GUIDANCE FOR ASSESSING THE LIKELIHOOD THAT A SYSTEM WILL DEMONSTRATE ITS RELIABILITY REQUIREMENT DURING INITIAL OPERATIONAL TEST.

1. INTRODUCTION

Purpose

The purpose of this white paper is to provide guidance to the personnel within Program Manager's office and operational test agencies who develop the test strategies to demonstrate the reliability requirements for the Department of Defense systems. It will serve as a guideline for assessing the likelihood that a system will achieve its operational reliability threshold during the initial operational testing (IOT). This paper will provide guidance for planning reliability testing and will discuss the implications of the test duration / number of trials and maximum acceptable number of failures on the probability of successfully completing operational testing.

Note: Throughout this white paper reliability threshold means reliability requirement and target reliability means goal reliability.

Policy Statement

With reference to the Directive – Type Memorandum (DTM) 11 – 003 – Reliability Analysis, Planning, Tracking, and Reporting, March 21, 2011, Project / Program Managers (PMs) and operational test agencies shall assess the reliability growth required for the system to achieve its reliability threshold during initial operational test and evaluation and report the results of that assessment to the Milestone Decision Authority at Milestone (MS) C.

General Observation

Approximately twenty-five to thirty-five percent of the programs entering Operational Test in 2009 demonstrated poor reliability (ref DOT&E Memorandum, SUBJECT: Next Steps to Improve System Reliability, dated 18 Dec 2009). It is not intuitively obvious that a program may need a target reliability value substantially greater than the program requirement to have a high (e.g., 80%) chance of successfully demonstrating the requirement during the operational testing. The target reliability for a test is a function of the test duration, acceptable number of failures, consumer risk, and producer risk.

For example, an existing Department of Defense (DoD) program had a required Mean Time Between Failure (MTBF) of 250 hours. The Initial Operational Testing (IOT) consisted of 3000 hours of testing with 8 or fewer acceptable failures. For this test design, the program target MTBF to have an 80 percent chance of success was 467 hours. Increasing the test duration would lower the target MTBF. Table 1 below provides a listing of other potential test designs for this program. As can be seen in this table, the target MTBF is inversely proportional to the test duration. Reliability test designers must trade-off the test duration with the programmatic reliability target to develop their reliability program plan. This paper will discuss some of the tools that can be used to support this effort.

Table 1. Sample Relationship of Test Duration and Target MTBF for IOT (MTBF Requirement = 250 hours)

Test Duration	Maximum Acceptable Failures	Target MTBF (80% chance of success)
3000	8	467
6000	19	371
10000	34	334
Note: Target MTBF is same as Goal MTBF.		

DOT&E has provided guidance specifying that OT plans will include evaluation of producer and consumer risks (ref DOT&E Memorandum, SUBJECT: Test and Evaluation (T&E) Initiatives, dated 24 Nov 2009). Implied but not explicitly stated is the need to carefully consider the proper balance of the two risks. Small values of producer risk may require unrealistically long/large test duration/number of trials to achieve the desired reliability that would demand impractical investments of time and money. Determining a reasonable producer risk requires balancing the costs of reliability growth and test duration against the penalties of test failure and the impacts of delayed fielding.

This document will act as a guide to explain the procedures and methodology for estimating and balancing consumer risk and producer risk against demands for minimal development costs and timely deployment.

Section 2 will discuss the process for assessing consumer risk and producer risk using Operational Characteristics curves. Section 3 will provide a discussion of the importance of the Reliability Growth Program in achieving the reliability necessary to have a high probability of passing IOT. In Section 4, the acceptable levels of consumer and producer risks will be discussed and how these risks may be used by the Program Managers in the decision making process. The balancing of these risks and their impact on the system reliability and adjustment of test duration/number of trials, if necessary, will be discussed. Appendix A provides techniques to calculate consumer risk and producer risk for the continuous and discrete systems for the fixed configuration test. The appendix provides estimation of the two risks for some real systems including the one which is quoted in the Introduction.

2. OPERATING CHARACTERISTIC (OC) CURVE AND CONSUMER AND PRODUCER RISKS

Program Managers, Evaluators, Testers, and other key acquisition personnel need to know the probability of acceptance for a test plan to design an appropriate test plan which will ensure demonstration of reliability requirement at the desired confidence level. The most commonly used tool for this purpose is the Operating Characteristic (OC) Curve. Figure 1 provides a sample OC Curve. This OC curve is generated for a fixed configuration test and displays the relationship between the probability of acceptance and MTBF based on test duration and acceptable number of failures. The OC curve is a tool to determine the probability of acceptance of a test plan corresponding to a given reliability requirement. The OC curve is used to quantify the consumer risk and producer risk associated with a given MTBF value for the associated test plan.

