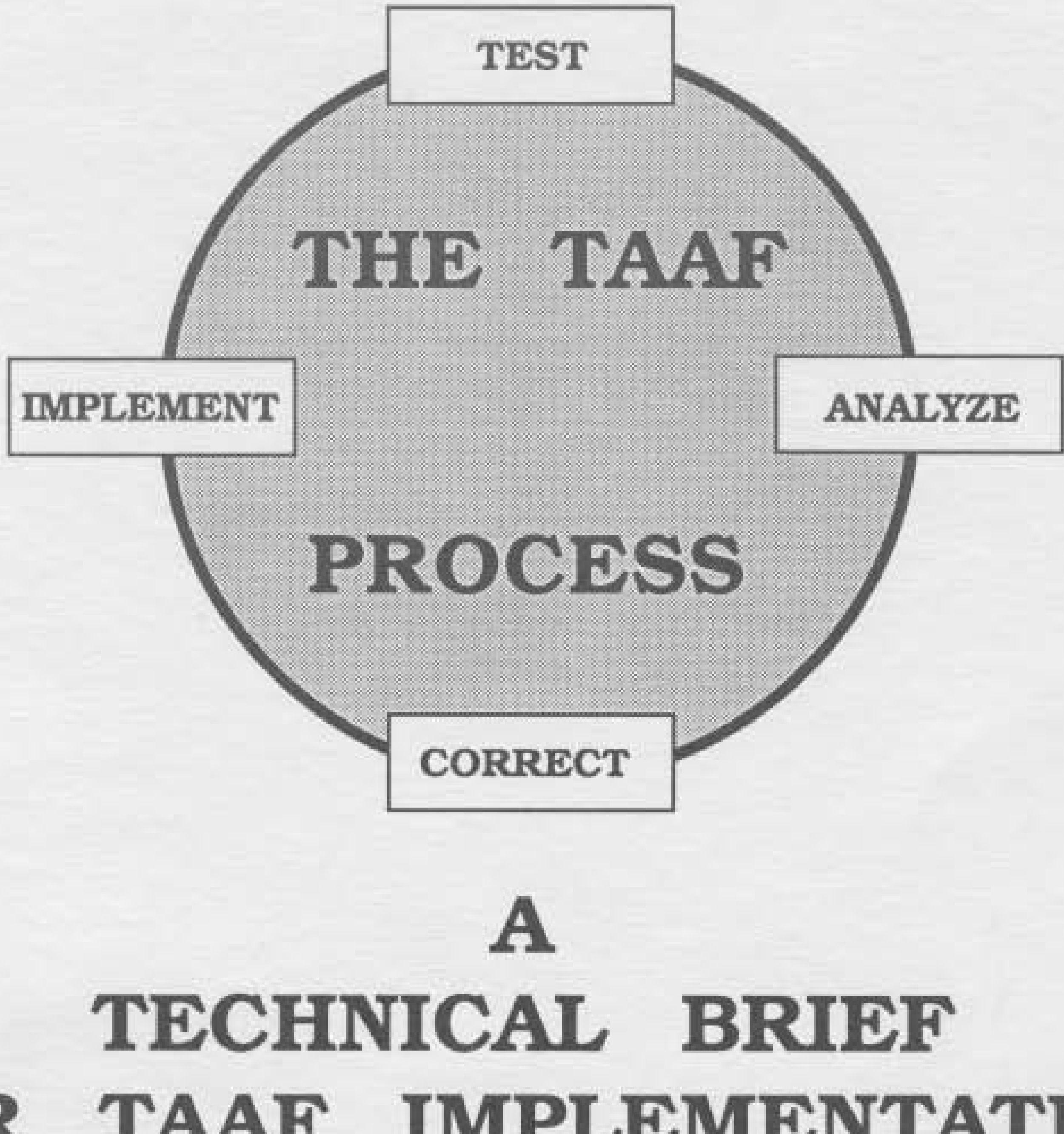


A TECHNICAL BRIEF FOR TAAF IMPLEMENTATION

JANUARY 1989

HQ AMC
QA

OASN
S&L

HQ USAF
LE-RD

EXECUTIVE SUMMARY

The Test, Analyze, and Fix (TAAF) process is a closed-loop reliability growth methodology. This technical brief provides in a single, concise source document TAAF program management methods, engineering practices, and suggested implementation contract language for the program manager and engineer. Because such a document has not been previously available, there is considerable misunderstanding regarding the purpose and scope of the TAAF process. Some, for example, equate TAAF with reliability demonstration testing. Nothing could be further from the truth. The purpose of TAAF is not to prove that a reliability goal has been met, but rather to deliberately search out and eliminate deficiencies. ***In TAAF, failures are welcome.***

The TAAF concept is necessary because, even with the very best of modern engineering methods, initial designs for mechanical or electronic systems that are complex or that involve new technology have reliability deficiencies that are difficult to fully detect and eliminate through design analysis. The TAAF process surfaces these problems early and eliminates them before rate production. Tenfold reliability improvements are not unusual.

Our goal in publishing this document is to help the program manager and engineer assure the design and delivery of reliable weapon systems.

FRANK S. GOODELL, BGen, USAF
Special Assistant for Reliability and Maintainability
SAF/AQ and DCS/LE

W.J. WILLOUGHBY, JR.
Director, Reliability, Maintainability, and Quality Assurance
Office of the Assistant Secretary of the Navy (Shipbuilding & Logistics)

S.J. LORBER
Deputy Chief of Staff for Product Assurance & Testing
Army Material Command
HQ AMC

TABLE OF CONTENTS

I. INTRODUCTION

A. PURPOSE

The Departments of the Army, Navy, and Air Force have combined their efforts in preparing this document because of concern over a common problem pervasive in military system acquisitions. Namely, the lack of uniform discipline and rigor in the planning and execution of the Test, Analyze, and Fix (TAAF) process. In varying degrees, both the military and industry are responsible for these problems.

This technical brief provides in a single, concise source document the methods most likely to result in a successful TAAF program. Provided are program management methods, engineering practices, and suggested contract language for the program manager and engineer.

Some TAAF programs have achieved significant reliability growth, some have not; the variability is attributable to the differing approaches and degree of management commitment. Although there are alternatives that might be "best" for your program, the preferred methods described in this pamphlet deserve your consideration.

B. BACKGROUND

Inconsistency impairs our efforts to use TAAF effectively today. It's an emotional issue. Some in the acquisition community remain unconvinced of TAAF's value and violently oppose it. There is a lack of direction in applying it--little real technical guidance is available at the present time. Programs that do use TAAF are as likely as not to do it wrong, and this doesn't help convince others. A lack of discipline has led to almost any test activity being called TAAF.

After reviewing a number of military acquisition programs it became apparent that the same problems were being repeated from one program to the next. The following is a list of some of these problem areas:

- Program office understanding and support of the need for and purpose of TAAF have been lacking.

- To levy contractual MTBF requirements at certain points in TAAF testing is counter productive since it will not encourage finding failures.

- Contractor performance must be tracked without providing a negative TAAF incentive. Techniques such as defining an acceptable growth range for reporting purposes should be used.

- Use of early hardware has drawbacks such as:

    - Hardware with tolerance or performance problems may be switched with TAAF hardware to allow performance tests to proceed.

    - Early hardware will contain early software; TAAF hardware may not be completely representative as software changes may be made which are not functionally compatible and easily installed in the older hardware. Test set and spare parts compatibility may also cause delays.

- o Holding onto assets is difficult and should be a major consideration in planning any TAAF Program.

- o Accumulation of TAAF hours may prove difficult due to factors such as:

  - Repair turnaround times--a lack of availability of spare parts, repair resources or failure analysis capabilities will greatly lengthen the repair cycle.

  - Test facility problems--if a new test facility will be used, chamber availability will probably be less than anticipated while chamber bugs are being worked.

- o TAAF progress is not briefed at weekly Program Managers meetings to keep management informed.

C. WHAT IS THE TAAF PROCESS?

The TAAF process is an iterative, closed-loop reliability growth methodology. TAAF is accomplished primarily during full-scale engineering development (FSED). The process includes testing, analyzing test failures to determine cause of failure, redesigning to remove the cause, implementing the new design, and retesting to verify that the failure cause has been removed.

TAAF is necessary because, even with the very best of modern engineering methods, initial designs for systems that are complex or that involve new technology have reliability deficiencies that are difficult to fully detect and eliminate through design analysis. The TAAF process should surface these problems early and eliminate them before rate production.

The heart of the TAAF process is the identification of reliability weaknesses. TAAF includes both formal and informal means for doing so. The formal aspect is called a Reliability Development Test (RDT), or sometimes a Reliability Development/Growth Test (RD/GT), and involves dedicated long term exposure of system equipment to ***simulated*** mission profile environments (See Appendix A - Glossary). The informal means is the systematic identification of reliability problems found during other activities such as systems integration, subsystem/component development testing, environmental qualification testing, and operational/field testing. Both means are essential to the TAAF process and are illustrated in Figure 1.

The RDT portion of the TAAF process, because it requires system hardware, test chambers, and similar resources, is necessarily a major investment. Thus, RDT is not a substitute for the disciplined design and design analysis process; it is a complement to it.

