Assessing the Adequacy of Ground Tests

Stephen B. Rogers

Prior to a missile flight test, ground tests are performed to give confidence that the ensuing mission will be successful. It can be difficult, however, to determine if a given set of ground tests is comprehensive enough to cover all mission requirements. It is also important to know if a successful ground test indeed indicates a high probability of a successful flight test. This article describes an objective, formal, systems engineering process developed by APL to assess the comprehensiveness and adequacy of a system-wide ground test program. To illustrate the process, ground tests for the Navy Theater Wide Aegis Lightweight Exo-atmospheric Projectile Intercept Project are assessed.

INTRODUCTION

Missile flight tests, particularly those in the area of Ballistic Missile Defense, are more costly and more politically high profile than ever. In light of these pressures, a comprehensive and effective ground test program prior to each flight test is crucial to minimize risk.

The Navy's Aegis and Standard Missile (SM) programs have a long history of comprehensive ground testing followed by successful anti-air warfare flight tests. In keeping with this legacy of successful testing, myriad ground tests are planned prior to each flight test of the Navy Theater Wide Aegis Lightweight Exo-atmospheric Projectile (LEAP) Intercept (ALI) Project currently under development. The ground testing needs to be reviewed to ensure the adequacy of this new exo-atmospheric, hit-to-kill operation of the Aegis Weapon System against Theater Ballistic Missile (TBM) targets.

It is important to know whether ground testing is comprehensive and covers all functional areas and whether the tests as planned will adequately reflect the required performance. Given that the planned tests are comprehensive and adequate, it is important to know how they are integrated to coherently show that the ground test program as a whole is adequate to ensure mission performance. This article describes an objective, formal, systems engineering process to assess the comprehensiveness and adequacy of ground testing a complex system. Examples show the assessment of the ALI system, which includes an Aegis ship, SM-3 missile, TBM target, and test range.

BACKGROUND

The Navy's ALI Project is intended to demonstrate the ability of the Aegis Combat System, integrated with SM-3 and the Vertical Launch System (VLS), to intercept a TBM target outside the atmosphere. Figure 1 shows the four major elements of the ALI System: ship, missile, target, and test range. The SM-3 is a four-stage missile consisting of a booster (first stage), dual-thrust rocket motor (second stage), third-stage rocket motor
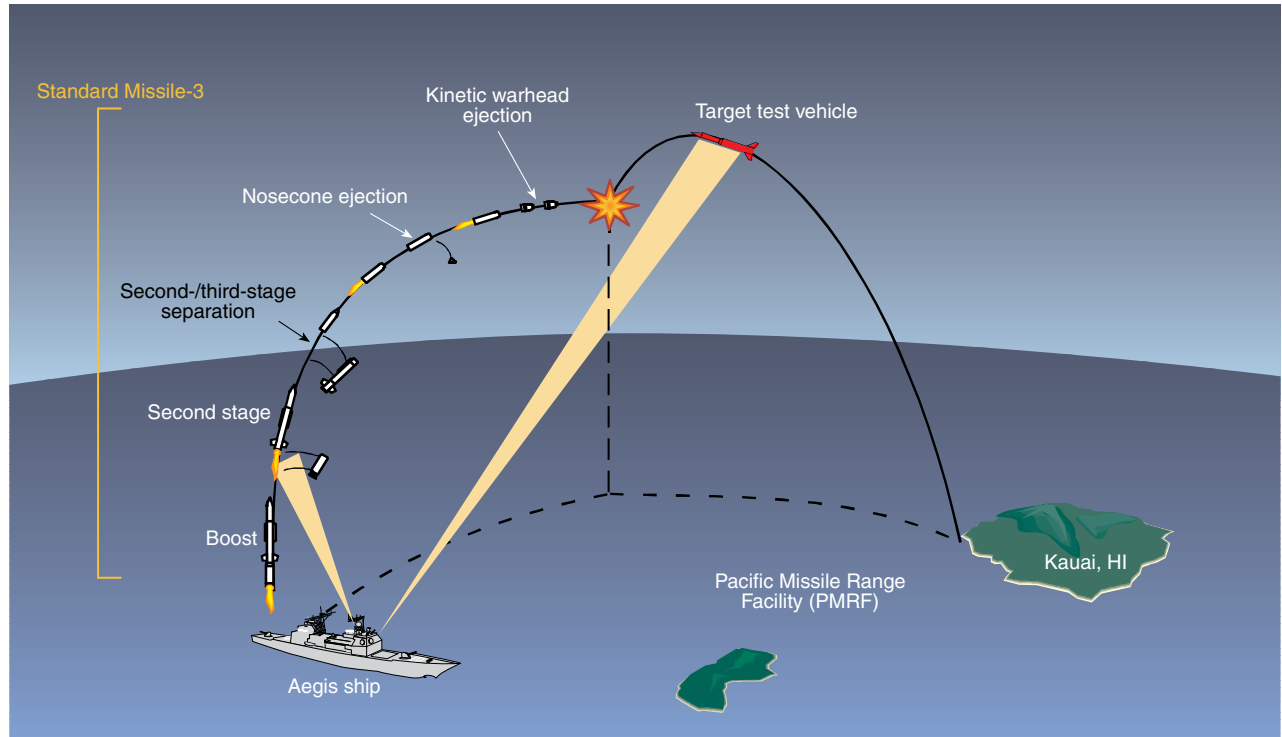
Figure 1. Aegis LEAP Intercept flight example.

(TSRM), and kinetic warhead (KW). The KW is a hit-to-kill warhead that has no explosives onboard but uses kinetic energy from its high velocity to destroy the target on impact. Although it has no propulsion (its velocity is generated by the previous three stages prior to KW ejection), it has the ability to divert its trajectory to create a direct hit on the target.

TEST DOMAINS

To develop a ground test adequacy assessment process, the different test domains must be defined: in-process integration, performance, and environmental testing. In-process integration tests are conducted on items that will be used in a flight test such as the flight missile hardware and software, the firing ship's systems, and the test range assets.

