finding ways to become more efficient, more cost-effective and quicker to market. Particularly critical was the challenge to produce extremely high quality products. U.S. industries have met the challenge with improved management styles, organizations, processes, methodologies and tools. They have established, individually and collectively, a new set of "Best Practices." Some of these practices are already embedded in Department of Defense (DoD) structures and operations. In fact, some originated there. However, as the Department downsizes and restructures, it is looking to industry for fresh management ideas and tools.

Study Objective

The Deputy Director, Developmental Test and Evaluation (DD, DT&E), Office of the Undersecretary of Defense (Acquisition, Technology and Logistics), is embarked upon a program to enhance the DoD Test and Evaluation (T&E) process. A key part of this program began with an initial study on industry T&E processes, *Best Practices Applicable to DoD Developmental Test and Evaluation,* published June 1, 1999. That study recommended government (1) reduce test cycle time through the use of streamlining and appropriate "fast track" procedures and (2) improve T&E processes. The study also provided a useful characterization of the general state of T&E best practices in industry. Based on that information, the DD, DT&E, sponsored a follow-on study with the objective of gaining additional levels of detail on commercial practices.

This follow-on to the 1999 study, sought to obtain the additional level of detail directed by the DD, DT&E. Science Applications International Corporation (SAIC) was tasked to conduct the study. SAIC formed a team of consultants consisting of former U.S. Government Senior Executives with extensive recent experience in both testing and evaluation. The team's challenge was to uncover T&E best practices of companies with commercial customers.

Once company T&E practices were identified as potential best practices, they were then examined for potential DoD application.

The team sought to cover multiple DoD applicable technology sectors. The focus was on companies well known for producing quality products, the underlying assumption being that quality T&E is a significant contributor to product quality. In that regard, the team visited each company identified in the Acknowledgment section of this report. All companies were accommodating, forthcoming and, somewhat unexpectedly, very interested in the government T&E processes.

Study Approach

The team sought to conduct on-site, one-to-two day visits with each company. The visits were preceded by extensive discussions to establish the basis for the visit and the team and company visit "ground rules." A package was sent to each company detailing the purpose of the study. The package included an extensive question set (see Appendix C) to frame the visit. Once the team was on site, each company proposed its method for transmitting its information. The sessions ranged from intense discussions with key people, to combination discussion and company tours, to briefings set within particular sectors of the operation. The assumption that the companies would know the best way to transmit their information proved useful. Another assumption was that short visits would minimize the impact on the companies and make the visit palatable to them. It was also assumed that if companies known for quality products were visited, trends in best practices would emerge. All assumptions proved correct.

The team met and discussed issues with senior corporate managers, engineers and technicians (see Appendix A). In all cases the companies were well prepared to discuss their T&E practices and processes. The team took extensive notes and, where requested, returned draft summaries to the companies for comment and review. All visits were unclassified with the understanding that the team would not receive company confidential design and production information. This did not prove to be a hindrance.

A critical ground rule was that in the body of the report companies would not be identified with a particular process or practice. This was an essential condition for admission into, and discussion with, most companies, as all were concerned about release of company proprietary information. Accordingly, all reporting is "normalized" to the twelve companies visited. A second condition was that each company receive the final report.

Selection of Companies

Companies producing quality products are not difficult to identify. The team's first step was to establish relationships necessary to gain entrance to those companies. The most difficult task, in most instances, was identifying the appropriate people and actually obtaining visit invitations. The team ultimately linked up with the very people who were most knowledgeable of T&E. Once entry was gained, very useful discussions ensued.

The team covered air, land and sea commercial systems, and to some extent, space operations. Also, from a DoD relevant technologies view point, the team visited companies involved in DoD aircraft and avionics systems, ground vehicle systems (including power plants), C^3I systems, software systems, computer technology systems, sea surface systems, information technology systems, and microelectronic systems. A cross-section of as many technologies as possible was desired to mirror the extensive nature of DoD developmental testing.

The goal was to visit twelve relatively high profile organizations. Since commencing the study, at least two mergers have resulted in the companies visited becoming sectors of a larger parent organization. However, because of the timing of the mergers, the operations within each sector remained sufficiently unique so as to permit different perspectives on the T&E process. More than thirty major companies were contacted. Responses ranged from no response, to no participation desired, to withdrawal of initial invitation, to acceptance. At the request of some companies, the team provided interim feedback for company use in their continuing quality reviews.

Focus

The team focused on gathering information in four distinct categories based on the DoD model for T&E. The first category was *Philosophy, Policy and Approach*, the second was *Test Investment*, the third was *Test Execution* and the fourth was *Test Evaluation*. While these categories did not map precisely with how industry conducts its T&E operations, it did permit a correlation of information and best practices to the general DoD model. During the visits, the team found the four categories intertwined in most organizations. The company discussions (companies were encouraged by the team to dialogue in their terms and within their organizational philosophy) often wove back and forth across the categories and the lines were more blurred than is normally the case in DoD. This free flow produced substantial information.

Caveats

The views and conclusions of this study are largely those of the study team as closely overseen by the DD, DT&E, and the senior management of the SAIC Test and Evaluation Group. In fact, the DD, DT&E, personally accompanied the team on selected visits. While a substantial number of "best practice" trends emerged, the team makes no claim that they have developed an exhaustive list.

The DoD 5000 series guidance, together with the services' implementing documentation, provides a basic standard for DoD T&E. No such documentation exists industry wide, although a number of practices are common, and industry does extensively use benchmarking and similar practices. Likewise, even with DoD standards, the Service's T&E structures and operations are less than homogeneous.

Certain practices in industry are proprietary and some T&E practices are so intertwined with the design, development and production processes that they are not suitable for full disclosure. The ground rule that practices would not be linked to specific companies within the report, and the report's subsequent review by several of the companies has, we believe, largely offset this disadvantage.

The canvass of companies was limited to twelve organizations. While more companies may have added a richness of detail, we believe that the frequency of emerging processes across several companies and across technology sectors suggests that the twelve were a reasonable sample.

There are some key differences between government and industry T&E structures. The government has two distinct types of T&E, Operational T&E (OT&E) and Developmental T&E (DT&E). The government also has two separate organizations at the DoD level for oversight and execution. Industry generally has only a single organization for all test and evaluation. That organization typically conducts both developmental and customer testing and does the assessment or evaluation. Often, the organization is not called by the title of T&E but may be referred to as a quality engineering or validation engineering group. The team attempted to correlate between the commercial sector entities and the government model as best possible in each visit.

Notwithstanding the differences, there are sufficiently distinct similarities between industry and government T&E that it is logical to take industry best practices and consider their applicability to the DoD.

SECTION II - CORPORATE PHILOSOPHY, POLICY AND APPROACH

Context

This report documents the investigation into the test and evaluation practices of companies that have successfully introduced quality products into the market place. It is important to broadly characterize some traits of the corporate environment that make them successful and allow best practices to develop and flourish. Best practices in this section follow.

Company Traits

Six particular company traits were identified, that when combined, frame the successful contributions of test and evaluation to product quality. The DoD T&E organization should try to emulate these factors.

- Stability - The corporate structures are almost uniformly stable. [BEST PRACTICE] There is relatively little turn over of key personnel and that, in turn, leads to consistency of policy and operations. The environment is conducive to reasoned long term solutions and investment in those solutions.

- Focus - There is top-down focus on introducing quality products into the marketplace as a prerequisite to corporate success. [BEST PRACTICE] While production of quality products is a source of company pride, it is obvious that the companies clearly understand that their products must be of high quality to compete. This puts tremendous pressure on companies to introduce products quickly into the marketplace while constraining product cost.

- Consistency - Companies emphasize global consistency of operations, including testing. Due to international market competition and reputation, a product produced by the company in one of their plants in another country must be of the same quality as one produced in the United States. This puts a premium on *consistent, documented, configuration controlled standards and processes*. [BEST PRACTICE]

- Commitment - There is continuing corporate commitment to a program once initiated. [BEST PRACTICE] There are few internal 'politics' to affect a program. Program managers may compete for support funding or capital investment money, but they do not compete their programs. Program termination occurs for rational reasons such as poor product performance or financial reasons. An example would be the entry of an additional competitor into the market creating a reduced market share and an unacceptable return on investment (ROI).

Program managers spend their time managing their programs, not restructuring their programs.

