

SANDIA REPORT

SAND2007-5171

Unlimited Release

Printed August 2007

Product Acceptance Environmental and Destructive Testing for Reliability

Michael A. Dvorack, Elmer W. Collins, and Thomas J. Kerschen

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico 87185 and Livermore, California 94550

Sandia is a multiprogram laboratory operated by Sandia Corporation,
a Lockheed Martin Company, for the United States Department of Energy's
National Nuclear Security Administration under Contract DE-AC04-94AL85000.

Approved for public release; further dissemination unlimited.



Issued by Sandia National Laboratories, operated for the United States Department of Energy by Sandia Corporation.

NOTICE: This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees, nor any of their contractors, subcontractors, or their employees, make any warranty, express or implied, or assume any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represent that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise, does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government, any agency thereof, or any of their contractors or subcontractors. The views and opinions expressed herein do not necessarily state or reflect those of the United States Government, any agency thereof, or any of their contractors.

Printed in the United States of America. This report has been reproduced directly from the best available copy.

Available to DOE and DOE contractors from

U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@adonis.osti.gov
Online ordering: <http://www.osti.gov/bridge>

Available to the public from

U.S. Department of Commerce
National Technical Information Service
5285 Port Royal Rd.
Springfield, VA 22161

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.fedworld.gov
Online order: <http://www.ntis.gov/help/ordermethods.asp?loc=7-4-0#online>



SAND2007-5171
Unlimited Release
Printed August 2007

Product Acceptance Environmental and Destructive Testing for Reliability

Michael A. Dvorack, Elmer W. Collins, and Thomas J. Kerschen
Reliability Assessment and Human Factors
Sandia National Laboratories
P.O. Box 5800
Albuquerque, New Mexico 87185-MS0830

Abstract

To determine whether a component is meeting its reliability requirement during production, acceptance sampling is employed in which selected units coming off the production line are subjected to additional environmental and/or destructive tests that are within the normal environment space to which the component is expected to be exposed throughout its life in the Stockpile. This report describes what these tests are and how they are scored for reliability purposes. The roles of screens, Engineering Use Only tests, and next assembly product acceptance testing are also discussed, along with both the advantages and disadvantages of environmental and destructive testing.

ACKNOWLEDGMENTS

The authors thank the Manager of the Reliability Assessment and Human Factors Department, Kathleen V. Diegert, as well as the Manager of the Reliability and Electrical Systems Department, Rena M. Zurn, for their review comments and clarifying points, which aided in providing a clear description of the topics discussed in this report. The authors also thank John F. Nagel, Jr., Manager of the W76-1/Mk4A Qualification Department, and Ken Pierce, of the Independent Surveillance Assessment & Statistics Department, for providing technical review of this document.

CONTENTS

1. Introduction.....	7
2. Environmental Testing.....	8
2.1 Definition of Environmental Tests.....	8
2.2 Purpose of E-Tests	8
2.3 Determination of E-Tests.....	8
2.4 How Reliability Uses Component and Next Assembly E-Tests.....	9
2.4.1 Component E-Tests During Product Acceptance.....	9
2.4.2 Component E-Tests During Next Assembly Product Acceptance.....	9
3. Destructive tests.....	11
3.1 Definition of Destructive Tests.....	11
3.2 Purpose of D-Tests.....	11
3.3 Determination of D-Tests	12
3.4 How Reliability Uses Component and Next Assembly D-Tests	12
3.4.1 Component D-Tests During Product Acceptance	12
3.4.2 Component D-Tests During Next Assembly Product Acceptance	12
4. Other Tests and Evaluations	14
4.1 Hostile Environment Testing	14
4.2 Destructive vs. Engineering Use Only Tests	14
4.3 Screens	14
4.3.1 HALT and HASS vs. E- and D-Testing	14
5. Multiple Exposures	16
6. Determination of the Number of E- and D-Test Samples.....	17
7. Advantages and disadvantages of E- and D-testing.....	19
7.1 Advantages of E- and D-Testing.....	19
7.2 Limitations of E- and D-Testing.....	19
8. Summary	20
9. References.....	21
Distribution	22

