ESA STUDY CONTRACT REPORT

ESA CONTRACT	SUBJECT	CONTRACTOR
N° 4000105946 /12/NL/RA	Analysis of Spacecraft qualification Sequence & Environmental Testing (ASSET)	**Thales Alenia Space Italia S.p.A.**
* **ESA CR () No**	**No of volumes: 5** **This is Volume No. 3**	**Study Manager:** P. Messidoro

ABSTRACT:

SUMMARY REPORT

This document summarizes the findings of the ESA ASSET (Analysis of Spacecraft qualification Sequence and Environmental Testing) study, in the form of a concise and informative conference or journal paper.

The work described in this report was done under ESA contract. Responsibility for the contents resides in the author or organization that prepared it.

Names of authors: ASSET Study Team

NAME OF ESA STUDY MANAGER:	ESA BUDGET HEADING
B. Laine	

Summary report

DRL/DRD: D3

Written by	***Responsibility*** + handwritten signature if no electronic workflow tool
M. Pasquinelli	Author for ASSET team 24/11/2014
Verified by	
V. Basso	Checker
G. De Logu	Configuration Control 28/11/2014
Approved by	
P. Messidoro	ASSET Study Manager 28/11/2014
Released by	
E. Vaccari	Documentation Manager 28/11/2014

Approval evidence is kept within the documentation management system.

ThalesAlenia Space

AIRBUS DEFENCE & SPACE

CHANGE RECORDS

ISSUE	DATE	§ CHANGE RECORDS	AUTHORS
01	10/SEP/2014	Document written by M. Pasquinelli on behalf of the ASSET team. First issue.	L. Pace, M.Pasquinelli, P.Messidoro, P. Giordano, S. Voglino, V. Ancona, F. Verges, J. Buffe, J. D'Add, P. Hugonnot, R. Werner

Analysis of Spacecraft qualification Sequence and Environmental Testing

B. Laine(1), J. Candé(1), P. Messidoro(2), M. Pasquinelli(2), P. Giordano(2), V. Ancona(2), V. Stefano(2), L. Pace(3), J. Buffe(4), P. Hugonnot(4), J. D'Add(4), F. Vergès(5), R. Werner(6)

(1) European Space Agency
email: {benoit.laine, jacques.cande}@esa.int

(2) Thales Alenia Space Italia
email: {piero.messidoro, mauro.pasquinelli, pietro.giordano, vittorio.ancona, stefano.voglino}@thalesaleniaspace.com

(3) Politecnico di Torino (resident in Thales Alenia Space Italia)
email: lorenzo.pace@polito.it

(4) Thales Alenia Space France
email: {jerome.buffe, patrick.hugonnot, jerome.d-add}@thalesaleniaspace.com

(5) Airbus Defence & Space S.A.S.
email: francois.verges@astrium.eads.net

(6) Airbus Defence & Space GmbH
email: reiner.werner@astrium.eads.net

1. Acronyms

ASSET	Analysis of Spacecraft qualification Sequence & Environmental Testing
Airbus DS	Airbus Defence & Space
ECSS	European Cooperation for Space Standardization
ESA	European Space Agency
FLA	(in-) Flight Anomaly
FM	Flight Model
ITT	Invitation To Tender
JAXA	Japan Aerospace eXploration Agency
JPL	Jet Propulsion Laboratory
MATED, MAT€D	Model And Test Effectiveness Database
MBSE	Model-Based System Engineering
Mech	Mechanical tests
NASA	National Aeronautics and Space Administration
NCR	Non-Conformance Report (used as alias for on-ground anomaly)
NCTS	Non Conformances Tracking System
PFM	ProtoFlight Model
RFT	Reduced Functional Test

TAS	Thales Alenia Space
TTE	Total Test Effectiveness
TV	Thermal Vacuum
w.r.t.	With Respect To

2. Introduction

The study here presented has been aimed to improve the effectiveness of the testing campaign to detect anomalies before flight on the basis of lessons learned from the past and in view of future project needs. Its general objectives are:

- identification of the user requirements to reach the wanted improvements
- survey of the available data potentially used in the study
- selection of the interesting cases to be analyzed
- analysis of the in-orbit anomalies oriented to testing improvements for avoiding their occurrence in flight
- analysis of the on-ground testing anomalies oriented to testing improvements for detecting them earlier
- analysis of the as-run test programs oriented to testing effectiveness improvements in comparison to actual standards
- synthesis of the study results including recommendations for test programs improvements and proposal for further activities

Inside the purpose of the study, the *test effectiveness* has been defined as the ability to minimize the number of anomalies (causing mission failure or mission performance degradation) discovered after launch, maximizing their discovery during ground testing activities [1].

To achieve the abovementioned general objectives, four questions have been considered as guidelines:

- Which problems that occurred in previous programs could have been avoided through a different test campaign (limited to environmental tests)?
- Considering the lessons learned of previous programs in terms of test effectiveness, are there any activities related to environmental testing (including test result analysis) that in future programs could be deleted, reduced, aggregated or optimized without impacting mission success?
- What could be an effective test sequences and model philosophies according to mission type, product type and acceptable risk, to reduce hardware related anomalies?
- Which type of methodologies and database(s) should be developed/improved to follow-up the test effectiveness topics?

The aim of the study has not been to necessarily find a direct answer to each question, rather to bring elements useful to build relevant analyses.