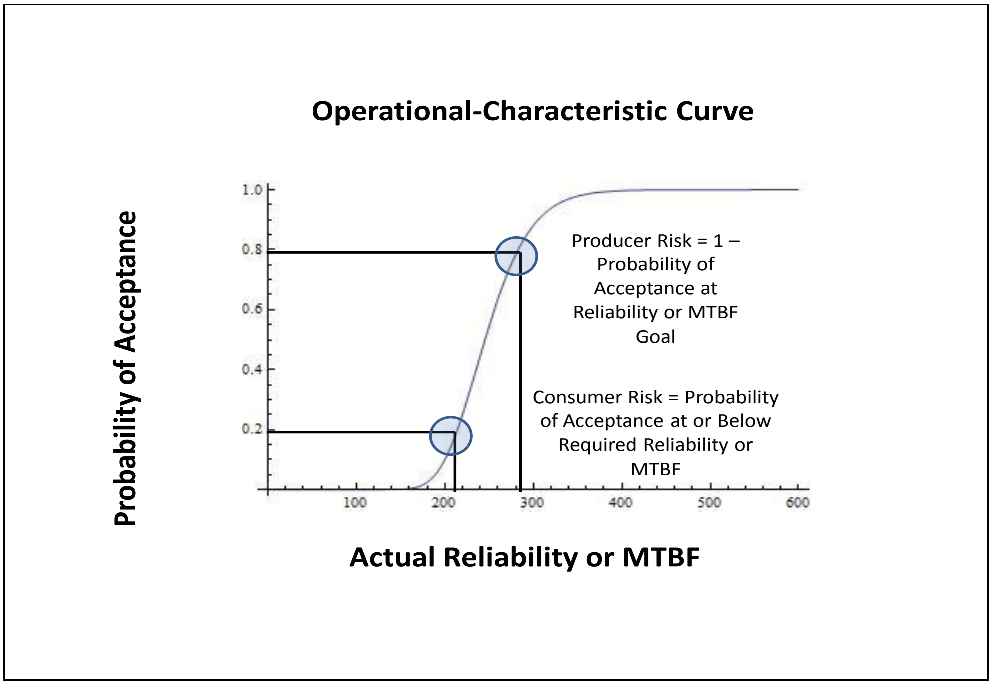

Figure 1: OC Curve Depicting of Decision Risks.

Reliability Risks

There are two types of decision risks which are of significant importance during the demonstration of reliability requirements. These risks are called Consumer Risk and Producer Risk.

1. **Consumer risk**: The probability that a level of system reliability at or below the requirement will be found to be acceptable due to statistical chance. This is depicted on the operational characteristic curve in Figure 1. We should endeavor to quantify and manage consumer risk because reliability below the requirement results in reduced mission reliability and increased support costs.
2. **Producer risk**: The probability that a level of system reliability that meets or exceeds the reliability goal will be deemed unacceptable due to statistical chance. This risk is also depicted in Figure 1. If the system is incorrectly deemed unsuitable, major cost and schedule impacts to the acquisition program may result.

An appropriate balance between the consumer risk and the producer risk is important to determine test duration/number of trials. If the consumer risk and producer risk are not balanced appropriately, the test duration/number of trials may be too short/small or too long/large. If the test duration/number of trials is too short/small, the reliability goal (target) for the test will be higher (test reliability requirement is inversely proportional to the test duration/number of trials). For short/small test duration/number of trials, one or both risks may be too high. If the test duration/number of trials is too long/large, it may be very costly to perform the test. The cost factor may lead to an unacceptable program burden.

Construction of Operating Characteristic (OC) Curve – Continuous System

Continuous System

A system which operates continuously during a specified period of test time is referred as a Continuous System. For example, a radar system operates continuously during a specified time period. Such a system is repairable and it may consist of various subsystems/components. During the test period, a continuous system may fail to perform its intended operation because of malfunction of one or more subsystems/components. In order to measure the reliability of a continuous system the number of failures are collected during the test time which is used to estimate reliability. The reliability metric could be in terms of mean time between failures (MTBF) or mean miles between failures (MMBF) depending on the type of system and the way it is used. For example, in the case of a radar system, failure times (hours) are recorded and for a truck, failure miles are recorded.

The OC curve displays both acceptance and rejection risks associated with all possible values of the reliability parameter and not merely the requirement and goal (target) reliability values. For continuously operating time, the OC curve is a plot of Probability of Acceptance versus MTBF. As an example, the family of OC curves is constructed as shown in Figure 2 for the test plans presented in Table 1 above.

