A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON "RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed
A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr.
N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri

•

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

e-mail: afries@ida.org
wpcherry@gmail.com
rgeaste@comcast.net
elsayed@rci.rutgers.edu
aparna@lanl.gov, scottv@lanl.gov
pajacobs@nps.edu
wqmeeker@iastate.edu
nachin@microsoft.com
pecht@calce.umd.edu
anandas@med.umich.edu
CCitro@nas.edu, MCohen@nas.edu, MSiri@nas.edu

ABSTRACT

This paper, extracted from National Research Council (2015), 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.

1   INTRODUCTION

Reliability – the innate capability of a system to perform its intended functions – is one of the key performance attributes that is tracked during U.S. Department of Defense (DoD) Acquisition processes. Although every system is supposed to achieve a specified reliability requirement before

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

being approved for acquisition, the perceived urgency to operationally deploy new technologies and military capabilities often leads to defense systems being fielded without having demonstrated adequate reliability. Between 2006 and 2011, one-half of the 52 major defense systems reported on by the DoD Office of the Director, Operational Test and Evaluation (DOT&E) to Congress failed to meet their prescribed reliability thresholds, yet all of the systems proceeded to full-rate production status.

Defense systems that fail to meet their reliability requirements are not only less likely to successfully carry out their intended missions, but also may endanger the lives of the Armed Service personnel who are depending on them. Such deficient systems are also much more likely than reliable systems to require extra scheduled and unscheduled maintenance and to demand more spare and replacement parts over their life cycles. In addition, the consequences of not finding fundamental flaws in a system's design until after it is deployed can include costly and strategic delays until expensive redesigns are formulated and implemented and imposition of operational limits that constrain tactical employment profiles.

Recognizing these costs, the Office of the Secretary of Defense (OSD) – through DOT&E and the Office of the Under Secretary of Defense for Acquisition, Technology, and Logistics (USD AT&L) – in 2008 initiated a concerted effort to elevate the importance of reliability through greater use of design-for-reliability techniques, reliability growth testing, and formal reliability growth modeling. To this end, handbooks, guidance, and formal memoranda were revised or newly issued to provide policy to lead to the reduction of the frequency of reliability deficiencies. To evaluate the efficacy of that effort and, more generally, to assess how current DoD principles and practices could be strengthened to increase the likelihood of defense systems satisfying their reliability requirements, DOT&E and USD AT&L requested that the National Research Council conduct a study through its CNSTAT. The Panel on Reliability Growth Methods for Defense Systems was created to carry out that study.

2   SCOPE AND CONTEXT

The panel examined four broad topics: (1) the processes governing the generation of reliability requirements for envisioned systems, the issuance of requests for proposals (RFPs) for new defense acquisitions, and the contents of and evaluation of proposals in response; (2) modern design for reliability and how it should be utilized by contractors; (3) contemporary reliability test and evaluation practices and how they should be incorporated into contractor and government planning and testing; and (4) the current state of formal reliability growth modeling, what functions is it useful for, and what constitutes suitable use.

The current environment for defense system acquisition differs from the conditions that prevailed in DoD in the 1990s and also differs from the circumstances faced by commercial companies. Compared to the past, today's DoD systems typically entail: greater design complexities (e.g., comprising dozens of subsystems with associated integration and interoperability issues); more dependence on software components; increased reliance on integrated circuit technologies; and more intricate dependencies on convoluted nonmilitary supply chains.

In commercial system development, all elements of program control are generally concentrated in a single project manager driven by a clear profit motive. In contrast, DoD acquisition processes are spearheaded by numerous independent "agents" – a system developer, one or more contractors and subcontractors, a DoD program manager, DoD testers, OSD oversight offices, and the military users – all of whom view acquisition from different perspectives and incentive structures. In addition, in the commercial sector the risk of delivering a poor reliability system is borne primarily by the manufacturer (in terms of reduced current and future sales, warranty costs, etc.), but for defense systems, the government and the military users generally assume most of the risk because the government is committed to extensive purchase quantities prior to the point where reliability deficiencies are evident.