Performance tests demonstrate that the design of an item is adequate to provide the functionality required to meet the mission objectives. In this assessment process, the term "performance" testing refers to testing that has flight-representative hardware and software integrated and working together as they would in flight while being subjected to realistic stimuli. Performance testing is contrasted with operational testing, which merely checks whether a component is operating within tolerances. Operational testing also includes parameter checking such as impedance or voltage measurements. Performance tests are common in design verification testing whereas operational tests are common in in-process

integration and acceptance testing where the ability to test performance is often limited. Examples of ALI performance tests are the live battery test and combat system engineering development site (CSEDS) testing. Examples of operational tests are verifying a check-sum on an embedded computer program or measuring the internal pressure of a coolant bottle.

Environmental tests demonstrate that an item can survive expected operating and nonoperating environments while adequately performing the required functionality to meet mission objectives. Examples of environmental tests are packaging, handling, storage, and transportation tests and electromagnetic environmental effects testing.

As shown in the upper portion of Fig. 2, these domains are not mutually exclusive and, in fact, there should be as much overlap as possible in these testing domains. An ideal test would have the flight unit demonstrate performance while being nondestructively subjected to expected environments. However, owing to practical constraints on flight hardware testing, environments may be present only at reduced levels or not at all. In fact, performance may not be measured or observed directly, but rather only core operating parameters may be measured, which merely indicate that the unit under test is operating as expected. Acceptable values and tolerances for these core operating parameters are determined either by previous performance testing or by analysis.

Because of the constraints on flight hardware testing, SM design verification tests are typically conducted on inert operational missiles (IOMs). IOMs use
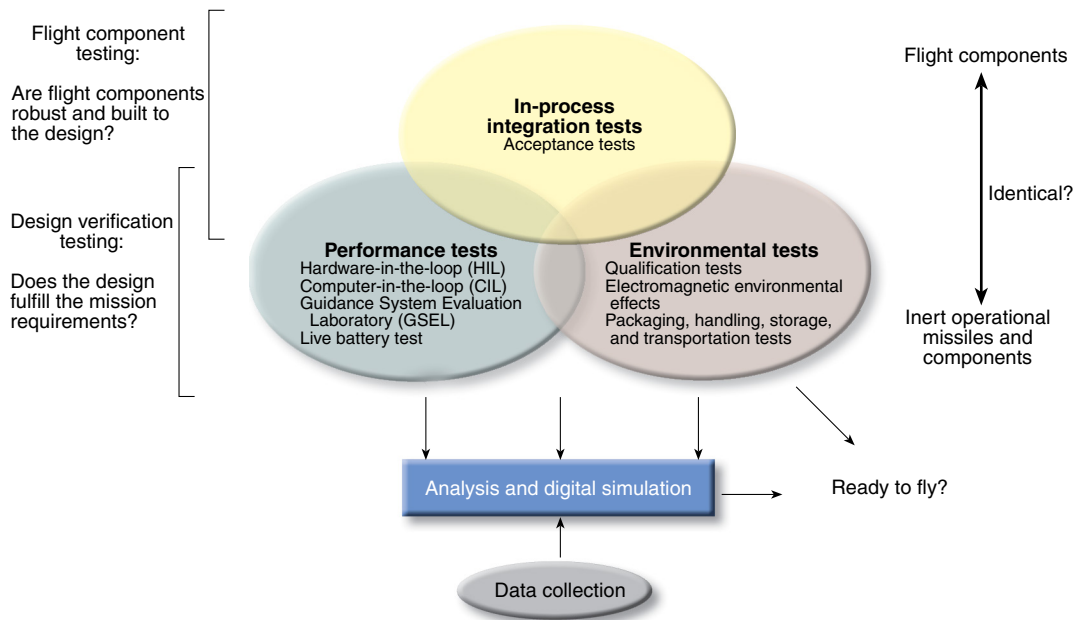
Figure 2. Ground test domains help assess readiness for flight testing.

flight-representative hardware and software, but do not contain live explosives such as warheads and rocket motors. This allows them to be safely tested in a laboratory environment and also avoids the overstressing of flight hardware. In the ALI Project, there are several IOMs of varying degrees of flight representativeness. For a particular mission function that is tested on an IOM but not on the actual flight round, the IOM must be configured identically to the flight round for that function for the results to be relevant.

In addition to the three test domains, Fig. 2 also shows the vital roles that analysis and simulation play in the overall determination of whether a system is ready to proceed to flight test. Although this article focuses on the assessment of ground testing, an area of future work is a rigorous examination of analysis and simulation in the overall assessment of flight test mission risk.

THE ASSESSMENT PROCESS

The process of assessing the adequacy of ground testing of a complex system involves two major steps: (1) a functional decomposition of the system, which is independent of ground testing, and (2) a ground test assessment, which evaluates how well the planned testing addresses the mission functions identified in the functional decomposition.

Functional Decomposition

By its very nature, a complex system is difficult to analyze when taken as a whole. To work with more manageable pieces, functional decomposition is used. This process decomposes a system into its component subsystems and defines a set of critical mission functions that capture all the mission requirements. These functions can then be mapped to the subsystems that are responsible or involved in the performance of each function. This not only allows the analysis to focus on a more manageable problem, it also tends to match how subsystems are actually designed and tested.

Using the ALI System as an example, the following paragraphs describe the six steps of functional decomposition.

1. **Partition the system**

The system is partitioned into its major subsystems, and the subsystem interfaces are identified. Figure 3 illustrates the ALI System partitioned into its major subsystems: Aegis ship, SM-3 missile, target test vehicle (TTV), and test range, including the flight termination system signal relay plane. Human operators involved in the system also are shown in the partitioning. The SM-3 missile is partitioned down to the section level, which corresponds to how it is assembled and tested. The partitioning of nonmissile elements is left at a higher level. To focus on the shipboard systems, it may be possible to partition the SPY radar or the Vertical Launch System (VLS) into relevant subcomponents to assess the testing of those elements.