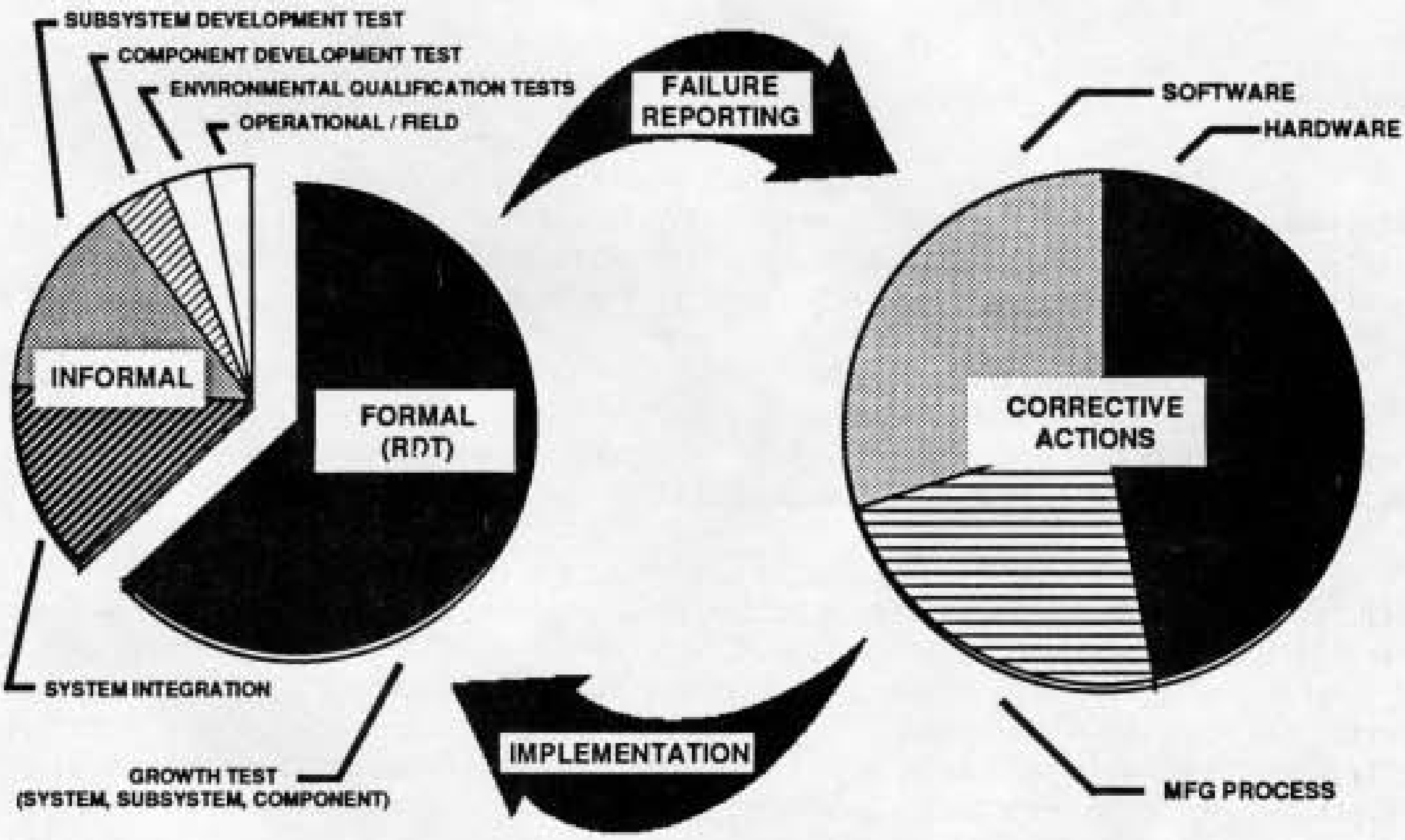
A well executed RDT program will get much better results than traditional reliability demonstration tests because the incentives are different. The purpose of RDT is not to "prove" that a mean-time-between-failure threshold has been met but rather to deliberately search out and eliminate deficiencies. ***In RDT, failures are analyzed and corrected, not scored.***

Tenfold improvements in reliability are not unusual for a well-executed TAAF program. However, the amount of reliability growth that the TAAF process will provide depends on the stage of development and the technology. The more immature the technology, the greater the need for RDT.

The TAAF process, as described in this guide, is principally intended to eliminate hardware reliability weaknesses by empirical means. A large percent of today's military equipment includes software. The general approach and acquisition methods presented in this technical brief should be a starting point for structuring a software TAAF process, but implementation will require adjustments.

FIGURE 1

THE TAAF PROCESS



D. SUPPORTING POLICIES, STANDARDS, AND SPECIFICATIONS

This document draws extensively from the policies and procedures detailed in the DOD directives, military standards, and individual service documents listed in the appendices. Where applicable, specific references to these documents are noted in the text.

II. TAAF PROGRAM MANAGEMENT

Successful implementation of TAAF has been spotty. Either a lack of understanding of the TAAF process or contractual constraints have resulted in inadequate efforts. The Defense Science Board Task Force on the Transition from Development to Production recognized in 1983 the need for TAAF as a key requirement to achieving adequate system reliability. Their recommendations have been incorporated into DOD Manual 4245.7-M and the Navy's "Best Practices" Manual NAVSO P- 6071. A major concept presented in these documents is that tailoring must be done to the specific program when considering the necessary effort for high risk components.

How should a manager assess his TAAF program? - - Ask the following - - Does everyone understand the benefits of TAAF? Are the required pieces of the TAAF program in place? Are adequate resources available and being applied? Is achieved versus planned growth being tracked throughout program development as well as at major milestones?

A. POLICIES AND PRACTICES

In the final analysis it is the management policies and commitment to funding, control, and integration of the related engineering efforts that determine the effectiveness of the TAAF process.

Minimum policies and practices are as follows:

- High quality parts (semiconductors, ICs, hybrids, etc) must be used and be at a 100 parts/million goodness level or better.
- For cost effective RDT resource utilization, the hardware and software to be tested must have been through a thorough, iterative design analysis process.
- A well conducted, documented TAAF program may obviate the need to conduct reliability demonstration tests.
- Formal RDT will require a Government/contractor agreed-to projected growth curve prior to starting the test.
- The reliability program plan will detail the implementation of the TAAF process including monitoring, assessment efforts, and verification of fixes.
- RDT will be performed at the optimum level (system, subsystem, assembly) considering high risk assemblies, growth requirements, test time, and test resources.
- RDT testing should be combined with other tests, if possible, to avoid duplication of resources.
- RDT environments will be based on the mission profile(s).
- TAAF should be terminated when further testing is likely to produce insignificant improvement.
- Environmental stress screening will precede TAAF where appropriate.

B. MANAGEMENT COMMITMENT

In addition to the initial planning for the program and allocation of resources, high-level management of reliability growth is necessary to make available all the options for difficult program decisions. For example, high-level decisions in the following areas may be necessary:

- o Increasing or decreasing the amount of testing.
- o Adding or reallocating program resources.
- o Funding additional development effort.
- o Revising the program schedule.

Some of these options may result in program schedule or cost impacts. However, because the TAAF process will surface unexpected reliability problems early, (if they exist) it will facilitate early and less expensive resolution.

To further assist in TAAF management, two additional avenues are open to the program manager. First, use the Checklist in Figure 2 (or an appropriately tailored version) to help assure the required characteristics are in place during initial TAAF planning as well as follow-on periodic program reviews.

FIGURE 2

TAAF PROCESS
CHECKLIST

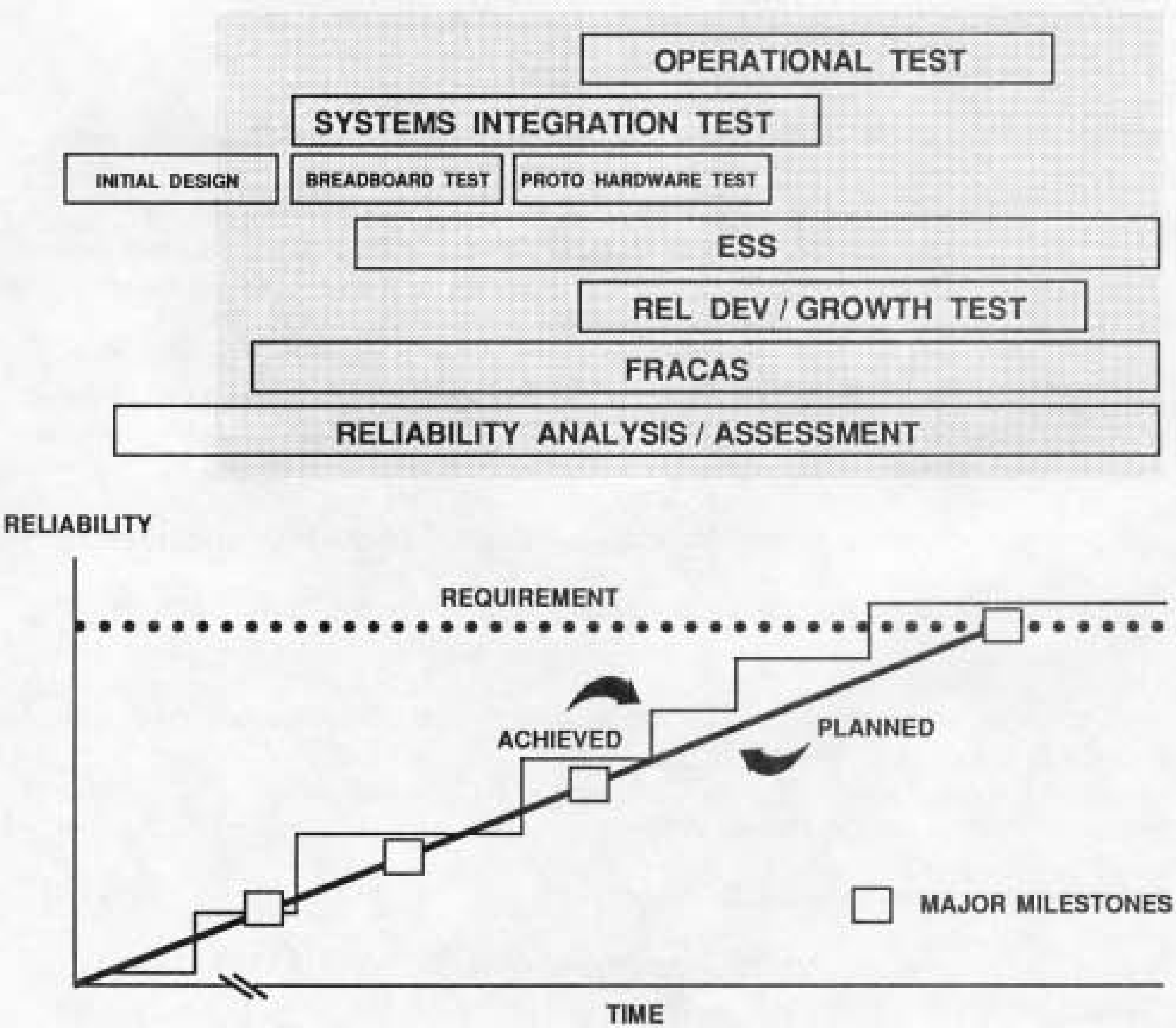
- o Is there evidence of Government commitment to the concept that reliability growth during FSED will be necessary and will be provided?
  - Are TAAF growth procedures used in place of total reliance on reliability demonstration tests?
  - Will the RFP require the offeror to incorporate a reliability growth plan in his proposal?
  - Will the RFP require line item costing corresponding to the plan?
- o Is there evidence of contractor commitment to the concept?
  - Is the proposed growth plan realistic?
  - Is it priced realistically?
- o Did the commitments survive the negotiation process?
  - Is the schedule essentially intact?
  - Is the funding essentially intact?
- o Are the commitments surviving the vicissitudes of the ongoing development program?
  - Has diversion of funds from the TAAF program been avoided?
  - Has diversion of other resources from the TAAF program been avoided? If thereis slippage in the RDT program schedule, is it no greater than the slippage of other aspects of the development program?
  - When diversion of TAAF resources is attempted, is the TAAF program defended by both Government and contractor program management?
- o Are the TAAF tests combined with other tests wherever possible?
- o Is design relatively complete prior to initiation of the RDT program?
- o Have environmental stress screens been completed, where appropriate, prior to the RDT?
- o Is an adequate and timely Failure Reporting Analysis and Corrective Action System (FRACAS) being implemented?
- o Are growth slopes used to monitor progress?
- o Will corrective actions be verified?
- o Will the TAAF process include early operational feedback data if the TAAF period includes early operational use?