NOMENCLATURE

DA	Design Agency
dB	decibel
D-Test	Destructive Test
E-Test	Environmental Test
EUO	Engineering Use Only
HALT	Highly Accelerated Life Test
HASS	Highly Accelerated Stress Screen
MC	Military Characteristic
PPI	Process Prove-in
STS	Stockpile-to-Target Sequence
UCL	upper confidence limit
WR	War Reserve

1. INTRODUCTION

Nuclear weapon reliability is defined to be the probability of achieving yield at the target across the normal Stockpile-to-Target Sequence (STS) environments throughout the weapon lifetime. As indicated in the definition, this probability must be estimated across all required normal environments and for the entire stockpile life of the weapon system. Concerns with system performance at end of life and in conjunction with normal environmental extremes arise from these considerations. The reliability engineer evaluates test conditions and test data according to whether they can be used to support this estimate. Thus, the basic question for the reliability engineer is how to relate test data and performance measurements to the expected performance of the weapon over the target. There is never sufficient data to fully characterize weapon performance across the entire test space defined in this manner.

Environmental (E-) and Destructive (D-) tests performed during War Reserve (WR) production are an important source of data that can be used to indicate whether or not the component or subsystem will function properly during and after experiencing normal STS environments.

E-tests and D-tests are intended to support assessment of the product's ability to meet design, process, and manufacturing requirements [1]. These two types of tests have the following objectives:

1. Provide an opportunity to test over the normal environment and usage state space,
2. Detect defects early in production so that they can be fixed, and
3. Provide data on representative hardware for reliability estimation.

Hence, the focus behind these tests is both on improving inherent reliability and obtaining data necessary for providing a reliability estimate for the component and, if necessary, identifying issues that may require production or design changes to improve the inherent reliability. E-testing and D-testing allow identification of design and production process issues that are apparent only when the subsystem or component is exposed to STS environments. Continued test success on randomly sampled units throughout the production run strengthens the assertion (beyond that provided by 100% testing) that the production process remains under control [2].

E-tests and D-tests provide the Reliability Engineer with the hard data required to either support or refute the original component event unreliability prediction(s). The reliability engineer will judge the applicability of every E- and D-test that the component experiences on how well completion of the test answers the following question [2]:

If a component selected for testing were instead put into a weapon, would it have worked in that weapon at any time during its stockpile life under any valid normal STS operating conditions for that component?

The following sections provide separate discussions that specifically relate to E-tests and D-tests, respectively, and their role in providing a reliability estimate of a component during WR production and acceptance. This report also describes hostile environments and screens, and discusses both the advantages and limitations of E- and D-testing.

2. ENVIRONMENTAL TESTING

2.1 Definition of Environmental Tests

Environmental (or E-) tests are *nondegrading* tests [1, 3] that evaluate the functionality of a component either during or after the application of one or more *normal* environments (defined below). The normal environments defined for E-testing represent the *range* of those (normal) environments that may be encountered during stockpile life, as opposed to ambient conditions only. E-tests are *not* a screen and are performed on units selected by means of a product acceptance sampling algorithm (or sampling plan). The E-test results help determine the final acceptance of the product, and E-tested units are returned to the production flow and used in the next assembly [3]. E-tests are meant to demonstrate environmental capability but not to degrade the ability of the device to experience the full normal STS environments and function properly.

A *normal* environment is an expected logistical storage and operational environment, as defined in the weapon's STS and Military Characteristics (MCs), which the weapon is required to experience without degradation in operational reliability [4].

Only multi-shot components and subassemblies (i.e., components and subassemblies that can be functioned multiple times) are candidates for an E-test. The E-test must not degrade the component to the point that it is unable, from a design standpoint, to survive a life in the stockpile at the completion of the test. Demonstration that any performance degradation from the E-test is within an acceptable limit must occur during the development and qualification program [2].

An E-test does not necessarily have to represent an entire life in the stockpile. Rather, it should provide sufficient environmental exposure to reveal any defect in the component that would have resulted in a stockpile use failure [2].

2.2 Purpose of E-Tests

E-tests are performed to monitor the quality and reliability of a component during WR production without incurring the cost of scrapping tested units. Selected units that undergo and successfully pass the E-tests can be yielded to either the next assembly or to WR stores, since no significant degradation of the units has occurred as a result of the testing. They can also be designated to undergo subsequent D-tests.

