Summary of Findings and Recommendations From the National Research Council Report on "Reliability Growth: Enhancing Defense System Reliability"

A. Fries (Panel Chair)[1], W.P. Cherry[2], R.G. Easterling[3], E.A. Elsayed[4], A.V Huzurbazar[5], P.A. Jacobs[6], W.Q. Meeker, Jr.[7], N. Nagappan[8], M. Pecht[9], A. Sen[10], S.V. Wiel[5], C. Citro[11], M.L. Cohen[11], M.J. Siri[11]

[1]Institute for Defense Analyses, Alexandria, VA 22311
[2]Science Applications International Corporation, McLean, VA 22102
[3]Cedar Crest, NM 87008
[4]Rutgers University, Industrial & Systems Engineering Dept., New Brunswick, NJ 08901
[5]Los Alamos National Laboratory, Statistical Sciences Group, Los Alamos, NM 87545
[6]Naval Postgraduate School, Operations Research Department, Monterey, CA 93943
[7]Iowa State University, Department of Statistics, Ames, IA 50011
[8]Microsoft Research, Empirical Software Engineering Group, Redmond, WA 98052
[9]University of Maryland, Clark School of Engineering, College Park, MD 20742
[10]University of Michigan, Department of Family Medicine, Ann Arbor, MI 48104
[11]The National Academies, Committee on National Statistics, Washington, DC 20001

Abstract

This paper summarizes the findings and recommendations from a recent report from the Panel on Reliability Growth Methods for Defense Systems, operating under the auspices of the Committee on National Statistics (CNSTAT) within the National Research Council (NRC). The report offers recommendations to improve defense system reliability throughout the sequence of stages that comprise U.S. Department of Defense (DoD) acquisition processes – beginning with the articulation of requirements for new systems and ending with feedback mechanisms that document the reliability experience of deployed systems. A number of these recommendations are partially or fully embraced by current DoD directives and practice, particularly with the advent of recent DoD initiatives that elevate the importance of design for reliability techniques, reliability growth testing, and formal reliability growth modeling. The report supports the many recent steps taken by DoD, building on these while addressing associated engineering and statistical issues. The report provides a self-contained rendition of reliability enhancement proposals, recognizing that current DoD guides and directives have not been fully absorbed or consistently applied and are subject to change.

Key Words: reliability, reliability growth, design, acquisition, test, evaluation

1. Background

Reliability is the innate capability of a system to perform its intended functions: it is one of the key performance attributes tracked during Department of Defense (DoD) acquisition. Yet the urgency to deploy new technologies and military capabilities often

leads to defense systems being fielded with- out having first demonstrated adequate reliability.

Defense systems with poor reliability are not only less likely to successfully carry out their intended missions, but they may also endanger lives. Deficient systems are also much more likely to require extra scheduled and unscheduled maintenance and to demand more spare and replacement parts over their life cycles. In addition, not finding fundamental flaws in a system's design until after it is deployed can lead to costly program delays, expensive redesigns, and the imposition of operational constraints.

Reliability Growth: Enhancing Defense System Reliability (2015), a report from the National Research Council, offers recommendations to improve defense system reliability throughout the sequence of stages that comprise DoD acquisition processes – beginning with the articulation of requirements for new systems and ending with feedback mechanisms that document the reliability experience of deployed systems. A number of these recommendations are partially or fully embraced by current DoD directives and practice, particularly with the advent of recent DoD initiatives that elevate the importance of design for reliability techniques, reliability growth testing, and formal reliability growth modeling. The report supports the many recent steps taken by DoD, building on these while addressing associated engineering and statistical issues. The report provides a self-contained rendition of reliability enhancement proposals, recognizing that current DoD guides and directives have not been fully absorbed or consistently applied and are subject to change.

2. Findings and Recommendations

The basis for this present section is an extract from the NRC report brief developed for the completed Panel study. Figure 1 presents an overview of the Panel's findings and recommendations. A detailed list of the Panel's 25 recommendations follows in Section 3.

