

# **First Crewed Flight: Rationale, Considerations & Challenges from the Constellation Experience**

*Carlos Noriega  
William Arceneaux  
Jeffrey A. Williams  
Jennifer L. Rhatigan*

*Formerly of the Constellation Program Office  
NASA Johnson Space Center, Houston, TX*

## THE NASA STI PROGRAM OFFICE . . . IN PROFILE

Since its founding, NASA has been dedicated to the advancement of aeronautics and space science. The NASA Scientific and Technical Information (STI) Program Office plays a key part in helping NASA maintain this important role.

The NASA STI Program Office is operated by Langley Research Center, the lead center for NASA's scientific and technical information. The NASA STI Program Office provides access to the NASA STI Database, the largest collection of aeronautical and space science STI in the world. The Program Office is also NASA's institutional mechanism for disseminating the results of its research and development activities. These results are published by NASA in the NASA STI Report Series, which includes the following report types:

- **TECHNICAL PUBLICATION.** Reports of completed research or a major significant phase of research that present the results of NASA programs and include extensive data or theoretical analysis. Includes compilations of significant scientific and technical data and information deemed to be of continuing reference value. NASA's counterpart of peer-reviewed formal professional papers but has less stringent limitations on manuscript length and extent of graphic presentations.
- **TECHNICAL MEMORANDUM.** Scientific and technical findings that are preliminary or of specialized interest, e.g., quick release reports, working papers, and bibliographies that contain minimal annotation. Does not contain extensive analysis.
- **CONTRACTOR REPORT.** Scientific and technical findings by NASA-sponsored contractors and grantees.
- **CONFERENCE PUBLICATION.** Collected papers from scientific and technical conferences, symposia, seminars, or other meetings sponsored or cosponsored by NASA.
- **SPECIAL PUBLICATION.** Scientific, technical, or historical information from NASA programs, projects, and mission, often concerned with subjects having substantial public interest.
- **TECHNICAL TRANSLATION.** English-language translations of foreign scientific and technical material pertinent to NASA's mission.

Specialized services that complement the STI Program Office's diverse offerings include creating custom thesauri, building customized databases, organizing and publishing research results . . . even providing videos.

For more information about the NASA STI Program Office, see the following:

- Access the NASA STI Program Home Page at <http://www.sti.nasa.gov>
- E-mail your question via the internet to [help@sti.nasa.gov](mailto:help@sti.nasa.gov)
- Fax your question to the NASA Access Help Desk at (301) 621-0134
- Telephone the NASA Access Help Desk at (301) 621-0390
- Write to:  
NASA Access Help Desk  
NASA Center for AeroSpace Information  
7115 Standard  
Hanover, MD 21076-1320



# **First Crewed Flight: Rationale, Considerations & Challenges from the Constellation Experience**

*Carlos Noriega  
William Arceneaux  
Jeffrey A. Williams  
Jennifer L. Rhatigan*

*Formerly of the Constellation Program Office  
NASA Johnson Space Center, Houston, TX*

Available from:

NASA Center for AeroSpace Information  
7115 Standard Drive  
Hanover, MD 21076-1320  
301-621-0390

National Technical Information Service  
5285 Port Royal Road  
Springfield, VA 22161  
703-605-6000

This report is also available in electronic form at <http://ston.jsc.nasa.gov/collections/TRS/>

# Contents

Abstract .....	1
Introduction .....	1
Why We Test: Perspectives on Testing Effectiveness.....	2
Verification and validation.....	2
Test like you fly .....	2
The “incompressible” test list .....	3
The Criterion for Human-Rating .....	3
Meeting the human-rating rationale .....	4
Constellation Overview.....	6
Flight Test Plan—Beginnings and Rationale .....	7
Flight test plan .....	7
Early development testing.....	9
Budget Challenges Leading to Test Erosion .....	10
Incremental erosion of the testing scope .....	11
Results of test erosion .....	12
The program’s response to the erosion.....	13
The Unknown Risk—Can You Understand it <i>a priori</i> ? .....	15
Conclusions .....	18
References .....	20

## **Acronyms & Abbreviations**

CAIB	Columbia Accident Investigation Board
FT	Flight Test
IC	Initial Capability
IOC	Initial Operational Capability
ISS	International Space Station
IVGVT	Integrated Vehicle Ground Vibration Tests
KSC	Kennedy Space Center
LAS	Launch Abort System
LC	Lunar Capability
NESC	NASA Engineering and Safety Center
PA	Pad Abort
PDR	Preliminary Design Review
T&E	Test and Evaluation
WSMR	White Sands Missile Range

## Abstract

NASA's Constellation Program has made the most progress in a generation toward building an integrated human-rated spacecraft and launch vehicle. During that development, it became clear that NASA's human-rating requirements lacked the specificity necessary to defend a program plan, particularly human-rating test flight plans, from severe budget challenges. This paper addresses the progress Constellation achieved, and problems encountered, in clarifying and defending a human-rating certification plan, and discusses key considerations for those who find themselves in similar straits with future human-rated spacecraft and vehicles. We assert, and support with space flight data, that NASA's current human-rating requirements do not adequately address "unknown-unknowns," or the unexpected things the hardware can reveal to the designer during test.

## Introduction

Deciding that a system is ready for first crewed flight requires evidence that the design avoids or controls human space flight safety and survival risks well enough to warrant flying a crew on that mission. In the face of repeated budget cuts,<sup>i</sup> the Constellation Program found it necessary on at least an annual basis to reduce the scope of its test and evaluation (T&E) plan and thus reduce the body of evidence expected to be produced in support of the decision to fly crew aboard the vehicle.

As a result, the expected increase in *understanding the risk* was hampered by ever shrinking test data availability. There was persistent debate among knowledgeable and experienced experts within and without the Program as to how to achieve sufficient readiness for the first crewed flight. In response to the Program's request, the NASA Engineering and Safety Center (NESC) prepared and released a Technical Assessment Report of its study on "Readiness for First Crewed Flight"<sup>ii</sup>.

The purpose of this paper is to supplement the NESC report findings with detail of the Constellation experience and to provide additional observations and recommendations. We begin by amplifying the NESC Report's Appendix on testing with a brief Constellation view of why we test, followed by discussion of the background and rationale for the Constellation flight test plan, and describe the incremental erosion of that flight test plan through a series of annual budget reductions. We discuss the Preliminary Design Review (PDR) timeframe study of a proposal aimed at recouping the ground lost via earlier reductions. We then address the shortfalls in NASA's current human-rating criterion from the perspectives of the Constellation experience. We conclude by recommending actions for future human space flight programs in this regard.