The base of data consists of 18 ESA-led space programs, of which TAS and Airbus-DS are the final integrators:

- **Science**: INTEGRAL, XMM-Newton, Mars Express, Rosetta, Venus Express, Herschel, Plank, Cluster 2 (Cluster 1 considered for qualification)
- **Human-rated**: ATV, Node 2, Node 3, Cupola
- **Telecommunications**: Artemis
- **Earth observation**: ERS 2, MetOP, GOCE, CryoSAT, MSG

A total of 48 models, of which 34 flight/protoflight models, have been considered. Further 86 flight/protoflight models have been considered for non-ESA programs.
Additional programs integrated by TAS and Airbus-DS but outside of the ESA perimeter have been considered without disclosing confidential information, in the form of statistics and anonymous cases (both FLAs and NCRs).

3. Methodology

As a first step of the study, a methodology has been built, consisting of some *definitions* that were given to synchronize the study team, *selection criteria* for the screening of NCRs and FLAs and of a set of *study questions*, which have been thought as specialization of the general objectives.

Definitions

The study team was composed of people from different companies and of varied background. To be sure that the used technical language was shared, a set of definitions have been produced, tracked and kept along the whole study:

- The cause of each anomaly (both ground/flight) shall be determined looking for the originating cause (root cause), and not at the primary effect
- As a reference for *test effectiveness*, TTE is used, as defined in MAT€D [4]
  - The TTE corresponds to the number of NCRs in the test of interest, divided by the sum of the total NCRs plus the FLAs in the early flight period.
- *Qualification* is done on specific hardware (therefore with qualification durations and levels)
- *Protoqualification* is done on flight hardware (therefore typically with reduced durations and levels, sometimes with qualification levels and acceptance durations);
  - Any FM on which the Protoqualification approach (i.e. one or more requirements are verified at qualification level on flight hardware) is applied for at least one requirement is to be considered PFM.
- When the issue is in *software*, if the issue is due to the fact that the software wrongly commands the thermal control system, this anomaly is considered linked to environmental testing, even if the root cause is considered software, if it is discovered only under environmental conditions.

Also, some aspects related to anomaly severity had to be specified, as described in the selection criteria.

Selection criteria

A total of 35843 on-ground anomalies (NCR) and 199 in-flight anomalies (FLA) were available from the selected space programs.

The selection criteria for NCRs and FLAs have been defined with reference to ECSS standard on testing. The perimeter is restricted to:

- Environmental tests:
  - Mechanical: Static Load, Transient, Vibration, Random Vibration, Sinusoidal Vibration, Acoustic, Modal Survey, Shock, Microvibration susceptibility/ emission, Spin
  - Thermal Cycling and Thermal Balance
  - Electro Magnetic Compatibility and Radio Frequencies (EMC/RF): Conducted, Radiated/auto-compatibility/RF, RF system test
- Verification levels:
  - Space Segment Module (i.e. embedded space segment element as payload modules, service modules, pressurized modules contained in a spacecraft);
  - Space Segment System (i.e. stand-alone space segment element as a satellite or a human rated infrastructure system)

This means that Space Segment Equipment (e.g. valves, batteries, individual electronic boxes) and Space Segment Subsystem (e.g. electrical power, attitude control, structure, thermal control, software) were not considered in ASSET.

- Verification stages:
  - Qualification: demonstration that the design, meets the applicable requirements including margins,.
  - Proto-qualification: qualification verification conducted on the flight item (protoflight approach)
  - Acceptance: demonstration that the flight item is free from workmanship errors and it is ready for subsequent operational use

Other stages (pre-launch, in-orbit, post-landing) were not considered.

- Severity (in case of NCRs, evaluated in term of potential in-flight consequence, not the impact on AIT program):
  - Critical:
    - FLA: Loss of Mission, Partial loss of functionality
    - NCR: it would have caused the total loss of the mission or a partial loss of functionality, in case it had not been identified during test
  - Major:
    - FLA: Switch to redundant unit, without impacting the system functions; Delay caused to the operations
    - NCR: it would have caused the passage to a redundant unity or a delay in the operations, in case it had not been identified during test;

Minor were not considered:

- FLA: Unexpected behavior, causing no delays or impacts on the system functions
- NCR: it would have not caused any delays or impacts on the system functions, in case it had not been identified during test.

Nevertheless, according to their sensibility, all the personnel involved in the study may consider including cases that, even if not exactly in line with these criteria, can contribute effectively to the results.

Beyond the anomalies, the data used to contribute to ASSET is also made up of the test programs and conditions of the selected spacecraft. In particular, the attention is focused on the sequence of environmental tests (mechanical and thermal) and on the levels of some tests (e.g.: number of cycles of thermal vacuum, number of axes for sinusoidal vibration).

Study questions

A set of 24 questions have been defined to guide the analysis of cases and as-run test programs. These questions have not been intended to be necessarily answered directly but rather to support the methodology.

Each question has been completed with details regarding how to proceed with data handling:

- Empirical/statistical analysis: each question may be answered following a case by case approach (i.e. deeply understanding what happened for each NCR and FLA considered), and this is called *empirical* approach; It can also be answered gathering quantifiable parameters on a number of cases, and this is called a *statistical* approach; it is possible that a question may be covered by both approaches.
- Filter: indications of type of cases (NCR/FLA) to be considered for that specific question; a filter could be used in well-structured databases. For this purpose, a MAT€D-like syntax and taxonomy are used for the study.
- Required exchanged data: describes what data must be shared among the partners.
- Required own data: describes what documents must be accessed, without the need to share all the proprietary information.

The study questions may be relevant for analyzing NCRs, FLAs or both. There are 19 questions with general purposes and 5 that are more specific to one single test.

The study questions are listed and numbered below. They are grouped together into 6 general objectives and further into 10 detailed objectives.