Second, track the achieved growth versus planned growth during the program development both continuously and at major milestones. There are two models, the Duane (for J.T. Duane) model and the AMSAA (for Army Material Systems Analysis Activity) model that are in widespread use. Both have their place; their application is discussed in Section III. Three important things need to be recognized when using a growth chart:

- o The reliability measure of merit selected to track (MTBF, captive carry reliability, etc.) must be both operationally relevant and measurable.
- o Achieved versus planned growth is a primary means of determining program health.
- o Growth will almost certainly be discontinuous (or stair-step) because fixes are normally introduced as block updates (Figure 3).

FIGURE 3

GROWTH TRACKING



C. PLANNING AND IMPLEMENTING THE TAAF PROCESS

1. Overview

TAAF is part of FSED. Hence, planning for TAAF occurs prior to FSED in demonstration/ validation. A generalized representation is shown in Figure 3. Three considerations are particularly important in planning and implementing the TAAF process:

- o A good TAAF process, rather than being a separate entity, is part of a comprehensive reliability and maintainability program that is in itself part of the systems engineering process. TAAF should be seamlessly integrated into the overall systems engineering process.

- o The TAAF end product is more reliable equipment and lower cost. However, the emphasis of TAAF, as stated earlier, is not on proving that the equipment is reliable but on identifying and removing reliability weaknesses. Hence, ***identifying and removing weaknesses is the primary activity that is being procured.***

- o All TAAF programs will require tailoring to the type of equipment, program timeliness, level of maturity, and degree of technical challenge presented by the reliability requirements and similar considerations. However, the "core" elements of test, analyze, determine corrective action, and implement will always be required.

2. Acquisition Strategy

In line with the overview above, the major aspects of the TAAF contract strategy are:

- o Demonstration by the contractor in his proposal (rather than during execution of the contract) that he can integrate TAAF into his program in a manner that will assure reliable equipment.

- o Demonstration by the contractor, during execution of the contract, that he is identifying and removing reliability weaknesses in a manner that will assure fielding reliable equipment.

- o Contract fee provisions linked to the identification and removal of reliability weaknesses.

In order to implement this strategy, the government's primary attention will be on Section L, "Instructions to Offerors," and the Statement of Work. Suggested contract RFP language is in Appendix B.

The reader of Appendix B will note that the TAAF RDT hardware is a priced line item in Section B of the RFP. If management is going to be committed to TAAF, then management (both government and contractor) needs to clearly see the investment.

This section has discussed the principal management issues. The next section will present preferred practices.

III. SPECIFIC GUIDANCE AND PREFERRED PRACTICES

This section, will cover the role of the Failure Reporting, Analysis and Corrective Action System (FRACAS), planning the RDT, and estimating the reliability growth.

A. FAILURE REPORTING, ANALYSIS AND CORRECTIVE ACTION SYSTEM

A disciplined and aggressive closed-loop FRACAS is an essential element in the TAAF process. The essence of a closed-loop FRACAS is that failures and faults of both hardware and software are formally reported, analysis is performed to determine failure cause, and positive corrective actions are identified, implemented, and verified to prevent further recurrence.

For a successful TAAF process, all failures must be analyzed to the extent needed to determine the root cause of the failure. In many instances, this will not require a detailed laboratory analysis because the cause of the failure, such as test procedure errors, wrong parts, overstressed parts, workmanship errors, etc., will be readily apparent. Likewise, the corrective actions for many of these failure types will be relatively straightforward and easily implemented. Determining the root cause of more complex failures and developing corrective actions early in the development process are crucial, even though the process may be costly and time consuming. However, the cost and time will be considerably less than waiting until later in the acquisition program to correct the failure.

Corrective action options and flexibility are greatest during the early design evolution when even major design changes can be considered in order to eliminate or to significantly reduce susceptibility to known failure models. These options and flexibility become more limited and expensive to implement as a design becomes firm. Early elimination of failure modes, and thus early implementation of a good FRACAS, has the following advantages:

- Cost and schedule savings.
- Ample time to assess corrective actions.
- Reduction of previously identified failure modes; reducing redundant data analysis.
- Adequate time to address all failures prior to full rate production (ie. prevention of corrective action backlog).

If RDT results are to be interpreted correctly, all test conditions and occurrences must be recorded accurately and completely. A solution and key complement to FRACAS is the test log because a major source of problems is the dynamic status of test item configurations. When multiple copies of an equipment item are under test this is especially significant. By definition, if reliability growth is taking place, the equipment is changing (normally both hardware and software). In addition, temporary repairs/replacements usually are permitted so that testing may continue while permanent fixes are being explored. The implications of a failure differ both quantitatively and qualitatively between repairs and fixes--***a repeat failure after repair simply provides more data on the same phenomenon, while a repeat after "fix" may invalidate the corrective action***. An accurate test log can prevent misinterpretation of results.

It is recognized that there are pragmatic limits to the resources in time, money, and engineering manpower to expend on an analysis of a particularly complex failure mode or the implementation of preferred corrective actions. These limits are determined by item priority, program urgency, available technology, and engineering ingenuity. The limits will vary from program to program. Management involvement is required to determine these limits.

B. PLANNING THE RELIABILITY DEVELOPMENT TEST

At the heart of the TAAF process is the formal RDT. Through this test the weapon system or critical subsystem is subjected to operational stresses for an extended period of time. The RDT is designed to expose the equipment to thousands of operational use cycles, with corrective actions incorporated and verified during the test. RDT is an accelerated test in that the dwell time in nonstress environments is limited to time needed for equipment to stabilize. The formal nature of this test also provides the program manager with a continuous assessment of the improvement in weapon system design maturity and a means of controlling reliability growth.

1. What must precede RDT?

As discussed above, considerable expense and resources are required for the RDT effort. The TAAF process by itself, however, is not the most efficient or economical means of achieving acceptable reliability. Proper emphasis must be placed on design fundamentals, quality parts and capable manufacturing processes. When this is done, then the expensive and time-consuming RDT, in particular, will not be overwhelmed and constrained by hardware and software problems that should not exist at that stage of development. Available resources can then be devoted to eliminate design problems that cannot be detected through analysis.

The pertinent details of a comprehensive reliability program are given in references AFP 800-7, AR 702-3, NAVSO P-6071, and MIL-STD-785 but a few will be discussed briefly to show their influence on RDT efforts:

- Reliability Prediction and Allocation. Using the best available failure rates with the appropriate reliability block diagrams and prediction models, the results should indicate the design's predicted reliability to be at least 125% of the required value. This should provide confidence of achieving a satisfactory reliability in spite of vagaries in the prediction process.

- Derating. Derating of components in accordance with published guidelines helps to keep associated failure rates low and reduce marginal operation.

- Failure Modes and Effects Analysis (FMEA). The FMEA identifies obvious redesign areas to improve reliability while still in the pre-hardware stage. These redesigns should be implemented in the hardware prior to RDT.

- Parts Selection. For years military contracts have required minimum quality levels for components used, e.g., Class B for ICs and JANTX for semiconductor devices. To ensure this quality actually exists in components used, the current requirements call for 100% rescreening as part of incoming inspection of all semiconductor components. The joint military services approach now is that 100% rescreening will continue "unless evidence exists" that the components in question have a quality level of no more than 100 defects per million.

- Environmental Stress Screening (ESS). ESS should be conducted on all electronic equipment scheduled for RDT. This helps prevent overloading the RDT with problems not directly related to design.