- Knowledge - Senior personnel, including the CEO in some cases, are knowledgeable about testing and understand the value and cost of testing. [BEST PRACTICE] *The acquisition managers understand testing, and testing is visible to at least the business sector Vice-President level. They view testing as an enabler for product success and as an engineering discipline of the design, development and manufacturing processes.*

- Metrics - Corporate management has intense interest in the schedule, cost and metrics established for program success. [BEST PRACTICE] They do not micromanage; rather, they delegate responsibility and hold program and test managers accountable for meeting the metrics.

- Market Driven - Successful commercial companies satisfy marketplace demands to remain competitive and profitable. [BEST PRACTICE] They must meet market opportunities promptly with high quality, competitively priced products, whether it be airplanes or microchips. T&E is an enabler to those ends.

Efficiency and Effectiveness

Successful commercial companies retain the competitive edge by efficiently and effectively conducting design, development, T&E and manufacturing. This means that their organizations, and particularly their processes, must collectively be excellent and continually improving. [BEST PRACTICE] In application, this requirement is particularly applicable to T&E in three key ways:

- Early Resource Commitment. *All corporate stakeholders, including T&E organizations, agree at program initiation on the necessary resources to develop and field a product*. A comprehensive and tightly documented plan addressing technology risk, market needs, schedule and resources drives the product process from design to manufacture. T&E is consistently part of the decision, planning and execution process.

- Essentiality of T&E. Corporate management recognizes the value and cost of T&E. T&E is incorporated as an essential enabler and driver in producing safe, high quality products. [BEST PRACTICE] Although product cost and schedule are critical elements, *commercial companies are more likely to increase T&E (test cycles and stringency) to assure quality than to decrease T&E for other reasons*.

- Ensuring Capability. Corporate management recognizes that human resources and physical capability are critical in achieving product quality. Management takes the responsibility to adequately fund both of these areas.

DoD should follow this efficiency and effectiveness best practice approach.

Success Engenders Best Practices and Vice Versa

The companies encourage their workforces to engender new ideas or best practices as essential to their continued success. They also recognize that success engenders enthusiasm for new ideas. [BEST PRACTICE] DoD T&E organizations should encourage development of best practices and identify where best practices have contributed to organizational success. Industry facilitates this practice by implementing internal web sites for the exchange of ideas and applications, the execution of internal benchmarking and the exchange and cross provision of data and processes. Information moved on such web sites includes models and simulations, test and evaluation methods, and archived data.

SECTION III - TEST INVESTMENT PROCESS

Context

This section addresses best practices pertaining to investment in, and acquisition of, instrumentation, equipment, facilities, data and personnel. All the companies possess facilities, people and data base systems. Two companies have relatively little investment in instrumentation and test specific equipment due to the product they produced and the testing it required.

Early Involvement

In virtually every company visited, *tester involvement occurs early in a program*. [BEST PRACTICE] Almost uniformly, they emphasize that test managers and test engineers are in partnership with development organizations from the very beginning of project initiation. *The companies consider it essential for the test and program organizations to reach consensus on the resources and equipment required to support a program. This is not a case of inviting the test organization to join the project; but rather it is an established process where the test organization is always part of the team from project initiation forward.* While this approach is sometimes followed within DoD, full and willing implementation should produce benefits such as ensuring the on-time availability of test equipment for test program support. Test provisioning planning should address all potential funding sources and the working relationship between test managers and program managers that will ensure the right capability is delivered. DoD needs to ensure the capital investment for test requirements is identified and planned early in each program.

Return on Investment Analysis

It is generally the case that companies require a thorough ROI analysis be performed prior to any major acquisition of instrumentation, equipment and facilities. [BEST PRACTICE] They require a rigorous definition of shortfalls as part of the ROI analysis. They also, as a rule, must formally consider alternatives to outright purchase of test capability(often Government test ranges). Technology is less a factor in acquiring instrumentation, equipment and facilities than is meeting a specified test capability need. Commercial companies replace standard instrumentation when it wears out or when it can no longer meet requirements. Both the test manager and the program manager must generally agree with the requirement and the ROI analysis. The analysis is typically approved at senior levels. The level of approval depends on cost of the investment. For any significant amount, it is usually approved at the Vice-President level.

In several companies, process discipline comes from showing the post-testing program ROI associated with the test equipment acquired. This practice was surprisingly similar in many companies and should be applied within the DoD T&E community.

Charge Back for Investment Cost

With one exception, *each company allocates equipment investment cost back to programs.* [BEST PRACTICE] When a large program needs new equipment, the large program is charged the entire cost of the capital investment. The cost, in turn, was allocated against the product cost. Where there are many small programs, the cost is allocated proportionally across the programs. This practice ensures that test equipment investment is a front-end investment that is recovered as part of the cost of producing the product rather than a company overhead item. This approach engenders an interesting behavior. Because employees have ownership in selling company products (market competitive), all are interested in keeping product cost as low as possible. Therefore, the testers, wishing to acquire new equipment, have an interest in exploring all alternatives and conducting a high quality ROI analysis.

There are some funding challenges in DoD that may not permit charging back over the life of the program. First, the funding structure does not allow for revolving funds for test instrumentation in RDT&E. Second, the accounting system is not set up to handle a charge back arrangement over the life of a program. Third, cost of the product for market sales is not a meaningful concept in DoD. This best practice should be considered, but may not be feasible.

Use of Outside Facilities

There are seemingly conflicting practices regarding companies' use of outside facilities. One practice used by virtually every company that builds products having DoD application, DoD common elements, or DoD common environments, is to use, to the extent possible, government test ranges/facilities. Companies using government ranges and facilities clearly appreciate the test capability and the investment in instrumentation, equipment and facilities that the government has made.

A general best practice is to certify component suppliers to test their own products and provide the test data at component delivery. The certification and quality control process for the suppliers is extensive and detailed. It includes certifying supplier equipment and processes, and the continuous tracking of delivered quality and timeliness of deliveries. Certification takes place by training the supplier in the acceptable standards and processes consistent with the methods by which the company conducts its T&E. That establishes a baseline for performance, and a historical record is tracked with investigation of deviations in product performance from the supplier. An additional certification check is often made by reviewing test plans in advance and almost always by review of the test report that is required to accompany the suppliers' products at delivery.

It may be useful to have DoD test organizations certify supplier test data. It is conceivable that this front-end effort may result in significant savings associated with a reduction of end item re-test.

Notwithstanding the occasional use of government test facilities and the use of certified supplier test data, nearly every company indicated a preference not to outsource testing. [BEST PRACTICE] They cite the same reasons that government test ranges cite in response to outsourcing testing of government equipment. First, they cite having priority control over test facilities and resources; being market driven, they are naturally very conscious of schedules. They do not believe they can consistently achieve schedule priority under an outsourcing arrangement. Second, they believe (supported in some cases by analysis) that outsourcing is more expensive. Third, they do not find that outsourcing provides the same test data consistency as data gotten from in-house testing. Finally, they cite pride in company and product as a reason to retain in-house testing. DoD should reverse the trend to outsource testing.

Tools

Several of the companies invest in and use simulation networks, automated databases, built-in instrumentation and multi-use platforms. Two companies are elements of the same larger organization and two companies are suppliers of other companies the team interviewed, the best practices, while remarkably similar, were clearly all developed independently.

Simulation nets in one form or another are used by many of the companies. These nets allow worldwide access by company test engineers, development engineers and appropriate others. [BEST PRACTICE] Discussion of the simulations is conducted over the company web, and documentation of the simulations (including appropriate changes) is posted. Several companies also include applications, and at least one company conducts company-wide "contests" to promote understanding of applications. It should be noted that the context of that practice in that company is a quality application to include the simulations network. A similar web-based approach should be possible and useful in cross discussion of simulation application within the DoD T&E community.

Automated database networks, with access limited due to concerns about proprietary information, exist in nearly every company visited. The databases are updated frequently during testing and all data are archived. Similar automated databases exist in many DoD locations. The DoD test community may wish to explore the accessibility, the commonality of structure, the interface of separate databases to facilitate analysis and the maintenance costs of such networks.

Most of the companies use on-board or built-in instrumentation. This is also done in the DoD test community. Likewise, several of the companies consider investment in System Integration Laboratories (SIL) to be very beneficial. [BEST PRACTICE] Again, this is not unusual within the DoD development community.

The multi-use platform is an investment tool under development or in use in at least half of the companies. [BEST PRACTICE] It permits test of systems and subsystems and multiples of systems for performance within varying environments. It allows assessment of product performance within multiple interfacing product environments, in co-site interference environments, and in countermeasure and other environments. For example, a series of tests might test a single system/item for specific functions and then test a sequence of functions in specific environments or a sequence of environments. Next a following series might test multiple common interacting systems in a series of varying environments. A final series of tests might test multiple, non-common systems interacting in multiple, varying environments. DoD should consider the broader use of the above tools.