2. **Partition the mission timeline**

This step is done only to break the mission time line into more manageable pieces (or phases) and is not required if the timeline is short or if there are few mission functions. Because of the sectional nature of SM and the allocation of mission requirements during
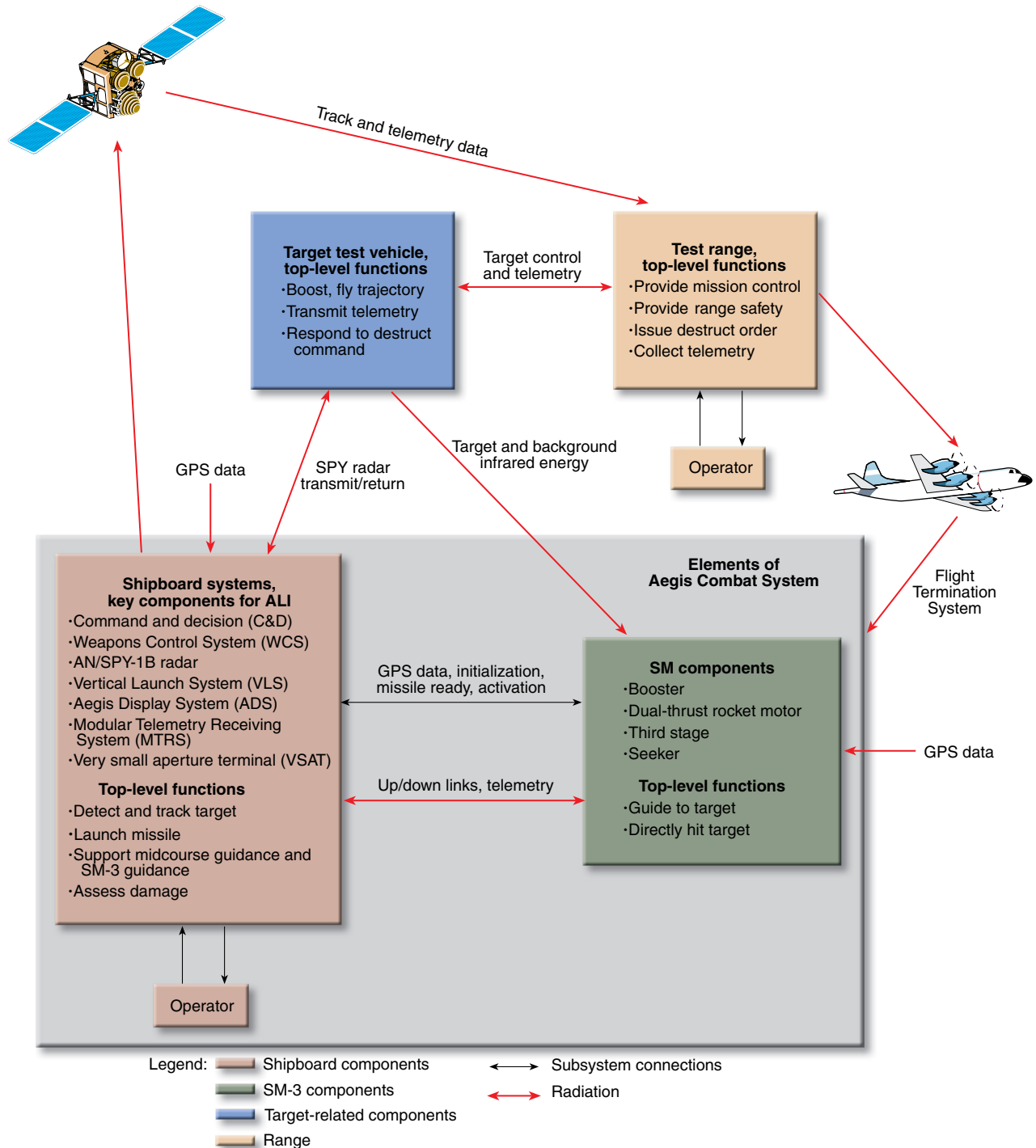
Figure 3. ALI partitioning (GPS = Global Positioning System).

system design, the ALI mission can be readily partitioned into six distinct phases: mission preparation, prelaunch, boost, endo-midcourse, exo-midcourse, and terminal.

3. Map mission functions to phases

The following paragraphs discuss each of the mission timeline phases for ALI and describe which mission functions correspond to each of the phases. In this partitioning, if a mission function is a continuous process covering more than one phase, it appears in the timeline when it is first functioning and is assumed to apply thereafter for its intended duration. For example, generation of missile telemetry begins prior to missile initialization and continues throughout the mission; however, that function is listed only in the prelaunch phase.

The Aegis and SM systems engineering communities have defined 110 mission functions for ALI, not including mission preparation. *Mission preparation* includes such functions as prior mission planning, clearing the test range for safety purposes, and getting the firing ship to the correct location to begin the mission. To date, the mission preparation phase has not been formally addressed, and mission functions for this phase have not been defined. Therefore, the 110 ALI mission functions are allocated among the remaining five mission phases:

- The *prelaunch* phase begins with the launching of the TTV, includes target acquisition and tracking by the Aegis SPY radar and missile selection, and ends with receipt of firing interlock by the VLS from the SM-3 missile. For ALI, 22 mission functions were identified in the prelaunch phase (Table 1). Examples include "Launch TTV," "Activate Missile Batteries," and "Initialize Missile."
- The *boost phase* begins with ignition of the SM-3 booster inside the VLS and ends with separation of the booster from the upper stage. This phase has 25 mission functions such as "Clear VLS," "Acquire Downlink Beacon," and "Separate Booster."
- *Endo-midcourse* begins with commanding dual-thrust rocket motor ignition and ends with Stage 2/3 separation. ALI has 18 endo-midcourse phase mission functions. Examples include "Perform Stage 2 Guidance" and "Actuate Tail Fins."
- *Exo-midcourse* begins with commanding ignition of the first pulse of the TSRM and concludes just prior to ejection of the KW. Exo-midcourse has 18 mission functions such as "Ignite TSRM Pulse 1" and "Eject Nose Cone."
- The *terminal phase* begins with KW calibration and ends with target intercept. This final phase consists of 27 mission functions.

4. Map functions to subsystems

The fourth step in the functional decomposition process is mapping the mission functions to the subsystem elements. This is done with a chart for each mission timeline phase as shown in Table 1. The mission functions for each phase are listed in the left-hand column of the chart, with the subsystem elements (defined in step 1) across the top. For each mission function on the left, an X is placed in each subsystem element column on the right if it is involved with the conduct of that function.

5. Characterize interfaces and control loops

This step captures the interactions between subsystem elements for each mission function. For each function, X's in multiple subsystem columns in the function mapping produced in step 4 (looking horizontally across the row) indicate that the function will involve an interface between the subsystems. To explicitly indicate interfaces and help identify data paths and control loops, it is useful to make a diagram to show which elements of the system are used in a specific function as well as control and data connections. For example, Fig. 4 shows such a diagram for the prelaunch mission function of "Initialize Missile."