All of the concerns voiced and decisions reported in this paper were documented "in the open" in the appropriate Constellation Program management and control forums; however, we have chosen not to cite the many applicable meeting minutes or decision packages, since they are not readily available in the open literature. This paper serves as the summary of relevant discussions and decisions at that time, and, by citation, is part of the lessons learned from the Constellation Program<sup>i</sup>.

## Why We Test: Perspectives on Testing Effectiveness

A number of measures and approaches can be used to establish and execute effective testing with the objective of addressing all risks, both known and unknown. This section addresses these approaches, and discusses areas where we believe the approaches have fallen short.

### ***Verification and validation***

Appendix A of the NESC report<sup>ii</sup> does a good job of describing the role of testing with respect to verification and validation. This view was shared among many within the Program.

We assert, however, that the Agency, and indeed the aerospace community as a whole, has become too fixated on verification of compliance to a specification. Verification objectives alone should not comprise a complete T&E plan, and do not sufficiently interrogate the system, allowing it to “speak” of issues not addressed in a specification. Verification has become a checklist merely confirming compliance with specified items. Granted, the good specifications are based on a mature heritage of testing failures and lessons learned, but cannot take the place of thoughtful design, risk identification, and a thorough T&E plan.

Validation activities are also part of the necessary work set forth in a T&E plan to create or sustain an “as-certified” baseline, even for a “one-off” system. Validation includes defining the capabilities and constraints of the specific design, characterizing the system’s behavior, confirming the control of hazards, confirming the accuracy and validity of predictive models, confirming operational plans and procedures, and proving that the system is suitable for its intended purposes. For context, human-rating certification examines the subset of the as-certified baseline (verification and validation data) related to crew safety and survival.

### ***Test like you fly***

What is understated in the NESC report, and ignored by many, is that there is another, possibly more important, role for testing—to provide opportunities to learn from flight-configured hardware and software in operational scenarios with appropriately relevant environments (*test like you fly*). This approach is critical in finding unknown-unknowns and even “unknown-knowns.”<sup>1</sup>

In an era where projects are often continuously challenged to do more with less, it is easy to rationalize moving individual preliminary test objectives from one distant test to another without recognizing a loss of test effectiveness (ability to find design flaws or latent defects) in an attempt to project a lower-cost test campaign. This “bucketing” activity—delaying testing until a higher level of assembly is reached where more objectives can be met with fewer tests—presupposes a more comprehensive understanding of the inherent risks than actually is possible during the early life of the program. This delaying approach, while attractive in reducing early costs, risks higher costs

---

<sup>1</sup> An example of an unknown-known was the rediscovery of tribo-electrification during the *Ares I-X* countdown. This is characterized as an unknown-known since the Space Shuttle was plagued by this phenomena several years prior.

and schedule impacts to correct problems when they are later found. More importantly, this way of thinking only deals with the “known-knowns” and “known-unknowns.” “Test like you fly” is a proven approach that should not be underestimated or discounted.

### ***The “incompressible” test list***

One recommended technique to sustain a test plan is to document an incompressible test list at the early stages of the program<sup>2</sup>. This can be difficult to develop for a complex multi-element program, but provides a means to establish and protect a program’s ability to deliver the test results needed to make the human-rating decisions. Constellation set out to identify the testing necessary to certify safety and survival-related functions (considered “incompressible”) as part of the integrated T&E plan delivered at PDR. This plan was based on requirements, design feature assessments, and hazard analyses to identify the critical content. Constellation established a plan to accomplish T&E plan development in lockstep with requirement, architecture, and design maturation; however, it was incomplete at PDR, as the program intended to further mature T&E planning at later milestones, concentrating instead on requirement and design issue resolution at PDR.

Program managers must strive to consider or forecast what the minimum data set may be for making the first crewed flight decision with limited information at the beginning of a program, and then work to protect it. The program’s frequent need to react to a changing budget outlook rippled through its T&E planning efforts and was also a factor in its inability to solidify its strategy.

In retrospect, some of the initially conceived T&E objectives being used for planning (those that accompanied the development and trades of the architecture and design concepts), could have been identified as “human-rating critical” when the early work could show a connection between those T&E objectives and the eventual need for the first-crewed-flight decision makers to know that the safety and survival critical functions, features, and controls would be available as intended to protect that crew. As a lesson learned, we offer that an approach like this would enable managers to more easily protect key objectives earlier as the program and its T&E plan mature and react to changes of any kind or from any cause.

## **The Criterion for Human-Rating**

NASA’s general guidance from NPR 8705.2B, Human-Rating Requirements for Space Systems<sup>iii</sup> is as follows:

*A human-rated system accommodates human needs, effectively utilizes human capabilities, controls hazards with sufficient certainty to be considered safe for human operations, and provides, to the maximum extent practical, the capability to safely recover the crew from hazardous situations. Human-rating consists of three fundamental tenets:*

---

<sup>2</sup> This is conceived as a minimum threshold listing of absolute needs to certify the program, developed early in the program. Its weakness can be that it is not necessarily a requirements-driven list, but is based on judgment, and is therefore little more effective in the face of budget threats.

(1) *Human-rating is the process of designing, evaluating, and assuring that the total system can safely conduct the required human missions.*

(2) *Human-rating includes the incorporation of design features and capabilities that accommodate human interaction with the system to enhance overall safety and mission success.*

(3) *Human-rating includes the incorporation of design features and capabilities to enable safe recovery of the crew from hazardous situations. Human-rating is an integral part of all program activities throughout the life cycle of the system, including design and development; test and verification; program management and control; flight readiness certification; mission operations; sustaining engineering; maintenance/upgrades; and disposal.*

NASA's human-rating criterion<sup>ii</sup> as stated by the NESC is "Given that the human spaceflight system is designed for human spaceflight, it is accepted that the objective is to fly humans when risks to the crew safety have been mitigated to the point where the need or benefit is worth the residual risks." NASA has the experience of human-rating prior launch vehicles: the Mercury, Gemini, and Apollo programs in the 1960s and the Space Shuttle Program in the early 1980s, all of which occurred before the career arc of the vast majority of the current NASA workforce.<sup>3</sup>

This criterion, while sensible and easy to understand in principle, is difficult to implement as certification guidance for the first human test flight because it can be broadly interpreted. It is left to program management to interpret into a design, development, analysis, and test plan that leads to certification for human space flight. Such broad interpretation can result in some otherwise necessary work being defined as not critical, and therefore more vulnerable to being cut in the face of budget challenges. Moreover, this approach does not necessarily address whether a minimum level of T&E is necessary to reveal unknown-unknown risks, since these are, by definition, not part of residual risk.