- Completeness of the test program (are we covering enough / too much)
  - Is the test program sufficient to detect all the mission critical and major anomalies?
    1) What is the correlation between completeness of the verification by test approach and presence of flight anomalies?
    2) Which test would have allowed to detect this FLA? Are there FLA that could have been discovered on ground? How the related test plan should have been performed?

- Consistency (e.g. with lower level tests, with model philosophy or with activities objective)
  - Are there missing tests at equipment level?
    3) Did we find at system level NCR that should have been seen at equipment level? Did we find FLA that should have been seen at equipment level?
    4) Is there any test duplication among equipment testing and Spacecraft level acceptance? What may go wrong if we skip Spacecraft level testing on already tested equipment/refurbished hardware?
  - Is the test outcome exploited correctly?
    5) Are there any anomalies that occurred on-ground and again in flight and why?
    6) Are there anomalies that may have been anticipated with different measurements or exploitation of test data? i.e. would better instrumentation/monitoring allow to detect anomalies?
  - What is the most effective model philosophy?
    7) What is the link between test effectiveness and the development/verification philosophy (i.e. model philosophy)?
    8) In a product line (satellite, equipment), on which model can a test be removed? Analyze decrease rate of a type of anomaly across the product line.
- Correctness/relevance (are we doing the right things)
  - Are different test conditions required?
    9) Are there FLA not seen on ground although the test has been performed, because of the low likelihood of occurrence (probabilistic anomalies)?
    10) Would different test conditions allow detecting issues? E.g. solar simulation, different boundary conditions…
    11) What is the relationship between the test effectiveness and the severity (number of TV cycles, levels…) of testing, the type of anomaly, and characteristics of related equipment?
    12) Are there NCR/FLA where a parameter is out of spec, but disposed easily because the allowable range for the parameter was too limited (margins too high) and the design works well outside of those boundaries?
  - Can we detect dangerous test to improve them?
    13) Which type of risk does exist associated to specific test, in terms of flight hardware degradation?
- Test sequence (are we doing the things in the right order?)
  - What is the most effective test sequence at Spacecraft (and lower levels)?
    14) Impact of having Mech/TV or TV/Mech on the number of anomalies detected in orbit/on ground, to answer: between Mech/TV and TV/Mech, is there a better sequence or is it indifferent?
- How to consolidate the standards
  - How to consolidate standards through feedback/advices?

15) What is the relationship between the test effectiveness and the ECSS test standard?
16) What is the typical test program per mission type, in comparison also with standards?
17) How can deviations from standards be justified, in different mission types and products?

- How to customize ECSS in function of the product (one-of S/C, product line, constellation)?

18) What is the added value (effectiveness) of each test (in particular for environmental tests) per type of mission/product?
19) How the test effectiveness varies with the verification stage (i.e. acceptance, qualification) for different types of products types?

- Specific questions:
  - What are the questions derived from lessons learned from previous programs?

20) Can we cancel shock test?
21) Can we cancel tests after transportation?
22) Can the duration of the Thermal Vacuum (TVAC) be reduced?
23) Can we limit the number of Axis for sine?
24) In which case can STM be cancelled?

4. Selected cases

118 NCRs and 35 FLAs, including anonymous non-ESA anomalies, have been selected as corresponding to the selection criteria and further analyzed. Automatic filtering has not been used systematically inside the utilized databases for those aspects that are not totally unambiguous (particularly severity); instead, a case by case assessment of those parameters has been carried out, with the result of including some cases that would have been nominally outside of the study perimeter.

An overview of screening results is provided with the following figures [5].