Investment Plans

Every company that makes an appreciable investment in test equipment, *instrumentation and facilities has a multi-year investment plan*. [BEST PRACTICE] The plan achieves consistency of investment programs over time and requires, where appropriate, a post-purchase documentation of cost savings/avoidance. Changes to the plans are carefully scrutinized. DoD also has multi-year plans and should ensure that continuing scrutiny of investment plans is an effective dimension of multi-year plans.

Control of Test Instrumentation and Equipment

Generally, companies place their test managers in control of test equipment, including its scheduled use. At two companies where testers, developers and program managers are tightly integrated, the responsibility for the priority use of equipment may temporarily shift depending on the phase of product development. However, as a rule, virtually all companies consider it a best practice for the testers to control test instrumentation, facilities and equipment. [BEST PRACTICE] This is currently the practice in DoD and should be maintained.

Investment in People

Several of the companies clearly related that investment in people as a “capital” resource, without question, is more important than investment in equipment. They believe that experienced, in-house people are critical to ensuring consistent and quality testing. They believe that organizational consistency is critical to ensure continuity of test processes and knowledge. They believe in investing heavily in education and training. [BEST PRACTICE] This training heavily emphasizes test and evaluation methodology and tools. They rely on the knowledge of their test workforce to test, to diagnose problems, to recommend solutions, and to meet schedules as critical skills in designing and producing quality products. DoD trains testers extensively in the acquisition process, but it should seek ways to incorporate more test and evaluation related material in that training for test personnel.

SECTION IV - TEST EXECUTION PROCESS

Context

This section deals with the best practices within the test execution process. By virtue of the wide variation of end items, actual testing varies widely in terms of scope, duration, instrumentation and facilities. Interestingly, in many instances, the best practices of the varied organizations are remarkably similar. This suggests that in DoD, common test practices can often be applied across the board.

Early Entry and Full Participation

In virtually every company visited, the test organization and test engineers get involved early. [BEST PRACTICE] Early involvement in commercial organizations literally means at the beginning. Test engineers are involved in program design, scope, structure, goals and cost determination. The test organization, test management and test engineers are essential elements of the program management team from program initiation. Testers are considered part of the solution, not part of the problem. Of all practices observed, this one was uniformly followed and the most strongly emphasized. *Testers are expected to develop good, tough and appropriate tests to evaluate product quality, while developers concentrate on fixes when testing reveals problems.* The key mission for the testers is ensuring product testability and capability to test.

The program managers drive the test programs, but it's a collaborative effort involving design, development and manufacturing. Nearly every company has a team arrangement that includes the test element. By whatever name they are called, shared feature teams, product teams, or multidisciplinary teams, they have the same underlying feature in that they are a critical and continuing link in the process.

DoD policy calls for early participation of testers in the acquisition process. In practice, the policy is not always followed and testers are not always viewed as full team members. The overwhelming support this practice receives in industry suggests that DoD should fully enforce its own policy on early participation of test and evaluation.

Cost of Testing

Generally, *the cost of testing is borne by development programs.* [BEST PRACTICE] Two companies visited pay for T&E from corporate or sub-corporate overhead. Even in those cases, it is allocated back through the accounting system to total program cost. This seems to be driven less by best practices than by the need in commercial organizations to define a product's production cost. It is theoretically possible to allocate all test cost to programs in the DoD. However, inconsistent workload at DoD facilities (especially at the change of the fiscal year) and the inability of program management budgets to accommodate the high cost of facilities discourages direct allocation of cost. This should not preclude DoD facilities from developing estimates of allocated institutional cost such that the DoD knows the full cost of programs.

Companies also seek to *determine the T&E cost estimate early in the program so that a credible cost estimate is incorporated into the corporate program plans*. [BEST PRACTICE] An interesting comment by at least three companies was that, at least to some extent, the cost of testing doesn't matter when the goal is to get the product working properly. Likewise, test cost is not as important as product quality and schedule. It is a major exception for them to release a product with a known problem. *Normally they will intensify test effort (and test cost) to fix the problem and maintain schedule, if possible.* DoD should seek to determine firm test cost estimates from the testers early in program development. These costs should be firm, negotiated estimates derived through negotiation between test activities and program managers. Both should concur in the final estimates.

Concurrent and Integrated Testing

Nearly all companies conduct concurrent testing. Test planning continually seeks to schedule non-interfering tests in a front-loaded arrangement. [BEST PRACTICE] GANT charts or similar tools were in evidence in nearly every organization. The drive to get to market makes this a necessity in commercial organizations. It is followed in DoD in many instances as well. Because of the value of this practice, it should receive additional emphasis in DoD.

Integrated testing is also the norm in nearly every organization. In several instances, corporate customers participate in product development by testing the product in an operational setting. This allows concurrent measurement of customer satisfaction with product performance and system technical performance. DoD is moving rapidly into combined developmental and operational testing. This practice should continue to receive DoD impetus.

Test Process Quality

Virtually all companies agree that the quality of the test process is everyone's responsibility and must be a daily way of life. One company expressed quality as a journey, not a goal. It cannot be inspected in. Time and cost savings are a must, but not at the expense of compromising a product's quality. Top management drives quality by maintaining a focus on finding and fixing problems and accepting the test process as a key engineering enabler. One of the companies outlined a simple but effective "quality drill" as follows:

- Measure the performance and process to know what you have.
- Analyze the problem using data available (if you measure, it will be available)

- Problem solve as a team. Look at it from many views.
- Develop an action plan to implement the solution.
- Institutionalize the solution as applicable.
- Document the solution and continue the process.

Test Organization Certification and Quality Process

Virtually all companies are quality oriented, most of them not International Organization for Standardization (ISO) 9000 series qualified (because of varying sets of ISO 9000 certification levels and standards/processes applicable to and pursued by different companies, this report will simply refer to the collective standards as ISO 9000 series). Some are also qualified or certified under other standards such as SEI Levels III and IV. Underwriters Laboratories has every laboratory and test organization accredited by some documented national or international standard or process. [BEST PRACTICE] While two organizations are not ISO 9000 series qualified, they have internal quality systems very similar to the formal systems. Although resource intensive in implementation, every organization that has undergone the ISO 9000 series process indicated it was a positive process in developing a quality culture in the organization, that it forced development of testing standards and that it forced review of all processes resulting in multiple efficiencies. At least one company conducts assessments against their ISO 9000 series standards every six months.

Implementing certification and accreditation programs in DoD test operations and organizations could be important in assuring international acceptance of U.S. test data (note the acceptance of UL as the authority on product safety). It could also be the key to ensuring the integration of testers with developers and the early acceptance and influence of testers in the development process. The credibility of a formally certified or accredited organization should be greatly enhanced. Efficiencies derivative of the process should become evident to both test and program managers.

Six Sigma (6s)

Several organizations are using or beginning to use Six Sigma (6s). [BEST PRACTICE] The 6s (or reliability of .99999) process (see Appendix B, Special Topics) is a tool to attain very high quality in design, production, testing and other operations. This tool is typically applied to all processes. Everyone in the organization is schooled in the Six Sigma culture. The goals are cycle-time reduction and customer satisfaction. Entire workforces take some level of formal statistical training. As with ISO 9000 series certification, the intent is to instill a quality culture from the CEO down. It becomes a daily way of life, and functional managers are required to find solutions to defects while proactively seeking workforce ideas and input. The process applies to manufacturing, procured components, product performance, testing – everything. While the DoD developmental test organizations do not produce a market product as such, application of the 6s culture and process to test processes and quality would be significant. The process can be applied to instrumentation output, report timeliness, test time estimates, test cost estimates and test delay reductions.

Outsourcing of Testing

The team expected to find outsourcing of testing to be a rather common practice. Surprisingly, just the opposite was found. None of the companies visited outsource extensively and most virtually not at all. Most of the companies *develop and use their own in-house test capability.* [BEST PRACTICE] They want test consistency and continuous development of in-house expertise and experience. They believe that it is essential to test in-house so they can control the schedule and cost, while taking pride in their accomplishments. They keep their test organization largely intact. They consider this practice essential and consistent with test philosophies such as "no room for mission error," "quality is everyone's responsibility" and "continuous triage." They accept supplier testing on components, but almost every company has a quality certification program with their suppliers to ensure consistent test execution and consistently high quality products.