E-test results generally contribute to the data base used by the Reliability organization for component assessment, determining whether the product is meeting its reliability requirement as it is produced.

2.3 Determination of E-Tests

As stated above, E-test environments represent the *range* of those normal environments that may be encountered during stockpile life in which there is no degradation of the component. Non-degrading environments that represent normal stockpile environments and are, therefore,

applicable for E-tests are determined from development and qualification testing, as well as Process Prove-in (PPI) [2].

An example of an E-test suite may be thermal cycling and random vibration tests at levels below the normal environment extremes, such as at those environmental levels that are, say, 3dB below the normal environment maximums. Such tests might be employed because the E-tests cannot be performed at the normal environment extremes without degradation. Thus, a compromise is made by coming down 3dB from these extremes, although reliance on tests at these levels introduces some uncertainty into the reliability estimates. [Note: 3 dB below the maximum level means the following mathematically — $(10)(\log_{10} P_1/P_0) = -3$, resulting in $P_1 = \frac{1}{2} P_0$, where P_1 represents the test and environment and P_0 represents the normal environment maximum.] Another example of an E-test suite is that used by the radar design groups whereby a suite of E-tests is defined such that the unit can survive undergoing the suite three times without degradation. For this particular program, it was judged that this suite of E-tests was acceptable.

For a reliability assessment of a component or subsystem, the ideal situation would be to have complete characterization across the normal STS environments. However, the reality is that it is not always possible to test at the normal environment extremes and still accept product for WR use. Thus, cost considerations drive the compromises made in developing the E-test suite. Because of the necessity of these compromises, there are no hard and fast rules regarding how E-tests are defined, other than that they do demonstrate environmental capability to the greatest degree possible while still not degrading the expected component or subsystem performance.

2.4 How Reliability Uses Component and Next Assembly E-Tests

2.4.1 Component E-Tests During Product Acceptance

During component WR production, a unit that is selected for E-testing undergoes a series of tests representing selected normal thermal and mechanical environments. ***This entire suite of E-tests constitutes one E-test for scoring reliability failure events.*** While this suite is indicative of device performance in all environments to which it is subjected, it is only one device and represents a selected range, among many, of the expected normal logistical and operational environments to which the component may be exposed. As stated above, the E-test results contribute to the data base used by the Reliability organization for assessing the component to determine whether the product is meeting its reliability requirement as it is produced.

As a matter of practical compromise between the number of parts and the amount of testing that can be reasonably performed, component E-tests are not always carried out at the lowest level, such as at the resistor or transistor level.

2.4.2 Component E-Tests During Next Assembly Product Acceptance

Next assembly E-testing may be performed in which the component is exercised with respect to its WR function. Such testing may provide the opportunity to more completely test the unit with proper interfaces and can, therefore, provide additional data for the reliability assessment of the component. On the other hand, there has not been a known situation in which a component

required next assembly testing to demonstrate that it met its reliability requirement. The E-tests (and D-tests) for a component that are called out in the component's product specification (PS) are used to help provide evidence as to whether the component is not meeting its reliability allocation on its own without requiring next assembly product acceptance testing as part of this demonstration. If the reliability requirement is too stringent, then the sample sizes used in E- and D-tests will be too small to demonstrate that the requirement is being met.

It has been suggested that savings could be realized by relying solely on the use of E-testing at the next assembly level, rather than at the component level, to demonstrate that component requirements are being met. This is not recommended for the following reasons:

1. E-testing at the next assembly level, rather than at the component level, means that it will take longer to find any problems. If a problem exists, then there is the potential that the entire next assembly will have to be discarded, rather than just the component in question. Also, more components would have been produced before discovering the problem. Rework and scrap are more expensive at the next assembly level rather than at the component level; thus, any savings resulting from reduced testing could be lost when problems are encountered.
2. Although potential interface issues may be identified at the next assembly level, the ability to make certain measurements on the component will be compromised.
3. The higher the level of assembly, the fewer number of tests available for reliability scoring at the component or subsystem level, since the sampling rate at the next assembly level is based on the lower reliability requirement at that level. Hence, the number of tests may be insufficient to demonstrate component performance with respect to the reliability failure event description under consideration. Because testing is imperfect, one test cannot be depended upon to detect all possible defects.