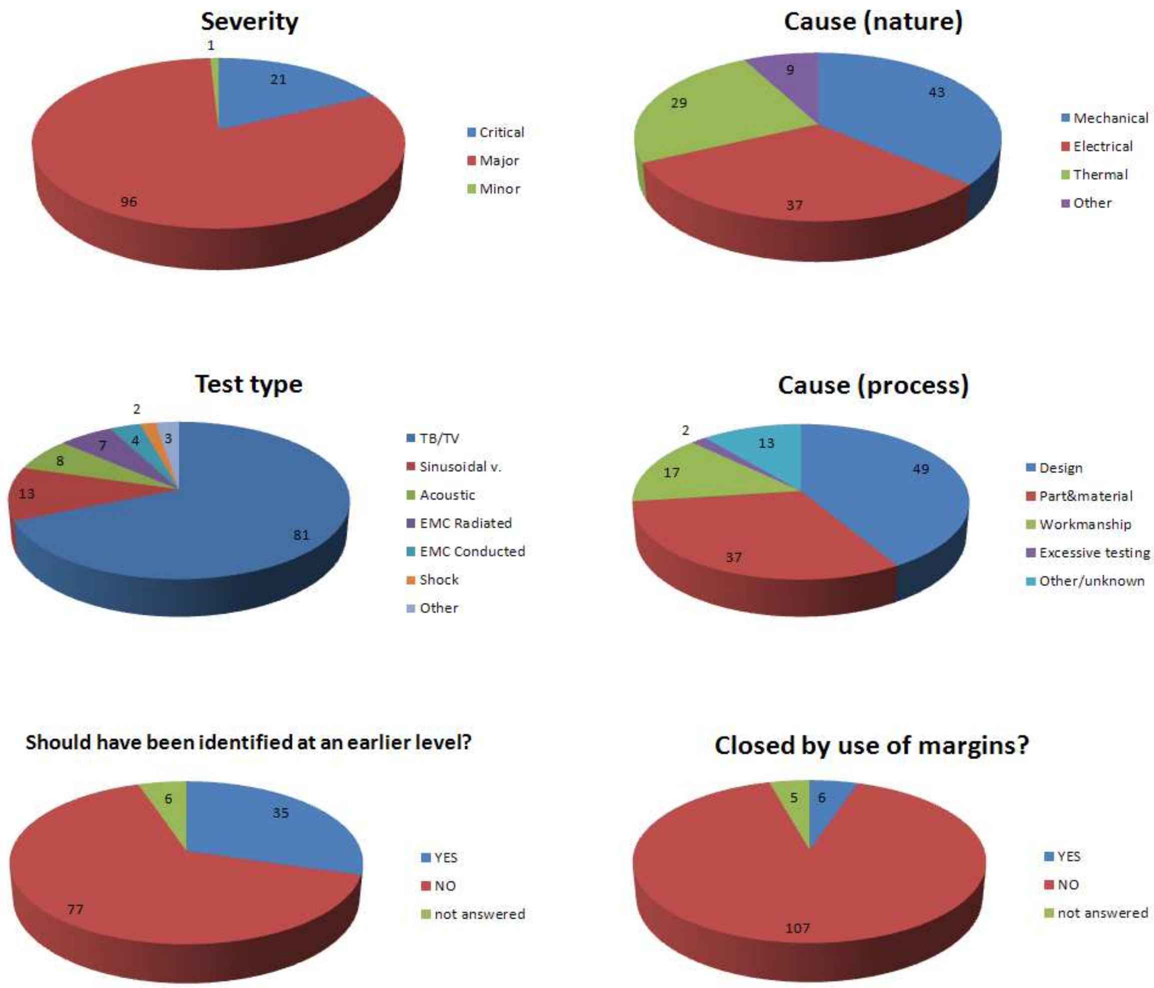

Figure 1: On-ground anomalies statistics (for a total 118 on-ground anomalies).

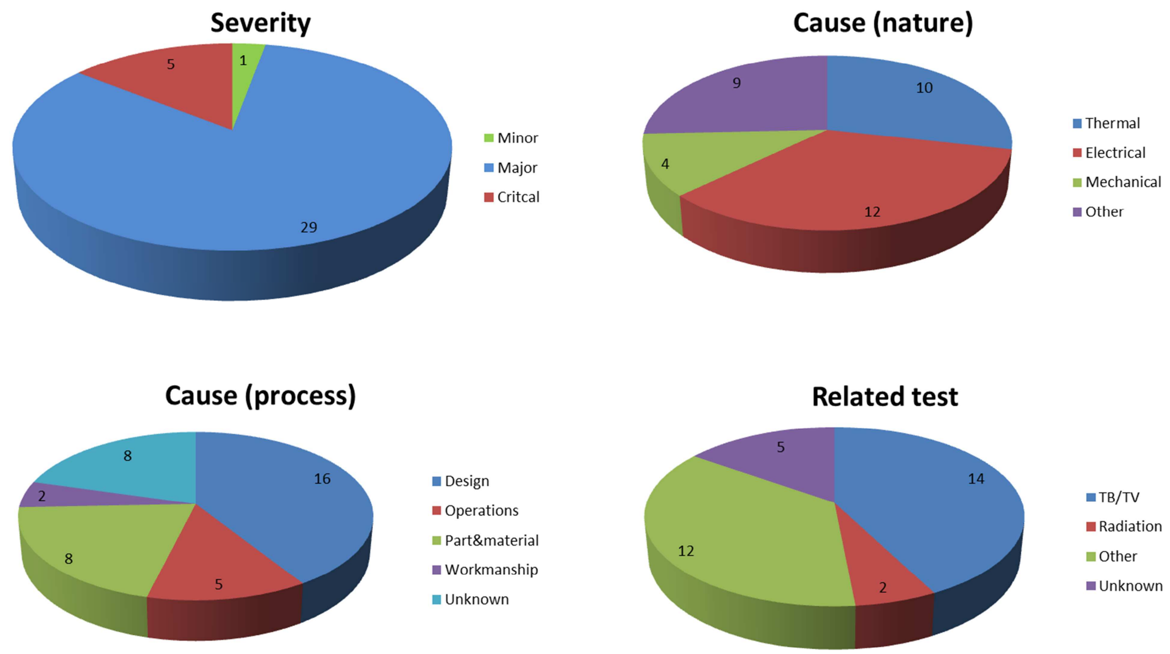

Figure 2 : In-orbit anomalies statistics (for a total 35 Flight Anomalies). Note that multiple choices are possible for *Cause (process)* and *Related test*. Note that one *minor* anomaly has been selected as reputed to have higher potential impact if occurred in different conditions.

The activities performed along this study revealed that it is very difficult to get detailed information, and even more difficult to get analysis on an anomaly (both on-ground and in-flight) some years after its occurrence. Therefore the aim is to learn from the past it is necessary to store reports in a well-structured format, with complete description of the events, enabling people that have not worked on the program to understand what has happened.

Also, the great influence of subjective judgment emerged from the analyses:

- it is not always possible to know exactly the dynamics at the root of every in-flight anomaly,
- the analysis of raw data required high effort (from hours to weeks), because current processes and related documentation (i.e. flight anomaly reports and related annexes, such as related test reports) are not conceived to support such second level analysis as in this study
- the same flight anomaly analyzed by different people may give different outcomes, both in terms of root cause and of suggested feedback. This situation gets worse if the analysis is made long after the anomaly occurred.
- the need for the application of very tight selection criteria led to the impossibility to use automatic filters on databases.