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

Over the past few decades, commercial industries have developed two basic approaches to producing highly reliable system designs: techniques germane to the initial design, referred to as design-for-reliability methods; and testing in development phases aimed at finding failure modes and implementing appropriate design improvements to increase system reliability. In contrast, DoD has generally relied on extensive system-level testing, which is both time and cost intensive, to raise initial reliabilities ultimately to the vicinity of prescribed final reliability requirements. To monitor this growth in reliability, reliability targets are established at various intermediate stages of system developmental testing (DT). Upon the completion of DT, operational testing (OT) is conducted to examine reliability performance under realistic conditions with typical military users and maintainers. The recent experience with this DoD system development strategy is that operational reliability has frequently been deficient, and that deficiency can generally be traced back to reliability shortfalls in the earliest stages of DT.

Central to current DoD approaches to reliability are reliability growth models, which are mathematical abstractions that explicitly link expected gains in system reliability to total accrued testing time. They facilitate the design of defensible reliability growth testing programs and they support the tracking of the current system reliability. As is true for modeling in general, applications of reliability growth models entail implicit conceptual assumptions whose validity needs to be independently corroborated.

DoD reliability testing, unless appropriately modulated, does not always align with the theoretical underpinnings of reliability growth formulations, such as that system operating circumstances (i.e., physical environments, stresses that test articles are subjected to, and potential failure modes) do not vary during reliability growth periods.

The common interpretation of the term "reliability" has broad ramifications throughout DoD acquisition, from the statement of performance requirements to the demonstration of reliability in operational testing and evaluation. Because requirements are prescribed well in advance of testing, straightforward articulations, such as mean-time-between failures (MTBF) and probability of success, are reasonable. Very often, the same standard MTBF and success probability metrics will be appropriate for describing established levels of system reliability for the data from limited duration testing. But there may be instances – depending on sample sizes, testing conditions, and test prototypes – for which more elaborate analysis and reporting methods would be appropriate. More broadly, system reliabilities, both actual and estimated, reflect the particulars of testing circumstances, and these circumstances may not match intended operational usage profiles.

3 PANEL OBSERVATIONS AND RECOMMENDATIONS

The Panel on Reliability Growth Methods for Defense Systems offered 25 recommendations for improving the reliability of U.S. defense systems. These are listed in entirety in Section 4 below. Here we first summarize the panel's primary observations that underlie the resultant recommendations. Then we highlight the content and substance of the individual recommendations. The panel's conclusions cover the entire spectrum of DoD acquisition activities:

- DoD has taken a number of essential steps toward developing systems that satisfy prescribed operational reliability requirements and perform dependably once deployed.
- Fundamental elements of reliability improvement should continue to be emphasized, covering:
  - operationally meaningful and attainable requirements;
  - requests for proposal and contracting procedures that give prominence to reliability concerns;
  - design-for-reliability activities that elevate the level of initial system reliability;
  - focused test and evaluation events that grow system reliability and provide comprehensive examinations of operational reliability;

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

- appropriate applications of reliability growth methodologies (i.e., compatible with underlying assumptions) for determining the extent of system-level reliability testing and the validity of assessment results;
- empowered hardware and software reliability management teams that direct contractor design and test activities;
- feedback mechanisms, spanning reliability design, testing, enhancement initiatives, and post-deployment performance, that inform current and future developmental programs; and
- DoD review and oversight processes.
- Sustained funding is needed throughout system definition, design, and development, to:
    - incentivize contractor reliability initiatives;
    - accommodate planned reliability design and testing activities, including any revisions that may arise; and
    - provide sufficient state-of-the-art expertise to support DoD review and oversight.

Support for the panel's recommendations that are put forward throughout the panel's report. Here we present the content of the recommendations in terms of four aspects of the acquisition process: (1) system requirements, RFPs, and proposals; (2) design for reliability; (3) reliability testing and evaluation; and (4) reliability growth models.