3. DESTRUCTIVE TESTS

3.1 Definition of Destructive Tests

D-tests are *destructive* or *degrading* tests (i.e., resulting in either destroying the component or losing design margin) that evaluate the functionality of a component either during or after the application of one or more *normal* environments. The normal environments defined for D-testing represent the *range* of those (normal) environments that may be encountered during stockpile life, as opposed to ambient conditions only. However, the particular D-tests represent (a) the normal environment extremes, (b) an increased exposure to selected normal environments, such as thermal cycling in which a particular D-test may consist of three times the E-test thermal cycling, or (c) an increased number of component functional WR operations, such as stronglink reset/enable operations, in a given normal environment or set of normal environments. D-tests may be an extension of the E-tests, in which a fraction of the samples that have successfully undergone E-testing is selected for D-testing, in accordance with a product acceptance sampling plan or algorithm [1]. In addition, any functional testing of one-shot devices (e.g. detonators, valves, ferroelectric firing sets, etc.) as part of production sampling are, by definition, D-tests.

D-tests are sometimes designed to represent all environments that a unit could be exposed to during the normal life of the component. D-test environment levels must not exceed normal environment levels, so there is no question of overtest-induced failures. However, since the most stressing environments for the component in question may not be known a priori, stressing across the entire normal STS environmental spectrum will therefore include the worst-case environments. Over time, these worst-case environments may become known as a result of fallouts.

As in the case of E-tests, D-tests are *not* a screen, and the D-test results can help determine the final acceptance of the product.

3.2 Purpose of D-Tests

D-tests are performed to monitor the quality and reliability of a component during the course of WR production. The units that undergo D-testing cannot yield to the next assembly or to WR stores. After testing, D-test units are occasionally disassembled to assess potential degradation. Thus, D-testing not only provides a demonstration of product life, but also provides a check on the manufacturing process.

D-test results contribute to the data base used by the Reliability organization for assessing the component reliability failure events to determine whether the product is meeting its reliability requirement as it is produced.

3.3 Determination of D-Tests

As stated above in the case for E-tests, D-tests also represent the *range* of those normal environments that may be encountered during stockpile life; however, these are degrading tests that can be, at the normal environment extremes, because of either increased exposure to selected environments or increased WR functional operations in selected normal environments. As with the E-tests, applicable D-tests are determined from development testing, as well as Process Prove-in (PPI). The goal of a D-test is to represent a life in the stockpile with functionality at the extreme(s) of the normal delivery environments. In addition, combined environments used in the D-test must be compatible; that is, the test levels must represent reality. For example, a flight vibration environment will not be seen at an STS low-temperature storage extreme.

An example of a D-test suite may be thermal cycling, followed by mechanical shock and random vibration tests at maximum normal STS environment levels. As stated earlier, the unit may be taken apart upon completion of the tests and evaluated for degradation.

3.4 How Reliability Uses Component and Next Assembly D-Tests

3.4.1 Component D-Tests During Product Acceptance

During component WR production, a unit selected for D-testing undergoes a series of D-tests representing selected normal thermal and mechanical environments. ***This entire suite constitutes one D-test for reliability scoring purposes***, since it is a single sample and represents a selected range, among many, of the expected normal logistical and operational environments to which the component may be exposed. As stated above, the D-test results contribute to the data base used by the Reliability organization as evidence whether the component is not meeting its reliability requirement as it is being produced.

3.4.2 Component D-Tests During Next Assembly Product Acceptance

Next assembly D-testing may be performed in which the component, which has not undergone any prior D-testing, is exercised with respect to its WR function. Such testing may provide the opportunity to more completely test the unit and can, therefore, provide additional data for the reliability assessment of the component. On the other hand, *as stated above for the E-tests, there has not been a known situation where assessment of a component required next assembly testing to accumulate sufficient data to demonstrate that it met its reliability requirement*. Both the E-tests and D-tests for a component that are called out in the component's product specification (PS) are used to assess component reliability.

It has been suggested that savings could be realized by relying solely on the use of D-testing at the next assembly level, rather than at the component level, to demonstrate component performance. This is not a good idea for the following reasons:

1. D-testing at the next assembly level, rather than at the component level, means that it will take longer to find any component-specific problems.