The instrument for assessing whether the program can achieve human-rating is the T&E plan. A sufficient plan *queries* the hardware and software for the known-unknown risks (part of the residual risk in the human-rating criterion parlance), *and also* affords opportunities for it to *reveal* problems that requirement verification, design validation, and risk mitigation did not.

### ***Meeting the human-rating rationale***

The very nature of the current human-rating criterion implies that the real decision to first fly crew is not made until the parameters of value and benefit can be evaluated and understood. Logically, the information required for this would not be available until later in a program's life cycle, approaching the time at which the decision must be made—that is, the point at which prerequisite testing and evaluation is complete or nearing completion, when hazards and other risks are identified and controlled to the extent practical, and when the value can be judged by the nation's

---

<sup>3</sup> This does not discount the valuable experience of sustaining the Space Shuttle's human rating, or the two major re-assessments of the human certification following the *Challenger* and *Columbia* accidents. Additionally, the ISS program provided recent workforce with experience in human rating spacecraft, but not launch vehicles. Historically, the greater risk to life lies with the launch vehicles.

needs *at that time*. The reasons for needing a crew on a given flight may change over the Design, Development, Test & Evaluation timeframe depending on the program's objectives and national priorities.

It is therefore necessary to assure that the information required by decision makers to assess the risk is available to them when that time comes. The processes that a program puts in place—be they design, verification, risk analysis, etc.—are equally as important as the T&E strategy it executes. Deficiencies in either will hamper the decision makers' ability to make the right call at that time. An insufficient plan makes an insufficient data set.

In reality, however, the decision maker will rarely have all the information he or she would like to have to make that call. As with Constellation, most programs will face tightening constraints in budgets and timelines that will reduce or eliminate the program's ability to provide comprehensive data for the decision.

Since safety is not a binary function with a yes or no evaluation, we postulate an interpretation of the certification criterion used by NESC: it is safe *enough* to fly humans *when* the benefit of flying the crew is greater than the residual risk. This implies that understanding *both* the benefit of flying the crew *and* the residual risk is necessary to make the decision.

#### *Understanding the benefit of flying the crew*

The Constellation Program followed the Space Shuttle Program in human-rating a new space transportation system. Lack of understanding of the initial launch decision for the Space Shuttle caused great debate almost 30 years later. Most had forgotten, or never knew, that the precedent of flying the crew on the first flight had been driven by the value of flying the crew to compensate for uncertainties in the entry handling characteristics (known-unknown) that could manifest themselves during reentry communications outages (known-known)<sup>iv</sup>.

Unlike the clear need for a crew to mitigate known risks to save an extremely valuable reusable asset, which existed on the Space Shuttle's first flight, the Constellation Program was in the midst of defining crew interfaces and procedures. The Program had yet to define scenarios where crew intervention was critical to improve mission success. The one value of flying the crew established to date was to meet the earliest declaration of initial operational capability (IOC) possible, enabling American flights to the International Space Station (ISS) after Space Shuttle retirement.

#### *Understanding the risk*

Constellation used the concept of Risk Informed Design, along with more traditional approaches<sup>i</sup>, to gain a better understanding of the risks in the design. In addition to expected preliminary design products, the projects had made significant progress in identifying hazards and control strategies, probabilistic risk assessments and reliability assessments, including failure modes and effects analyses. These products were to mature as detailed designs matured and test data validated assumptions and models. Initial drafts of the integrated test plans, though preliminary and driven

primarily for the purpose of scoping cost, appeared to be modest, yet sufficient to human rate the Initial Capability (IC) systems.

The known residual risk is well discussed in the NESC report<sup>ii</sup>. While the NESC authors touch on it, one of the greatest challenges for Constellation was to get the larger team, beyond those daily involved in risk analysis, to better understand the need to seek opportunities to find the unknown risks.

## Constellation Overview

NASA formed the Constellation Program in 2005 to achieve the objectives of maintaining American presence in low Earth orbit, returning to the moon for purposes of establishing an outpost, and laying the foundation to explore Mars and beyond in the first half of the 21st century. The Constellation Program's heritage rested on the successes and lessons learned from NASA's previous human space flight programs: Mercury, Gemini, Apollo, Space Shuttle, and the ISS<sup>v</sup>.

Following the loss of *Columbia*, NASA established the Columbia Accident Investigation Board (CAIB) to perform an in-depth review of the Space Shuttle Program. As a result of this review, the CAIB concluded that it was in the best interest of the U.S. to develop a replacement for the Space Shuttle. The CAIB concluded that it should be possible using past and future investments in technology to develop the basis for a system, "significantly improved over one designed 40 years earlier, for carrying humans to orbit and enabling their work in space."<sup>vi</sup>

The Program was charged with achieving an order-of-magnitude improvement in risk to crew and mission over that of the Space Shuttle. Probabilistic risk assessment at that time put the risk of loss of crew for the Space Shuttle on the order  $10^{-2}$ . Constellation was challenged by the Agency to improve this to the order of  $10^{-3}$ .

Since exploration of the moon and beyond was the overarching goal of the Constellation Program, all elements were initially designed to perform the lunar missions, while also being capable of performing missions to the ISS. The Program was planned as a stepwise capability build-up largely based on Space Shuttle heritage components. The IC comprised elements necessary to service the ISS by 2015 with crew rotations, including the *Orion* Crew Exploration Vehicle, the *Ares I* Crew Launch Vehicle, and the supporting ground and mission infrastructure to enable these missions. The Constellation Lunar Capability (LC) added the *Ares V* Cargo Launch Vehicle, the *Altair* Lunar Lander, and spacesuits designed for partial gravity exploration. Lunar outpost elements and capabilities were to follow, including mobility elements such as rovers, permanent or semi-permanent habitats, and power and communication elements to support a sustained exploration presence.

The remainder of this paper focuses on the erosion of test flight planning for the IC, since it had the more immediate need, although similar debate was taking place in the initial planning for the LC

components and configuration.<sup>4</sup> The following discussions illustrate the challenges encountered by the Constellation Program in implementing the human-rating criterion to certify the IC configuration for human flight, and how further clarity could enhance future human-rating certifications.