2.1 A Challenging Endeavor

Today's DoD systems typically entail greater design complexities, more dependence on software components, increased reliance on integrated circuit technologies, and more intricate dependencies on convoluted nonmilitary supply chains than at any time in the past. Moreover, unlike industrial system development with a single project manager driven by a clear profit motive, DoD acquisition involves many "agents" – a system developer, one or more contractors and subcontractors, a program manager, testers, oversight offices, and military users. Also unlike the commercial sector, where reliability risks are borne primarily by the manufacturer, for defense systems the government generally assumes most of the risk.

2.2 Fundamental Elements for Improving System Reliability

2.2.1 Requirements

Reliability requirements should be grounded in terms of operational relevance, explicitly linked to the costs of acquisition and lifetime sustainment, technically feasible, and measureable and testable. Reliability should be designated as a "key performance parameter."

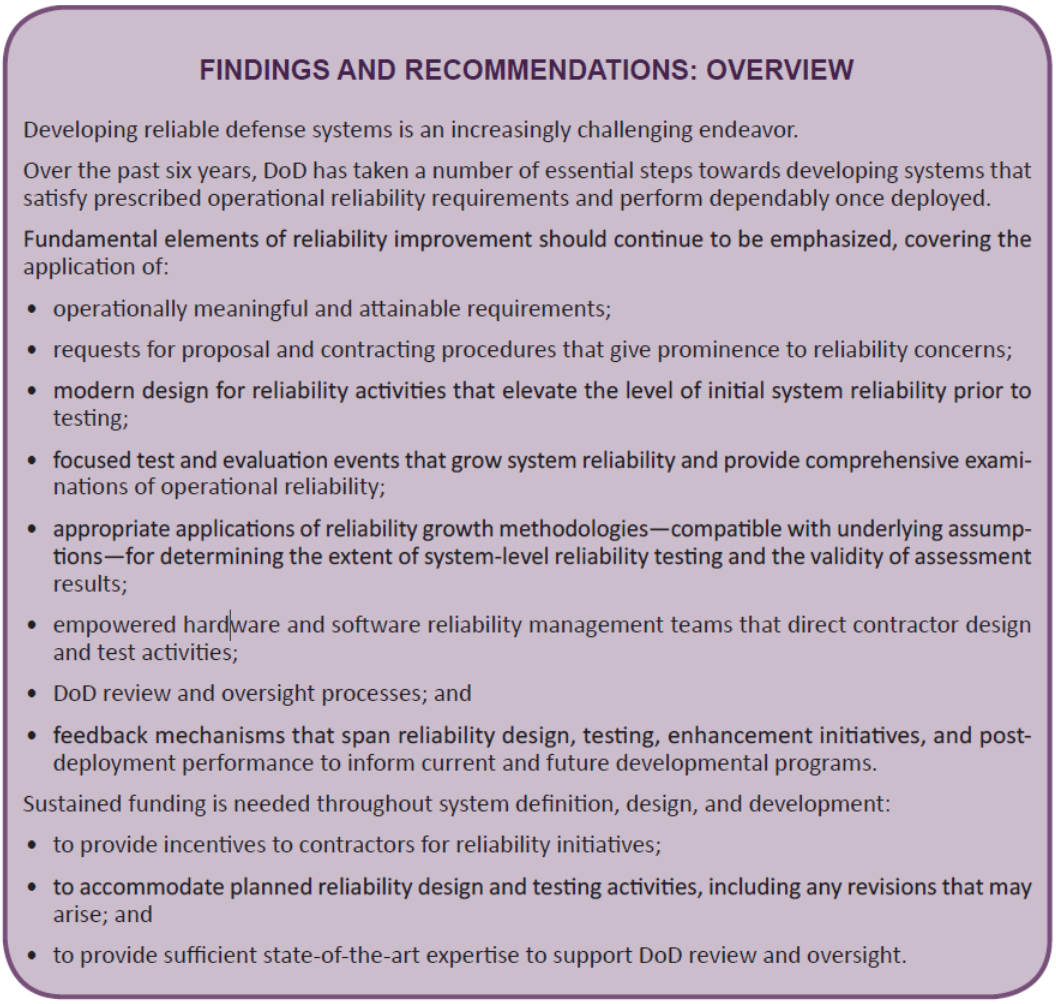

Figure 1: Overview of NRC Panel's Findings and Recommendations

2.2.2 Request for Proposals (RFP)

The government's RFP should contain sufficient detail for contractors to specify how and at what cost levels they would design, test, develop, and demonstrate system reliability.