2. There may not be enough D-test samples at the next assembly level to help demonstrate reliability because of the cost of the next assembly unit vs. that for the component.
3. Although potential interface issues may be identified at the next assembly level, the ability to make component-level measurements is compromised.
4. The higher the level of assembly, the fewer number of tests available for reliability scoring. Hence, the number of tests may be insufficient to demonstrate performance with respect to the reliability failure event description under consideration. Because testing is imperfect, one test cannot be depended upon to detect all possible defects.
5. Rework and scrap at the next assembly level is more expensive than at the component level. Thus, any savings resulting from reduced testing could be lost if any problems are encountered.

4. OTHER TESTS AND EVALUATIONS

4.1 Hostile Environment Testing

In hostile environments, the unit is expected to survive and function “without severe degradation in reliability.” If there is no numerical requirement associated with this definition, hostile environment tests cannot be scored directly for reliability purposes, but may provide insight regarding the reliability of the component. Hostile environment testing is usually part of what is typically referred to as Engineering Use Only (EUO) tests.

4.2 Destructive vs. Engineering Use Only Tests

D-tests are not to be confused with EUO tests, which are used to determine the product design margin. Any EUO tests that are planned must be done at the end of D-testing for the following reason: If an EUO test is performed before D-testing and a product failure occurs during the subsequent D-tests, it cannot be determined whether the failure was because of the D-test(s) or the EUO test.

4.3 Screens

Screens, such as 100% tests, provide a means of assuring product quality and the stability of the production processes by weeding out infant mortality failures. However, they are generally not considered useful for reliability scoring purposes, since failures are normally rejected and only successful units accepted. Thus, the test results characterize the incoming population *to* the screen, but not the out-going population *from* the screen. Furthermore, screens are generally not intended to demonstrate that the unit will work in the use environments.

4.3.1 HALT and HASS vs. E- and D-Testing

Another type of screen is referred to as Highly Accelerated Stress Screens (HASS), which are used in the screening of production items on a 100% basis using stresses substantially higher than those experienced in normal use, including shipping and storage. The accelerated stresses are applied in production to shorten the time to failure of the defective units and, therefore, shorten the corrective action time and the number of units built with the same flaw. HASS may also include stresses that do not occur in expected use if these stresses help locate flaws that would occur in an expected field environment [5]. HASS is generally not advisable unless a comprehensive Highly Accelerated Life Test (HALT) has been performed during the design phase of a product, in which every stimulus of potential value is used under accelerated test conditions to find the weak links in the design and fabrication processes [6]. Without the confidence in product robustness that HALT provides, the acceptable stress levels in production screens must be limited, preventing the large accelerations of flaw precipitation and time compression [7].

The stresses used in HALT and HASS include, but are not restricted to, all-axis simultaneous vibration, high-rate broad-range temperature cycling, power cycling, voltage and frequency

variation, humidity, and any other stresses that may expose design or process problems. No attempt is made to simulate the field environment — one only seeks to find design and process flaws by any means possible. The stresses used generally far exceed the field environments in order to gain time compression; that is, to shorten the time required to find any problem areas. When a weakness is discovered, only the failure mode and mechanism are of importance; the relation of the stress used to the field environment is of no consequence [8].

HASS results are not used for reliability scoring purposes for the following reasons [9]:

1. *HASS does not appear to offer opportunities to collect key data for reliability estimation purposes.* HASS is defined to be a 100% screen. Since it is a 100% screen whereby faulty or weak manufacturing product is removed prior to shipment, the data from these screened units cannot be used for estimating reliability because they are not representative of the performance of material in the stockpile. E- and D-tests are an important source of reliability data, and HASS, as it is strictly defined, does not replace it.
2. *HASS may stimulate defects that are unique to non-representative environments.* Because Design Agency (DA) components and systems are assessed over the intended stockpile life and the range of normal environments, it is not clear how HASS results could be used in performing a reliability assessment, since there would be no production test data pertaining to normal STS environments. Any observed defects would require investigation to ensure that they were not the result of an overtest. This is also a concern for E- and D-tests, but to a lesser degree because of the more severe HASS environments.
3. *HASS may induce product degradation because of exposure to environments beyond the normal STS.* HASS employs stresses that go beyond the normal field environments (i.e., beyond the normal STS environments). Hence, HASS may employ stresses considered to be outside of normal environment requirements in which operational reliability is not expected to be maintained, thereby potentially degrading or destroying the unit. Although “Safety of HASS” testing is performed to mitigate this risk, it is generally done on only one unit. Again, this is a challenge also faced in the E-test program, but HASS environments will likely be more challenging in this regard.
4. *Many Sandia products are one-shot devices that do not lend themselves to HASS techniques.* HASS is geared towards products used repeatedly, as opposed to some DA hardware used once (one-shot devices).
5. *The ability of HASS to emulate the effectiveness of current testing vis-à-vis defect detection must be proven.* E- and D-tests fulfill an important role in the overall testing program for DA components. They allow for the examination of product performance in, or following exposure to, various normal environmental extremes, thereby providing an important opportunity to detect defects that would otherwise go unnoticed. HASS would have to be carefully designed to be able to perform the same function as the E- and D-tests. In addition, E- and D-tests can identify defects (e.g., the presence of particles) in the product at an early stage; it would have to be demonstrated that HASS can also perform this function.