2. Determining the Test Conditions

The initial step is the development of a representative Mission/Life Cycle Profile, by determining the most stressful environmental conditions and durations of stress that the system will experience during its life cycle. For example, for large, multi-cabinet systems the requirement should identify the assembly level (drawer, module, etc) subjected to an operational environment and how close that test environment is to actual, including interfaces the system would see in an operational situation.

For systems that have multiple missions and/or environmentally distinct mission phases, thorough environmental simulation in the RDT often is impractical. In these cases, the choice of mission(s), phases(s) or combined mission profiles to be simulated should be driven by the combination of operational relevance and feasibility of measurement. Almost always, the RDT combined environments should include temperature cycling and vibration as a minimum, along with operational stresses such as on/off cycling and variation in inputs (e.g. line voltage) and outputs (e.g. loads).

3. Estimating the Reliability Growth Curve

Reliability growth planning is defined in Task 103 of MIL-STD-781D. One of the primary tools in reliability growth planning is the reliability growth curve. Previous studies of successful reliability growth programs have shown that growth curves exhibit approximately a straight line function when plotted on log-log graph paper. In such curves, a measure of reliability, such as MTBF, is projected as a function of total test time. For program planning, the amount of test time required to achieve a given reliability threshold may be estimated after the selection of an initial starting point and the establishment of a growth slope.

Figure 4 illustrates two separate planning curves, the first with a starting MTBF equal to 10% of the predicted MTBF (i.e. a high risk program) and the second with a starting MTBF equal to 30% of the predicted MTBF (i.e. a low risk program). Both curves use a growth slope of .49 and an initial test time of 100 hours. The reason for the 100 hour starting time is to provide a sufficient period of operation for assuring that test procedures, facilities, and test operators are all functioning properly. As can be seen, it will take approximately 6000 additional test hours to achieve the required 300 hour MTBF, using the 10% (high risk) starting point. This illustrates, in a dramatic fashion, the need for a comprehensive reliability design program, as described in III.B.1., prior to commencement of the RDT.

a. The Starting Point

In the absence of historical data, engineering assessment of the design and of early development test results may be used to estimate the reliability starting point. An arbitrary selection of a starting point for RDT is not recommended, although traditionally values from 10% to 30% of the predicted MTBF have been used. ***Techniques for determining a more realistic value are given in MIL-HDBK 189 and are based on design analysis and experience on similar systems.***

b. The Growth Rate

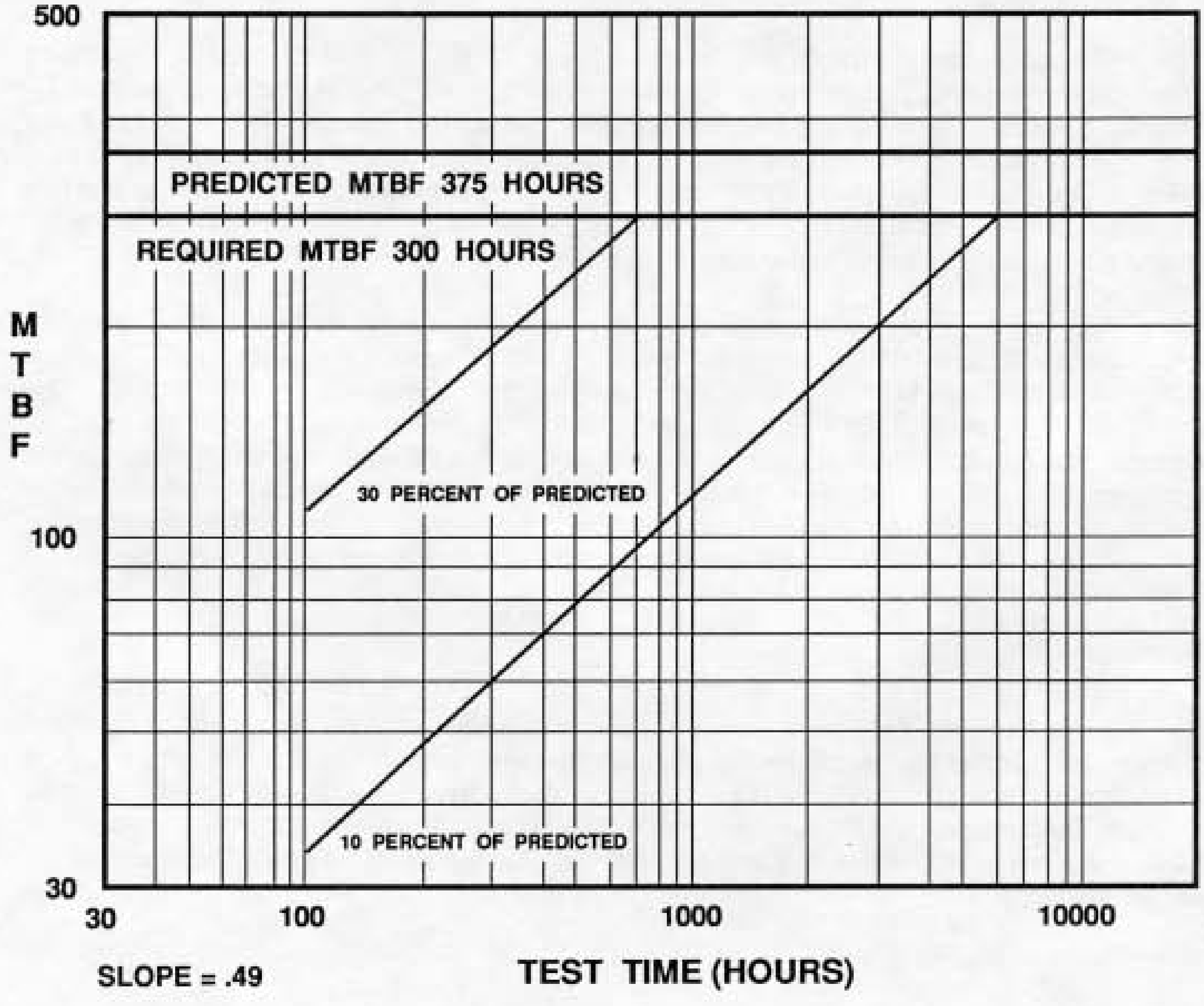
The growth rate is a function of the amount of control, rigor, and efficiency by which failures are corrected. An arbitrary selection of a growth rate for RDT is not recommended. Analytical techniques exist for predicting more realistic rates as a function of equipment attributes and development program characteristics.

When Duane or AMSAA models are used, rates in the vicinity of 0.3 are common, 0.5 is achievable if resources are applied intensively, and values as low as 0.1 (or less) and as high as 0.7 have been reported. Early tests can lead to relatively fast growth as the most glaring design errors become evident. Later tests usually are less productive.

A range of acceptable growth rates may be established (e.g., growth rates from .45 desired, to .35 minimum) as a means of tracking contractor performance, with the minimum rate being used to trigger more aggressive actions needed to improve unsatisfactory growth.

FIGURE 4

RDT PLANNING EXAMPLE



c. Hardware to be Tested

To determine the optimum number of equipment/systems to be allocated for the RDT, trade-off analyses must be made. Of primary concern is the availability of systems, subsystems, or even functional modules for the extended test periods involved. Ideally, two or three systems would be dedicated for the required test duration. This is rarely possible, particularly for large or complex systems, because resources are not allocated for a sufficient number of FSED systems to accomplish all test objectives as well as other development tasks. The logical solution, in order to optimize the use of equipment and test time, is to carefully perform pre-RDT analyses and select high risk equipments for emphasis in the RDT. As mentioned before, this may involve management decisions to build and test more of the high risk equipment but this is necessary to assure the system will have the desired reliability. Similarly, items such as test chamber availability, performance monitoring capabilities, personnel availability, other test program requirements, and the overall program schedule must be considered in the trade-off analyses.

Because of unavoidable program constraints, you may need to test systems that are incomplete or not configured with realistic operational relationships among system elements. In such cases, meaningful reliability measurement is not possible unless surrogate performance measurements at interfaces have been defined and implemented.

Use of early hardware has drawbacks such as:

- o Hardware with tolerance or performance problems may be switched with RDT hardware to allow performance tests to proceed. Holding onto assets is difficult and should be a major consideration in planning any RDT program.

- o Early hardware will contain early software; RDT hardware may not be completely representative as software changes may be made which are not functionally compatible or easily installed in the older hardware. Test set and spare parts compatibility may also cause delays. If early hardware is used, plans should call for the cycling in of up-to-date hardware.

d. Test Time

Reliability growth takes place when weaknesses are corrected by changes in designs or processes. Such changes must be preceded by identification of weaknesses, development of corrective actions, implementation of these actions, and verification. The rate at which this process can progress in calendar time is constrained by the rate at which weaknesses can be identified (affected by test time per unit and by the starting reliability) and by the speed of development and implementation of corrective actions (affected by engineering capability and management policy). Growth therefore is non-uniform and discontinuous. Variations are discussed in considerable detail in MIL-HDBK-189.

The overall test hours needed for a RDT may be estimated by using the reliability growth curve planned starting point, growth rate, and required final MTBF. The calendar time needed to complete the RDT will be dependent on the number of units under test and the overall test efficiency, sometimes as low as 50% or less. The following are some of the many factors which must be considered when estimating test efficiency:

- o Downtime for maintenance and repair of test chambers or debugging of new test chambers and test setups.

- o Availability of spare and repair parts.

- o Availability of resources for troubleshooting and failure analysis.