## **Flight Test Plan—Beginnings and Rationale**

The Constellation Program was developing NASA's first new human-rated spacecraft in three decades, and a comprehensive series of flight test activities was planned in order to find and fix any design problems, and to certify and document vehicle performance capabilities (including flight crew safety and survival designs) before human flights began.

Flight test operations are unpredictable by nature, and this is especially true when new designs and new hardware are being tested for the first time in extreme operational environments (see Table 1, Figure 2, and related discussion). Indeed, recent historical launch vehicle test flight experience indicates that overall confidence in the engineering design will be established after the 7<sup>th</sup> launch<sup>vii</sup>. While the Program worked to design the safest human space system ever developed, its test plan was success oriented given the constraints of the budget and schedule<sup>v</sup>.

### ***Flight test plan***

The initial flight test plan<sup>viii</sup> is described below and illustrated in Figure 1. Two of these flight tests successfully launched prior to Program cancellation (as noted by check marks in Figure 1).

- The pad abort (PA) flight tests were designed to test the Launch Abort System (LAS) and parachute system in a pad-abort scenario (no launch vehicle ignition). PA-1 was an early pathfinder intended to provide design insight in the PDR timeframe using a mock-up of the crew module. It successfully launched from the White Sands Missile Range (WSMR) in April 2010. While similar, PA-2 was planned with higher fidelity (flight-like) hardware.
- The ascent abort tests (AA series) were to use surplus Air Force Peacekeeper first stage and/or second stage motors. This solid-fuel booster would launch the crew module to an altitude high and fast enough for the LAS to operate at supersonic speeds and during periods of ascent profile maximum dynamic pressure (max Q) and unstable (tumbling) flight modes. The plan included a high-altitude abort simulating failure of the system's upper stage at the point of stage separation. All abort flight test activities were designed to take place entirely within the WSMR.
- The first developmental flight test of the *Ares-I* vehicle (designated *Ares I-X*) was a highly successful uncrewed launch from the Kennedy Space Center (KSC) in October 2009. Test flight objectives focused on first-stage flight dynamics, controllability, and separation of the first and

---

<sup>4</sup> While the baseline plan for the *Ares V* was to fly cargo only (the Earth Departure Stage), options were retained to human-rate this vehicle if later studies showed launching *Orion* along with the Earth Departure Stage (ala Apollo) was feasible and desirable.

upper stages, along with ground operations and first stage recovery. *Ares I-X* tested the integration and performance of a simulated *Ares/Orion* “stack” prior to Critical Design Review so that resulting design changes could be incorporated before production of flight articles. *Ares I-X* utilized a four-segment solid rocket booster excessed from the Space Shuttle Program, with a mass-simulator for the fifth segment. It included mass/dynamics simulators for the upper stage and the *Orion*.

- The second uncrewed, developmental flight test of the *Ares-I* vehicle (designated *Ares I-Y*) was to also be launched from KSC; this flight test was to consist of a five-segment booster with a flight-like upper stage and simulated engine. The *Ares I-Y* flight was intended to validate the operation of the *Ares I* five-segment first stage, demonstrate the ability to prepare the (unpowered) second stage for flight after first stage separation, and demonstrate performance of a high-altitude abort of *Orion* after separation of the first stage.
- The first orbital flight test of the integrated *Orion/Ares-I* vehicle was designated *Orion-1*. This would be an uncrewed first test flight of a complete *Ares-I* first stage and operational upper stage, paired with an operational *Orion*. The *Orion* would be inserted into an orbit that enabled rendezvous with the ISS (although there were no plans to dock with the ISS on this first flight) to test onboard systems such as the solar panels, reaction control system thrusters and main engine, and rendezvous systems. A water landing and recovery off the coast of California was under study for this mission.
- The first flight of the *Orion/Ares-I* vehicle to carry astronauts was designated *Orion-2*<sup>5</sup>. This would be a two-crew mission that would rendezvous and dock with the ISS. After the *Orion-2* mission, the *Orion/Ares* vehicle would have achieved IOC and would begin approximately two flights per year to the ISS to support crew rotations and to have the *Orion* spacecraft docked for 180-day intervals as an emergency crew return vehicle. At this point, when it was proven that the Constellation systems could operationally support the *Orion* vehicle with routine operations to the ISS, the Program had planned full operational capability. Initially, the Program was committed to full operational capability with *Orion 4* in early 2015.

This test plan was conceived at the time that the IOC designs were conceptual, and was therefore immature. The flight test plan was expected to mature and evolve with the design, and as understanding of risks and uncertainties developed. Indeed, one of the earliest changes (not indicated in Figure 1) was to delete a planned AA-4, a high-altitude abort test that was to be performed at WSMR. It was determined that these test objectives could be included with the *Ares I-*

---

<sup>5</sup> The early program contained an unpressurized cargo version of *Orion*, intended for ISS resupply. The flight test manifest contained an additional uncrewed, unpressurized cargo flight that was to perform automated rendezvous and docking with ISS, prior to first flight with a crew. The unpressurized cargo and automated rendezvous and docking requirements were dropped after the first year, primarily due to budget pressures, and the associated flight was deleted. This constitutes the first ‘erosion’ of the number of flights before first crewed flight.

Y flight. This had the advantage of avoiding potential range issues at WSMR for such a long down-range flight, and also afforded the opportunity to add water landing and recovery to the objectives.

As discussed below, the expected maturation of the test flight plan was complicated by the budget challenges to the Program. Moreover, as the Program matured, the Agency was unable to fully clarify either what minimum body of evidence was required for the first flight of humans on a new spacecraft or how much testing was enough to provide unforeseen problems sufficient opportunity to be revealed<sup>6</sup>.

### ***Early development testing***

The Program also planned and executed early development tests for identified risk areas. For instance, a plan to incrementally build and test the LAS innovative steering motor was implemented, and the motor later performed flawlessly in the PA-1 flight test. Multiple parachute tests for both the *Orion* crew module and the *Ares* first stage were performed at the Yuma proving grounds.

The second stage of the *Ares I* was liquid oxygen/liquid hydrogen fueled, powered by a modified J-2 engine designated J-2X. The modifications were necessary to increase performance of the engine. The J-2X engine was also to be used on the upper stage of *Ares V* for the LC. Unlike past rocket engine developments, far fewer test starts and far less total run times were planned to qualify the engine (less than one tenth of the starts and runs used to qualify the Space Shuttle Main Engine, for instance). Given that the J-2X was an engine modification rather than full-scale development, and that the design and analysis software currently used didn't exist for any prior U.S. human space flight engine development, engineers and managers felt this test campaign baseline would adequately qualify the engine. The engine development was not immune to test article deletion; however, the deletion of spare test articles was assessed to be predominantly a schedule risk in that J-2X production assets planned for flight could be diverted to back-fill the test plan if necessary.