Some MAT€D stored data was questioned during the analyses; this could mean mainly that different background hypotheses or interpretations were made, thus building a different analysis.
The fact that some mistakes have been found in databases is an important clue: it does not mean that databases are wrong and thus useless, rather that only using them there is the possibility to improve their content and reach the possibility to have reliable analyses and fruitful return of experience.

5. Results of the analyses

This section summarizes the results of the study, in the form of recommendations and findings.
Recommendations have not to be interpreted as mandatory dispositions, but as useful hints that have emerged from real data and are aimed to make environmental testing more effective.
As well, the findings shall be spread inside the testing community, to make everyone aware of this *big picture* and help the possibility to build new considerations.

It must be noted that the proposed recommendations, findings and conclusions emerge from real facts; they summarize what past on ground and in flight anomalies, together with test plans and conditions, tell about the environmental testing program. Anomaly databases, document repositories and MAT€D have been used for retrieving useful data; experts of the analyzed programs have been interviewed (or directly involved when possible) to avoid misinterpretation of reports and to expand the understanding of past programs also through the access to additional documentation.

General recommendations

Lifetime qualification at equipment level

6 out of 35 FLA and 3 out of 118 NCR are connected with lifetime. Issues related to long-term behavior are often not sufficiently covered at equipment level.
Therefore a first outcome would be to suggest more lifetime testing. This is mainly addressable to mission-specific equipment.
Mission-critical and mission-specific instruments testing campaign should consider a complete lifetime test on qualification model at equipment level for specific technologies requiring full equipment representativeness (i.e. full prototype), which cannot be covered by component qualification, especially in case of little heritage or no heritage.
The recommendation is limited at equipment level lifetime testing, qualification stage, where no analysis is possible due to lack of past data on that specific technology/design, limited to items that are essential for mission success and are not redundant.
It is also proposed to continue life test on QM even after QR to learn more about lifetime.

Impact of system level testing on equipment life

Analyzing the anomalies that are dealing with lifetime issues (as in the preceding recommendation), it is possible to derive that, during acceptance, the sum of equipment and system level testing duration could detect a lifetime issue.
The current trend to reduce the system level tests reduces test duration and increases the risk not to discover such problems. Test minimizations at system and equipment level are independent, while cumulative effect is relevant for lifetime.
It is recommended to take into consideration, when minimizing the number or tests at system level (and thus the duration of the test), the contribution of system level tests to lifetime.
The recommendation is focused on the acceptance stage.
A study on the topic may be envisaged.

Satellite product lines

The number of NCR and FLA decreases along the consecutive Flight Models of a product line, except in case of major design changes.
Some variability aspects shall be taken into account, as shown by one analyzed case, where variability connected with the nature of used material (quartz) is to be considered a significant risk source which still deserves thorough testing.
For satellite product lines, in case trend in anomalies number stabilizes (without any mission-relevant anomaly), it could be allowed to reduce environmental testing at system level, on flight models after the firsts (provided that adequate acceptance testing is provided at equipment level for critical items).
Only workmanship-related tests are necessary, if no retrofit on design from satellites already on orbit has been applied.
This recommendation is aimed to system-level acceptance-stage testing of satellite product lines.

Testing after transportation to launch site

Out of 32541 NCR on 86 S/C only one anomaly has transportation of the S/C to launch site as a possible cause. Cases when the shock monitoring system detects loads outside the allowed envelope have not been considered. Found NCR was very specific and root cause could not really be transport-related.
RFT after transportation should be recommended in ECSS only in case that some load out of design envelope occurs, as long as the load monitoring system is reliable (with *load* covering the complete environment : temperature, humidity, EMC, mechanical loads, etc., i.e. as long as the environment is controlled as required).
The recommendation applies only on transportation of the whole spacecraft, not equipment or part of the S/C.

Test data analysis

In some cases test data could have been exploited more effectively : the anomalous behavior was already measured but the related measurements were not exploited correctly (analysis of records "a posteriori" detected the problem and escaped post processing even if they have actually occurred). This is mainly due to spurious events, which are difficult to detect and reproduce.
The recommendation is to consider using automated diagnostic tools to improve analysis of test data. Example of such tool is the Dr. MUST diagnostic tool developed by ESOC and available to the European community [6].
Some side suggestions for such tools are:

- Pay special attention to spurious data
- Evidence coupled parameters out-of-specification.
- Evaluate trends

The recommendation could be applied to all test records.
Application in the short/mid term is disputable. However, in some case automated screening of test data is already performed.

Assessment of trends

2 out of 35 flight anomalies show that analysis of test data limited to comparison with margins is not always sufficient.
In some cases, analysis of trends or out-of-place measurements (detectable from a chart built on the purpose) may have detected the anomalous behavior beforehand.
When checking that the test record falls inside the specification, it could be advantageous to evaluate that the trend w.r.t. time (consecutive measurements) is not divergent (when applicable). This should be limited to critical parameters, agreed between customer and supplier.
The recommendation could be applied to all test records.

Mech/TV - TV/Mech sequence

The analysis of as-run-test programs and related anomalies had the specific purpose to analyze the differences between different sequences. Quite surprisingly, even if in the previous version of ECSS standards the Mechanical/Thermal approach was favored, a large majority (9 out of 13 FMs/PFMs) of analyzed test programs is based on Thermal/Mechanical approach. Analyzed test programs do not show difference in test effectiveness among the two approaches.
The last version of the ECSS leaves free in the domain of test sequence (Mechanical/Thermal or vice versa) [3].
It is recommended that, in the ECSS handbook, some rationale is given to both Mechanical/Thermal - Thermal/Mechanical sequence, with pros and cons.
This recommendation is relevant for all space programs, given that both thermal and mechanical verification are accomplished by test. It is applicable both to qualification and acceptance stages, as well as for protoqualification.