Modeling and Simulation

Modeling and simulation (M&S) is used extensively. [BEST PRACTICE] Typically, the companies insist on verification, validation, and accreditation (VV&A) for the proper application of the models, and they document the VV&A process prior to using models or simulations for system test. However, while all are committed to M&S, the actual use of M&S varies with company and by product. In aircraft testing for example, extensive use is made of M&S to design the aircraft, its components and for avionics and electrical systems interoperability testing. *Flight-test programs are not simulated. The goal is not to replace end-item testing but to reduce the problems found in end-item testing to preclude expensive retest.* One company commented that the key to structural validation (equations backed by science) is rigorous analytical review of input based on historical data from physical validation testing and of the input assumptions that must precede validation of the output.

On the other hand, virtually all testing of some electronic systems is simulated using tools such as interchangeable virtual instrumentation platforms. One company, not an electronics manufacturer, has a goal of increasing test M&S use from a current 40 per cent to more than 70 per cent.

Test Procedures and Standards

Virtually all of the companies visited have or are working to *achieve consistency and integrity of test procedures and standards*. This is a particularly important effort in recently merged companies with worldwide test operations. Many have corporate level teams working on the problem. In many cases, there is a corporate organizational element that serves as the repository configuration controller for test procedures and standards. All have formal documentation in print, on the web, or both. The "Inspection Manual Instruction Document" and the "Corporate Purple Book" are examples. Underwriters Laboratories (UL), for example, has 750 safety standards in publication. Another publication, UL1998, contains software safety standards.

Data Collection

Nearly every company uses some form of automated data collection and archival system. [BEST PRACTICE] Data is made available through secure internal web sites that feature firewalls, encryption codes and accreditation requirements. One company uses an artificial intelligence based system. Some company's archival systems include the test requirements, the test standards, or both. Standard architectures and central configuration control of the databases are also being pursued to ensure consistency of structures. Links are being pursued to on-board test instrumentation or internal test data hooks to permit real-time visibility of test data. DoD should develop interfacing web-based data collection and archiving systems among the DoD test organizations.

Benchmarking

Virtually every company visited does benchmarking. [BEST PRACTICE] Externally, the key is to benchmark to the "best in class," to both customer need and to industry benchmarks. Companies participate in industry wide quality standards meetings, symposia, or workshops focused on quality tools and practices. All view operating to industry standards as essential to ensuring customer satisfaction and public credibility.

Several companies also conduct internal benchmarking by exchanging information, often through internal web site knowledge navigators. The intent is to promulgate ideas, indicate successful performance against standards, seek input to standards, provide company-wide feedback on internal or external quality-based operations, and chart application measures of effectiveness. Benchmarking should be the norm in DoD testing. Benchmarking should be conducted relative to the best in industry and among DoD test organizations.

Measuring and Establishing Metrics and Metric Criteria

A practice used by several organizations, especially those with a Six Sigma or quality driven culture, is to *measure everything*. [BEST PRACTICE] One senior manager, responsible for supplier quality, stated that by simply starting to measure performance and keeping statistics, performance increases (defects decrease) by 25 percent without any other initiative. There are instances of this happening in DoD test organizations as well. A senior test manager emphasized that quality is everyone's responsibility, and it begins with knowing what you have, knowing how you are performing and knowing where you are in comparison to past and future. Measuring performance, quality, defects, timeliness, schedule deviation and every other significant parameter ought to be a continuing process in DoD test operations.

In conjunction with measuring everything, commercial companies stress the need for meaningful metrics. An institutional best practice in at least one organization, and which other organization are now applying, is developing good metrics. They suggest the following criteria for good metrics:

- The definition and purpose are clear.
- The baseline and goals are easily set.
- The goals are aggressive but achievable.
- The ownership is clear, unambiguous and understood.
- The internal quality program can be applied.
- The sources of data are clearly defined.
- The data can be easily collected and reported.
- The metrics address at least one of the following characteristics:
    - Provides clear warning or pre-warning of a problem.
    - Correlates to key customer issues and satisfaction.
    - Improves or drives a process.
    - Fixes a problem.
    - Correlates to profit.

DoD test organizations should begin the process of establishing metrics and measure against those metrics. Effectiveness and efficiency should begin improving immediately.

Reduction in Software Testing

Companies producing software intensive products are exploring the potential for reducing software testing. [BEST PRACTICE] Solutions are not yet complete, or are otherwise unavailable due to their proprietary or competition sensitive nature. Software test reduction goals of one to two orders of magnitude were discussed, and the application of Physics of Failure to software testing is being explored by more than one organization. (See Physics of Failure discussion in Appendix B, Special Topics.) Reducing the number of software tests and the magnitude of testing has significant potential and should be a key technical initiative for DoD software testers.

Master Test Plans

Most companies believe *a detailed, complete and documented master test plan is essential to efficient program execution*. [BEST PRACTICE] Master test plan content varies by company and test complexity. The companies also have program management plans. Master test plans follow the program plans with milestones common between the two. The master test plans are more detailed than a DoD Test and Evaluation Master Plan (TEMP) but broader in scope than a DoD test facility detailed test plan. They appear to be somewhat analogous to DoD Live Fire Test plans but cover the complete range of program testing. Test management organizations and personnel develop the test master plans and is responsible for (1) ensuring consistency with the program plan, (2) coordinating with supporting and program elements, and (3) achieving consensus with program management, design management and evaluators. It is important that, uniformly, the testers and the test planners are one in the same and part of the program team from program inception.

The test master plans are also where, typically, the test engineers document the consideration of the use of external or non-company facilities. Even when companies do not have test master plans as such, a relatively formal consideration of the use of external facilities and capabilities occurs.

Companies execute their plans and tightly adhere to cost, schedule and product quality requirements. Product quality is the top priority. One test manager commented that the market does not accept bad quality. Another emphasized that, with regard to schedule, they really don't talk about "time to market" as much as "time to quality." A third discussed it in terms of "time to market structure," or getting a quality product out on time at a competitive price. DoD should consider better ways to integrate the test execution planning with the TEMP and program plans. Some DoD programs have done this, and the practice should be expanded at least to major programs.

Measure of Test Program Effectiveness

Virtually every company measures test program efficiency using some methodology. Several of the companies use, at least in part, what is known as 'cost of warranty' or 'cost of coverage.' [BEST PRACTICE] As products are fielded, warranty accruals give an indication of "fielded" product quality, and based on the nature of the defects could spotlight shortfalls in the testing process. A less direct but important measure is market share because if testing isn't adequate and quality slips, market share drops.

Additional practices include tracking of return rates, parts replacement and fielded performance against product success criteria. The product success criteria are developed in conjunction with the testers and are established as early as the requirements phase. Tracking performance of fielded products has two purposes. The purposes are to determine design performance and to compare field performance with test performance. If product performance indicates a test shortfall, the test process problem is fixed. At least two companies instrument sets of high use fielded systems (providing customers incentives) to get rapid feedback on system performance. A particularly important systemic and effective practice in use by one company is to initiate the next program with an intensive self-critique and post mortem of the design, development and especially the testing of the last program. They emphasize that nothing is sacred in the post mortem and that includes the program, development and test teams.

Systemic field performance feedback to testers in DoD for comparison with test performance and conditions should be done.

Risk-Based Testing

Three companies in three different technology sectors use risk-based testing. [BEST PRACTICE] As implemented by commercial companies, risk based testing is characterized by:

- Rapid and comprehensive parallel analysis of problems by the entire team (developers, safety engineers and testers) with a single risk assessment and proposed solution.

- Rapid decision cycles (usually no more than three levels) once a proposed solution is presented and a willingness to assume the potential consequences of increased test risk.

- Willingness to utilize "workarounds" to continue testing until problem solutions are implemented.

- Risk management through frequent reports (daily/weekly), which focus on status of problem solution, schedule adherence, cost, test progress using workarounds.

- Intensive management focus at all levels.

Closed-Loop Process

Several companies emphasized the need to create and intensely manage open defect reports. Commercial companies consistently use closed-loop systems for configuration management and fixing defects. [BEST PRACTICE] This practice ensures that all faults are corrected or addressed appropriately and works to the ultimate quality of the product. DoD organizations are now using this approach in many cases and the practice should be continued and reinforced.

SECTION V - TEST EVALUATION PROCESS

Context

This section describes best practices concerned with evaluation planning and data analysis processes. Uniformly the companies produce high quality products and their rigorous evaluation methods reflect this goal.