The Program also responded to the inevitable development problems with testing. High ascent acoustic loading and uncertainties surrounding the aero-thermal behavior of the LAS motor exhaust plumes and the potential effect on the boost protection cover led to a series of additional wind tunnel tests, including hot gas tests, to develop empirical data for refinement of analytical models.

The well-publicized thrust oscillation issue<sup>i</sup> was addressed with both analysis and testing. Testing investigated the predicted oscillation's potential impact on the crew's ability to monitor displays and evaluated several possible oscillation reduction design changes. The Space Shuttle solid rockets were instrumented to provide validation for transfer functions. From these, a design solution was selected, along with further options, should further testing reveal more severe oscillations than expected.

---

<sup>6</sup> Such requirements do exist for payloads and reflect varying levels of test of new commercial flight vehicles based on the level of government insight and importance of the payload.

Stability of the *Orion* crew module after water landing, and safe recovery of the crew in challenging sea states was undertaken in a series of water-tank tests that then proceeded to the open ocean.

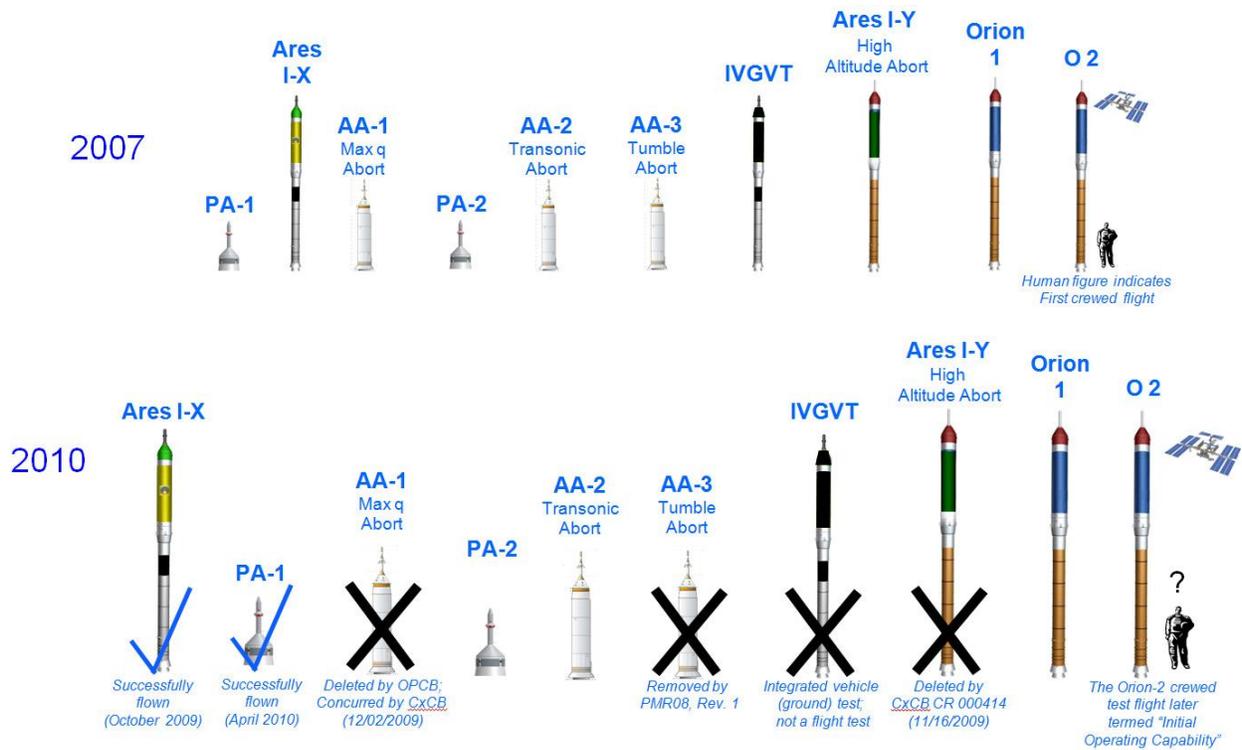
A series of *Ares I* scale model acoustic tests were also undertaken to better understand the expected ignition overpressure pulse and liftoff acoustics, and to characterize the effectiveness of several sound suppression mitigation strategies, including water deluge, rainbirds, and water bags.

In these and other cases, the Program directed testing resources to mitigate initially identified and emerging development risks. This can be characterized as an expected allocation of development budget and contingency reserves and reflected the efforts by the Program to focus dwindling resources on the most critical tasks. These are only highlights of the developmental test plan; more information related to early developmental testing is available.<sup>ix</sup>

## **Budget Challenges Leading to Test Erosion**

*Constellation Program's Lessons Learned, Vol. I: Executive Summary*<sup>i</sup> begins with the observation: "Funding for the Constellation Program was inconsistent and unreliable from its initial formulation through its cancellation." Indeed, budget cuts were levied in the first fiscal year, and cuts were made in all subsequent years. Initially the Program slipped dates to preserve content (mission performance), but the continually decreasing annual buying power eventually caused requirement scrubs focused on ground and flight tests (rather than performance requirements) and further schedule impacts. Figure 1 illustrates the incremental test erosion that occurred in the test flight plan.

As explained in *Constellation Program's Lessons Learned, Vol. I: Executive Summary*<sup>i</sup>, these cost and schedule problems were not driven by technical challenges that the Constellation Program and its government/contractor team faced, but were a direct result of the budget profile and the erosion to that budget.



**Figure 1: Erosion of Constellation's planned integrated flight test plan, 2007 to 2010.**

### ***Incremental erosion of the testing scope***

Unfortunately, continuing budget cuts caused delays of procurements, and shifted test dates past the point where flight tests could feed data into design, without delay of the targeted IOC date.

One test that was repeatedly scrutinized for potential deletion was *Ares I-Y*. The baseline test flight did not have a fully functioning upper stage, which was paced by availability of the upper stage engine (the J-2X). Supporters of the flight test argued that an early ground test version of the J-2X engine could be added to the upper stage to expand the *Ares I-Y* flight test objectives. This would address one of the Program's top technical risks—the ability to successfully detach and maintain guidance control of the upper stage—and also assure that the critical J-2X “start box” conditions could be supplied by the upper stage main propulsion system after staging. The conundrum was that this addition would potentially delay the test even further, so although test planners from the outset promoted the upper stage conditioning objectives, it was never adopted as part of the Program baseline.