The recommendations include a few "repeats" – endorsements of earlier CNSTAT and DoD studies, as well as reformulations of existing DoD acquisition procedures and regulations. These are presented to provide a complete self-contained rendition of reliability enhancement proposals, and because current DoD guidance and governance have not been fully absorbed, are inconsistently applied, and are subject to change.

3.1 System Requirements, RFPs, and Proposals

Prior to the initiation of a defense acquisition program, the performance requirements of the planned system, including reliability, have to be formally established. The reliability requirement should be grounded in terms of operational relevance (e.g., mission success) and be linked explicitly (within the fidelity available at this early stage) to the costs of acquisition and sustainment over the lifetime of the system. This operational reliability requirement also has to be technically feasible (i.e., verified to be within the state-of-the-art of current or anticipated near-term scientific, engineering, and manufacturing capabilities). Finally, the operational reliability requirement needs to be measureable and testable. The process for developing the system reliability requirement should draw on pertinent previous program histories and use the resources in OSD and the services (including user and testing communities). Steps should be reviewed and supplemented, as needed, by external subject-matter experts with reliability engineering and other technical proficiencies relevant to the subject system. [Recommendations 1, 2, 24, and 25]

The reliability requirement should be designated as a key performance parameter, making compliance contractually mandatory. This designation would emphasize the importance of reliability in the acquisition process and enhance the prospects of achieving suitable system reliability. During developmental testing, opportunities to relax the reliability requirement should be limited: it should be permitted only after high-level review and approval (at the level of a component acquisition authority or higher), and only after studying the potential effects on mission accomplishment and life-cycle costs. [Recommendations 3 and 5]

The government's RFP should contain sufficient detail for contractors to specify how they would design, test, develop, and qualify the envisioned system and at what cost levels. The RFP needs to elaborate on reliability requirements and justifications, hardware and software considerations, operational performance profiles and circumstances, anticipated environmental load conditions, and definitions of "system failure." The preliminary versions of the government's

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

concept for a phased developmental testing program (i.e., timing, size, and characteristics of individual testing events) should also be provided. The government's evaluations of contractor proposals should consider the totality of the proffered reliability design, testing, and management processes, including specific failure definitions and scoring criteria to be used for contractual verification at various intermediate system development points. [Recommendations 1, 2, 4, 7, and 16]

3.2 Design for Reliability

High reliability early in system design is better than extensive and expensive system-level developmental testing to correct low initial reliability levels. The former has been the common successful strategy in non-DoD commercial acquisition; the latter has been the predominantly unsuccessful strategy in DoD acquisition.

Modern design-for-reliability techniques include but are not limited to: (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. The appropriate mix of methods will vary across systems. At the preliminary stages of design, contractors should be able to build on the details offered in RFPs, subsequent government interactions, and past experience with similar types of systems. [Recommendation 6]

The design process itself should rest on appropriately tailored applications of sound reliability engineering practices. It needs not only to encompass the intrinsic hardware and software characteristics of system performance, but also to address broader reliability aspects anticipated for manufacturing, assembly, shipping and handling, life-cycle profiles, operation, wear-out and aging, and maintenance and repair. Most importantly, it has to be supported by a formal reliability management structure and adequate funding (possibly including incentives) that provides for the attainment and demonstration of high reliability levels early in a system's design and development phases. If a system (or one or more of its subsystems) is software intensive, then the contractor should be required to provide a rationale for its selection of a software architecture and management plan, and that plan should be reviewed by independent subject-matter experts appointed by DoD. Any major changes made after the initial system design should be assessed for their potential impact on subsequent design and testing activities, and the associated funding needs should be provided to DoD. [Recommendations 6, 7, 15, and 18]