5. MULTIPLE EXPOSURES

How multiple tests (or exposures) are scored for reliability assessments is generally a gray area that is component-dependent. Decisions in this area are usually made as failures are found, rather than a priori.

One approach to this issue is to take the position that units undergoing multiple (or cascading) tests are not to be counted multiple times; that is, a unit that undergoes two E-test suites, for example, is not counted the same as two units undergoing one E-test suite each. Also, if an E-test unit goes on to D-testing, it is counted as only one test (i.e., the E-test suite is considered to be a part of the D-test suite for the unit). An example of the first situation would be where a unit undergoes E-testing at the component level, followed by E-testing at the next assembly level. In this situation, the next assembly E-testing can be considered a continuation of the component-level E-testing. The value of the next assembly E-tests is whether damage to the unit was incurred from insertion into the next assembly, as well as whether or not the unit works under next assembly conditions. The basis for this approach is that the next assembly E-test can be thought of as a continuation of the component-level E-test if different environments, or different environmental levels, are used. If a failure occurs, the ensuing investigation will determine what data are used to evaluate the probability of occurrence of the failure, as well as why the unit passed at the component level but failed at the next assembly level.

The alternate approach to this issue is to take the position that cascading tests provide more opportunities for component failure; therefore, they should be counted individually rather than as only one suite of tests. As stated earlier, this is a component-specific decision.

Another point to remember is that cascading E-tests (i.e., E-tests at succeeding levels of assembly) may have the potential of degrading a component by the time that it is yielded to the system. Therefore, testing environments for higher and lower level assemblies must be planned carefully to ensure agreement and that unacceptable degradation will not occur [2].

6. DETERMINATION OF THE NUMBER OF E- AND D-TEST SAMPLES

During development, a reliability allocation, or requirement, is assigned to a component that supports an overall system reliability allocation. A reliability prediction, on the other hand, is based on inference or extrapolation from available data on similar designs. The allocation is used only if there is nothing else to use for reliability purposes; otherwise, the prediction is used. Note that the prediction might not be the same value as the allocation.

As indicated in the Introduction, the reliability engineer attempts to make the best estimate possible of the probability that the hardware will function properly in use conditions. For a Bernoulli variable (i.e., pass/fail data), the point estimate is the maximum likelihood estimate. For reliability failure events having a low probability of occurrence, a considerable number of tests may be conducted without observing a failure. Thus, the point estimate of the failure probability, p , will be zero. An assessment of a zero failure rate may be accurate, but it could also be data-limited. Standard practice at Sandia National Laboratories is to use the initial prediction until sufficient data are accumulated to support a different assessment. This could occur in one of two ways:

1. A failure occurs, in which case the point estimate is used for the reliability assessment, or
2. Sufficient data without failure are collected to indicate that the initial prediction was overly conservative.

In the second case, sufficient data are generally taken to be that number of tests such that if the true reliability were any worse than the prediction, then there would be at least a 50% chance of having seen one or more failures, that is, the 50% confidence bound. Because of the discrete nature of the binomial distribution, this value is called the 50% upper confidence limit (UCL) for p [3].