Since *Ares I-X* had no functioning upper stage and utilized a four-segment Space Shuttle solid rocket booster, elimination of *Ares I-Y* would have resulted in only a single flight prior to flying the first crew, and elimination of the high-altitude abort test. This was reported to the Program to be unacceptable by the safety and mission assurance community. Because the budget situation drove annual reviews of content, the test flight program was always under review for affordability. *Ares I-*

Y was preserved on the manifest during several annual budget-driven attempts to delete it. Eventually, *Ares I-Y* was eliminated in late 2009 because the achievable flight date had slipped with respect to the progress of *Ares I* upper stage development. After the deletion of the *Ares I-Y* test flight, the Program was not able to achieve consensus on which flight would be the first to fly a crew. The Program designated *Orion 2* for planning purposes as the target flight for the first crew and initiated studies to determine whether the program would be ready, and planned to designate the actual first crewed flight later in the life of the Program after the value of flying the crew was better defined and the residual risk was better understood.

As the Program progressed toward PDR, concern about the erosion of the abort tests and the cuts of ground and flight tests was evident in the increasing number of risks and issues discussed about the adequacy of the ground and flight test campaign. The targeted first human flight had also become tagged IOC, a significant program developmental and contractual milestone. These actions fueled debates about the adequacy of testing specific systems, as well as a general argument fueled by precedent for flying a minimum number of times before flying a crew<sup>vii</sup> (see also<sup>x,xi</sup>). The NESC report was commissioned because these debates could not be concluded.

At the time Constellation was cancelled, the development ascent test designated *Ares I-X* and the development pad abort test designated PA-1 had been successfully completed, while only a second pad abort test and an ascent abort at maximum dynamic pressure remained in the plan. Several recorded open issues related to the deletions were being worked that would have constrained certification for crewed flight.

### ***Results of test erosion***

While the Program was directed to mitigate or “buy down” the risk to the first crewed flight through ground and flight test, cuts to ground tests and flight test continued to occur. Due to the budget situation, every test was assessed for deletion as long as the preliminary test objectives could be relocated to a remaining test, or if rationale could be developed for elimination of the test.

At the Program’s PDR, known untenable gaps existed, such as the lack of a first stage separation test article. Significant parts of the originally conceived ground and flight test strategy were never put in the baseline or were later cut from the Program planning references or deferred until after first crewed flight. These known gaps and critical emerging test requirements were being recorded as risks, but many contained mitigation plans that Constellation did not have available funds to implement. There was growing doubt expressed that the existing plans could support placing humans on the targeted flight. The Program’s shrinking resources were unlikely to rectify this<sup>7</sup>.

---

<sup>7</sup> In the historical context, many senior managers had experienced a similar erosion of testing in the ISS program. Integrated testing was later added back to the program when the political situation improved and international partner schedules slipped, which allowed its inclusion. Many in the Constellation program held out hope that testing would ultimately be restored when the Agency was past retirement of the Space Shuttle and subsequent funding was freed from this pressure.

As a result, the expected increase in *understanding the risk* was hampered by ever-shrinking test data availability. More problematic was the challenge that most *unknown* risks would have had only a single opportunity to manifest before first crewed flight<sup>8</sup>. This concern is further explored later in this paper. Another problematic aspect of the minimized test plan was that much integrated ground testing was planned with prototype (or even brass board) hardware prior to flight tests, which when coupled with compressed “green light” scheduled sequences that mandated successful results on all objectives of all tests to achieve the IOC date may have risked a higher cost and schedule impact to complete the work than using higher-fidelity test hardware.

### ***The Program’s response to the erosion***

Late in 2009, the Program reacted to the slips and erosion with a proposal that would shift the flight test plan from late, “validation class” testing of ground-qualified systems to earlier, “development class” flight testing of lesser fidelity systems. Several motives prompted this proposed shift: including the desire to recapture objectives that had been lost in the erosion of test flights over time, the desire to shift some of the ground test content to flight test for earlier and more representative feedback that would more readily influence design than later findings could, and the desire to fly earlier for stakeholder support.

After the president’s budget in Feb. 2010 proposed Program cancellation, the Agency management encouraged the Program to continue working on the test plan reformulation to add value to those portions of the Constellation architecture that were thought likely to be used in pursuit of the President’s proposed “beyond low Earth orbit” objectives. Although thoroughly studied, the work did not reach final, actionable conclusions and was never baselined as part of the Program. But including discussion of this proposal is relevant in that it illustrates the efforts to address the lateness and omissions present in the eroded plan. Moreover, this proposal also rejoined the discussion of how much testing was enough before flying the first crew.

The proposal set forth three flight tests to be flown from KSC, with no further change to the plan to fly a second pad abort test and an ascent abort test from WSMR. The three tests were referred to in the study as Flight Tests (FTs)-2, -3, and -4 (indicating the successful *Ares I-X* as having been FT-1). The proposal was based on the premise that in development flight tests, while continuing to design for the eventual final configuration, test article configurations do not have to exactly match the eventual design intended to be certified, and that each flight could be certified for its own unique test flight reference mission based on the specific test objectives and the specific flown configuration. This was expected to be a smaller, “narrower” allowable flight envelope than would be permitted for the eventual end design. These flights were intended to produce data that would allow the empirical validation of models and reduction of uncertainties necessary to progressively increase, or “open up” the allowed flight envelope for subsequent flights.

---

<sup>8</sup> The Space Shuttle history documents well the phenomena of unknown risks “speaking” through the hardware well into operational flight, as documented in “Significant Incidents and Close Calls in Human Spaceflight”, March 8, 2010, version of the rapid information page, PAS-2009-003, SAIC [Science Applications International Corporation], and the Johnson Space Center Flight Safety Office.