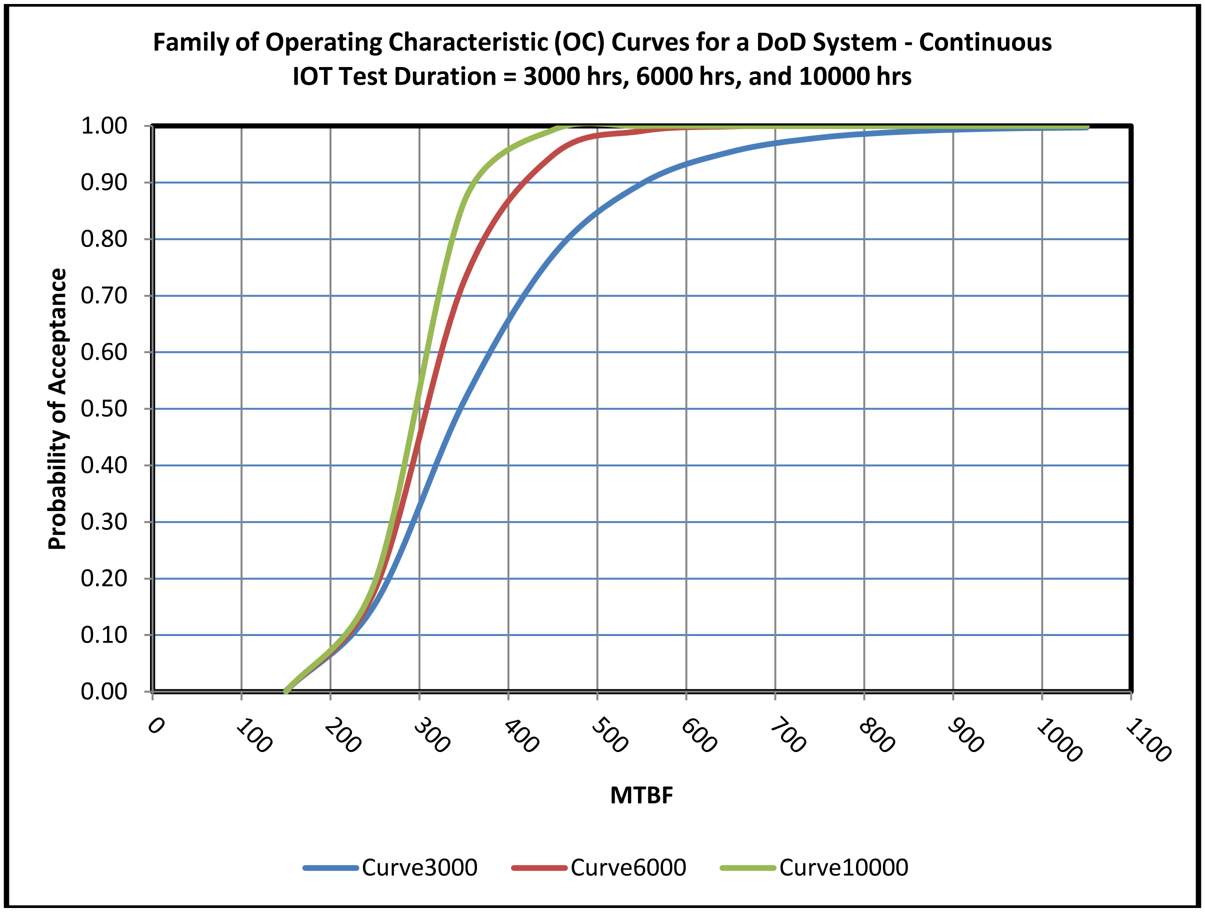

Figure 2. OC curves for the IOT test plans – Continuous System.

The consumer and producer risks associated with each OC curve in figure 2 are listed in Table 2.

Table 2. Consumer and Producer Risks for OC Curves – Continuous System

OC Curve	Max. Acceptable Failures	MTBF Requirement	Consumer Risk	Target MTBF	Producer Risk
Curve3000	8	250	0.16	467	0.20
Curve6000	19	250	0.18	371	0.20
Curve10000	34	250	0.19	334	0.20
Note: Target MTBF is same as Goal MTBF. Curve3000 is the OC curve for IOT duration of 3000 hours. Curve6000 is the OC curve for IOT duration of 6000 hours. Curve10000 is the OC curve for IOT duration of 10000 hours.					

Construction of Operating Characteristic (OC) Curve – Discrete System

Discrete System

A one-shot system is referred to as a discrete system because the time of operation is not continuous. For example, a missile system does not fire missiles continuously during a specified period of time. Such a system may consist of various subsystems/components. Even though some subsystems/components may operate continuously, the success of the system is measured in terms of successful firing of the missiles. For a missile system, the reliability is measured in terms of probability of successful firing of a missile. In general, the reliability of a discrete system is measured in terms of probability of successful occurrence of a discrete event known as a trial.

Similar to the case of a continuous system, the OC curve displays both acceptance and rejection risks associated with all possible values of the reliability parameter and not merely the requirement and goal (target) reliability values. For the discrete system, the OC curve is a plot of Probability of Acceptance versus Probability of Success per Trial (i.e., Reliability). As an example, the family of OC curves is constructed as shown in Figure 3 for the test plans presented in Table 3 below.

Table 3. Sample Relationship of Number of Trials and Target Reliability for IOT (Reliability Requirement = 0.66)

Number of Trials	Maximum Acceptable Failures	Target Reliability (80% chance of success)
22	5	0.815
32	8	0.79
48	13	0.77
Note: Target Reliability is same as Goal Reliability.		