The recommendation is not aimed to force one approach or the other (as it is addressed to ECSS handbook and not testing standard), rather to give guidelines for both choices.

Usage of Reduced Functional Tests

During the selection of on-ground anomalies and the analysis of as run test programs and their comparison with standards, special attention was given to functional tests after environmental tests, which are able to discover issues caused by the related conditions. A variety of cases have been found and in some cases Reduced Functional Test (RFT) is not always performed, without any obvious impact on test effectiveness. The ECSS-E-ST-02C requires RFT after each environmental test block.
A common definition of test block was not clearly identified.
In the ECSS handbook, clarification should be given in which case RFT has to be performed before and after each environmental test block.
Also, an unambiguous definition for “test block” should be provided.
The recommendation is focused on the environmental test campaigns for acceptance and protoqualification stages.

Recommendations related to model philosophy

Design qualification with PFM approach

During the analysis of as-run test programs and their comparison with the current standards, some discrepancies were found between current practices and ECSS requirements. ECSS standard for testing asks for flight hardware manufacturing to start at the end of the qualification phase. This is obviously impossible to be followed with a protoflight approach. ECSS testing standard at section 4.5.2c says: "Qualification testing: Qualification testing shall be completed and design improvements or modification incorporated and qualified prior to authorization for the flight product manufacturing".
The request is to modify the abovementioned requirement in the ECSS testing standard, considering the exception of PFMs.
Details should be given in the standard (or in the handbook) about how the PFM is managed.

Hybrid model philosophy

Hybrid model philosophy is by far the most commonly used (15 programs, versus 2 pure QM+FM and 4 pure PFM), with one or more QMs (intended as models representative of a portion of aspects, not complete qualification models, mostly STM/SM and/or EM) and the PFM.
Those programs adopting a hybrid model philosophy do not show any evident impact on mission success.
ECSS should reflect the practice putting adequate attention to this kind of model philosophy. The request is to detail in ECSS-E-HB-10-02 the Hybrid model philosophy QM(s)+PFM.
It is suggested to indicate discipline by discipline if a test is to be performed with qualification or acceptance level/durations. As a consequence, the S/C may be PFM w.r.t some aspects and FM

w.r.t others. A way to handle this would be to put verification objectives against how they are verified (e.g. a matrix view with “what is to be done” w.r.t. “in which model/activity…”)

Recommendations for improvement of Thermal tests

Thermal transients in Thermal Vacuum test

Significant anomalies (3 FLA out of 35, at least one of critical impact) do not occur at extreme conditions, but at intermediate temperatures. Some anomalies are linked only to the extreme temperatures and use-as-is disposition is decided after long investigation.
In thermal vacuum, it is recommended to add functional test during transients (temperature change) and at intermediate temperatures at steady state. A minimization of the tests performed on plateaus may be envisaged in parallel.
Speed of increase/decrease of temperature should be aligned (if not in line) with the real physical phenomena (if possible).
This recommendation applies on thermal vacuum test at acceptance or protoqualification stage, at spacecraft level.

Thermal Vacuum cycles

Spacecraft with just one TV cycle performed do not show difference in test effectiveness.
None of the programs considered shows any relevant NCR at the second cycle. However, one case was reported by an expert to be actually happened at the second cycle (with the corresponding functional test already performed at the first cycle).

As an additional hint, it is interesting to know that NASA/JPL requires a minimum of 2 cycles if the orbital environment leads to temperature cycling (e.g. low orbit), otherwise 1 cycle if environment does not cycle (e.g. interplanetary transfer), but the landscape is varied at international level.
There is a JAXA study ongoing on this topic, scheduled to be completed by the end of 2014, which is focused on thermal vacuum on units (the number of thermal cycles in unit thermal vacuum/cycle tests has been specified as eight in JAXA test standard **Error! Reference source not found.**).

Unfortunately, analysis of reports of on-ground anomalies shows that most of times there is no available track of the exact cycle in which the anomaly happened.
It is suggested to consider reduction of the minimum amount of TV cycles to 2 (for acceptance and protoqualification) at system level.
This recommendation shall be associated with the suggestion to produce a deeper analysis on the subject. It is important to improve the understanding of such cases in the future.

Recommendations for improvement of Mechanical tests

Vibration tests

The deviation to avoid random vibration if acoustic is performed is used most of the times at system level.

At system level it is assumed that the random vibrations coming from the launcher interface are not a major contributor for most of the equipment.
Therefore it is commonly accepted and justifiable to run solely acoustic test at system level.
For a unit, the random environment monitored at its interface during the acoustic test at system level is a good starting point to define the unit random environment. Composite methods are used in practice.

It is recommended to make ECSS standard consider the possibility to avoid random vibration at system level if acoustic is performed, while at the moment this is regarded as a deviation. This does not mean that such practice should be preferable.