6. Identify function observables

The final and most difficult step in functional decomposition is identifying function observables. These are quantities or qualities that can be measured or observed in a test to indicate the proper execution of a function. Identifying a complete list of these necessary elements of correct performance is the key to subsequently conducting a ground test assessment. It is important to generate an objective, independent, and complete list of observables, regardless of whether they are known to be observed in planned tests. It is vital to avoid bias or complacency based on known test capabilities so that holes or inadequacies in the existing test regimen can be uncovered. Table 2 includes a list of the function observables for "Initialize Missile."

Ground Test Assessment

By identifying mission functions and their associated measures of performance (observables), the foundation has been laid for now assessing ground test

Table 1.   Examples of prelaunch function mapping.

Mission function	Ship					Missile				Range		Target	
	Navigation	VSAT/ comms	MTRS	SPY, C&D, WCS, ADS	VLS	Missile stages				PMRF	Relay plane	TTV	Operator
						1	2	3	KW				
Launch TTV										X		X	X
. . .													
Activate Missile Batteries						X	X	X	X				
Initialize Missile	X			X	X	X	X	X					

SPY

MTRS C&D

Navigation → WCS → VLS → | Stage 1 — Stage 2 — Stage 3 | KW

Communications Operator Relay plane PMRF TTV

Operator

Legend:
■ Component used by function
□ Component not used by function
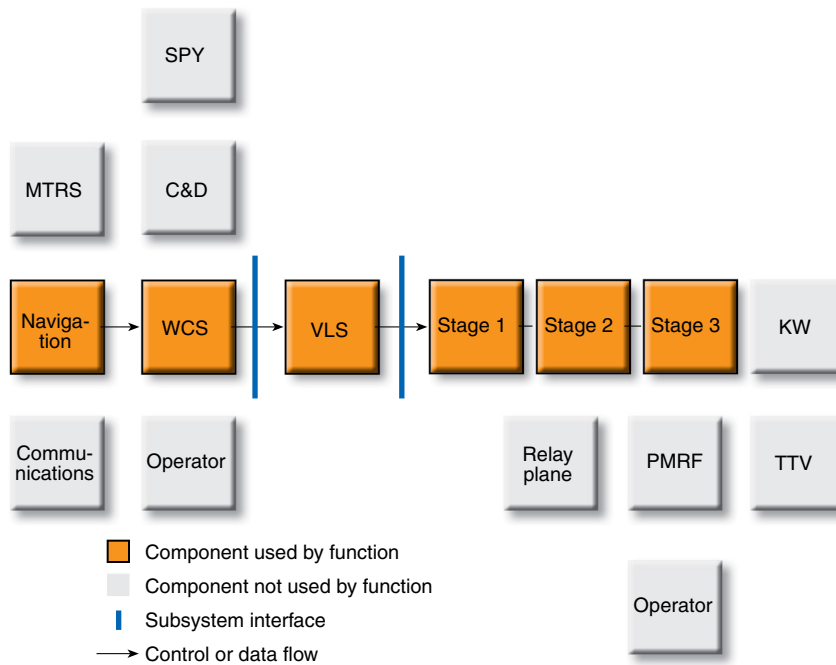| Subsystem interface
→ Control or data flow

Figure 4. Interfaces and control/data flows for the function "Initialize Missile."

adequacy. The ground test assessment is a formal, six-step process designed to show whether there are ground tests for all mission functions and whether each planned test is of high enough fidelity to provide confidence in the subsequent performance of the mission.

Using the ALI System ground testing as an example, and the results of the functional decomposition performed as described in the previous section, the six steps of ground test assessment are as follows:

1. **Acquire and review ground test plans**
To assess the comprehensiveness and adequacy of testing, all planned ground tests must be reviewed. These test plans should define the test configurations, including which

Table 2. "Initialize Missile" test assessment sheet.

Initialize Missile	In-process integration testing						Performance (design verification) testing					
							CSEDS					Flight
Prerequisite functions (none)	Missile check-out (MCO) 3rd stage	MCO upper/ round	WSMR 3rd stage	WSMR round	WSMR MAF VLS	WIT (ship)	WIT	IOM	CIL	HIL	GSEL	experience
Observables — Fidelity (max.=12)	8	8	8	8		12	12	8	8	8	8	12
Visual inspection of interstage connections	X	X	X	X								
Continuity and isolation of electrical path from booster umbilical connector through guidance section (GS)				X								
Continuity and isolation of electrical path from canister connector through GS					X							
Timing from initialization reset to initialize message (IM)	X	X	X	X			X	X	X	X	X	X
Bad IM handling	X	X	X	X				X	X	X		
Correct interpretation of message data (seen on theater missile)	X	X	X	X			X	X	X	X	X	X
Generation of MISSILE READY signal	X	X	X	X			X	X	X	X	X	X
Confirm correct content of IM						X	X	X				X
Timing from IM to MISSILE READY	X	X	X	X			X	X	X	X	X	X
Components												
SM-3	✓	✓	✓	✓			✓	✓	✓	✓	✓	✓
VLS						✓	✓					✓
Ship navigation						✓					✓	
WCS						✓	✓		✓			✓
Interfaces												
Ship navigation to WCS						✓	✓					✓
WCS to VLS						✓	✓		✓			✓
VLS to SM-3						✓	✓					✓
Environmental influences on function												
None												

elements of the system are present in the test and which elements are simulated or emulated. They should also include which parameters are measured in the test and which values are expected or acceptable. Review of the test plans should be augmented as necessary with discussions with system and test engineers to address any questions.

Some of the primary test plans and procedures relevant to assess ALI ground testing are

- CSEDS test plan for the Navy Theater Wide ALI
- Waterfront integration test (WIT) plan
- Acceptance test procedure for the SM-3 guidance section (GS)
- SM-3 live battery test plan
- SM-3 missile check-out (MCO) procedures
- Test range firing day procedures

In addition to these kinds of primary documents, it is necessary to review several assembly procedures and other specific test documents. A comprehensive listing and description of the planned ALI ground testing can be found in the Integrated Ground Test and Evaluation Plan for the ALI Flight Demonstration Program.[1]

2. Map mission functions to planned tests

For each mission function in the timeline, ground tests that exercise that function are identified. For functions that are evaluated in numerous tests, an attempt is made to identify the most thorough tests of the function. Test activities that meet the definition of performance testing given earlier in this article are subsequently evaluated for test fidelity in step 4. The guiding principle in mapping tests to functions is that it is generally best to test functionality at the highest level of integration possible. Because of this and to keep the test assessment process manageable, lower integration level tests are assessed only when necessary. For example, if a missile function is not tested at the round level, section-level tests will need to be examined. If the function is not tested at the section level, it should be tested at the plate level, and so on, but it should always be tested at the highest level possible.