2.2.3 Modern Design for Reliability

Building in high reliability early in system design is better than relying on extensive and expensive system-level testing later in development and post-deployment to correct low initial reliability levels. Modern design for reliability techniques include appropriate mixes of (1) failure modes and effects analysis, (2) robust parameter design, (3) block diagrams and fault tree analyses, (4) physics-of-failure methods, (5) simulation methods, and (6) root-cause analysis. For electronic components, current reliability pre- diction handbooks should be eschewed in favor of system-specific physics-of-failure methods and validated estimates.

At the preliminary stages of design, contactors should be able to build on the details offered in RFPs and subsequent government interactions. Software-intensive systems and subsystems should be subject to special scrutiny, and holistic design methods should be used to integrate hardware, software, and human factors elements to address potential interaction failure modes.

2.2.4 Testing

Reliability test plans, both hardware and software, should be regularly reviewed (by DoD and the developer) and updated as needed, especially at major design reviews. Attention should be given to contractual requirements, reliability goals, and what remains uncertain about component, subsystem, and system reliability. Reviews need to consider testing conditions, especially since results from non-operationally representative environments can inflate reliability estimates.

The primary goal of early developmental test and evaluation should be to identify and address substantive reliability deficiencies early on, when they are least costly to address. For hardware components and subsystems, there are numerous "accelerated" testing approaches available to identify, characterize, and assess failure mechanisms (including long-term operational usage issues such as material fatigue, aging, and environmental effects) within the limited time available in early testing. They include exposing test articles to controlled nonstandard overstress environments and invoking physically plausible models to translate observed results to nominal use conditions. For software, contractors should be required to test the full spectrum of usage profiles and to implement meaningful performance metrics to track software completeness and maturity.

When system prototypes (or actual systems) are produced, system-level reliability testing can begin, but that should not occur until the current system reliability is demonstrated to be compatible with the prescribed target in the program's reliability demonstration plan. Individual test phases should be used to explore system performance capabilities under different combinations of environmental and operational factors. System-level testing should incorporate elements of operational realism to the extent feasible. At a minimum, a single full-system, operationally relevant developmental test event should be scheduled near the end of developmental testing and evaluation – with advancement to fully realistic operational testing and evaluation primarily contingent on satisfaction of the system operational reliability requirement or other supportable justification (e.g., combination of proximate reliability estimate, well-understood failure modes, and tenable design improvements).

2.2.5 Reliability Growth Methodologies

Currently, every developmental system is required to establish an initial reliability growth curve and to revise the curve as needed when program milestones are achieved or in response to unanticipated testing outcomes. The current strategy is to bring the system's operational reliability at the end of developmental testing to a satisfactory point, thus supporting stand-alone operational testing and evaluation, with acceptable statistical performance characteristics. This strategy is eminently reasonable.

Reliability growth models can be used to synthesize data from different tests and to track and project progress towards attaining intermediate and final reliability target values. However, care must be taken to ensure that underlying model assumptions are not violated.

2.2.6 Developer's Reliability Management

The execution of a developer's reliability testing program should be overseen and governed by a formal reliability management structure that is empowered to make reliability an acquisition priority, retains flexibility to respond to emerging results, and

comprehensively archives hardware and software reliability test designs, data, and assessments. Complete documentation should be budgeted for and made available to all relevant program and DoD entities.

2.2.7 DoD Oversight Processes

DoD oversight spans the complete spectrum of acquisition activities, beginning with the formulation of reliability requirements. The processes for designing and developing a reliable system should draw on pertinent previous program histories and incorporate contributions from user and testing communities. Implementations should be reviewed and supplemented, as needed, by external subject-matter experts with relevant reliability engineering and technical proficiencies.