The number of samples needed to demonstrate the required reliability failure probability, without failures, at the 50% upper one-sided binomial confidence limit, is derived as follows:

The one-sided upper binomial confidence limit on the failure probability, p , is the value that satisfies the following equation:

$$1 - \gamma = \sum_{x=0}^c \binom{n}{x} p^x q^{n-x}$$

where γ = confidence level, p = failure probability, c = no. of failures, n = no. of samples, and $q = 1-p$.

The 50% upper confidence limit on the failure probability, for zero failures, is, therefore, using the above equation, where $\gamma = 0.5$, $c = 0$:

$$1 - 0.5 = \sum_{x=0}^0 \binom{n}{x} p^x q^{n-x}$$

which reduces to:

$$0.5 = \binom{n}{0} p^0 q^{n-0} = (1)(1)q^n = (1-p)^n$$

The value for p that satisfies the above equation is the 50% UCL for p. Now take the natural log on both sides:

$$\ln(0.5) = \ln(1-p)^n = n[\ln(1-p)]$$

For small values of p, $\ln(1-p) \approx -p$. Also, $\ln(0.5) = -0.693 \approx -0.7$.

Substituting yields:

$$\ln(0.5) \approx -0.7 = n(-p)$$

$$\text{Then } n = 0.7/p$$

This last equation provides the required number of both E- and D-test samples to demonstrate p (or q) at the 50% upper confidence limit, assuming zero failures.

Thus, for a reliability requirement of 0.997, this means that $p = 0.003$ and therefore $n = 0.7/0.003 \approx 233$ units are needed to demonstrate a reliability of 0.997 at the 50% upper confidence limit, assuming zero failures. Note that $(0.997)^{233.3} = 0.496 \approx 0.5$.

Note that an exact solution would be $\ln(0.5) = n[\ln(q)] = n[\ln(.997)]$.

Then we get $n = [\ln(0.5)]/[\ln(.997)] = -0.693/-0.003 = 231$. Note that $(.997)^{231} = 0.49955 \approx 0.5$.

It should be pointed out that, although the Binomial distribution is truly correct when the sample is drawn randomly from an infinite population, it provides a useful approximation in most cases [3].

When there are sufficient test data to show that the assessed reliability is better than that of the prediction, the assessed value is then reported. Furthermore, if there are no failures in the test data, then the 50% upper confidence limit for the failure probability is usually reported if it is better than the predicted value.

7. ADVANTAGES AND DISADVANTAGES OF E- AND D-TESTING

7.1 Advantages of E- and D-Testing

It is perhaps helpful to discuss first the role played by E- and D-testing during production. This testing is one element of an entire weapon life cycle evaluation process. The ultimate objective of this entire suite of tests is to provide the opportunity to detect all classes of defects that may be present in the product. Because reliability is assessed over a range of operational capabilities and environments, and across the lifetime of the weapon, it is desired to test in such a way as to explore this state space efficiently. Note that statistically significant quantities of testing at any one environment are rarely obtained, even across the life of the program. Testing during production offers the opportunity to examine product performance at the component level, where more monitoring points tend to be available and environmental exposure is easier. One does not generally have the opportunity during production to examine synergisms between components or interface issues—these are the focus of the system-level testing done as part of the Stockpile Evaluation Program (although some limited information with respect to the interfaces is obtained during production tests at the next assembly level). Testing during production also provides an early opportunity to fix design or production problems that are detected through the tests; this helps to improve the inherent reliability of the product. Finally, production testing is a source of data to support reliability estimation.

E-testing and D-testing have been designed to help achieve these objectives and to complement the other elements of life cycle stockpile stewardship by fulfilling an important role in the overall testing program for DA components. They allow for the examination of product performance in, or following exposure to, various realistic STS normal environmental extremes, thereby providing an important opportunity to detect defects that would otherwise go unnoticed, such as those defects (e.g., the presence of particles) that are in the product at an early stage.

7.2 Limitations of E- and D-Testing

No single test program can provide a check on all the important characteristics, such as aging effects, material incompatibilities, and combined environment behavior. Thus, E-testing and D-testing have their limitations, as well, which include the following [3]:

1. D-tests expend hardware, which is expensive.
2. The E- and D-tests may not stimulate defects that arise because of dormant storage or material aging.
3. Synergistic effects may not be detected.
4. Test requirements may exceed use requirements.
5. Not all environments can be tested.
6. Combined environment effects may not be simulated.