Examples of ALI test activities considered for assessment include

- Round-level MCO at Tucson and WSMR Assembly Facility (MAF)
- Third-stage acceptance tests
- HIL
- CIL
- GSEL
- CSEDS
- WIT

Once mission functions have been mapped to specific tests, a test assessment worksheet as shown in Table 2 can be completed. A worksheet that shows all the testing conducted for a function at the highest levels of system integration is made for each mission function. The list of function observables determined in the last step of functional decomposition is given on the left-hand side of the table. Across the top of the worksheet are the test activities that involve that function. The rest of the worksheet is described in the following paragraphs.

3. Assess test comprehensiveness

After mapping the mission functions to planned tests, one can make a general assessment of ground test comprehensiveness. If there are any mission functions not tested in any planned tests, there is obviously a deficiency in the ground test program. The performance of every mission function should be demonstrated via ground test prior to flight. If it is impossible to test a function except via the actual flight test, this should be noted and covered by analysis or digital simulation. This analysis should also include an assessment of the risk involved in demonstrating a mission function via analysis or simulation instead of by actual test.

For each function observable, X's are placed in the worksheet in each column that has a test that covers that observable. If the test is a performance test, the box is shaded and the performance test fidelity is assessed as described in step 4 below. If only core operational parameters are tested, rather than performance, the box is left unshaded. Note that there are two groups of test activity columns on the worksheet. On the left are in-process tests that are conducted on the actual flight hardware. On the right are design verification tests that are typically conducted on flight-representative IOMs. Performance testing, as defined earlier, can be conducted in either category.

Using the diagram generated in step 5 of functional decomposition (Fig. 4), which depicts all the elements or components and subsystem interfaces involved with a mission function, these components and interfaces are listed down the left-hand side of the worksheet. If a diagram has not been previously generated, a complete list of all the components and interfaces relevant to the function is made. A check mark is placed in the column of each test activity that has a flight-representative version of that component or that exercises that interface in its test.

Now looking horizontally across the row for each observable and for each component, each is assessed as to whether it is adequately represented and tested. This assessment is done as discussed in step 5 by answering core questions.

The worksheet shown in Table 2 focuses on in-process integration and function-specific design verification performance testing. However, it can address environmental effects, which directly influence the performance of a specific function. This worksheet fully

covers the in-process and performance test areas, including those that overlap the environmental test area (Fig. 2). A method for assessing the adequacy of shock and vibration environmental tests that are not function-specific is discussed in step 4.

To include relevant environmental concerns in this functional assessment, all environmental influences that directly affect the specific function are listed. This list is not to include overall environmental conditions that affect the entire missile, such as electromagnetic environmental effects or transportation vibration, but rather function-specific effects such as aerodynamic heating on the nosecone or triband antenna and rocket motor plume attenuation on uplink and downlink signals. In the example for Initialize Missile shown in Table 2, there are no environmental influences of concern.

4. Assess fidelity of individual test activities

The test fidelity of each mission function and each test activity that involves a performance test must be assessed. This is done for each mission function because a specific test activity may be designed for high-fidelity testing of certain functions but have low-fidelity capability for testing other functions.

This assessment comprises a set of questions about the test activity for that mission function. A highest-fidelity test has flight-representative hardware and software, all subsystem interfaces and interactions present, and realistic sensor stimuli and is executed in a real-time sequence with other relevant functions. Starting with a value of 12, a test activity's fidelity rating is decremented when any of these important characteristics is missing. Figure 5 shows a logic diagram of the test fidelity questions and the relative decrement weighting for any negative responses. This process provides for a range of fidelity values from 0 to 12, with 12 being the highest fidelity. Any values below 7 should be considered of insufficient fidelity to adequately assess that mission function.

In the example of Initialize Missile shown in Table 2, the preliminary fidelity assessment of the performance tests is high (fidelity ratings of 8 to 12). The ratings of 12 are for the WITs that have a flight-representative IOM for the missile and use the actual firing ship systems for the initialization. The other test activities received a rating of 8 for missing a relevant component (such as VLS or the navigation system) and for emulating a subsystem interface (such as missile to VLS).

Note the provision for two green "flags" in the rating process (Fig. 5). The partial test flag allows for assessing the fidelity of subsystem tests separately whose combined results adequately cover a function. For example, the function of uplink from the ship to the missile requires
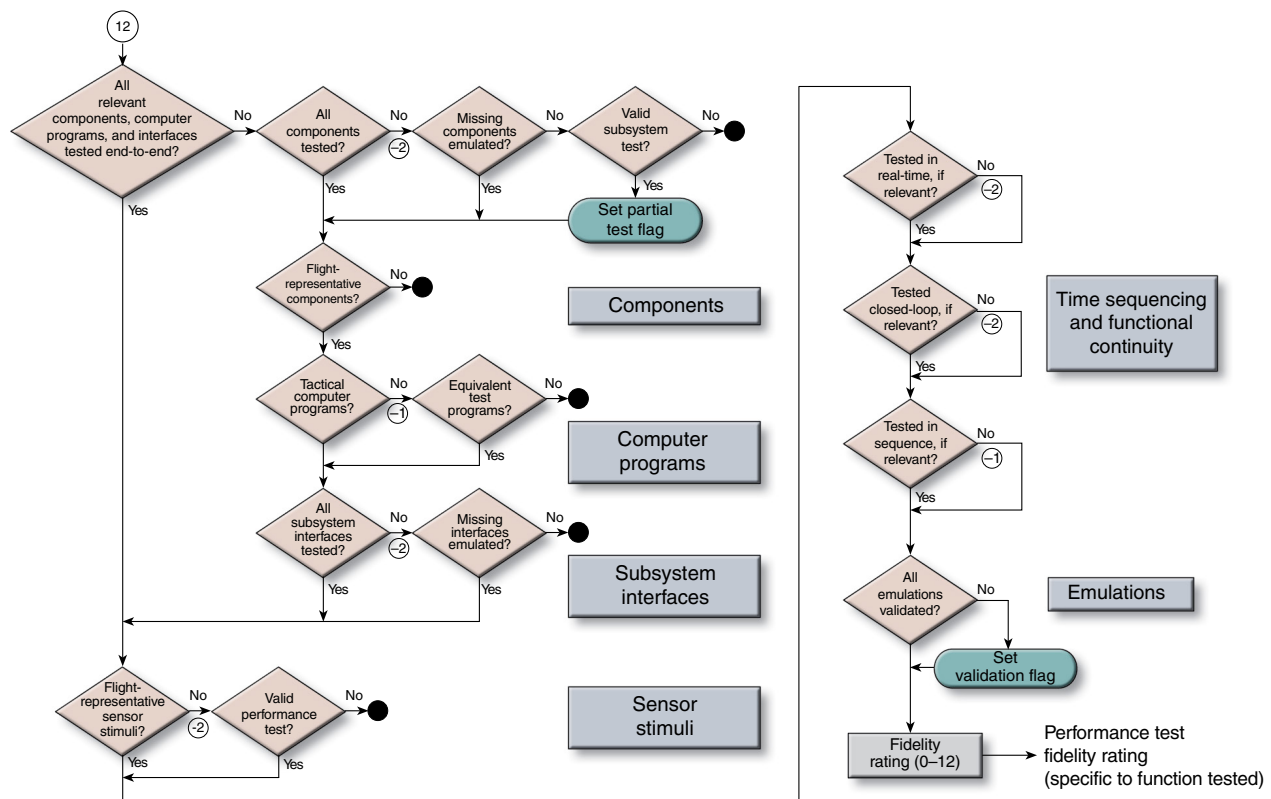