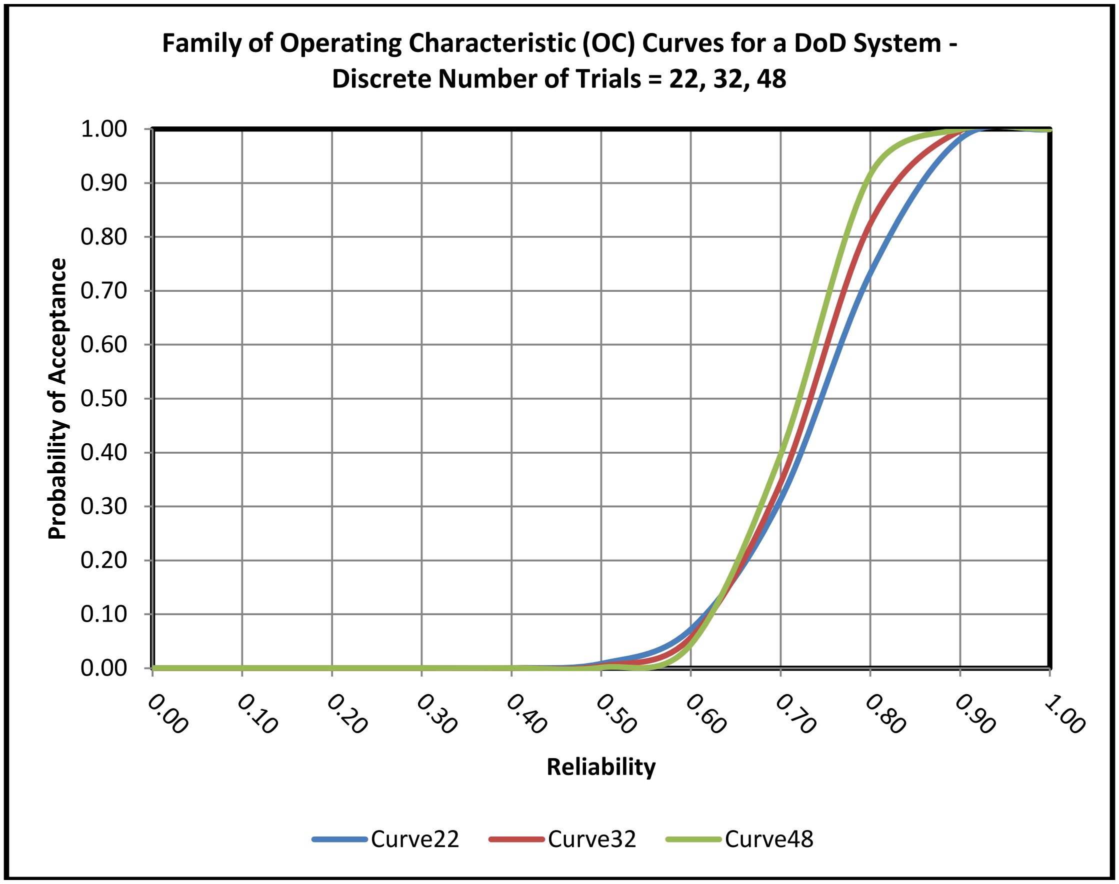

Figure 3. OC curves for the IOT test plans – Discrete System.

The consumer and producer risks associated with each OC curve in Figure 3 are listed in Table 4.

Table 4. Consumer and Producer Risks for OC Curves – Discrete System

OC Curve	Maximum Acceptable Failures	Reliability Requirement	Consumer Risk	Target Reliability	Producer Risk
Curve22	5	0.66	0.19	0.815	0.21
Curve32	8	0.66	0.19	0.79	0.22
Curve48	13	0.66	0.20	0.77	0.20
Note: Target Reliability is same as Goal Reliability. Curve22 is the OC curve for 22 trials during IOT. Curve32 is the OC curve for 32 trials during IOT. Curve48 is the OC curve for 48 trials during IOT.					

Phases of the Acquisition Process

In this white paper the focus is on the consumer and producer risks that a system will encounter during IOT. However, the system has to go through different phases of the acquisition process under DT before entering IOT. Therefore, the two risks during IOT will be impacted by the way the resources and test activities in each phase of the acquisition process are managed.

The phases of the acquisition process are shown in Figure 4.

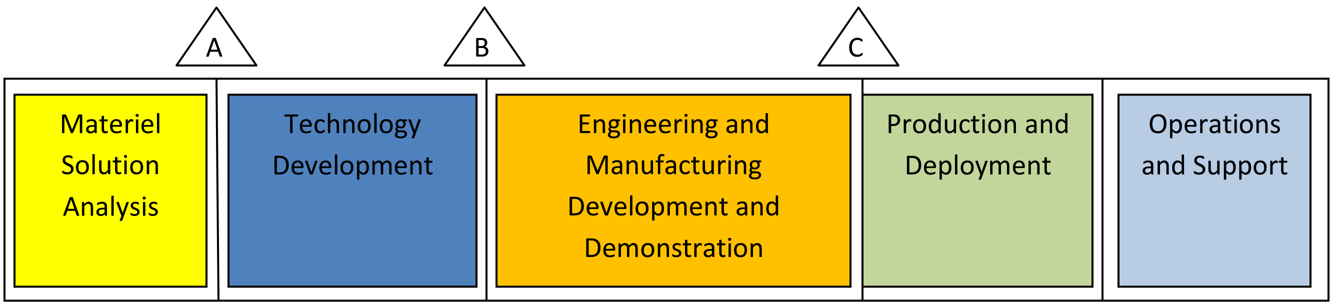

Figure 4. Phases of the Acquisition Process.

Initially, the consumer and producer risks for IOT test phase should be calculated as early as possible and well before MS B. These risks should then be routinely updated as programmatics (e.g., number of test assets, test scope, test durations/number of trials, etc.) change. The estimation of these risks depends on:

1. Reliability requirement
2. Reliability goal (target)
3. Test duration/Number of trials
4. Maximum acceptable failures

Using the information listed above, the consumer and producer risks can be determined using the techniques described in Appendix A. A graphical tool which readily provides the consumer and producer risks is the OC curve as mentioned in Section 1 above.