Recommendations for improvement of engineering feedback and lessons learned

Availability of program data

The survey performed during the ASSET study revealed that the effort to get relevant pieces of information from ended programs is very high. Especially for very old projects, where experts are not further available or cannot remember, the information acquisition is very difficult. This lack of knowledge capture could lead the European space sector to miss important lessons learnt, reproducing the same errors in future programs, or missing the occasion to improve the current processes. One key aspect has been that typical repositories lack a semantic structure or that the semantic structure is badly filled and are not meant for re-use after the program ends.
It is recommended to improve the knowledge capture across programs and for future re-use.
As a potential solution, for space programs led by ESA, a dedicated activity (e.g. in a dedicated work package) should be included aimed at gathering (e.g. inside MAT€D, NCTS, ARTS) data from AIV (model philosophy, activities performed, test conditions), NCR and FLA of the commissioning phase. Such work package should be independent (not included in the proposal nor in the financial plan) and should be not negotiable. At least, awareness shall be raised within the project- and operation teams to the importance of correctly and clearly capturing the data that can be used as lessons learned.
As an alternative potential solution, an ESA-controlled group should keep control of such activity across programs. The application of such recommendation has a foreseeable impact. This should be seen as an investment for the improvement of the return of experience.

Exploitation of test and flight data

At least 1 FLA out of 35 revealed that same problems occur again in later programs because the past experience is difficult to access while handling very specific domains. This seems to be a symptom of the lack of proper feedback from operations to engineering.
Experience flow from test and flight operations to engineering should be improved. This can be realized by a well-structured base of data for each discipline and related models.
Possibly, the application of this recommendation could also help the exploitation of test data and individuation of cause of in-flight anomalies, linking such base of data to automated tools.
This proposal can be envisaged as part of a larger MBSE logic applied to design and verification.
The use of databases built focused on the understanding of the relationships among items, testing activities and flight operations (as MAT€D) could help in take advantage easily from what has already happened. Today MAT€D is used as a research tool by a small group, but its involvement should be extended at least in term of methodology.
Enhanced data management methodologies (e.g. proposed by the "Space System Data Repository" or MBSE initiatives) may help to link discipline/related models to the system level data and activities, allowing also easier inter-disciplinary exploitation of flight/test data.

Anomaly severity analysis

When analyzing anomalies (both on-ground and in-flight), the assigned severity was found to be determined with inhomogeneous criteria, mainly linked to the impact on the current operations.
NCR are often evaluated w.r.t. their impact on schedule.
The suggestion is to consider the impact on mission of the occurred anomalies, both on-ground and in-flight, while determining the severity of an anomaly on its own report. NCR should not be evaluated only w.r.t. their impact on schedule but also w.r.t. their potential impact in-flight.

FLA should be examined in detail in order to assess if their impact has been reduced due to fortuitous conditions, especially when providing precious input to future programs (where the same situation could be mission critical).

As a potential solution, as a first step, it is suggested to add in Non-Conformances Reports and in Flight Anomaly Reports a "potential impact on the mission" field.
This should be associated to what actually happened during flight or ground operations, to provide feedback to next programs.

One general example for FLA can be as follows: the anomaly is easily overcome thanks to a very robust design (e.g. a backup of the failed item is provided by another item) and results in little or no impact on mission. Without that backup the impact would have been higher, so a good lesson could be learned for future programs. Even if Minor, this anomaly could be evidenced for return of experience.

One general example for NCR can be as follows: a mechanical fixation is found to be decoupled. It is rapidly repaired without any relevant impact on schedule or test program and therefore classified as Minor; if not spotted, such anomaly could have had a large impact on flight operations (so the right classification should have probably been Major or Critical).

NCR in TV and cycles

Surveys performed in ASSET revealed that some details in the nonconformance reports are not fully clear or accessible, as the number of cycle a NCR has occurred in.
This could be linked to the investigation of the optimal number of TV cycles and could help to determine whether such reduction in number of cycles is justified
This could be seen as coupled with the recommendation above, as it invites to complete the anomaly reports with an eye on the return of experience, giving the possibility for technicians to understand the events even if not directly involved on that specific program.
The recommendation is to include in the NCR found during TV the cycle in which the NCR occurred and detail if the test has been performed on previous cycles with a correct result (to identify if it could have been found before - from a functional viewpoint). Same for structural failures during acoustic and random, note at what time it happened.

Completeness of anomaly reports

The recommendation above evidences one of the major points of TV-related NCR analysis, but may be extended to a more general sentence, covering the completeness of all anomaly reports: anomaly reports shall provide all the details needed to understand what has happened, its impact and the resolution put into effect. The provided pieces of information shall be self-sufficient for the complete comprehension of the event.
This should be already done, at least theoretically.

***

The following findings are not strong enough to derive recommendations, but they constitute interesting matter for future consideration and discussion.

General findings

Equipment level testing as feedback from system level

6 out of 35 Flight anomalies, 21 of 118 on ground anomalies evidences problems that should have been detected at equipment level.
The reasons for non-detection are widely spread and not linked to a specific activity or root cause, might be linked to e.g. test duration, representativeness of interfaces, exploitation of test results (problems were present in the test results, but they were not detected by analysis of test results during equipment level testing).

The prescribed tests were performed in almost all the cases. In some cases it is problem of qualification (bringing even to re-design) or acceptance.

It is not possible to derive a general recommendation, because of the widespread cases, but care should be taken on the definition of qualification representativeness of test at equipment level and on the exploitation of data. Also, feedback from NCR and FLA should be provided to equipment supplier, so to improve design of equipment level testing.

A specific type of unit (or provider) may be prone to a kind of anomaly. To achieve better specific recommendations for type of equipment, lessons learned for equipment level anomalies should be organized in order to be successfully exploited for future test programs (e.g. structured database).

Long term perspective: extend the coverage of testing through discipline analysis models

STM has been found to be less used in favor to larger utilization of discipline virtual models. As a long term trend, it is possible to foresee that qualification is going to be more and more fulfilled through virtual simulation models, with the exception of new technologies or bad/no inheritance, while environmental testing cannot be avoided for acceptance and protoqualification purposes. Such virtual models will continue to need a validation performed through a smaller program of environmental tests.