Early Evaluation Planning

As with test resource application and planning, evaluation methods are applied from the product concept stage forward. Test and evaluation has continuous input to the design process. Most companies make extensive use of models and simulations in evaluating design alternatives, and they establish standards for data display. Their subcontractors are required to provide quality data using the same display standards. Suppliers are also held to the ISO 9000 series standards to reinforce quality, and suppliers are required to update their data each month.

Methods

Leading companies have in-house advanced training facilities that help to develop skills in modeling and data analysis. Workshops with suppliers are conducted to establish metric quality and applications.

Simulation is applied to design, production, components, assembly and logistics aspects of product development, manufacture, delivery and employment. Simulation, including virtual prototyping, is used to determine what to test, how to reduce dynamic testing and how to extend test results. Predictive techniques are used in many areas, such as electromagnetic interference testing, aerodynamic performance testing, reliability/durability testing and costs of warranties. Simulation is used to stimulate and workload software.

Among the most promising evaluation tools is Physics of Failure. [BEST PRACTICE] The applications of Physics of Failure models and techniques have been applied mainly during system design (see Appendix B, Special Topics). The focus is on product reliability. These methods can reduce operations and support costs, and ultimately, improve system performance.

Leading companies know that higher quality leads to lower costs. Along with Physics of Failure methods, which helps to produce robust designs and quality production methods, the Six Sigma quality process enhances staff creativity. One company's intermediate goal is a 10-fold improvement in product reliability every two years.

Systems integration involves broad simulation environments during design and product development. Modeling supports component testing with exact geometric models of detail such as gas flow, heat dissipation and spring rates. In most cases, the people who know their systems best – the design, development and test teams – employ the simulations. Thus the experience gained from the use of models and simulations enhances the team's ability to evaluate the implications of test results. VV&A of models and simulations is consistently practiced and the model validation results are often shared with customers. [BEST PRACTICE] This sharing occurs by providing the customer the VV&A report documentation and by providing historical data samples that validate the continuing model results.

Sound metrics are chosen, and both the metrics and their methods of display are made standard across a company and its suppliers. Pareto analysis is used to identify problem categories and sources.

Evaluation is keyed to the engineering developers who are charged with product quality and marketability. The test operation evaluates performance to specifications and standards, but the "so what" questions are issues for developers and senior managers. Just as companies use the closed-loop process of comparing faults to solutions in order to ensure comprehensive and complete test execution, company engineers conducting evaluations use the same data from the close-loop process to determine system effectiveness and production readiness. [BEST PRACTICE] Typically the companies share evaluation methods across the corporation, and in some instances, simulations and models along with results are shared with customers and suppliers. Internal web sites are used to facilitate sharing. Virtually every company has the objective of increasing the role of simulation in their evaluation processes.

SECTION VI - CONCLUSIONS AND RECOMMENDATIONS

CONCLUSIONS

With regard to the four areas investigated (Philosophy/Policy/Approach; Test Investment; Test Execution; and Test Evaluation), it is concluded that successful commercial companies have the following best practices that are applicable to DoD. The practices in bold print are likely the highest value for priority of effort.

The most fundamental conclusion is that there is a continuous corporate commitment by leading companies to support and recognize T&E as a key enabler for the development of quality products. This fundamental conclusion cuts across all areas investigated and is seen separately and collectively in the best practices identified in this paper.

Best Practices

Philosophy, Policy, Approach

- **Recognize that testing is a way to identify and solve problems early in the process in order to control time, cost and schedule late in the process.**
- Recognize that best practices generate success and vice versa.
- Stabilize corporate leadership and test staff and commit to T&E as a key enabler.
- Focus on quality of product and process to drive the efficiency and effectiveness of T&E.
- **Develop consistent processes to ensure consistent products. Understand the value and cost of T&E.**
- Implement efficient and effective test processes in order to compete. Keys:
    - Ensure T&E is consistently part of the decision, planning and execution process.
    - **Early commitment by all stakeholders on required T&E resources.**
    - Certification of T&E processes and organizations (~ISO 9000).
    - Ensuring capital capability.
- **Increase T&E to assure product quality rather than reduce it to save T&E cost.**
- Use metrics and quality control processes to understand how well the test process is operating.

Test Investment

- **Ensure early determination of the investment costs to acquire new capability for program support.**
- Require analytically sound ROI analysis for test investments.
- **Ensure cohesive (year-to-year) investment plans.**
- **Charge cost of test investment to program.**

Test Execution

- **Involve testers and evaluators very early:**
    - **Ensures testers know test requirements.**
    - **Ensures developers know requirements for test.**
- **Capture test costs at program initiation.**
- Emphasize concurrent and integrated T&E.
- Institute formal quality check processes.
- Use System Integration Laboratories and embedded instrumentation.
- Give proper consideration to the use of external test capability in test planning.
- Ensure testers control test planning, equipment, facilities, instrumentation and test resources.
- Continue to increase the use of modeling and simulation to expand the test process.
- Do not generally support the outsourcing of testing and evaluation.
- Frequently use the Six Sigma (6s) or similar quality processes.
- Automate data collection and archiving.
- Benchmark in-house and within industry.
- **Use measurements and metrics.**
- Initiate programs to seek ten-fold reductions in the number of software tests required.
- **Integrate Master Test Plans and test execution with program resources and milestones.**
- **Charge full cost of testing to the program.**
- **Establish measures of effectiveness.**
- Quantify risk for management decision when considering reduced testing.
- **Train the in-house test workforce in test engineering disciplines.**
- Emphasize multi-use T&E platforms.

Test Evaluation

- Continue to increase the use of modeling and simulation to expand the evaluation context based on verified test data.
- Correlate faults and solutions in a closed loop process to ensure problems are resolved
- Use Physics of Failure as a tool to predict and analyze system performance and shortfalls.

Test Philosophy/Process/Evaluation
(A combined category)

- Establish corporate internal web based sites for exchange of ideas, benchmarks, data, applications and processes. Address:
    - Data collection retrieval/archiving.
    - Modeling and Simulation.
    - Test and Evaluation methods.

RECOMMENDATIONS

With regard to the conclusions drawn from this study, the following recommendations are made:

- Implement or reinforce bolded Best Practices in DoD as soon as possible.
- Develop implementation or reinforcement strategies for the remainder of the Best Practices using DoD stakeholder teams.
- Present the results of this study to the DoD Acquisition and T&E communities.

APPENDIX A - ORGANIZATION PARTICIPANTS AND CONTACTS

Boeing Commercial Airplane Group
P.O. Box 3703 (MC 14-RE)
Seattle, WA 98124-2207
Mr. G. J. (Jerry) Zanata
Director, Test Programs

Boeing Phantom Works
2401 East Wardlow Road (MC C078-0821)
Long Beach, CA 90807-5309
Mr. Bob G. Zimmerman
Director, Materials and Process Technology

Cummins Engines Corporate Technical Center
1900 McKinley Avenue
Columbus, IN 47201
Mr. Mike Galarno
Director, Test Technology Engineering

DaimlerChrysler Corporate Technical Center
800 Chrysler Drive
Auburn Hills, MI 48326
Mr. Allen White
Senior Manager
Power Train Testing & Development

INTEL Performance Microprocessor Division
2111 N. E. 25th Avenue
Hillsboro, OR 97124-5961
Mr. John D. Barton
PMD Validation Director

Lucent Technologies
6200 East Broad Street
Columbus, OH 43213
Mr. Douglas Mix
Development Engineer
Network Wireless Systems

Motorola, Inc.
1301 E. Algonquin Road (IL02-SH5L)
Schaumburg, IL 60196
Mr. Jack Murch
Senior Manager
Global Customer Relations

Microsoft Office Systems
One Microsoft Way
Redmond, WA 98052-6399
Mr. Grant N. W. George
Vice President, Office Systems
US-OFF Test.

Northrop-Grumman Corporation Integrated Systems Sector
2000 West NASA Blvd
Melbourne, FL 32902-9650
Mr. Royce G. Grones
Chief Test Pilot and Director T&E
J-STARS Surveillance and BMS
Electronics & Systems Integration Division

Northrop-Grumman Ship Systems – Ingalls Operations
Northrop Grumman Corporation
Ship Systems Full Service Center
P.O. Box 149 Pascagoula, Mississippi 39568-0149
Mr. L. Teno Henderson
Vice President, Fleet Support

Underwriters Laboratories
333 Pfingsten Road
Northbrook, IL 60062
Mr. Gordon Gillerman
Director, Government Relations

Xerox Corporation
800 Phillips Road
Webster, NY 14580
Mr. Lou Schneider
Deployment Manager
Corporate Engineering Center

APPENDIX B - SPECIAL TOPICS

Physics of Failure

Because of the frequent reference to *Physics of Failure* in the sections above, more complete discussion of the tool seems appropriate.