3. RELIABILITY GROWTH PROGRAM

During the developmental test of a system the reliability growth planning and assessment models are the useful tools to manage reliability improvement of the system. The level of consumer and producer risks during IOT depends on how mature the system design is, that is, how closely the reliability growth of the system follows the planning curve and how well the reliability improvement is tracked during the test. The reliability improvement of the system under test will be expected to follow closely the planning curves if (a) test plan is designed / executed adequately, (b) the failure modes are documented and analyzed appropriately, (c) corrective actions are implemented efficiently, and (d) design modifications are performed appropriately. Development of the planning curve at early stages, that is, pre-MS B of the acquisition program is important because the planning curve acts as the baseline for the system reliability growth.

A Reliability Growth Program (RGP) shall be included in the Systems Engineering Plan at MS A and updated in the Test and Evaluation Master Plan beginning at MS B. An important part of the RGP is development of reliability growth planning curve. The growth curve serves as a baseline to determine whether the system reliability growth is on track. This is a smooth idealized curve and system reliability will grow along this curve only if the corrective actions (fixes) are incorporated successfully as soon as a failure occurs. In reality, it is not possible to incorporate fixes instantly. Therefore, a finite number of test phases are defined over the entire test duration and the RGP states the proportion of all the failures during a test phase which will be fixed at the end of test phase. Since, it takes time to collect, analyze, and fix the failures which occur during a test phase, the RGP indicates an appropriate lag time for each test phase. Figure 5 illustrates the idealized planning curve, test phases and lag time during developmental test (DT).

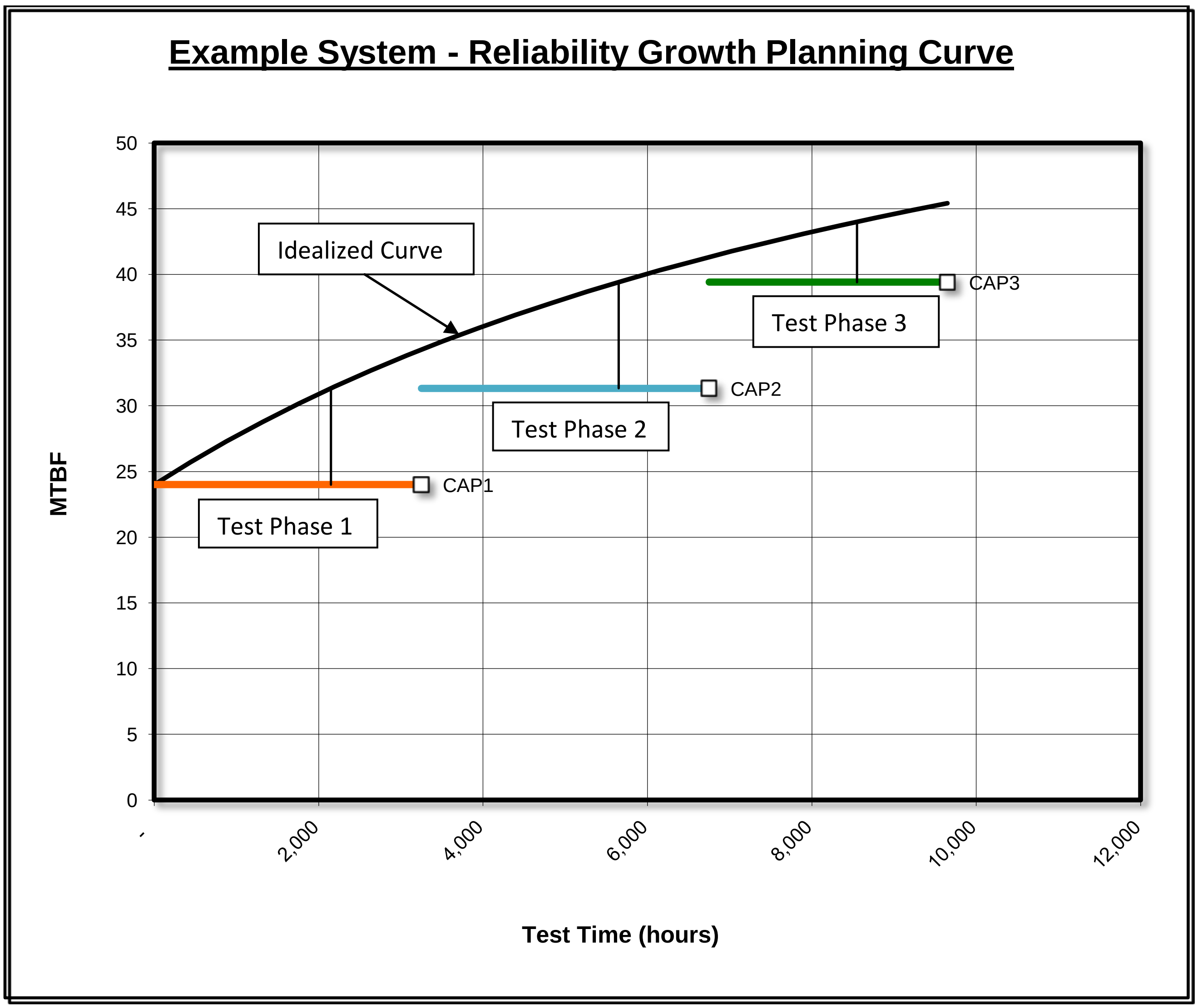

Figure 5. Planning curve, test configurations and lag time during DT.