The estimation of total test time for RDT becomes significant because of the resources involved. Contractual requirements, as spelled out in Appendix B, require that the proposal submitted in response to a RFP details all tasks proposed to perform a satisfactory TAAF. Included would be specifics on quantity and type of equipment, test environments, total test time, starting point, projected growth rate, growth model used, and FRACAS. ***These details, including any government/contractor improvements, are then incorporated in the contract.***

C. EXECUTING THE RELIABILITY DEVELOPMENT TEST

As mentioned above, the test articles should represent a "mature" design. No further engineering changes should be pending prior to the RDT. The RDT supplements the design reliability effort but does not take its place. Similarly, it is too expensive to be a quality screen. Therefore, the test articles should be constructed with quality parts (rescreened for 100 ppm or less defect rate). The equipment should be subjected to environmental stress screening, when appropriate, prior to the RDT to find and eliminate workmanship defects.

The conduct of the RDT is defined in Task 202 of MIL-STD-781D. This task in turn identifies Task 102, Task 103 and Task104 of MIL-STD-785 as essential elements. Other essentials listed include specification of the required value for each reliability parameter to be measured, specification of the combination of environmental test conditions and levels, and specification of performance parameters to be measured and frequency of measurement.

1. Test Reviews

a. Test Readiness Review. A test readiness review should be held approximately one week prior to the start of the RDT to assure that all test items and supporting elements are ready to begin the test. This review should include the following, as a minimum:

(1) Results of latest reliability predictions.

(2) Status of design

(3) Results of previous tests

(4) Review of all open problems and failures

(5) Availability of approved test procedures

(6) Readiness status of test equipment and test chambers.

b. Monthly Status Reviews. After the start of the RDT, monthly tracking meetings should be held to review test progress. These reviews should address the following, as a minimum:

(1) Current reliability growth assessments and projections based on test results and illustrated by reliability growth plots.

(2) Open failures and schedule for correction.

(3) Results of failure analyses and corrective action recommendations.

(4) Effectiveness of previously implemented corrective actions.

(5) Results of other tests which might affect RDT test hardware.

c. Other Reviews

(1) Weekly Status Reviews. After the start of the RDT, test progress should be briefed at weekly Program Managers meetings. Test progress should be shown through the use of RDT growth plots and any ongoing or anticipated problems should be discussed.

(2) Special Status Reviews. When the reliability growth plot shows a growth rate less than the minimum required, special reviews should be held to discuss options for correcting this situation. Such options might include more aggressive failure analysis and corrective action activity by the contractor, or suspension of the RDT and a major redesign effort.

2. Failure Actions

Upon the occurrence of a failure, the equipment under test should be repaired and placed back on test. The failed parts, circuit card, etc., can then be analyzed off-line, without adversely impacting the test schedule. Otherwise, the RDT could be stopped for weeks or months while corrective actions were being developed and implemented. When corrective actions are ready for implementation they should be installed in all equipments under test as logical block changes. These changes must then be tested adequately to ensure the original failure mode has been corrected and no new failure modes introduced

D. ASSESSING RELIABILITY GROWTH

In RDT, as in the entire TAAF process, the primary focus is on accomplishing reliability improvement rather than demonstrating reliability achievement. However, because of the relatively formal control, RDT is better suited to quantification than other elements of the TAAF process. RDT can help answer the quantitative questions of where are we? (currently achieved reliability) and how rapidly are we progressing? (growth rate). In this section, we discuss the most important models, interpretation of data, and the kinds of decisions you may need to make that growth assessment will support.

The only direct, model-independent measure of current reliability is the elapsed time between the two most recent failures. The uncertainty associated with a sample of one is so large as to make that measure virtually useless. Therefore, it is necessary to take the accumulated data in RDT into account, which in turn requires choosing a model for the growth process.

MIL-HDBK-189 discusses a number of growth models, with emphasis on the AMSAA model. Section 4.3 of MIL-HDBK-781 indicates the procedures to be followed in employing the AMSAA and Duane methods. The AMSAA model is essentially the stochastic extension of Duane's method. Both model reliability growth as a power function of test time. The Duane method, with its heavy reliance on graphic techniques, is well suited for quick (and "dirty") analyses and for detection of discontinuities in the growth process. A Duane plot requires only logarithmic graph paper and the calculation of ratios of cumulative test time to cumulative failures. The AMSAA model is essential for determination of confidence bounds and for objectivity in parameter estimation. It should always be used in connection with contractual quantification requirements. Use of the AMSAA model requires a few minutes with a calculator or a few seconds on a personal computer.

Both models imply continuity of growth, which requires that testing be suspended after each failure until the corresponding fix has been implemented in all test items. In the real world, discontinuous growth is the norm. In fact, subparagraph 202.2.2.3 of MIL-HDBK-189 explicitly permits the procuring activity to authorize replacement of failed items so that the test can continue while failures are being investigated. Figure 5 (adopted from MIL-HDBK-189, Figure 5.29) illustrates the effects of such test continuation. Figure 6 (from MIL-HDBK-189, Figure 5.30) also shows the impact of noninstantaneous fixes on the relationship between test time and calendar time. Although such discontinuities interrupt the smooth theoretical progress of reliability growth versus test time, it is common (and reasonable) practice to make estimates as if a smooth power function were the appropriate model unless there is clear evidence that distinct test phases need to be analyzed separately. Plotting methods are well covered in MIL-HDBK-189.

Reliability growth produces inherently very "noisy" test data--even under ideal conditions. This fact is reflected in the wide confidence intervals tabulated in Appendix C of MIL-HDBK-189. Appendix C of this technical brief presents additional information not covered by MIL-HDBK-189. After 5 failures, 10 percent of the MTBF estimates can exceed the true value by factors greater than 2.6, and 10 percent at less than one half of the true value. Even at the 30th failure--implying a fairly lengthy test--the factors can be approximately 1.4 and 0.7, respectively.

Growth rates are also difficult to measure accurately. They can be in error by a factor of 3 or more in either direction for reasonable test times. It is possible for the growth to appear as negative (that is, worsening) when the true trend is positive and vice versa. The possibility for error varies with growth rate (relatively much smaller for higher growth rates) (See Appendix C). Thus, estimated reliability and estimated growth aid your judgment--they do not substitute for it. Attempts at quantification of RDT results for contractual purposes should be avoided unless a test program of substantial length is implemented.

The types of decisions you may need to make and for which growth assessment is an aid are the following:

- Early termination of RDT on the basis of compelling evidence (on the growth curve) that reliability objectives have been reached early.
- Early termination of RDT when quantitative and qualitative results (tha is, types and patterns of failure) indicate a need for major redesign (followed by a TAAF restart) rather than a series of fixes.
- Extension of RDT when there is both substantial doubt that the objectives have been reached and confidence that growth will continue--for example, when there is evidence of growth, but at a lower rate than planned.
- Extension of RDT when additional time is needed to verify the effectiveness of the most recent fixes.
- Test multiple, identical units in parallel to increase the opportunity of identifying failure modes.

This and the previous section discussed management issues and preferred practices. Appendix B presents suggested RFP language to implement these practices. A list of related references is contained in Appendix D.