For software-intensive systems and subsystems, a contractor's development of the software architecture, specifications, and oversight management plan needs to be reviewed independently by DoD and external subject-matter experts. Automated software testing tools and supporting documentation should be developed and reviewed by an outside panel of subject-matter experts appointed by DoD.

Exhibited reliabilities should be monitored and tracked to gauge progress towards achieving formal operational reliability requirements. Of critical importance is the scored reliability at the beginning of system-level testing, a direct reflection of the quality of the system design and production processes. If by the end of operational testing the attainment of adequate system operational reliability has not been demonstrated with satisfactory confidence, DoD should not approve the system for full-rate production and fielding without a formal review of the likely effects that deficient reliability would have on mission success and system life-cycle costs.

2.2.8 Feedback Mechanisms

DoD should encourage the establishment of information-sharing repositories that document individual reliability program histories and are made available to support future system acquisitions. Documentation should include demonstrated reliability results and underlying conditions from developmental testing, operational testing, and post-deployment operation. In developing and using this database, DoD needs to ensure that the data are fully protected against the disclosure of proprietary and classified information.

2.3 Funding and Resources

Planning for and conducting a robust testing program for increasing system reliability requires that sufficient funds be allocated for design, testing, and oversight activities and that the funding be dedicated so that it cannot be redirected for other purposes. Early investments in reliability are typically more than regained in the form of reduced life-cycle costs. Decisions about proposals, awarding contracts, and performance incentives for contractors all should consider long-term program costs.

To perform at a level consistent with best industrial practices, DoD needs to develop and maintain expertise in a number of domains – reliability engineering, software reliability engineering, reliability modeling, accelerated testing, and the reliability of electronic components – through combinations of in-house hiring, consulting or contractual agreements, and the training of current personnel.

3. List of Recommendations

RECOMMENDATION 1 The Under Secretary of Defense for Acquisition, Technology, and Logistics should ensure that all analyses of alternatives include an assessment of the relationships between system reliability and mission success and between system reliability and life-cycle costs.

RECOMMENDATION 2 Prior to issuing a request for proposal (RFP), the Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics should issue a technical report on the reliability requirements and their associated justification. This report should include the estimated relationship between system reliability and total acquisition and life-cycle costs and the technical justification that the reliability requirements for the proposed new system are feasible, measurable, and testable. Prior to being issued, this document should be reviewed by a panel with expertise in reliability engineering, with members from the user community, from the testing community, and from outside of the service assigned to the acquisition. We recognize that before any development has taken place these assessments are somewhat guesswork and it is the expectation that as more about the system is determined, the assessments can be improved. Reliability engineers of the services involved in each particular acquisition should have full access to the technical report and should be consulted prior to the finalization of the RFP.

RECOMMENDATION 3 Any proposed changes to reliability requirements by a program should be approved at levels no lower than that of the service component acquisition authority. Such approval should consider the impact of any reliability changes on the probability of successful mission completion as well as on life-cycle costs.

RECOMMENDATION 4 Prior to issuing a request for proposal (RFP), the Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate the preparation of an outline reliability demonstration plan that covers how the department will test a system to support and evaluate system reliability growth. The description of these tests should include the technical basis that will be used to determine the number of replications and associated test conditions and how failures are defined. The outline reliability demonstration plan should also provide the technical basis for how test and evaluation will track in a statistically defendable way the current reliability of a system in development given the likely number of government test events as part of developmental and operational testing. Prior to being included in the request for proposal for an acquisition program, the outline reliability demonstration plan should be reviewed by an expert external panel. Reliability engineers of the services involved in the acquisition in question should also have full access to the reliability demonstration plan and should be consulted prior to its finalization.

RECOMMENDATION 5 The Under Secretary of Defense for Acquisition, Technology, and Logistics should ensure that reliability is a key performance parameter: that is, it should be a mandatory contractual requirement in defense acquisition programs.

RECOMMENDATION 6 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that all proposals specify the design-for-reliability techniques that the contractor will use during the design of the system for both hardware and software. The proposal budget should have a line item for the cost of design-for- reliability techniques, the associated application of reliability engineering methods, and schedule adherence.