In Figure 5, the smooth curve is the idealized planning curve. The three horizontal steps represent test phases and the lag time for a test phase (step) is from the vertical line to end of the step. The usage of lag time means that, for example, over the first step, failures between the starting point of the step and the vertical line were examined and fixed at end of first step. The CAP1, CAP2, and CAP3 represent corrective action periods which are shown as white squares at end of each step. The CAPs may extend over a few months. In Figure 5, a jump from one step to the next indicates reliability improvement assuming that fixes were incorporated successfully.

The risk associated with the reliability growth plan can be assessed by using the following parameter risk matrix as shown in Table 5.

Table 5. Reliability Growth Curve Risk Matrix

Category	Low Risk	Medium Risk	High Risk
MTBF Goal (DT)	Less than 70% of Growth Potential	70 – 80% of Growth Potential	Greater than 80% of Growth Potential
IOT&E Producer Risk	20% or less	20^+ - 30%	Greater than 30%
IOT&E Consumer Risk	20% or less	20^+ - 30%	Greater than 30%
Management Strategy	Less than 90%	90-96%	Greater than 96%
Fix Effectiveness Factor	70% or lower	70^+ - 80%	Greater than 80%
MTBF Goal (DT) / MTBF Initial	Less than 2	2-3	Greater than 3
Time to Incorporate and Validate Fixes in IOT&E Units Prior to Test	Adequate time and resources to have fixes implemented & verified with testing or strong engineering analysis	Time and resources for almost all fixes to be implemented & most verified w/ testing or strong engineering analysis	Many fixes not in place by IOT&E and limited fix verification
Corrective Action Periods (CAPs)	5 or more CAPs which contain adequate calendar time to implement fixes prior to major milestones	3 - 4 CAPs, but some may not provide adequate calendar time to implement all fixes	1- 2 CAPs of limited duration
Reliability Increases after CAPs	Moderate reliability increases after each CAP result in lower-risk curve that meets goals	Some CAPs show large jumps in reliability that may not be realized because of program constraints	Majority of reliability growth tied to one or a couple of very large jumps in the reliability growth curve
Percent of Initial Problem Mode Failure Intensity Surfaced	Growth appears reasonable (i.e. a small number of problem modes surfaced over the growth test do not constitute a large fraction of the initial problem mode failure intensity)	Growth appears somewhat inflated in that a small number of the problem modes surfaced constitute a moderately large fraction of the initial problem mode failure intensity	Growth appears artificially high with a small number of problem modes comprising a large fraction of the initial problem mode failure intensity

Note: $^+$ indicates strictly greater than.

This risk matrix can help determine if the program can likely attain the reliability goals (targets) consistent with the growth curve.

Tracking Progress during DT and Operational Test Readiness Review (OTRR) for IOT&E.

During DT, reliability tracking and projection tools are used to track a system reliability growth program. The overall risk of the reliability growth program can be assessed based on the evaluation results of the reliability growth tracking and projections against the planned reliability growth curve.

At the conclusion of the DT, the OTRR is conducted to ensure that the system can proceed into IOT with a high probability of success, and that the system is effective and suitable for service introduction. Since IOT is usually conducted by users in a harsher environment than DT, there is typically a drop-off in the reliability experienced during IOT from the projected growth based on DT. Therefore, the reliability goal (target) for the reliability growth curve is selected high enough during DT to compensate for the degradation anticipated during IOT.

4. ACCEPTABLE LEVEL OF CONSUMER AND PRODUCER RISKS, IMPACT OF LOW RELIABILITY, AND DECISION MAKING.

In order to determine the acceptable level of consumer and producer risks it is important to understand the sensitivity of the risks to other parameters involved and the way the two risks vary. The following observations about these risks are made based on a number of actual and sample datasets.

a. The consumer and producer risks are sensitive to test duration/number of trials and acceptable number of failures.
b. If the consumer and producer risks are balanced (i.e., risks are close to each other) increasing the test duration/number of trials will reduce both risks.
c. The system reliability growth program also significantly impacts the consumer and producer risks. Some of the key factors which play an important role in calculating the risks are the requirement, goal (target), and test duration/number of trials.

Demonstration of Lower Reliability and its Impact

If the projected or demonstrated reliability of the system is lower than the requirement for the IOT, then the following factors can be considered:

1. Improved design for reliability (DFR)
2. Adjust reliability growth program (additional testing, CAP(s), higher average fix effectiveness factor and/or management strategy)
3. Increased IOT test duration/number of trials
4. Reduced mission duration
5. Reduced reliability requirement (as the last option)

It is the most economically sound approach to have reliability designed and built into the system at the earliest possible stages of the acquisition process. Therefore, DFR should be examined and improved early in the acquisition process to improve the initial reliability for the reliability growth program. The higher initial MTBF will allow the system reliability to grow to a higher MTBF at the completion of the developmental testing, providing a better chance of successfully completing IOT.