FIGURE 5

EFFECT OF DEFERRING REDESIGN

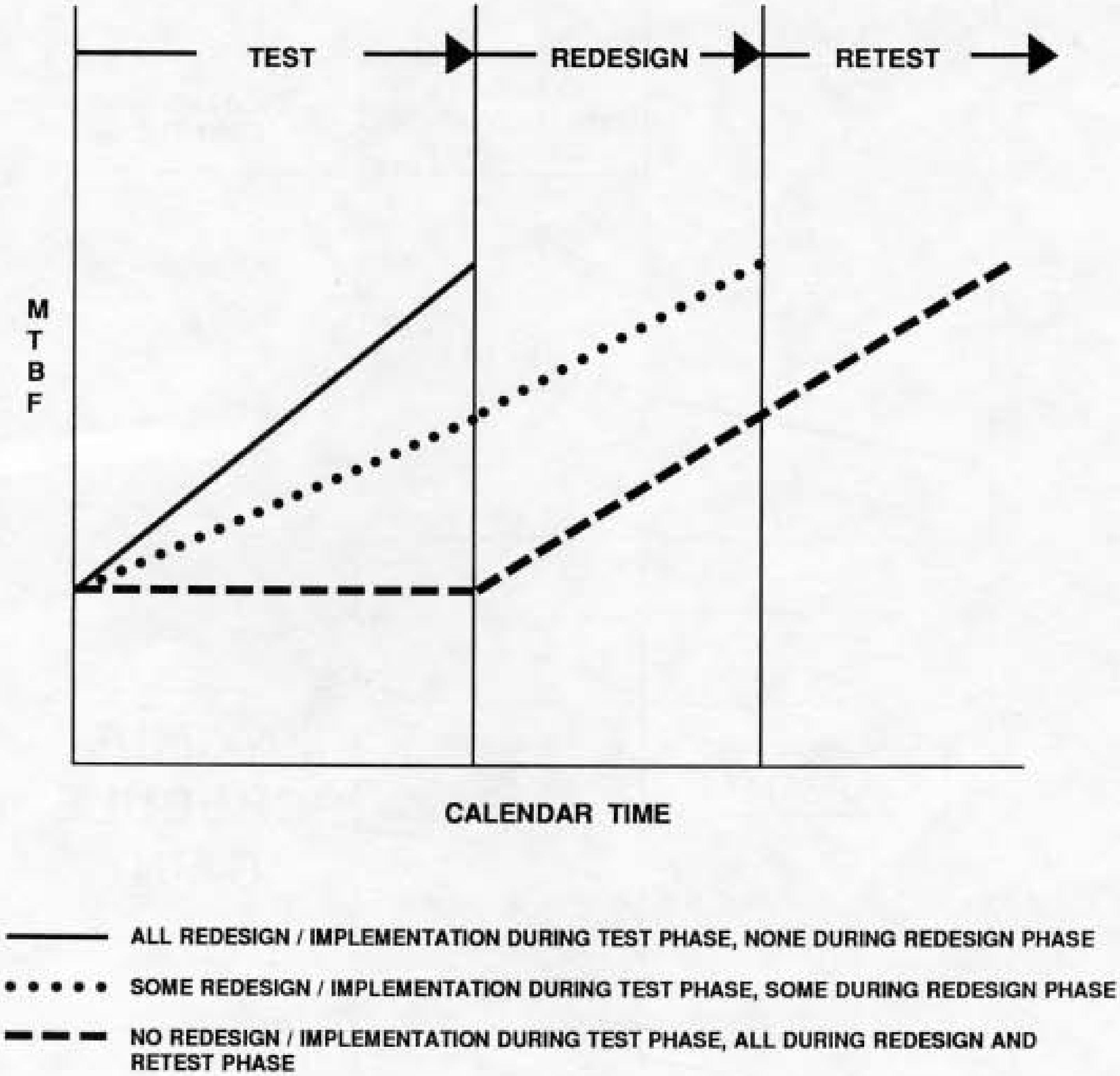
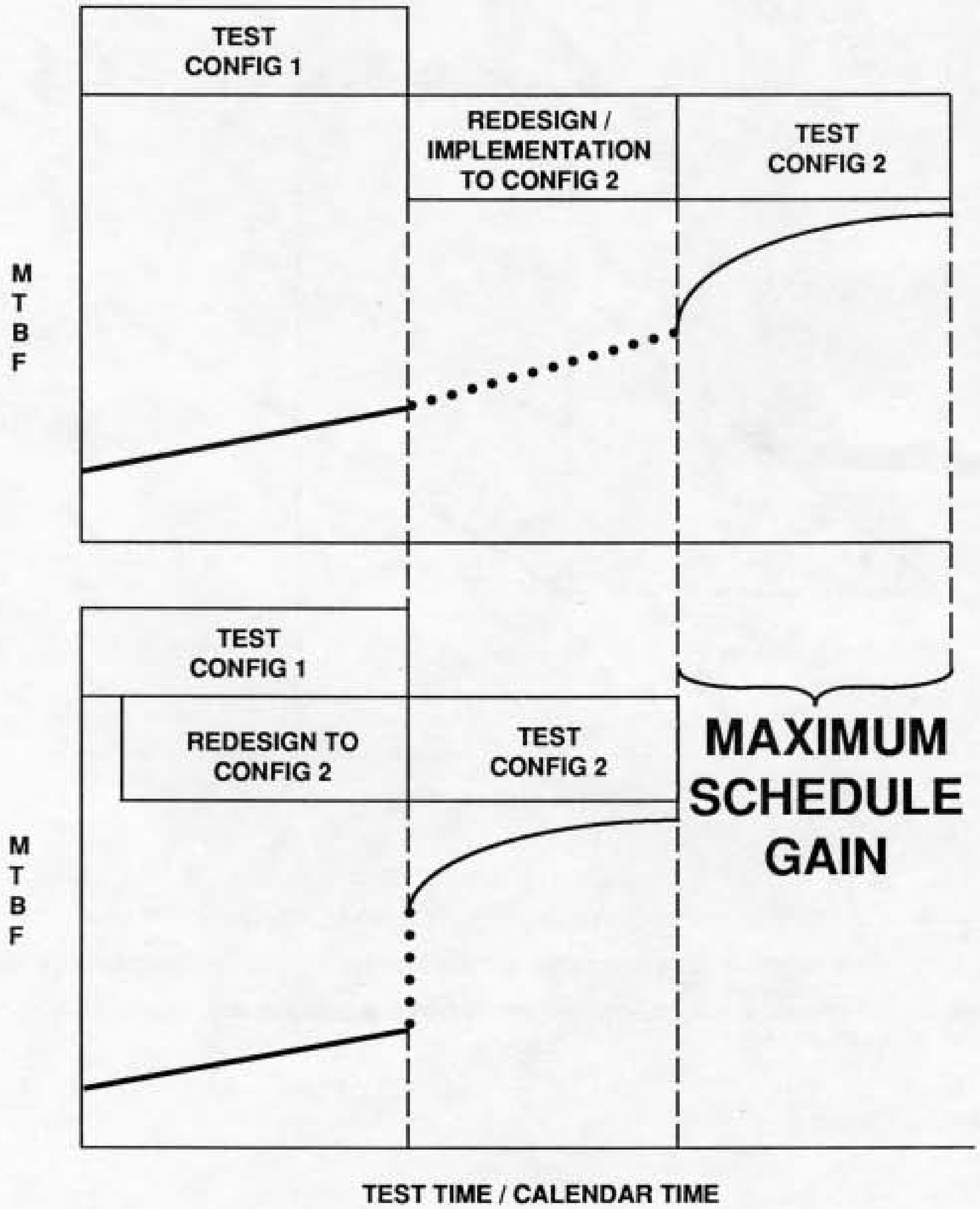


FIGURE 6

ACCOUNTING FOR CALENDAR TIME REQUIRED FOR REDESIGN



APPENDICES

APPENDIX A

GLOSSARY

AMSAA MODEL	A reliability growth model developed at the U.S. Army Materiel Systems Analysis Activity by Larry H. Crow. Can be viewed as a stochastic extension (a nonhomogeneous Poisson process) of the Duane model.
DUANE MODEL	A reliability growth model developed by J.T. Duane based on the empirical observation of straight-line decrease in failure rate versus cumulative test time when plotted on log-log paper.
ESS	Environmental Stress Screening: A ***stimulation*** process intended to induce early detection and correction of workmanship flaws, weak parts, etc., by the application of thermal and vibrational stresses to electronic products.
FRACAS	Failure Reporting Analysis and Corrective Action System.
FSD	See FSED.
FSED	Full-Scale Engineering Development: The acquisition phase immediately preceding production.
RD/GT	Reliability Development/Growth Test: A formal environmental test ***simulating*** operational parameters in which failures are welcome to achieve reliability growth via the TAAF process. In such a test, growth also may be analyzed; for example, by use of the AMSAA model.
RDT	Reliability Development Test: See RD/GT.
RELIABILITY GROWTH	The improvement of reliability as a result of maturation--that is, the elimination or reduction of specific, individual failure causes by improvements in design or processes.
RGT	Reliability Growth Test: See RD/GT.
TAAF	Test, Analyze, and Fix: The process by which growth is achieved.

APPENDIX B

EXAMPLES OF TAAF REQUEST FOR PROPOSAL LANGUAGE

Table B-1 portrays the sections of a government RFP. This appendix presents, by applicable section, sample language to implement TAAF. Where appropriate, a discussion of the rationale for the language is included. The sections are presented below in alphabetical order. It is important to tailor this language to the specific contract circumstances.

SECTION B. SUPPLIES/SERVICES AND PRICES/COSTS

Item no. xxxx Deliver xx shipsets of Prime Mission Equipment (PME) for Reliability Development Test.

Rationale/Comment:

Management commitment and involvement is essential to TAAF. Establishing the TAAF hardware as a line item deliverable makes the investment visible to management. The government will accept the equipment on a form DD-250 and then bail it back to the contractor for test. If desired, Section B could also contain line items for partial shipsets where subsystem RDT is appropriate and refurbishment (correction of wear and tear, update to most recent configuration) of the equipment so that it can be placed in the inventory or used for other purposes.

SECTION C. DESCRIPTION/SPECIFICATIONS/WORK STATEMENT

x.x.x.x The contractor shall implement a Test, Analyze, and Fix (TAAF) process to assure reliability of equipment designed under this contract. The TAAF process shall be fully integrated with the rest of the reliability program and the system engineering process.

Rationale/Comment:

Intentionally minimal implementing language for TAAF. The instructions to offerors (Section L) encourages the offering contractors to recommend changes and/or additions to this statement of work paragraph.

A separate TAAF data item (i.e., TAAF "plan") is not specified for the contract data requirements list (CDRL). This item would be in opposition to integrating TAAF into the overall reliability program. Some procuring activities will place the statement of work in an attachment to this section.

SECTION H. SPECIAL REQUIREMENTS

a. The following is a recapitulation of the Award Fee amounts for the xxxx contractor for each evaluation period during the basic contract. The award fee has a base value of 0% and a maximum fee of xx%.

Evaluation Period	Maximum Available Fee
Through milestone X	$ TBD
Through milestone XX	$ TBD

The total maximum Award Fee allotted is $ TBD.

b. The maximum award fee available for any evaluation period shall be based on the percent of the total estimated cost of implementing the RDT that falls prior to each RDT milestone and that has elapsed, if applicable, since a prior RDT milestone.

c. The Award Fee Determination Plan is in accordance with attachment x.

Rationale/Comments:

Unlike traditional Reliability Demonstration Tests, the intent of TAAF is to find and remove reliability design defects rather than "prove" that the design meets a reliability threshold. It is important, therefore, to provide an incentive for finding and removing defects. Ideally, contractors would be reimbursed on confirmation that redesign has removed the defects identified through test. However, it is impractical to decide prospectively how many defects found (or the rate at which they are found) are evidence of proper attention; similarly, the rate of removal of defects is also impractical to decide in advance. Assessment of the contractor's progress is necessarily subjective rather than objective. For this reason, an award fee is suggested. It is, however suggested that, where a single contract exists for both development and production, final award fee determination and payment should occur after the item is fielded and reliability performance is measured. "The amount of the award fee to be paid is determined by the Government's judgmental evaluation of the contractor's performance..." (FAR 16.404-2).

It should be noted that implementation of award fee provisions for TAAF will require substantial technical involvement on the part of the government to ensure fair evaluations. In particular, it will require the government to prepare an award fee determination plan to spell out the evaluation criteria. In this case, the criteria should principally focus on technical, schedule, and management performance on the RDT element of the TAAF process.