System reliability is an area where savings can be realized by applying design and assessment methodologies that address the root causes of failure. Cost savings in manufacturing, testing, fielding and sustainment of new systems can be attained using science-based reliability techniques early in the design process. One such technique is Physics of Failure, an analysis process applied during the design of new products to improve component and system reliability.

Physics of Failure is a science based, proactive approach that models the physical process of failure so as to predict, identify and understand potential failure mechanisms—the how and why of material failure. It is an approach to understand the interaction, or physics, between a product's materials within specific use environments. A basic premise is that while it is important to understand how equipment works, it is equally important to understand how and why equipment fails in its intended environment. (Intended environment, as used within the Physics of Failure context, encompasses a system's life cycle from start of manufacture to disposal.) Knowledge distilled from life-cycle loads and stresses, product architecture, and potential defects and failure mechanisms is used iteratively to create robust design and manufacturing practices, thereby significantly reducing the occurrence of failures during manufacturing, testing and field operations.

Traditionally, Physics of Failure methods have focused on laboratory testing and physical analysis of components to fully understand why they fail. Within recent years, great strides have been made in Physics of Failure analysis using computer simulations to identify first-order failure mechanisms prior to physical testing. Physics of Failure tools model the stress-failure relationship for the dominant environmentally induced mechanisms (mechanical, chemical, vibration, radiation, and temperature cycles). Once these relationships are developed, the expected life can be computed and compared to requirements.

Evaluation methods help to produce reliable and durable systems that perform well in tests and in combat. All the companies that contributed to this study predict component failure rates through the use of mathematical models of their design concepts, along with databases that describe the physical performance and compatibility of materials and combinations of materials that are under consideration in a specific design.

Then they introduce simulations of the intended integration of the designs into completed assemblies and into the planned operational environment. They study the physical interactions of materials in their designs and predict the failure rates. Ideally, the databases contain the results of historical experiences with material exposure to environments such as vibration, temperature and temperature cycles, pressure, corrosion and the frequency of occurrence of operations of the system under study. This allows them to consider material and manufacturing costs versus performance trade-offs as well.

Most of the companies studied refer to the application of this process as the analysis of the *Physics of Failure*. The increasing availability of material property physics and chemistry data, along with increased computer power, has made it possible to study failure mechanisms via computer simulation. The elements of failure assessment include product modeling, stress assessment, failure model evaluation, and display. By designing failures out of a system prior to test and fielding, these analyses have helped to reduce the cost of systems in development, the costs of development (and operational) testing, and the costs of logistics support.

Physics of Failure analyses have been successfully applied to Joint STARS, F-22, Comanche, Theater Air Defense, the Improved Chemical Agent Monitor and other programs. Conservatively, a 16:1 return on investment has been shown.

Technology growth contributes to more and more complexity in our hardware and software, and test and evaluation processes and methods have become increasingly more complex as well. Testing of intricate system designs is routinely supported and augmented with simulations and other analytical methods that:

- Emulate the environment in which the system and/or its components will be employed.
- Create various workloads.
- Help to extrapolate system performance and survivability for a wider range of conditions than are subjected to physical tests.
- Help to predict system reliability and durability.

System performance and reliability are linked. A system may contain redundant components, but whether there are redundant components or not, a system may have the ability to do its job in some degraded sense despite the loss of some component due to a reliability or durability failure. It is important to understand how a new system may perform in a wider range of workload cycles and environments than can be addressed in an affordable series of tests.

In the case of complex and expensive systems, there are never enough test samples for all the desirable performance or reliability/durability testing. We have seen significant growth in simulation support of system performance testing in the past decade. A rigorous process of verification, validation and accreditation must support simulation methodology and applications. Similar use of simulation to support reliability/durability testing of end items (whole systems) is growing.

Limited applications of Physics of Failure methods have resulted in reduced development test costs along with improved reliability and reduced costs of support to fielded systems. The potential to save the DoD significant test and support costs is great.

Six Sigma (6s)

The reference to the Six Sigma tool as a best practice in the sections above make a short description appropriate.

"Six Sigma" or "Six Sigma Breakthrough" is a term coined at Motorola Inc. in the 1980's. This came after years of continuously improving product process quality. The tools and techniques derive from basic quality management documented by Dr. Joseph M. Juran and were key to the launch of the Japanese quality revolution of the 1970's and 1980's.

Sigma (s) denotes the standard deviation of a process (standard deviation measures the variation of spread about the process mean). A "Six Sigma" (6s) capability means having twelve standard deviations of process output between the upper and lower specification limits. Essentially, process variation is reduced so that no more than some specified number of parts per million (depending on the s value such as 1s, 2s...6s) falls outside of the specification limits. The higher the sigma value the better the process output. It is also referred to as "Five Nines" in that 6s will result in a process yielding a reliability probability of .99999. As each higher number of s is applied to a process, the improvement is tenfold.

Some companies continuously seek a tenfold process improvement over some specified period such as every two years. The 6s sets goals and allows clear measurement of progress toward those goals. Juran Institute estimates that, on average, one 6s project will save an organization in excess of $150,000.

The "Six Sigma" term also refers to a philosophy, goal and objective used to drive out waste and improve quality, cost and time performance of any business.

- Source: The Juran Institute
- http://www.juran.com/consulting.html

Comparison of Commercial and Government T&E funding

During the course of the conduct of this study, questions arose about the comparison of commercial and government funding of T&E. This sub-section contains answers to those questions and will perhaps provide some additional detail to the T&E cost paragraph in Section IV. It is provided for information purposes as a derivative product of the study.

How is commercial funding of T&E different from DoD?

Once funded with full corporate commitment, change in T&E funding occurs only with significant program change such as unanticipated product failure or unanticipated major market change. Because there is a comprehensive business or program plan built at product initiation, costs, technology and market are relatively well known in the best commercial companies. This business plan based allocation of resources, with all parties involved and with hard information, contrasts with DoD, where programs may be initiated without a total business plan, where the technology may be speculative and where the requirements (market) float. Thus, in commercial industry, the funding is far better understood and analytically based. It also includes the required T&E funding and schedule from the outset.

If corporately funded, T&E is considered part of the engineering process and an essential capability. In most commercial organizations, it is also an identifiable entity. However, its funding is not divorced from the corporate commitment to the product development. The T&E budget is tied to the corporate research, development and engineering budget. It is consistent with the resource projection for T&E at project initiation.

With regard to T&E competition for funds, it appears commercial T&E competes on the margin of its budget. The basic budget driver is the resource defined at the outset of projects, and that allows the T&E budget to be relatively stable year to year. Competition occurs for budgeting of one-time facility or equipment investment such as an opportunity to procure new test technology. It also occurs if the product under test does not perform, and additional or unanticipated testing is required. By contrast, DoD T&E budgets are not necessarily product driven and may vary somewhat independently of the project requirements.

If a program is funded, the T&E organization must “sell” the program or product manager on the resources required in the initial planning. Adjustments are also considered based on changes such as new equipment benefits or product performance. The program/product manager makes the resource decisions.

In DoD, the program or system manager often sees T&E as a hurdle because he is often forced to meet artificial deadlines to ensure program viability. Schedule is extremely important to the commercial program manager in order to meet market requirements but the commercial program manager has direct responsibility to ensure a properly tested product. Commercial product failures in the field are unacceptable and product managers and corporate executives are held directly accountable. That accountability is key to how the T&E is funded. Commercial program managers clearly understand the need for T&E as part of the engineering effort and thus, good rational for increase in funding is generally accepted. Cost is less important than time to market and quality product.

DOD Funding of the Major Range and Test Facility Base (MRTFB)

The DoD MRTFB is a national asset; sized, operated, and maintained to support T&E of DoD acquisitions. There are 20 facilities in the MRTFB. The MRTFB represents a $30+B investment employing some 13,000 Military, 18,000 Civilian, and 20,000 Contractor personnel and encompassing 21,000 sq miles of land space, 221,000 sq miles of air space, and 243,000 sq miles of sea space.