Adjustment can be made to the reliability growth program to improve the growth achieved by the system. Additional test time/number of trials and CAP(s) will allow for additional growth. A higher average fix effectiveness factor and a higher management strategy (percentage of initial failure intensity to be addressed by corrective actions) will also allow for more growth. However, it should be noted that each of these adjustments will have cost/schedule impacts associated with them (test resources, test duration/number of trials and CAP duration, engineering support, etc.).

Reliability goal (target) should be set according to reliability growth potential. The growth potential is the theoretical upper limit of system reliability and it is calculated from the initial reliability, management strategy and fix effectiveness factor. If the goal (target) is set too close to growth potential, it may require extremely long test duration to reach the goal (target) which will not be cost-effective. As stated earlier, the reliability growth goal (target) at the end of DT necessary to achieve a higher chance of acceptance during IOT can be reduced by increasing the duration of IOT. Therefore, if the reliability growth program indicates a high risk in growing to the goal (target) MTBF, then increasing the IOT duration is an option for reducing the risk associated with the goal (target) MTBF.

Reducing a system's mission duration and/or the reliability requirement are also options for improving a system's chances for passing IOT, this could be done if requirements are deemed unreasonable and/or unachievable. However, changes to these values, require coordination amongst several organizations and should be pursued as early as possible.

Conclusions

1. This paper provides the approach to assess the consumer and producer risks associated with the reliability requirement for a specific test program during IOT.
2. It is important to know the probability of acceptance for a test plan to demonstrate the reliability requirement at the desired confidence level.
3. This paper provides guidance for planning reliability testing for a program with the emphasis on achieving the desired level of consumer and producer risks during IOT.
4. This paper emphasizes the importance of introducing a reliability growth program at MS A and subsequent updates at the beginning of MS B of the acquisition process. As part of the reliability growth program, the development of reliability growth planning curve at MS A will serve as a baseline to effectively track reliability improvement of the system and hence achieving the desired levels of consumer and producer risks during IOT. Reliability growth models are the useful tools for designing the reliability growth program.
5. The reduction of both consumer and producer risks depends on the adjustment of the test duration/number of trials and acceptable number of failures.

APPENDIX A

CALCULATION OF CONSUMER RISK AND PRODUCER RISK

This appendix describes how to calculate the consumer and producer risks for fixed configuration tests.

Techniques to Calculate Consumer and Producer Risks for Fixed Configuration Test (Continuous Systems)

For continuously operating systems, the consumer risk and the producer risk can be calculated by using either the Exponential distribution or the Poisson distribution for a fixed configuration test. It is a common practice to calculate these risks by using the Poisson distribution equations.

Poisson Distribution Equations

These equations use the following notation:

T = Total test duration
θ = True MTBF
c = Maximum acceptable number of failures
θ_0 = Required MTBF (also called lower test MTBF)
θ_1 = Goal (Target) MTBF (also called upper test MTBF)
α = Consumer risk
β = Producer risk

Prob(ac | θ) = Probability of accepting the system assuming the true MTBF is θ.
Prob(rej | θ) = Probability of rejecting the system assuming the true MTBF is θ.

The probability of acceptance that no more than c failures will occur can be calculated by the equation:

$$\text{Prob(ac} \mid \theta) = \sum_{k=0}^{c} \frac{\left(\frac{T}{\theta}\right)^k}{k!} e^{-\frac{T}{\theta}} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \text{(1)}$$

This is the Cumulative Poisson distribution equation.

Equation (2) can be used to calculate the Consumer Risk.

$$\alpha = \text{Prob(ac} \mid \theta=\theta_0) = \text{P(c or fewer failures} \mid \theta=\theta_0) = \sum_{k=0}^{c} \frac{\left(\frac{T}{\theta_0}\right)^k}{k!} e^{-\frac{T}{\theta_0}} \quad \ldots\ldots\ldots\ldots\ldots\ldots \text{(2)}$$

Equation (3) can be used to calculate the Producer Risk.

$$\beta = \text{Prob(rej} \mid \theta=\theta_1) = \text{P(more than c failures} \mid \theta=\theta_1) = 1 - \sum_{k=0}^{c} \frac{\left(\frac{T}{\theta_1}\right)^k}{k!} e^{-\frac{T}{\theta_1}} \quad \ldots\ldots\ldots\ldots \text{(3)}$$

The equations (2) and (3) can be used to determine the complete test plan. Generally, a test plan is specified by the test duration T and the maximum acceptable number of failures c. For the specified values of T, c, θ_0, and θ_1, the consumer and producer risks can be calculated from equations (2) and (3). On the other hand, if only consumer and producer risks are specified along with θ_0 and θ_1 values, then test duration T and maximum acceptable number of failures c can be calculated from equations (2) and (3) to describe the test plan completely.

Note: An alternate way to calculate these risks is by using Poisson function built in the Excel software. The built-in Poisson function provides the probability of acceptance. The risk levels are calculated as follows:

α = Consumer risk = POISSON(c, T/ Θ_0, TRUE) ………………………………. (4)
β = Producer risk = 1 – POISSON(c, T/ Θ_1, TRUE) ...………………………….. (5)

Following example illustrates how the estimates of the two risks calculated by equations (2) and (3) compare with the risks estimated by equations (4) and (5).