8. SUMMARY

E-testing and D-testing provide data to the reliability engineer that is used to help determine whether or not a component is meeting its reliability requirement as it is being produced. D-testing, in particular, not only provides a demonstration of product life, but also provides a check on the manufacturing process.

The topics discussed in this report are intended to provide general guidelines to both the reliability engineer and the component designer. As stated in this report, there are no hard and fast rules regarding the nature of the E- and D-tests, other than that E-tests are not to be degrading to the component, and neither type of test must exceed the normal STS environment extremes. Judgments pertaining to the development of E- and D-tests, as well as how they are to be used for reliability scoring, must be done on a component or subsystem-specific basis.

9. REFERENCES

1. Richard R. Prairie, *Guidelines for Generating Acceptance Sampling Plans for Continuous Production (U)*, SAND88-3185, Sandia National Laboratories, Albuquerque, NM (Unlimited Release), March 1992.
2. J. F. Nagel, *W76-1 Reliability Assurance: Background & Environmental Acceptance Test Plan Guidance (U)*, March 18, 2003.
3. *Nuclear Weapon Reliability Evaluation Methodology Guide (U)*, Second Edition, Sandia National Laboratories, May 2007. Document is marked Official Use Only.
4. *Glossary of Nuclear Weapons Materiel and Related Terms*, Technical Publication 4-1, August 20, 1991.
5. Gregg K. Hobbs, *Accelerated Reliability Engineering: HALT and HASS*, John Wiley & Sons Ltd., Chichester, England, 2000, p. 77.
6. Hobbs, *Ibid.*, pp. 4, 82-83, 86, 100.
7. Hobbs, *Ibid.*, p.8.
8. Hobbs, *Ibid.*, p. 9.
9. Memo to Distribution from M. A. Dvorack, MS 0830 (12335) and R. L Bierbaum, MS 9202 (8205), *Highly Accelerated Stress Screens vs. Environmental and Destructive Tests*, October 28, 2003.

DISTRIBUTION

1	MS0311	Louis S. Weichman	02626
1	MS0319	Robert W. Boney	02613
1	MS0319	Todd N. Hinnerichs	02613
1	MS0328	Clifton E. Briner	02625
1	MS0328	Patrick A. Smith	02625
1	MS0328	James A. Wilder Jr.	02625
1	MS0428	Todd R. Jones	12330
1	MS0447	J. Douglas Mangum	02111
1	MS0479	Dennis R. Helmich	02138
1	MS0481	Steven G. Barnhart	02132
1	MS0481	Randy J. Harrison	02132
1	MS0481	Christopher R. Landry	02132
1	MS0481	Danny L. Thomas	02132
1	MS0481	John F. Nagel, Jr.	02137
1	MS0483	Corey A. Cruz	02112
1	MS0525	John D. McBrayer	01732
1	MS0537	Tedd A. Rohwer	05353
1	MS0537	Richard E. Heintzleman	05353
1	MS0537	Norm Padilla	05353
1	MS0637	Everett L. Saverino	12336
1	MS0638	Nick J. DeReu	12341
1	MS0638	Peter H. Chauvet	12341
1	MS0638	Michael A. Plowman	12341
1	MS0638	James C. Dalton	12342
1	MS0638	Coyte S. Julian	12342
1	MS0829	Janet M. Sjulín	12337
1	MS0830	Kathleen V. Diegert	12335
1	MS0830	Douglas H. Loescher	12335
1	MS0830	Jeffrey A. Mahn	12335
1	MS0830	Michael J. Mundt	12335
1	MS0830	Roger W. Plowman	12335
1	MS0830	Max L. Terchila	12335
1	MS9007	Rena M. Zurn	08205
1	MS9007	Gloria J. Christensen	08205
1	MS9007	Constanzo A. LaJeunesse	08205
1	MS9007	Jennifer E. Robles	08205
1	MS9007	Elizabeth C. Wichman	08205
1	MS9013	Alfredo McDonald	08231
1	MS9013	Ronald D. Sauls	08231
1	MS9014	James O. Harrison	08238
1	MS9014	Raphael M, Molle	08238
1	MS9034	Ming K. Lau	08221
1	MS9154	Robert E. Oetken	08346
2	MS9018	Central Technical Files	09536
2	MS0899	Technical Library	09536