Three specific aspects of design for reliability warrant emphasis. First, more accurate predictions of reliabilities for electronic components are needed. The use of Military Handbook (MIL-HDBK) 217 and its progeny have been discredited as being invalid and inaccurate: they should be replaced with physics-of-failure methods and with estimates based on validated models. Second, software-intensive systems and subsystems merit special scrutiny, beginning in the early conceptual stages of system design. A contractor's development of the software architecture, specifications, and oversight management plan need to be reviewed independently by DoD and external subject-matter experts in software reliability engineering. Third, holistic design methods should be pursued to address hardware, soft- ware, and human factors elements of system reliability – not as compartmentalized concerns, but via integrated approaches that comprehensively address potential interaction failure modes. [Recommendations 6, 8, and 9]

3.3 Reliability Testing and Evaluation

Increasing reliability after the initial system design is finalized involves interrelated steps in planning for acquiring system performance information through testing, conducting various testing events, evaluating test results, and iteration. There are no universally applicable algorithms that precisely prescribe the composition and sequencing of individual activities for software and hardware developmental testing and evaluation at the component, subsystem, and system levels.

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

General principles and strategies, of which we are broadly supportive, have been espoused in a number of recent documents introduced to and utilized by various segments of DoD acquisition communities. While the reliability design and testing topics addressed in these documents are extensive, the presented expositions are not in-depth and applications to specific acquisition programs have to draw upon seasoned expertise in a number of reliability domains – reliability engineering, software reliability engineering, reliability modeling, accelerated testing, and the reliability of electronic components. In each of these domains, DoD needs to add appropriate proficiencies through combinations of in-house hiring, consulting or contractual agreements, and training of current personnel.

DoD also needs to develop additional expertise in advances in the state-of-the-art of reliability practices to respond to challenges posed by technological complexities and by endemic schedule and budget constraints. Innovations should be pursued in several domains: the foundations of design for reliability; early developmental testing and evaluation (especially for new technologies and for linkages to physical failure mechanisms); planning for efficient testing and evaluation and comprehensive data assimilation (for different classes of defense systems); and techniques for assessing aspects of near- and long-term reliability that are not well-addressed in dedicated testing.

Finally, to promote learning, DoD should encourage the establishment of information-sharing repositories that document individual reliability program histories (e.g., specific design and testing and evaluation initiatives) and demonstrated reliability results from developmental and operational testing and evaluation and post deployment operation. Also needed are descriptions of system operating conditions, as well as manufacturing methods and quality controls, component suppliers, material and design changes, and other relevant information. This database should be used to inform additional acquisitions of the same system and for planning and conducting future acquisition programs of related systems. In developing and using this database, DoD needs to ensure that the data are fully protected against the disclosure of proprietary and classified information. [Recommendations 22, 23, 24, and 25]

Planning for and conducting a robust testing program that increases system reliability, both hardware and software, requires that sufficient funds be allocated for testing and oversight of contractor and subcontractor activities. Such funding needs to be dedicated exclusively to testing so that it cannot be later redirected for other purposes. The amount of such funding needs to be a consideration in making decisions about proposals, in awarding contracts, and in setting incentives for contractors. 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 (beginning with system design options), retains flexibility to respond to emerging insights and observations, and comprehensively archives hardware and software reliability testing, data, and assessments. Complete documentation should be budgeted for and made available to all relevant program and DoD entities. [Recommendations 6, 7, 9, 12, 15, 16, 17, and 18]

The government and contractor should collaborate to further develop the initial developmental testing and evaluation program for reliability outlined in the RFP and described in the contractor's proposal. Reliability test plans, both hardware and software, should be regularly reviewed (by DoD and the developer) and updated as needed (e.g., at major design reviews) – considering what has been demonstrated to date about the attainment of reliability goals, contractual requirements, and intermediate thresholds and what remains uncertain about component, subsystem, and system reliability. Interpretations should be cognizant of testing conditions and how they might differ from operationally realistic circumstances. [Recommendations 4, 7, and 11]

The objectives for early reliability developmental testing and evaluation, focused at the component and subsystem levels, should be to surface failure mechanisms, inform design enhancement initiatives, and support reliability assessments. The scope for these activities, for both hardware and software systems, should provide timely assurance that system reliability is on track with expectations. The goal should be to identify and address substantive reliability deficiencies at

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON "RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

this stage of development, when they are least costly, before designs are finalized and system-level production is initiated.