Example: A communication system has requirement of 100 hours MTBF, and a goal (target) of 150 hours MTBF. Determine the consumer and producer risks given that the system is ready to enter initial operational test (IOT). The proposed test duration for IOT is 2000 hours and the test plan allows no more than 15 failures.

T = IOT test duration = 2000 hours
Θ_0 = Required MTBF = 100 hours
Θ_1 = Goal (Target) MTBF = 150 hours
c = Maximum acceptable number of failures = 15

If we use equations (2) and (3) the estimated risk levels are as follows:

α = Consumer risk = 0.16
β = Producer risk = 0.27

The Excel functions produce the same results as shown below.

α = Consumer risk = POISSON(c, T/ Θ_0, TRUE) = POISSON(15, 20, TRUE) = 0.16
β = Producer risk = 1 – POISSON(c, T/ Θ_1, TRUE) = 1 – POISSON(15, 13.33, TRUE)
= 1 – 0.73 = 0.27

Techniques to Calculate Consumer and Producer Risks for Fixed Configuration Test (Discrete Systems)

Suppose the system under test is a single-shot system. In this case the outcome is either success or failure. For such systems the Binomial Model will be used to calculate system reliability and the associated consumer and producer risks.

Binomial Distribution Equations

The following notation is used:

n = Number of trials
p = Probability of failure for any one trial.
$b_{n,p}$ (k) = Probability of k failures out of n trials.
$B_{n,p}$ (k) = Probability of k or fewer failures out of n trials.

The probability of k or fewer failures can be calculated as follows:

$$B_{n,p}(k) = \sum_{i=0}^{k} b_{n,p}(i) \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (6)$$

where,

$$b_{n,p}(i) = \binom{n}{i} p^i (1-p)^{n-i} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (7)$$

The plan for a binomial test will be completely determined if the following quantities are known.

P_0 = Maximum acceptable proportion of failures - Requirement.
P_1 = Desired proportion of failures – Goal (Target).
α = Consumer risk
β = Producer risk

The test plan consists of a sample size n and the maximum number of acceptable failures c.
The test plan will be completely determined if the consumer and producer risks are known. It is usually not possible to construct a plan which attains the exact values of the consumer and producer risks. The exact procedure to determine test plans for the four values listed above involves the following equations:

$$\alpha = \text{Prob}(\text{ac} \mid p=p_0) = \text{Prob}(\text{c or fewer failures} \mid p=p_0) = \sum_{k=0}^{c} \binom{n}{k} {p_0}^k (1-p_0)^{n-k} \quad \ldots\ldots \quad (8)$$

$$\beta = \text{Prob}(\text{rej} \mid p=p_1) = \text{Prob}(\text{more than c failures} \mid p=p_1) = \sum_{k=c+1}^{n} \binom{n}{k} {p_1}^k (1-p_1)^{n-k} \quad \ldots\ldots \quad (9)$$

Where

Prob(ac | p) = Probability of accepting the system assuming the true proportion of failures is p.
Prob(rej | p) = Probability of rejecting the system assuming the true proportion of failures is p.

The pair of equations (8) and (9) involves lengthy computations, especially for large n, to find values of α and β for specified values of n and c. On the other hand, if values of α and β are known then n and c can be determined from equations (8) and (9). However, it may be easier to use the alternate procedure (shown below) for large sample size n.

Alternate Procedure to Determine n and c for Specified α and β

For large sample size n the Normal Distribution provides a good approximation to Binomial Distribution. When the two risks α and β are known and when values of p are in the range $(0.1 \leq p \leq 0.9)$ n and c can be determined as follows:

$$n = \frac{(z_\alpha^2)(p_0 - p_0^2) + (z_\beta^2)(p_1 - p_1^2) + 2(z_\alpha)(z_\beta)\sqrt{p_0 p_1 (1-p_0)(1-p_1)}}{(p_1 - p_0)^2} \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (10)$$

$$c = (z_\beta)\left(\sqrt{np_1(1-p_1)}\right) + np_1 - 0.5 \quad \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \quad (11)$$

When p_0 and p_1 are very small, that is, less than 0.05, this procedure is not recommended (Ref. DoD Test and Evaluation of System Reliability, Availability and Maintainability, A Primer, March 1982).

To apply this procedure, consider the following example:

Minimum acceptable reliability - Requirement is 0.85.
Contractually specified reliability – Goal (Target) is 0.95.
Consumer risk is α = 0.11.
Producer risk is β = 0.11.

Then p_0 = 1 – 0.85 = 0.15, p_1 = 1 – 0.95 = 0.05, z_α = 1.225 for α = 0.11, z_β = 1.225 for β = 0.11, n = 49.6, c = 3.9

The values of z_α and z_β are obtained from the normal distribution table. The values of n and c are obtained from the formulas (10) and (11).

Since, the values of n and c are expected to be integers we round these values to
n = 50 and c = 4.

Determination of Consumer and Producer Risks for Specified Values of n and c

If the Program Staff already has the sample size n and the maximum acceptable number of failures c based on historical data then the consumer risk α and producer risk β can be obtained from standard binomial tables. For example, if n = 20, c = 3, p_0 = 0.15 and p_1 = 0.05 then α = 0.6477 and β = 1-0.9841 = 0.0159 from the binomial tables.