Figure 5. Performance test fidelity rating questions logic diagram (black circles = fidelity inadequate for performance testing).

both subsystems, and exercises an interface between them. There may be a high-fidelity missile system test that covers the function of uplink by use of flight (or flight-representative) missile hardware and software with a validated ship and interface emulation. There may also be a high-fidelity ship system test that covers uplink by use of real ship systems with a validated missile and interface emulation. Although each of these tests by itself is not sufficient to adequately cover the function of uplink, their combined results may be. However, each test's fidelity score will suffer from the emulation of missing components and interfaces and will naturally result in a lower rating than a comparable end-to-end test that exercises all relevant components and interfaces together.

The validation flag is used to indicate the need to validate all emulations used in testing. Although it does not result in a lower score, all validation flags must be removed to have adequate confidence in the test results. Any negative response that leads to a black

circle indicates a fidelity that is inadequate for performance testing.

Environmental test assessment. For assessing shock, vibration, and acoustical environment tests that affect a test article as a whole and not just a specific function, a separate adequacy rating flowchart was developed. Because of its complexity, it was divided into two figures. Figure 6 shows the flow of questions that are common to shock, vibration, and acoustical test adequacy, and Fig. 7 shows additional questions that are specific to each type of testing. These charts were developed to assess missile testing and so begin at the round level; however, other types of test articles can be substituted beginning at the highest level of integration.

Following these flowcharts, one can assess whether each type of environmental testing is adequately covered for a test article. After each of these test areas is assessed, it must be determined whether the flight-representative test article is subjected to the correct sequence of cumulative environments. Testing piecemeal may not provide
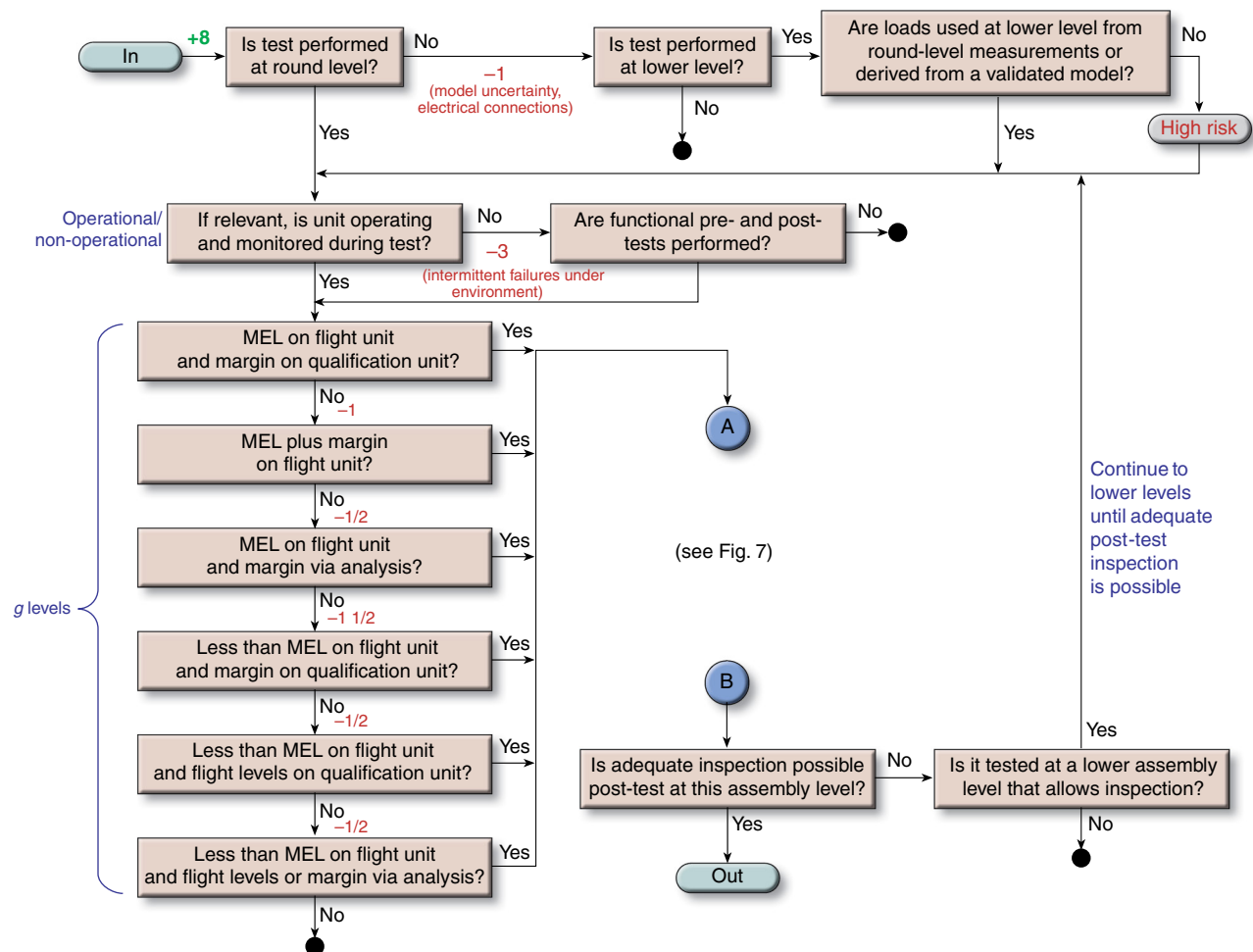