DoD Directive 3200.11 is the governing MRTFB regulation and a 1974 revision established "uniform cost reimbursement guidelines." The original basic premise was that *users pay for direct costs associated with a test while the "institution" pays infrastructure or indirect costs* and the policy distinguishes between DoD users, other Federal Agencies, and non-federal agencies. Over time a portion of "indirect" costs has also been shifted to the user, "on a pro-rata" basis by policy change.

Direct Budget Authority (DBA) and Reimbursable Budget Authority (RBA) are two terms used to describe how MRTFB facilities are funded and how customers (range users) are billed. DBA represents costs to provide infrastructure and these costs are borne by the "institution". RBA represents costs that users pay for direct cost associated with test. The types of direct costs that users pay depend on whether they are DoD users, other Federal agency users, or commercial or non-US users.

DBA or "Institutional Funds" are often referred to as "indirect costs." They include labor, material, minor construction, utilities, equipment & supplies associated with management test center operations; and other costs which are not normally identifiable to a program.

Current MRTFB policy provides guidance as follows regarding RBA. DoD component users shall reimburse the MRTFB for direct costs which can be readily identified with the particular program support excluding military labor costs.

- The term direct cost applies to all costs incurred due to user workload, and includes the cost of modifications specifically performed for a user to conduct test and evaluation activities and the costs of labor, material, facilities, utilities, equipment, supplies, and any other resources damaged or consumed during test and evaluation activities or maintained for a particular use.
- Non DoD users are charged for indirect costs such as military pay, depreciation, management, administration, etc. "Indirect costs" include all other costs related to operating and maintaining the test and evaluation facilities, including the cost of unused capacity, overhead, general and administrative, and investment. When customers are charged indirect costs, those costs must be charged on a prorated basis.

A 1994 change to policy continued the requirement that commercial organizations be charged "direct costs" but allows Range Commanders "to determine the appropriate indirect costs to charge." The intent was to allow MRTFBs to charge commercial customers "less" in order to attract their business during periods of reduced DoD workload.

When is commercial T&E funding developed and how is it different from DoD?

T&E funding is defined at two basic critical points. First is at product or program initiation. At that point a good business case is developed to ensure adequate resources of all activities to include T&E. The T&E (and other) budgets are updated prior to each execution year. The fundamental philosophy is not different from the DoD, but the implementation has several significant differences.

At program/product initiation, there has already occurred a sound technology assessment, a sound financial and cost assessment and a sound marketing assessment. Thus, when the project is initiated, the corporate executives in all areas know what the risk is, what the costs should be and what the market requires. In contrast, within DoD the technology is often accepted as incomplete or high risk, cost estimates are often inadequately based or overly optimistic based on assumptions of technical success and the market (stated requirements) change as the program evolves. In commercial industry, it is a corporate decision at the project initiation by the corporate accountable stakeholders

APPENDIX C - QUESTION SETS FOR INDUSTRY

Preamble to Question Sets

Attached is a set of questions concerning Best Practices for Test and Evaluation. The questions are arranged in three broad categories: Investment in T&E instrumentation and facilities; Process and Organization for the execution of Testing and Evaluation; and use of test data. Some questions are repeated from one category to another, but the context changes according to the broad categories. The authors have broad experience in the conduct of T&E within the DoD. Our task is to learn what the DoD can do more effectively and efficiently. We believe that you can offer innovative practices, methods and processes that can and should be implemented within the DoD Test and Evaluation structure. To that end, the questions serve as a mechanism to "peel the onion" back and learn the detail of innovative and effective practices in industry.

Investment Process Questions

The following questions are designed to help us understand the detail of your investment process as applicable to Test and Evaluation. We seek to understand the development of your needs for T&E investment, the decision process that culminates in investment, how the investment is "financed" and who has the responsibility for operation and maintenance of the investment once made:

1. What initiates the investment process with respect to instrumentation, equipment and facilities (IEF)?

2. What is the process followed to determine what the needs and solutions are?

3. Once needs and solution alternatives are identified, who decides on the investment requirement?

4. What is the process for management decision on the approval of the investment requirement?

5. What are the sources of funding that are used to implement the investment requirement and how are they amortized against the investment?

6. Do you seek return on investment in IEF or to recoup costs or to allocate against product or project and accept as a sunk cost of doing business?

7. Are the processes for instrumentation and equipment different from the decision and funding processes for facilities? How and for what reasons?

8. How do you determine what the payback or ROI is for investments and how do you balance need versus ROI?

9. Do the investments in IEF become part of product costs and how is that allocated?

10. Are control and maintenance of IEF separate from the test planners and implementers or part of the same organization? What are the benefits of current organizational structure? Is this structure the result of initial organization or have the organizational changes been evolutionary?

11. What are the ages of your current capital plant for T&E IEF, and what do you see as the optimal age and term of use?

12. How do you implement technology change into the IEF capital plant? Is this a proactive or reactive process? How do you measure technology insertion benefits?

13. What are your most essential IEF and by what measure or standard?

14. To what extent do you outsource your testing and to whom? What is the outsource decision process and how do measure the benefit tradeoff? Do you compete or use known sources and do you invest in T&E capital projects outside of your organization?

15. To what extent do you use or seek to use government IEF to support your testing? How could government processes be streamlined or reconfigured to assist industry partners in use of government IEF? What benefits have you derived from use of government IEF?

16. How do you determine the rates of use and the rates of cost for use of your IEF? Do you market your IEF to others, and what benefits have you derived from that process?

17. What do you consider to be the most unique, innovative or beneficial IEF currently in your capability and how do you measure that benefit?

18. What do you consider your best payback on investment in IEF? Why and how measured?

19. Do you conduct early internal research on development of T&E potential investments or developments of IEF? Can you measure the payback or benefit of that internal T&E R&D?

20. Do you have IEF investment partnerships with other industrial organizations? What are the benefits you have derived from such partnerships?

21. In all cases of benefits derived, can you quantify and by what process do you quantify those benefits?

Execution Process Questions

The following questions are designed to achieve an understanding of the organization with which and the process by which you conduct Test and Evaluation. We wish to understand at a detailed level how you staff your organization, efficiency of the reporting and decision processes, how you plan and execute, what your reporting mechanisms are and how you determine how much testing is sufficient:

1. At what point does testing enter the picture in your development process?

2. By what process do you arrive at the decision to initiate test planning and testing?

3. Who/How do you determine the level of test resources to be used on a specific project?

4. What defines a test project?

5. How is the institutional test and evaluation structure organized?

6. How do you pay for test infrastructure?

7. How are test project costs allocated?

8. How much modeling and simulation do you use, and who makes the M/S decisions as to use and acceptability?

9. By what process do you develop your test program/planning for a particular project or product?

10. Do you have documented test procedures and methodology? Who maintains the document file and by what process?

11. Organizationally, where is your in-house test organization and expertise?

12. What % of your testing is in-house and what % is outsourced. What is the economic (or other criterion) that determines whether you test in-house or outsource?

13. Who are the evaluators of the test data?

14. How automated is your test data collection, and archive system? Who has access and what safeguards do you have?

15. What data are archived and how are those data used from product-to-product or project-to-project?

16. Where do you obtain your test personnel/expertise? In-house, trained, outside hire?

17. How and who defines the purpose of your test program as an organization?

18. Who, institutionally, drives the test program? Testers, Developers or Evaluators?

19. What unique processes, methodology, procedures or equipment do you have in your test program? What are the benefits and the MOE you use to determine the payback/payoff to those unique factors?

20. How do you measure product effectiveness as you proceed? What are the feedback mechanisms in the test to decision process?

21. Do you consider your process as schedule or event driven?

22. How do you measure program efficiency and effectiveness? What are the MOEs? What are your "customer" feedback mechanisms?

23. Who has the authority to make a stop test decision as a function of the level of decision? Safety in test/product performance/sufficient test data?

24. What are your test safety procedures and who is responsible?

25. What do you consider the most effective part of your test program and how do you know that?

26. What is the least effective part of your test program, how do you know that and what are you doing to improve?

27. What interface do you have with the government test process and facilities? How does the government test process and structure mesh with yours?

28. In your view, what should the government change with respect to their test process and why. How would that improve the overall test process relative to bringing government requirements to fruition?

29. Is your internal test process fully integrated with your development process? If so, what do you see as the benefits? If not, what are the benefits?

30. What are your internal test milestones? Are they fixed from project to project or are they a function of the particular project?

31. Who within the total organization is involved in the test process and decisions, and what are their roles?

32. What level of internal visibility does the test process and do the results have in your company?

33. To what extent and how do your use the test results in your external marketing? How do your allocate test costs to the marketing process?