- FT-2, would redirect an *Ares* first stage five-segment test article from a planned horizontal test firing to serve as the first stage for the flight test. The high-altitude abort test objectives were allocated to FT-2, and the proposal called for using a production-like *Orion* crew module with those portions of the system executing the abort (such as the LAS and associated avionics) would be as flight-like as practical. FT-2 included a mass and dynamic simulator for the upper stage and service module. FT-2 objectives included first stage flight, flight-like frangible joint separation, recovery of the five segment booster, high-altitude abort, entry descent and landing, and recovery of the crew module.
- FT-3 added a functional upper stage with a qualification J-2X engine, and a partially outfitted production *Orion*. The definition of “partially” in this context was still under study. The FT-3 first stage objectives were identical to FT-2. Staging, engine ignition, and delivery to the required insertion point were among the primary objectives for the upper stage. The *Orion* vehicle was to perform orbital insertion and maneuvering, perform system checkouts, and perform a nominal entry, descent, landing, and recovery.
- FT-4 was to be a full production system with non-critical exceptions. This flight was proposed as the first crewed flight, and orbital flight test objectives included rendezvous and docking with the ISS.

The proposal sought to reduce or eliminate some ground testing by replacing the ground tests with flight tests. Though most of the proposal was found feasible, there were unresolved issues. For instance, the proposal called for elimination of the integrated vehicle ground vibration tests (IVGVTs). Constellation had planned to perform IVGVTs in both the first stage (fully fuelled and depleted), and the second stage configurations. In the new proposal, static, modal, and (where applicable) proof pressure testing remained for each structural element, along with the random vibration testing planned for *Orion*. The proposal also called for a “twang” test similar to what had been done on *Ares* I-X and Apollo to characterize the structural response for validation of guidance algorithms. The proposed elimination was based on the conceptual plan to address the resultant structural load uncertainties by certifying to fly only the flight test profile rather than certifying to fly the full envelope, enabling narrowly “placarded”<sup>9</sup> early test flights. This approach would rely on flight test results to progressively expand the as-certified operating envelope to required parameters. Although this plan was heavily studied, no agreement was reached. Primary objections were twofold: the structural load forcing functions and vehicle responses could not be separated and independently used to validate the models necessary to expand envelopes without what was argued to be “sufficient” ground truth data provided by the IVGVT; and that the very low energy twang test, while sufficient for guidance algorithm validation was not energetic enough to characterize the nonlinear vehicle response to the much higher forcing-function loads expected in flight.

---

<sup>9</sup> Placards are used to indicate limitations in an operational envelope (e.g. limits to speeds, pressures, winds, sea states, etc.). They can be based on known limitations, or used to indicate thresholds beyond which a system has not been tested. Thus narrow “placarding” indicates a narrow window in which the system can safely operate.

The Program was committed to resolving issues such as the IVGVT before committing a crew to fly on FT-4. Yet the debate regarding how much was enough flight testing was renewed by this plan and remained unresolved. There were those who argued not to fly a crew on the first full production vehicle in the design target configuration because it would be the first to have been delivered through simple acceptance tests, manufacturing controls, and assembly checkouts. The essentially complete design would not have the chance reveal the unknown-unknowns without a flight test of a production vehicle.

## **The Unknown Risk—Can You Understand it *a priori*?**

Of particular concern to the defenders of a more robust test plan was the need to provide opportunities for the hardware to reveal problems in the actual operational environment (natural and induced). This was a subject of great debate within the program, and is somewhat addressed by the NESC report in discussion of “repeatability”. Of greatest concern were dynamic events (primarily stage separations and engine starts) in the actual flight environment. While no one argued to fly enough tests to have the traditional statistically significant data set, there was no agreed-to predetermination of how many uncrewed flight demonstrations would provide sufficient confidence to begin crewed flights. In the absence of such, program decisions eroded the flight test manifest with the rationale that the ultimate determination of crewed flight readiness would occur later, closer to the date for first crewed flight. These decisions essentially traded the known risks associated with budget reductions for unknown risks down the road. Such future risks could manifest themselves in multiple ways, from schedule and cost hits (adding tests or redesigning systems) to actual failure in crewed flights, or not at all.

Relevant to this discussion is the fact that a recent analysis of the risk progression (or rather, the progression of risk understanding) for the Space Shuttle found that the risk of losing the crew was underestimated by 2 to 3 orders of magnitude<sup>xii</sup>. Instead of the assumed 1:1,000 to 1:10,000, the reanalyzed risk to the first crew was 1:9. This underestimation of residual risk was a combination of optimistic engineering judgment of only the known risks with insufficient acknowledgement of the unknown.

Figure 2 illustrates the failure history for uncrewed launch vehicles since 1980, plotted from data in Table 1. Only ascent failures from experienced developers that would have resulted in loss of crew or necessitated an ascent abort, had there been a crew aboard the vehicle, were included here. This history, derived from numerous countries with experience in launch vehicle development and manufacturing indicates significant learning between the first two flights of a new system and the next two. While the curve doesn’t begin to level out until around flight seven, risk drops significantly beginning with third flight. The data suggests that one test flight does not “buy-down” as much risk as three or even two test flights.

## Historical Aggregate Failure History Experienced Developers Post-1980

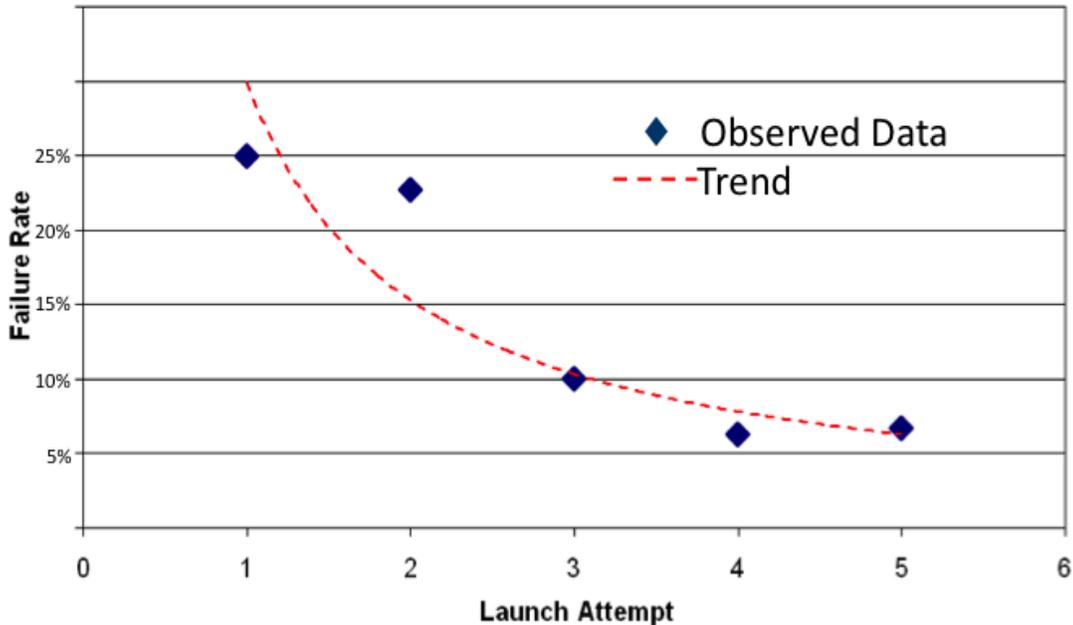


Figure 2: Historical aggregate failure rate for uncrewed launch vehicles built by experienced developers, post - 1980 (Source data Table 1).