Figure 6. Logic diagram showing fidelity rating questions common to shock, vibration, and acoustical environment tests. (MEL = maximum expected levels; black circles = fidelity inadequate for performance testing; "A" and "B" indicate the position of the logic diagram shown in Fig. 7.)
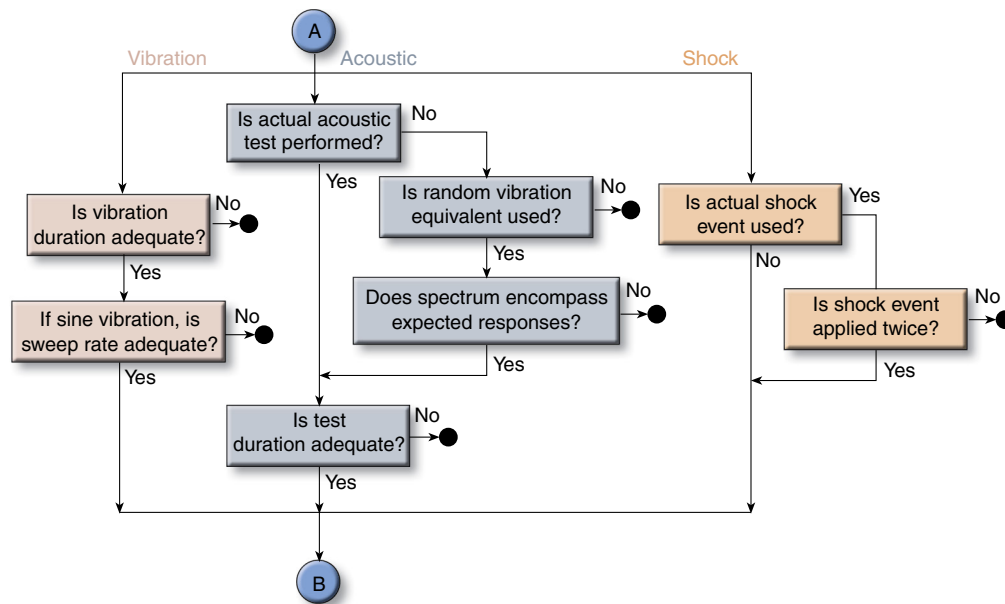
Figure 7. Logic diagram showing additional fidelity rating questions specific to shock, vibration, and acoustical environment tests (black circles = fidelity inadequate for performance testing).

the confidence that a flight test article will be able to survive the cumulative environment it will face in performing its mission.

The philosophy behind the vibration, acoustic, and shock assessment flowcharts is that the test article should be subjected to an environment at the highest level of integration possible that allows adequate posttest inspection and, if relevant, monitoring during the test. That is, if the test article is required to operate when subjected to the expected environment during the flight test mission, then it should be operating during the test and functional parameters should be measured during the test. If a test article cannot be adequately inspected after being subjected to an environment, there is risk that a failure occurred that will not be detected. Likewise, if a test article is expected to be operating under a particular environment (e.g., flight vibration), then it must be operating and monitored during the test so that intermittent failures under that environment can be revealed. A post-environment test under ambient conditions could miss these intermittent failures, which would cause a failure during flight. With this philosophy in mind, the assessment begins (as shown in Fig. 6) with a test effectiveness rating of 8, and this value will be reduced whenever a test effectiveness deficiency is found. This method will provide a final score that ranges from 8 to 0 (highest to lowest test effectiveness, respectively).

In the following explanation of the environmental test assessment questions, it is assumed that missile testing is being assessed. However, any test article such as an Aegis ship, automobile, or even a pogo stick may be substituted for a missile round.

The first question is whether the environmental test is performed at the highest level of integration or assembly. If not, there is a slight penalty because all the interfaces and interconnections are not in place and the loads presented to lower-level test articles will need to be derived from an analytical model that introduces uncertainty. If a test is not performed at the highest level of integration, or any other lower level, then the testing is deficient and is immediately rated as unacceptable (shown as a black circle on the flowchart). If tests are performed at a lower level, input levels should reflect responses measured from flight or round-level testing or from mathematical model predictions. If a math model is used to derive test inputs, it must be validated or else the results are subject to high risk.

Given that a test is performed, the flowchart then presents the question of whether the test article is operating and monitored during the test or if ambient pre- and posttesting is performed instead. If the test article is required to operate in an environment, but is not operating and monitored during the environmental ground test, then a significant penalty is incurred. A penalty is warranted because a failure may only manifest itself intermittently under that environment and not at all under ambient conditions. An inability to detect this condition during ambient post-testing will lead to a high risk of failure during its subsequent flight or mission. If there is no monitoring during the test, nor any pre- and post-testing, then the testing is unacceptable.

The next series of questions relate to the *g* levels or forces to which the test article is subjected. The reasoning applied here is that it is important to demonstrate that each flight article can handle at least the full

expected flight loads and that a margin needs to be demonstrated in some way. The best situation is with full flight loads tested on the flight article with the margin demonstrated on an identical qualification unit. This method provides confidence in the flight unit without overstressing it. The flowchart includes less desirable combinations of test levels with additional penalties accruing down the list. A large penalty occurs if the flight unit is tested to less-than-maximum expected flight levels. This penalty indicates a higher risk of failure for that particular unit in flight. If flight levels are not tested at all, the testing is rated as unacceptable. Implicit in this discussion is that the frequency range of the test input spectrum should encompass the frequency range of the expected levels.