RECOMMENDATION 7 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that all proposals include an initial plan for system reliability and qualification (including failure definitions and scoring criteria that will be used for contractual verification), as well as a description of their reliability organization and reporting structure. Once a contract is awarded, the plan should be regularly updated, presumably at major design reviews, establishing a living document that contains an up-to-date assessment of what is known by the contractor about hardware and software reliability at the component, subsystem, and system levels. The U.S. Department of Defense should have access to this plan, its updates, and all the associated data and analyses integral to their development.

RECOMMENDATION 8 Military system developers should use modern design-for-reliability (DFR) techniques, particularly physics-of- failure (PoF)-based methods, to support system design and reliability estimation. MIL-HDBK-217 and its progeny have grave deficiencies; rather, the U.S. Department of Defense should emphasize DFR and PoF implementations when reviewing proposals and reliability program documentation.

RECOMMENDATION 9 For the acquisition of systems and subsystems that are software intensive, the Under Secretary of Defense for Acquisition, Technology, and Logistics should ensure that all proposals specify a management plan for software development and also mandate that, starting early in development and continuing throughout development, the contractor provide the U.S. Department of Defense with full access to the software architecture, the software metrics being tracked, and an archived log of the management of system development, including all failure reports, time of their incidence, and time of their resolution.

RECOMMENDATION 10 The validity of the assumptions underlying the application of reliability growth models should be carefully assessed. In cases where such validity remains in question: (1) important decisions should consider the sensitivity of results to alternative model formulations and (2) reliability growth models should not be used to forecast substantially into the future. An exception to this is early in system development, when reliability growth models, incorporating relevant historical data, can be invoked to help scope the size and design of the developmental testing programs.

RECOMMENDATION 11 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that all proposals obligate the contractor to specify an initial reliability growth plan and the outline of a testing program to support it, while recognizing that both of these constructs are preliminary and will be modified through development. The required plan will include, at a minimum, information on whether each test is a test of components, of subsystems, or of the full system; the scheduled dates; the test design; the test scenario conditions; and the number of replications in each scenario. If a test is an accelerated test, then the acceleration factors need to be described. The contractor's budget and master schedules should be required to contain line items for the cost and time of the specified testing program.

RECOMMENDATION 12 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that contractors archive and deliver to the U.S. Department of Defense (DoD), including to the relevant operational test agencies, all data from reliability testing and other analyses relevant to reliability (e.g., modeling and simulation) that are conducted. This should be comprehensive and include data from all relevant assessments, including the frequency under which components fail quality tests at any point in the production process, the frequency of defects from screenings, the frequency of defects from functional testing, and failures in which a root-cause analysis was unsuccessful (e.g., the frequency of instances of failure to duplicate, no fault found, retest OK). It should also include all failure reports, times of failure occurrence, and times of failure resolution. The budget for acquisition contracts should include a line item to provide DoD with full access to such data and other analyses.

RECOMMENDATION 13 The Office of the Secretary of Defense for Acquisition, Technology, and Logistics, or, when appropriate, the relevant service program executive office, should enlist independent external, expert panels to review (1) proposed designs of developmental test plans critically reliant on accelerated life testing or accelerated degradation testing and (2) the results and interpretations of such testing. Such reviews should be undertaken when accelerated testing inference is of more than peripheral importance – for example, if applied at the major subsystem or system level, there is inadequate corroboration provided by limited system testing, and the results are central to decision making on system promotion.

RECOMMENDATION 14 For all software systems and subsystems, the Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that the contractor provide the U.S. Department of Defense (DoD) with access to automated software testing capabilities to enable DoD to conduct its own automated testing of software systems and subsystems.

RECOMMENDATION 15 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate the assessment of the impact of any major changes to system design on the existing plans for design-for-reliability activities and plans for reliability testing. Any related proposed changes in fund allocation for such activities should also be provided to the U.S. Department of Defense.

RECOMMENDATION 16 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that contractors specify to their subcontractors the range of anticipated environmental load conditions that components need to withstand.

RECOMMENDATION 17 The Under Secretary of Defense for Acquisition, Technology, and Logistics should ensure that there is a line item in all acquisition budgets for oversight of subcontractors' compliance with reliability requirements and that such oversight plans are included in all proposals.