Success criteria and success conditions definition

During the analysis of non-conformances report and the screening of flight anomalies, many anomalies which were considered at a first glance as major, after investigations were disposed as "use-as-is". Many times the disposition was due to a too stringent success criteria, connected e.g. to a too stringent requirement. The current feedback process is more connected with reporting critical issues than minor issues, and many times such reports are simply disposed and kept just by the experience of the project team, while they may be a good source to reduce the requirements in future projects, and capitalized.

Test success criteria are not always consistent with mission success conditions. The definition of very tight test conditions (and success criteria) could be a tendency of test specifiers (but also AIT managers) and has an impact by generating “false” NCRs (anomaly solved with "use as is" because does not impact mission capability/safety) which causes delays, extra activities and cost.
When assigning a "use as is" outcome on an NCR:

- clarify why the specification was too tight at the start
- assess and keep track of the impact of relaxing afterwards
- assess the combination with other applied "use as is" specification relaxations to check that they will not result into a real problem

Keep such track as a feedback from AIT to engineering.

Findings related to specific activities

Radiation test

The analysis of FLA found out that there is no room for relaxation on the radiation environment, and the issue remains of critical importance.
Experts suggested improving modeling capabilities (both radiative environment vulnerability and reaction of the system to a SEU); then this could help to test correctly at part level. It is proposed to recommend improving representativeness and early availability of radiation models and radiation testing of parts.

Thermal balance test

A thermal balance test monitoring just the behavior at steady state gives no clue on behavior during thermal transients, which may be important in case transient behavior is considered critical for the mission and no or limited significant on-orbit experience is available.

The duration of the eclipse phase has a strong impact on the temperature excursion of some units and the associated power needs for the heaters. Or the thermal time constant may result into temperature gradients incompatible of the functional needs of the subsystem.
In general, if thermal dissipation of a unit is dependent on operational mode a transient thermal test could be useful, or when the thermal inertia is important (e.g. batteries) and it is necessary to evaluate performances of new developed items.

ECSS standard on testing [3] specifies Thermal Balance with gradient for dissipative equipment and for equipment with important thermal gradient at vacuum conditions.
A similar recommendation is already included inside ECSS standard on Thermal control general requirements **Error! Reference source not found.**[9].

Thermo-elastic distortion during thermal test

The thermo elastic distortion topic is hard to manage just by analysis, so monitoring during thermal test is to be considered an added value. One analyzed flight anomaly has been proven to be caused by thermo-elastic distortion.

Shock test at qualification stage

Shock test is done only once on the majority of the S/C with no impact on the FLA.
The study has identified from ESA experience some risks related to the shock test that would push to limit them to the maximum extent.
Note that in general the shock test (and similarly the others) is used at system level to validate the specification of the lower level testing.

Two Shock test firing is a requirement from launcher authority. A recommendation in the sense of reducing the requested firings from two to one could be addressed to launcher authorities. ECSS-E-ST-10-03C considers one firing for shock test for acceptance and protoqualification, and a minimum of two for qualification.

Further investigation is suggested: the main need is to remove from ECSS the request of shock test on flight hardware (acceptance/protoqualification).

6. Conclusions

The study has proven that environmental testing is effective in individuating anomalies so that they do not occur in flight. Nevertheless, some aspects to be improved have been individuated. Relevant recommendations and findings are going to be submitted to pertinent authorities and departments, such as ECSS committee members, AIV and AIT teams and engineering boards. Some topic has showed to be worth more specific research and could be subject of further study.

References

[1] P. Messidoro, M. Pasquinelli, F. Vergès, P. Desiré, B. Laine, J. Candé: *ASSET Study on the improvement of the environmental tests*, 4th International Workshop on Verification and Testing of Space Systems, Torino, Italy, May 2013
[2] P. Messidoro, M. Pasquinelli, L. Pace, F. Vergès, B. Laine, J. Candè: *ASSET – Analysis of Spacecraft Qualification Sequence and Environmental Testing*, FADAT, June 2013, ESA ESTEC
[3] ECSS-E-ST-10-03C, Testing, ECSS Secretariat, June 2012
[4] P. Messidoro, O. Brunner, A. Cerrato, E. Giorgio: *The MAT€D initiative after ten years: main test improvement results and potentialities*, 25th Aerospace Testing Seminar, Manhattan Beach, California, 2009
[5] M. Pasquinelli, P. Messidoro, Stefano Voglino, L. Pace, B. Laine, J. Candé, F. Vergès: *Analysis of Spacecraft qualification Sequence and Environmental Testing: mid-term results*, 28th Aerospace Testing Seminar, Los Angeles, USA, 25-27 March 2014
[6] Martinez-Heras, J. A., Yeung, K., Donati, A., Sousa, B., & Keil, N. . *DrMUST: Automating the Anomaly Investigation First-Cut*. In the IJCAI-09 Workshop on Artificial Intelligence in Space. Pasadena, California, USA, July 2009.
[7] *Thermal Test Practices at Assembly and System Levels*, Jet Propulsion Laboratory, California Institute of Technology. Savannah, Georgia: International Conference on Environmental Systems (ICES), July 2009
[8] T. Niwa, D. Takahashi, Q. Shi: *Thermal cycle assessment by test effectiveness point of view*, 28th Aerospace Testing Seminar, Los Angeles, USA, 25-27 March 2014
[9] ECSS-E-ST-31C, Thermal control general requirements, , ECSS Secretariat, November 2008

END OF DOCUMENT