It is important to assure that RDT award fee provisions work in harmony (rather than at cross purposes) with other contract incentive clauses. For example, it is possible that the contract as a whole might include an award fee provision.

SECTION L. INSTRUCTIONS TO OFFERORS

Relevant Past Experience:

Offerors are advised that their performance of similar Test, Analyze, and Fix (TAAF) work will be an evaluation criterion in source selection. To this end, the offeror shall include in the proposal specific relevant past TAAF performance. Past performance means quality of work, essentially comparable to work contemplated by this RFP, completed under and in accordance with a contract. This information should also include the name and phone number of the cognizant contracting office technical representative, description of work, and discussion of the similarities between this previous experience and the TAAF requirements of the statement of work. Offerors having no specific relevant past performance will not be scored under that criterion but must state in their proposal that they have no relevant past performance.

Statement of Work:

If a government approach to the Test, Analyze, and Fix (TAAF) process is implied, suggested, or required, the prospective contractor is not limited to such an approach for equal or preferred consideration. In fact, streamlined and alternative, more efficient methods are encouraged. However, any variation from the approach suggested by the government must be justified in the proposal.

Test, Analyze, and Fix Process:

This solicitation includes a requirement for a Test, Analyze, and Fix (TAAF) process. The offeror shall include and identify the milestones, events, and criteria in the proposed System Engineering Master Schedule (or equivalent document) necessary for TAAF activities. The offeror shall include in any revision he proposes to the SOW and identify in the technical volume of the proposal, by paragraph number, all the proposed SOW tasks necessary to accomplish the proposed TAAF process. He shall also detail the specific equipments, quantities thereof, environments, test times, starting points, growth rates, integration with the Failure Reporting, Analysis and Corrective Action System (FRACAS), design analysis methods, and statistical analysis methods (Duane, AMSAA, etc). The offeror's combined approach to the specification, statement of work tasks, and system engineering master schedule is expected to reflect both a clear understanding of the goals of the TAAF process and an engineering and managerially sound approach to fulfilling those goals.

TABLE B-1

REQUEST FOR PROPOSAL SECTIONS

<u>SECTION TITLE, COMMENTS</u>

I (Schedule)

A	Solicitation
B	Supplies/Services and Prices/Costs, "Line Items"
C	Description/Specifications/Work Statement (in some cases)
D	Packaging and Marking
E	Inspection and Acceptance
F	Deliveries or Period of Performance
G	Contract Administration
H	Special Requirements, "Special provisions"

II (Clauses)

I	Clauses: FAR "general provisions"

III (Documents, Exhibits, and Attachments)

J	Attachments: Includes SOW (in some cases) Includes CDRL May include instructions for preparing proposals

IV (Representations and Instructions)

K	Representations and Certifications
L	Instructions to Offerors
M	Evaluation Factors

APPENDIX C

UNCERTAINTY OF MTBF AND GROWTH ESTIMATES

Growth processes are rather "noisy". In the development of this document, to determine the probable uncertainties for MTBF and growth estimates, Monte Carlo simulations were run using the AMSAA model. For each trial, all failure times up to the 30th failure were recorded, and estimates of the growth rate and current MTBF were made at the 5th, 10th, 15th, 20th, 25th, and 30th failure. Figure C-1.A indicates the band containing 80 percent of the simulation results in terms of the ratio of estimated MTBF to true MTBF versus the number of failures. After 5 failures, 10 percent of the MTBF estimates would be expected to exceed the true value by factors greater than 2.6 and 10 percent would be less than 0.45 of the true value. Even at the 30th failure--implying a fairly lengthy test--the factors were approximately 1.4 and 0.7, respectively.

Again using the Monte Carlo simulations, Figures C-1.B,C,and D reflect varying dispersions of growth estimates depending upon the true growth rate. The greater the true growth rate, the smaller the band of dispersion. Additionally, regardless of the true growth rate, as testing continues and failures increase the dispersion band narrows. There are risks of both gross overestimation and substantial underestimation of growth rates--to the extreme of seemingly negative growth, especially in the case of low true growth and limited test results. The point of Figure C-1 is that because of the considerable uncertainty, estimates of MTBF and growth rate are indicators only. Do not make hard decisions on the basis of these indicators alone - use engineering judgement.

While Figure C-1 illustrates how it is possible to observe a wide range of apparent MTBFs for a given true MTBF and, similarly, the range of apparent growth rates that could be observed, a program manager may well be interested in approaching the question from a different direction. That is: for a given estimated (or calculated, observed, etc.) growth rate, what is the possible range of the true MTBF and how does that range compare to the required MTBF?

Figure C-2.A is a graphic version of the confidence interval tables found in MIL-HDBK-189 and assists in answering the manager's question. Suppose that the required MTBF is 1000 hours and that after 5 failures the current estimate is 1500 hours. Figure C-2.B takes a "slice" through the Figure C-2.A confidence band at the 5 failure point and shows the results in MTBF terms rather than ratios. Even though the amount of test data is small, the manager would justifiably have confidence in his progress: with a range from 640 to 5080 hours, there is relatively little likelihood that the true MTBF is under 1000 hours. The same cautions apply as stated above for Figure C-1.

FIGURE C - 1

UNCERTAINTY OF PROSPECTIVE MTBF AND GROWTH RATE ESTIMATES

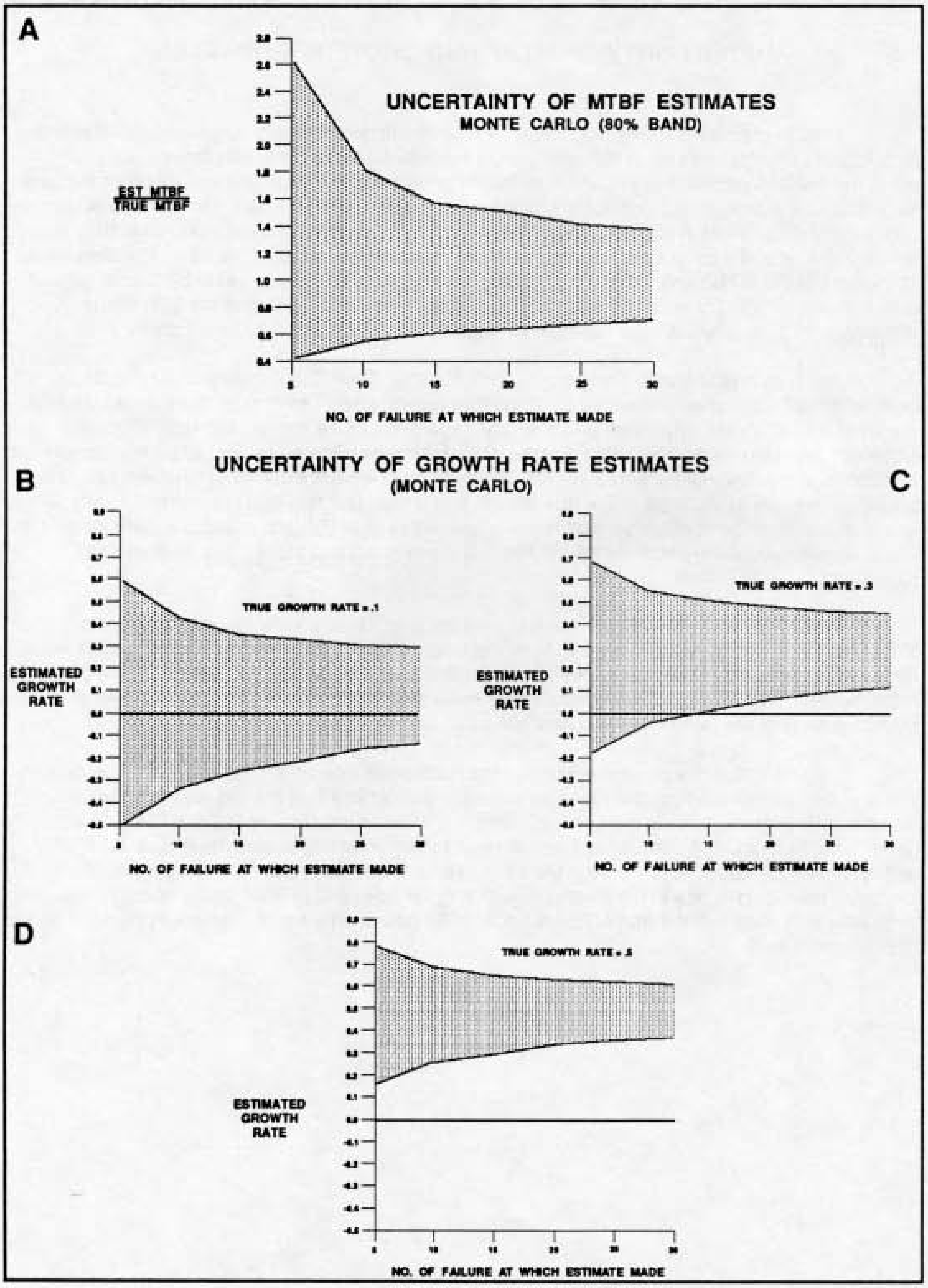
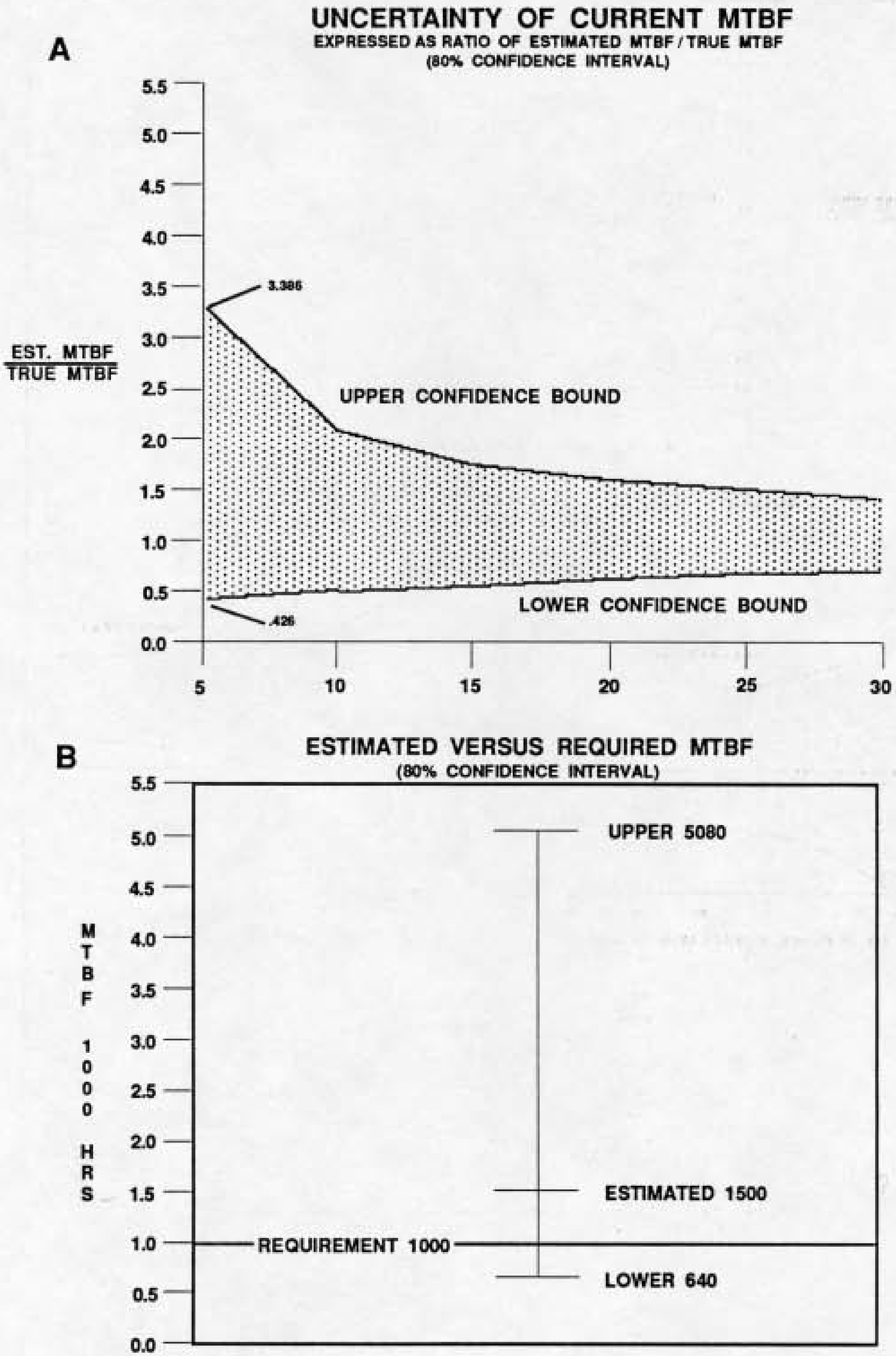