RECOMMENDATION 18 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that proposals for acquisition contracts include appropriate funding for design-for- reliability activities and for contractor testing in support of reliability growth. It should be made clear that the awarding of contracts will include consideration of such fund allocations. Any changes to such allocations after a contract award should consider the impact on probability of mission success and on life-cycle costs, and at the minimum, require approval at the level of the service component acquisition authority.

RECOMMENDATION 19 The Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate that prior to delivery of prototypes to the U.S. Department of Defense for developmental testing, the contractor must provide test data supporting a statistically valid estimate of system reliability that is consistent with the operational reliability requirement. The necessity for this should be included in all requests for proposals.

RECOMMENDATION 20 Near the end of developmental testing, the Under Secretary of Defense for Acquisition, Technology, and Logistics should mandate the use of a full-system, operationally relevant developmental test during which the reliability performance of the system will equal or exceed the required levels. If such performance is not achieved, then justification should be required to support promotion of the system to operational testing.

RECOMMENDATION 21 The U.S. Department of Defense should not pass a system that has deficient reliability to the field without a formal review of the resulting impacts the deficient reliability will have on the probability of mission success and system life-cycle costs.

RECOMMENDATION 22 The Under Secretary of Defense for Acquisition, Technology, and Logistics should emplace acquisition policies and programs that direct the services to provide for the collection and analysis of post-deployment reliability data for all fielded systems, and to make that data available to support contractor closed-loop failure mitigation processes. The collection and analysis of such data should be required to include defined, specific feedback about reliability problems surfaced in the field in relation to manufacturing quality controls and indicate measures taken to respond to such reliability problems. In addition, the contractor should be required to implement a comprehensive failure reporting, analysis and corrective action system that encompasses all failures (regardless whether failed items are restored/repaired/replaced by a different party, e.g., subcontractor or original equipment manufacturer).

RECOMMENDATION 23 After a system is in production, changes in component suppliers or any substantial changes in manufacturing and assembly, storage, shipping and handling, operation, maintenance, and repair should not be undertaken without appropriate review and approval. Reviews should be conducted by external expert panels and should focus on impact on system reliability. Approval authority should reside with the program executive office or the program manager, as determined by the U.S. Department of Defense. Approval for any proposed change should be contingent upon certification that the change will not have a substantial negative impact on system reliability or a formal waiver explicitly documenting justification for such a change.

RECOMMENDATION 24 The Under Secretary of Defense for Acquisition, Technology, and Logistics should create a database that includes three elements obtained from the program manager prior to government testing and from the operational test agencies when formal developmental and operational tests are conducted: (1) outputs, defined as the reliability levels attained at various stages of development; (2) inputs, defined as the variables that describe the system and the testing conditions; and (3) the system development processes used, that is, the reliability design and reliability testing specifics. The collection of these data should be carried out separately for major subsystems, especially software subsystems.

RECOMMENDATION 25 To help provide technical oversight regarding the reliability of defense systems in development, specifically, to help develop reliability requirements, to review acquisition proposals and contracts regarding system reliability, and to monitor acquisition programs through development, involving the use of design-for-reliability methods and reliability testing, the U.S. Department of Defense should acquire, through in-house hiring, through consulting or contractual agreements, or by providing additional training to existing personnel, greater access to expertise in these five areas: (1) reliability engineering, (2) software reliability engineering, (3) reliability modeling, (4) accelerated testing, and (5) the reliability of electronic components.

Acknowledgements

Permission from the NRC to reproduce the Panel report brief and list of recommendations is gratefully acknowledged. Any opinions, findings, conclusions, or recommendations expressed in this paper (as well as in the original NRC publication) are those of the authors and do not necessarily reflect those of the Panel's Office of the Secretary of Defense co-sponsors (the Office of the Director of Operational Test and Evaluation and the Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics).

Reference

National Research Council, 2015. *Reliability Growth: Enhancing Defense System Reliability*. Washington, DC, USA: National Academies Press [http://www.nap.edu/catalog/18987/reliability-growth-enhancing-defense-system-reliability]