For hardware components and subsystems, there are numerous "accelerated" testing approaches available to identify, characterize, and assess failure mechanisms and reliability within the limited time afforded in early developmental testing and evaluation. They include exposing test articles to controlled nonstandard overstress environments and invoking physically plausible models to translate observed results to nominal use conditions. To manage software development in this early phase, contractors should be required to test the full spectrum of usage profiles, implement meaningful performance metrics to track software completeness and maturity, and chronicle results. For software-intensive systems and subsystems, contractors should be required to develop automated software testing tools and supporting documentation and to provide these for review by an outside panel of subject-matter experts appointed by DoD. [Recommendations 7, 9, 12, and 14]

When system prototypes (or actual systems) are produced, system-level reliability testing can begin, but that should not occur until the contractor offers a statistically supportable estimate of the current system reliability that is compatible with the starting system reliability requirement prescribed in the program's reliability demonstration plan. System-level reliability testing typically proceeds, and should proceed, in discrete phases, interspersed by corrective action periods in which observed failure modes are assessed, potential design enhancements are postulated, and specific design improvements are implemented. Individual test phases should be used to explore system performance capabilities under different combinations of environmental and operational factors and to demonstrate levels of achieved reliability specific to the conditions of that test phase (which may or may not coincide precisely with operationally realistic scenarios). Exhibited reliabilities, derived from prescribed definitions of system hardware and software failures, should be monitored and tracked against target reliabilities to gauge progress toward achieving the formal operational reliability requirement. Of critical importance is the scored reliability at the beginning of system-level developmental testing, which is a direct reflection of the quality of the system design and production processes. A common characteristic of recent reliability deficient DoD programs has been early evidence of demonstratively excessive observed failure counts, especially within the first phase of reliability testing. [Recommendations 7 and 19]

Inadequate system-level developmental testing and evaluation results in imprecise or misleading direct assessments of system reliability. If model based estimates (e.g., based on accelerated testing of major subsystems) become integral to demonstrating achieved system reliability and supporting major acquisition decisions, then the modeling should be subject to review by an independent panel of appointed subject-matter experts. To enhance the prospects of growing operational reliability, developmental 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 operational testing and evaluation contingent on satisfaction of the system operational reliability requirement or other justification (e.g., combination of proximate reliability estimate, well-understood failure modes, and tenable design improvements). [Recommendations 13 and 20]

In operational testing, each event ideally would be of a sufficiently long duration to provide a stand-alone statistically defensible assessment of the system's operational reliability for distinct operational scenarios and usage conditions. When operational testing and evaluation is constrained (e.g., test hours or sample sizes are limited) or there are questions of interpretation (e.g., performance heterogeneity across test articles or operational factors is detected), nonstandard sophisticated analyses may be required to properly characterize the system's operational reliability for a single test event or synthesizing data from multiple developmental and operational test events. Follow-on operational testing and evaluation may be required to settle unresolved issues, and DoD should ensure that it is done. If the attainment of an adequate level of system operational reliability has not been demonstrated with satisfactory confidence, then DoD should not approve the system for full-rate production and fielding without a formal review of the likely effects that the deficient

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

reliability will have on the probability of mission success and system life-cycle costs. [Recommendation 21]

The glimpses of operational reliability offered by operational testing are not well suited for identifying problems that relate to longer use, such as material fatigue, environmental effects, and aging. These considerations should be addressed in the design phase and in developmental testing and evaluation (using accelerated testing), and their manifestations should be recorded in the post deployment reliability history database established for the system. [Recommendation 22]