34. Who has the authority to make the "test stop" decisions, particularly when the product is not performing or when your have "enough" test data to make the decisions required?

35. How do you maintain procedure integrity from test program to test program?

36. What are the largest efficiency benefits in your program? How do you measure it and how are the benefits documented?

37. Have you downsized your test capability and for what purpose? How do you measure the efficiency and benefit of whatever restructuring you have done?

38. Who are your most reliable and cost effective test outsource providers?

39. What do you do that the government should do to improve the government test cost effectiveness, efficiency and or benefit?

40. Is there anything else you would like to add to the discussion?

Evaluation Process Questions

The following questions are designed to assist us in understanding how your organization uses information gained from testing to evaluate the product tested. This includes how data are collected and organized, how the data support simulations, how the analyses are conducted and by whom and how decisions are achieved based on the data:

1. What evaluation methods do you use?

2. What kinds of testing are used at component levels? At unit (end item) levels?

3. What component tests do simulations support?

    - For test design planning
    - For extending the results of tests

4. What end item tests do simulations support?

    - For test design planning
    - For extending the results of tests

5. At what point do you design to facilitate future testing? Do simulations help to support this decision?

6. What are the priority considerations in controlling test costs? Are you using simulations to reduce the total test effort? In what areas?

7. How are product quality goals established? Who approves them?

8. How do you address test effort planning (dollars and time) to satisfy product quality goals?

9. Are product quality goals rigid? Are simulations used to explore the impact of changing product quality compared to end item cost to produce? Cost to support?

10. How do you decide, and who makes the decision, regarding test effort and market risks? Are simulations used to help bound the decision envelop?

11. In what areas do you rely on test and evaluation results feedback from the government? Do you obtain test and/or evaluation support from other commercial sources?

12. What other uses are made of simulations and other predictive techniques?

    - For choosing among design alternatives
    - For avoiding or limiting certain tests
    - For deciding to enter a bid competition

13. How are test results incorporated in evaluations? In validating simulation design?

14. What uses of evaluation techniques are included in making a decision to bid?

15. Do your customers readily share evaluation methodologies?

16. Are you able to re-use past simulation results? How are they archived?

17. What predictive techniques are used for study of the following factors?

    - Electromagnetic interference
    - Automotive performance
    - On-board electronics

- Reliability
- Durability
- Human factors
- Safety
- Interoperability among system components and across systems

18. What types and applications are made of component testing instruments?

19. Do the instruments you use for component testing find applications in fielded maintenance systems?

20. Do you have evaluation methods that would support a DOD decision to buy your instruments and not impose existing instrument criteria concerning attributes such as durability, electronic hardening, and multi-purpose applicability?

21. What do you consider your most useful applications of predictive techniques or simulation methods and other evaluation methods?

22. In the following categories, what are the weakest areas concerning simulation and evaluation methods? What are you doing to address these areas? Which areas are strongest?

- Performance and Reliability
  - Hardware } During the design processes
  - Software } Test and evaluation design and interpretation of results

- Environment
  - Climate
  - Electronic

- Supportability

- Interoperability

- Safety

- Human Factors

- Uses of simulation
  - Item level
  - Force-on-Force

- Test design

- Test cost trade-offs

APPENDIX D - REFERENCES

Anderson, Arthur, "DaimlerChrysler Digital Design," *Army RD&A*, May-June, 1999.

Army Material Systems Analysis Activity, www.amsaa.army.mil/rad/pofpage.htm (Physics of Failure).

Azani, Cyrus, "The Test and Evaluation Challenges of Following and Open System Strategy," *ITEA Journal of Test and Evaluation*, Vol. 22, No. 3, September/October 2001

Babicz, Gillian, "ISO 9004: The Other Half of the Consistent Pair," *Quality Magazine*, June 2001.

Battershell, A. Lee, "Technology Approach: DOD Versus Boeing (A Comparative Study)," *Acquisition Review Quarterly*, summer, 1995.

Birchard, Pamela and Frederic, William, "Digital Collection and Analysis Review System," *ITEA Journal of Test and Evaluation*, Vol. 22, No. 3, September/October, 2001.

Breslin, Frederick, C., "Accrediting Simulation Based on Very Limited Test Data," *ITEA Journal of Test and Evaluation*, Vol. 22, No. 3, September/October, 2001.

Burge, Albert R., "Test and Evaluation Based on Metrics, Measures, Thresholds and Indicators," *ITEA Journal of Test and Evaluation*, Vol. 17, No.3, 1996.

Cozby, Rick, "Smart Test and Evaluation: The Virtual Proving Ground," *Army RD&A*, May-June, 1999.

Cunningham, J., Valentin, R., Hillman, C., Dasgupta, A. and Osterman, M., "A Demonstration of Virtual Qualification for the Design of Electronic Hardware," *ESTECH 2001 Proceedings*, April 2001.

Deckert, Michael, "Physics of Failure, A Science-Based Approach to Ultra-High Reliability," *Program Manager*, September-October 1994.
Department of Defense Directive 5000.1, *Defense Acquisition*, 1996.

Department of Defense Regulation 5000.2-R, *Mandatory Procedures for Major Defense Acquisition Programs (MDAPs) and Major Automated Information System (MAIS) Acquisition Programs*, 1996.

Duesterhaus, David A., "Government Test Facility Supports Commercial Aircraft Engine Testing," *ITEA Journal of Test and Evaluation*, Vol. 17, No. 4, 1996.

Ferren, Brian, "Some brief Observation on the Future of Army Simulation," *Army RD&A*, May-June, 1999.

General Accounting Office, "A More Constructive Test Approach is Key to Better Weapon System Outcomes'" *GAO/NSIAD – 00-199*, July 2000.

General Accounting Office, "Best Practices – Successful Application to Weapons Acquisitions Requires Changes in DOD's Environment," *GAO/NSIAD- 98-56*, 1998.

General Accounting Office, " Best Practices – DOD Can Help Suppliers Contribute More to Weapons System Programs," *GAO/NSIAD- 98-87*, 1998.

General Accounting Office, "Test and Evaluation – DOD Has Been Slow in Improving Testing of Software Systems," *NSIAD- 93-198*, 1993.

Harry, Michel, and Schroeder, Richard, *Six Sigma,* New York, Doubleday, 2000.

Hyatt, Charles S. and Hannam, Andrew, "TULIP-Overcoming Interoperability Testing Limitations," *ITEA Journal of Test and Evaluation*, Vol. 22, No. 2, June/July 2001.

Juran Institute, *Six Sigma (6s),* http://www.juran.com/consulting/html.

McKinzie, Gordon A., " Working Together: Team Approach Used to Develop and Test Boeing 777," *ITEA Journal of Test and Evaluation*, Vol. XV, No 3, 1994.

Oberhettinger, D., , "Physics of Failure," The Jet Propulsion Laboratory Reliability Engineering Office, *http://rel.jpl.nasa.gov/Org/5053/qual/physics.html*

Osterman, M..and Stadterman, T., "Failure Assessment Software for Circuit-Card Assemblies", *Proceedings of Annual Reliability and Maintainability Symposium,* January 2000.

Pande, Peter S., Neuman, Robert P., and Cavanagh, Roland R., *The Six Sigma Way*, New York, McGraw-Hill, 2000.

Porter, Alex, "Quality Begins with Good Data", *Quality Magazine*, May 2001.

Raytheon, www.raytheon.com/rts/rtsproducts/rtsasc/ascrci.htm , (Physics of Failure).

Relex Software Corporation, www.relexsoftware.com , (Physics of Failure).

Sanders, Patricia, "Simulation Based Acquisition: The Revolution is Coming," *Army RD&A*, May-June 1999.

Science Applications International Corporation, *Best Practices Applicable to DOD Developmental Test and Evaluation,* June 1, 1999.

Stadterman, T., and Peltz, A., U.S. Army Materiel Systems Analysis Activity, Connon, Wm., U.S. Army Aberdeen Test Center, Choi, K., University of Iowa, Freeman, J., University of Tennessee, "Dynamic Modeling & Durability Analysis from the Ground Up," April 2001. http://www.amsaa.army.mil/rad/mpof.htm

Westbrook, Darrel E. III, "Reliability Versus Availability in Software Intensive Systems," *ITEA Journal of Test and Evaluation*, Vol. 21, No. 4, December 2000/January 2001.

Zerr, John J., "T&E From a Business Perspective," *ITEA Journal of Test and Evaluation*, Vol 19, No. 3, 1998.