Individual vibration, acoustic, and shock test assessment. Once the test level is determined, the next set of questions depends on the type of testing being assessed. As shown in Fig. 7, there are separate questions for vibration, acoustic, and shock testing. For a vibration test, two additional issues must be addressed. First, the duration of the test must be adequate to address any mechanical fatigue issues. Also, for a sine vibration test (as opposed to random vibration), the sweep rate and frequency range must be adequate to reflect the expected flight environments.

For an acoustic test, it must be determined whether it is performed using actual acoustic energy or whether a random vibration equivalent is used. Failure to perform an acoustic or random vibration test to meet the acoustic environment requirement is unacceptable. If a random vibration equivalent test is performed as a substitute for the acoustic test, the input spectrum for the test should encompass the expected vibration response from the required acoustic environment or else the testing is inadequate. The duration of the acoustic or random vibration test must be representative of flight.

For shock testing, it is significant whether the actual shock mechanism is used in the test or whether the shock spectrum is provided via a shaker or a high-impact shock machine. A shock spectrum requirement usually encompasses all possible responses from actual shock actuation. A single shock actuation will generally provide only a subset of the environment for that one test article. To overcome this, two shock actuations are required if the actual shock mechanism is actuated in the test. If a shaker or high-impact shock machine is used, which provides the full shock spectrum, then only a single test is required.

After questions relating to the test conditions are addressed, it must be determined whether adequate post-test inspection is possible at this assembly level to discover all test-induced failures. If not, then inspection must be performed at a lower assembly level until adequate post-test inspection is possible. If testing is not done at a low enough assembly level to allow adequate post-test inspection, then the testing is rated as unacceptable since there is a risk that a failure occurred during testing that will not be detected until flight.

The use of these test assessment flowcharts, combined with the final question of realistic cumulative environments, allows the assessment of shock, vibration, and acoustical environment testing. Formal procedures to assess other types of environments such as temperature, humidity, and electromagnetic effects have yet to be developed.

5. **Perform overall ground test assessment**

The overall ground test assessment can now be made by answering a series of core questions relating to how well the function is being tested across all test activities. As mentioned previously, environmental testing is not yet addressed fully in this process. This section describes the assessment procedure for in-process integration and design verification testing (Table 3).

The first set of questions related to in-process integration addresses whether this particular flight round was built correctly. They are designed to verify that the components of the round function correctly and that nothing is broken during the buildup throughout the integration process. These questions assess whether there is adequate testing to ensure that the flight round is built to the design and that the workmanship is good. In addition to the built-right assessment, to ensure mission success, it is necessary to evaluate whether the missile is designed properly. This second set of questions assesses the body of design verification testing to ensure that the design is able to meet the mission function requirements.

Note that this assessment process cannot be divorced from engineering judgment and be distilled to a mere checklist. Based on the answers to each set of questions, the assessment provides a green, yellow, red, or gray rating for the in-process and performance test categories.

Green means that the ground testing is adequate for that mission function. Yellow indicates that a mission function is marginally assessed and that there are deficiencies in testing. This condition may be due to a lack of testing that can be covered by additional tests or by modifying existing tests. Red indicates that a mission function is not adequately assessed. This may be due to an omission in the test planning that must be addressed to have confidence prior to flight. A rating of gray is reserved for the category of functions that by nature cannot be tested prior to flight (such as one-shot devices). This rating indicates there is risk to be considered since it is not evaluated as green; however, there are no additional tests or test modifications that would reduce that risk. A worksheet similar to Table 3 for capturing the overall environmental testing assessment is being developed.

Table 3. "Initialize Missile" core questions and evaluations.

Core questions—Was unit under test built according to design?	Answer	Comments	Impact
1. Are components/interfaces tested in the flight configuration?	Partially	Nonflight IOM at WIT; VLS emulator used in MCO	Low
2. Are components/interfaces tested in all expected operating states?	Yes		
3. Are components/interfaces subjected to appropriate stimuli?	Yes		
4. Are sufficient parameters/operations measured to assure retention of operations throughout the integration process?	Yes		

In-process integration rating
Initialize Missile — R, Y, G — Rationale

Core questions—Does the design execute the intended purpose?	Answer	Comments	Impact
1. Are sufficient function observables measured?	Yes		
2. Are function observables tested to adequate fidelity?	Yes		
3. Are emulations validated?	Yes	Note: No emulators used at WIT	
4. Are off-nominal conditions examined?	Yes		
5. Have environmental influences on function been tested?	NA		

In-process integration rating
Initialize Missile — R, Y, G — Rationale

Overall rating
Initialize Missile — R, Y, G — Rationale

6. Document results

A book of worksheets is created as this process is continued for all of the mission functions. With this assessment, feedback can be provided to the test community to address any test deficiencies identified. Because this process relies heavily on engineering judgment in the assessment, there will be disagreement on some issues. However, the process provides a traceable means to document the relevant issues and rationale for decisions so they can then be constructively discussed and debated.

CONCLUSIONS

The process presented here provides a formal, objective, and traceable means to assess the comprehensiveness and adequacy of ground testing. By its structured nature, this process focuses and documents the use of engineering judgment to assess the adequacy of planned tests. In its use of key questions that focus on what is important in testing, the process provides a means for test engineers to better plan their tests from the beginning.

To successfully test a complex system requires a comprehensive set of ground tests. This method provides a framework in which to unify a comprehensive ground test program and to show that all functions are tested adequately. Note that this assessment of test adequacy is independent of the results of the tests. The test adequacy assessment is used together with the test results to give an overall assessment of whether a particular ship system and flight round are capable of meeting all mission functions.

REFERENCE

[1]*Integrated Ground Test and Evaluation Plan for the ALI Flight Demonstration Program*, G684621, Raytheon Missile Systems Co. (Feb 1999).

THE AUTHOR

STEPHEN B. ROGERS received B.S. and M.S. degrees in computer engineering from Clemson University in 1986 and 1992, respectively. Before joining APL in 1999, he spent 5 years working on Standard Missile systems engineering at the Naval Surface Warfare Center in Dahlgren, Virginia, and 6 years working on Standard Missile development and testing at NSWC's Port Hueneme Division in Port Hueneme, CA. As a member of ADSD's Theater Missile Systems Engineering Group, his current work includes Naval Theater Ballistic Missile Defense test and evaluation and systems engineering. His e-mail address is steve.rogers@jhuapl.edu.