3.4 Reliability Growth Models

DoD applications of reliability growth models, focused on test program planning and reliability data assessments, generally invoke a small number of common analytically tractable constructs. The literature, however, is replete with other viable formulations – for time-to-failure data and discrete success/failure and both hardware and software systems (code). No particular reliability growth model is universally dominant for all potential applications, and some data complexities demand that common modeling approaches be modified in nonstandard and novel ways. [Recommendations 10, 11, and 19]

Within current formal DoD test planning documentation, each developmental system is required to establish an initial reliability growth curve (i.e., graphical depiction of how system reliability is planned to increase over the allotted developmental period) and to revise the curve as needed when program milestones are achieved or in response to unanticipated testing outcomes. The curve can be constructed from applying a reliability growth model, incorporating historical precedence from previous developmental programs, or customizing hybrid approaches. It should be fully integrated with overall system developmental test and evaluation strategies (e.g., accommodating other nonreliability performance issues) and retain adequate flexibility to respond to emerging testing results – while recognizing potential sensitivities to underlying analytical assumptions. The strategy of building the reliability growth curve to bring the system operational reliability at the end of developmental test and evaluation to a reasonable point supporting the execution of a stand-alone operational test and evaluation, with acceptable statistical performance characteristics, is eminently reasonable. Some judgment will always be needed in determining the number, size, and composition of individual developmental testing events, accounting for the commonly experienced DT/OT reliability gap, and in balancing developmental and operational testing and evaluation needs with schedule and funding constraints. [Recommendations 10 and 11]

Reliability growth models can be used, when supporting assumptions hold, as plausible "curve fitting" mechanisms for matching observed test results to prescribed model formulations – for tracking the development and maturity of software in early developmental testing, and for tracking the progression of system reliability during system-level testing. When overall sample sizes (i.e., numbers of recorded failures across multiple tests) are large, modeling can enhance the statistical precision associated with the last test event and support program oversight judgments. No elaborate modeling is needed, however, when the initial developmental testing experiences far more failures than anticipated by the planned reliability growth trajectory – indicative of severe reliability design deficiencies. [Recommendations 9, 10, and 19]

Standard applications of common reliability growth methods can yield misleading results when some test events are more stressful than others, when system operating profiles vary across individual tests, or when system functionality is added incrementally over the course of developmental testing. Under such nonhomogeneous circumstances, tenable modeling may need to require the development and validation of separate reliability growth models for distinct components of system reliability, flexible regression-based formulations, or other sophisticated analytical approaches. Without adequate data, however, more complex models can be difficult to validate: in this circumstance, too, reliability growth modeling needs to recognize the limitations of trying to apply sophisticated statistical techniques to the data. The utility and robustness of

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

alternative specifications of reliability growth models and accompanying statistical methodologies can be explored via simulation studies. The general caution against model-based extrapolations outside of the range of the supporting test data applies to projections of observed patterns of system reliability growth to future points in time. One important exception, from a program oversight perspective, is assessing the reliability growth potential when a system clearly is experiencing reliability shortfalls during developmental testing – far below initial target values or persistently less than a series of goals. Reliability growth methods, incorporating data on specific exhibited failure modes and the particulars of testing circumstances, can demonstrate that there is little chance for the program to succeed unless major system redesigns and institutional reliability management improvements are implemented (i.e., essentially constituting a new reliability growth program). [Recommendations 10 and 19]

4 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.

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

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.

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

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.

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

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

A. Fries (Panel Chair), W.P. Cherry, R.G. Easterling, E.A. Elsayed, A.V. Huzurbazar, S.V. Wiel, P.A. Jacobs, W.Q. Meeker, Jr., N. Nagappan, M. Pecht, A. Sen, C. Citro, M.L. Cohen, M.J. Siri - SUMMARY OF THE NATIONAL RESEARCH COUNCIL REPORT ON"RELIABILITY GROWTH: ENHANCING DEFENSE SYSTEM RELIABILITY"

RT&A # 04 (39)
(Vol.10) 2015, December

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.

5   ACKNOWLEDGMENT

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]