FIGURE C - 2



A
UNCERTAINTY OF CURRENT MTBF
EXPRESSED AS RATIO OF ESTIMATED MTBF / TRUE MTBF
(80% CONFIDENCE INTERVAL)
EST. MTBF / TRUE MTBF
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
3.386
UPPER CONFIDENCE BOUND
LOWER CONFIDENCE BOUND
.426
5
10
15
20
25
30
B
ESTIMATED VERSUS REQUIRED MTBF
(80% CONFIDENCE INTERVAL)
MTBF 1000 HRS
5.5
5.0
4.5
4.0
3.5
3.0
2.5
2.0
1.5
1.0
0.5
0.0
UPPER 5080
ESTIMATED 1500
REQUIREMENT 1000
LOWER 640

APPENDIX D

SELECTED BIBLIOGRAPHY AND REFERENCES

DOD PUBLICATIONS

3235.1-H	Test & Evaluation of System Reliability, Availability, and Maintainability
4245.7-M	Transition From Development to Production
5000.40	Reliability and Maintainability

This directive establishes policies and responsibilities for the reliability and maintainability of defense systems,subsystems, and equipment. It implements the principles of DOD Directives and Instructions for major system acquisition and for test and evaluation.

MILITARY HANDBOOKS AND STANDARDS

MIL-HDBK-189	Reliability Growth Management
MIL-HDBK-217E	Reliability Prediction of Electronic Equipment
MIL-HDBK-781	Reliability Test Methods, Plans, and Environments for Engineering Development, Qualification, and Production
MIL-STD-781D	Reliability Testing for Engineering Development, Qualification, and Production
MIL-STD-785	Reliability Program for Systems and Equipment Development

AIR FORCE PUBLICATIONS

AFP 800-7	USAF R&M 2000 Process, January 1989
UNNUMBERED	RADC Reliability Engineer's Toolkit, July 1988

ARMY PUBLICATIONS

AR 702-3	Army Material System Reliability, Availability, and Maintainability

NAVY PUBLICATIONS

NAVSO P-6071 — Best Practices: How to Avoid Surprises in the World's Most Complicated Technical Process

SECNAVINST 4490.2 — Transition from Development to Production, 13 Mar 1987

This instruction implements the policy and procedures for the discipline of risk management in the transition from development to production contained in DOD 4245.7-M

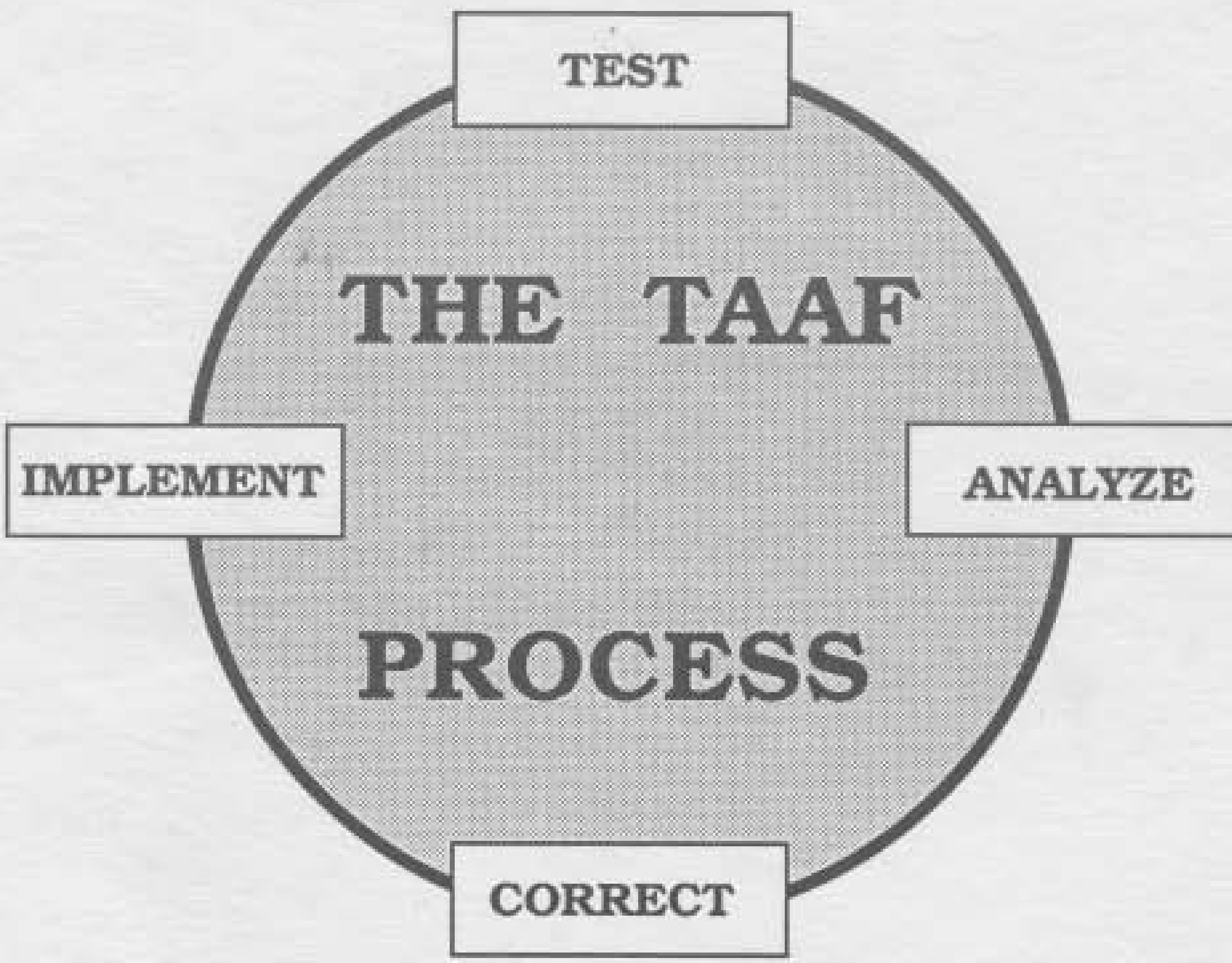
OTHER PUBLICATIONS

IES — Reliability Growth Through Test, Analyze, and Fix (TAAF) and Environmental Stress Screening

The proceedings of the October 2, 1987, seminar sponsored by the Institute of Environmental Sciences, Orange County Chapter. A compendium of formal papers and presentation material.

IES — Selected References on Reliability Growth

A 1988 compilation of papers from IES and IEEE sources available from the Institute of Environmental Sciences, 940E. Northwest Highway, Mount Prospect, IL 60056.



TEST, ANALYZE, AND FIX