The Constellation transportation architecture could be expected to follow a similar curve in reliability growth once integrated flights commenced. This formed the basis of the recommendation to retain the *Ares I-Y* test flight, albeit an incomplete flight test vehicle. Along with an uncrewed *Orion-1* flight test, *Ares I-Y* was believed to help move the integrated launch system further down the risk curve. However, no one could predict *a priori* where (at what level of reliability) the Program would begin its growth curve toward that maturity. Nor could anyone predict the shape of such a curve.

Whatever growth curve Constellation might have tracked (had it pressed toward operational capability), it would likely have been driven by whether previously unknown failure scenarios or interactions between systems (or between systems and environments) manifested themselves during integrated flight tests or operations. (This is based on a reasonable assumption that any known deficiencies would have been assessed and corrected prior to flight.) The uncovering of unknown-unknowns and their correction prior to subsequent flights is what primarily yields reliability growth and the growth curve in the early stages of new space launch systems as illustrated in Figure 2.

We assert, based on Figure 2 and the associated data in Table 1, that the unknown-unknowns cannot be understood *a priori* with the current state of risk analyses, and must be discovered

through well-planned testing. Moreover, the consideration of residual risk must include the potential for unknown-unknowns as related to the number of opportunities the system has been afforded to exhibit them.

**Table 1: Historical success and failure record for uncrewed launch vehicles, experienced developers, post-1980. Data provided by The Aerospace Corporation<sup>vii</sup>. Failures only include ascent failures that would have resulted in loss of crew or necessitated an ascent abort, had there been a crew aboard the vehicle. (LV Fail=launch vehicle failure; TV Fail=transfer vehicle failure )**

	Launch 1	Launch 2	Launch 3	:Launch 4	Launch 5
<b>United States</b>					
Delta I	Success	Success	Success	Success	Success
Titan 34D	Success	Success	Success	Success	Success
Atlas H	Success	Success	Success	Success	Success
Atlas G	TV Fail	Success	Success	Success	Success
Delta II	Success	Success	Success	Success	Success
Titan IV	Success	Success	Success	Success	Success
Pegasus	Success	LV Fail	Success	Success	Success
Atlas I	Success	TV Fail	Success	TV Fail	Success*
Atlas II	Success	Success	Success	Success	Success
Taurus	Success	Success	LV Fail	Success	Success
Pegasus XL	LV Fail	LV Fail	Success	Success	Success
Athena I	LV Fail	Success	Success	Success	
Athena II	Success	LV Fail	Success		
Delta III	LV Fail	LV Fail	Success		
Minotaur	Success	Success	Success	Success	Success
Atlas III	Success	Success	Success	Success	Success
Atlas V	Success	Success	Success	Success	Success
Delta IV	Success	Success	Success	Success	Success
Taurus XL	Success	LV Fail			
<b>USSR/Russia</b>					
Proton	Success	Success	Success	Success	Success
Energia	Success	Success			
Start-1	Success	LV Fail	Success	Success	Success
Shtil	Success	Success			
Rocket KM	Success	Success	Success	Success	Success
Dnepr	Success	Success	Success	Success	Success
Soyuz	Success	Success	Success	Success	Success
Soyuz 2	Success	Success	Success	Success	Success
VOLNA R29R	LV Fail				
<b>China</b>					
CZ-2C	Pre-1980	Pre-1980	Pre-1980	Success	Success
CZ-3	LV Fail	Success	Success	Success	Success
CZ-4	Success	Success			
CZ-2E	Success	Success	LV Fail	Success	LV Fail
CZ-2D	Success	Success	Success	Success	Success
CZ-3A	Success	Success	Success	Success	Success
CZ-3B	LV Fail	Success	Success	Success	Success
CZ-1D	Success				
CZ-2C/SD	Success	Success	Success	Success	Success
CZ-4B	Success	Success	Success	Success	Success
CZ-2F	Success	Success	Success	Success	Success

	Launch 1	Launch 2	Launch 3	:Launch 4	Launch 5
KT-1	LV Fail	LV Fail			
CZ-4C	Success	Success	Success		
<b>ESA</b>					
Ariane 1	Pre-1980	LV Fail	Success	Success	LV Fail
Ariane 3	Success	Success	Success	Success	LV Fail
Ariane 2	LV Fail	Success	Success	Success	Success
Ariane 4	Success	Success	Success	Success	Success
Ariane 5	LV Fail	LV Fail	Success	Success	Success
<b>India</b>					
SLV-3	Pre-1980	Success	LV Fail	Success	
ASLV	LV Fail	LV Fail	Success	Success	
PSLV A	LV Fail	Success	Success	Success	Success
GSLV	LV Fail	Success	Success	LV Fail	Success
<b>Japan</b>					
Mu-3	Success	Success	Success	Success	Success
N-1	Success	Success			
N-2	Success	Success	Success	Success	
H-I	Success	Success	Success	Success	Success
H-II	Success	Success	Success	Success	Success
J-I	Success				
M-5	Success	Success	LV Fail	Success	Success
H-IIA	Success	Success	Success	Success	Success
<b>Ukraine</b>					
Zenit-2	LV Fail	LV Fail	Success	LV Fail	Success
Zenit-3SL	Success	Success	LV Fail	Success	Success

\*Conditional success: payload failed to achieve orbit.

## Conclusions

A sufficient T&E plan *not only queries* the hardware and software to *confirm* that the foreseeable problems that can be identified by requirement verification, design validation, and risk mitigation planning activities have been avoided or eliminated, it *also interrogates* the hardware and software for the unknown-unknown risks that even the best planning will not identify. In other words, the T&E plan must afford opportunities for the system to reveal unforeseen problems. Testing nominal and contingency operational scenarios in representative environments using sufficiently flight-like configurations (test like you fly), is one among several mitigations that provide the appropriate opportunities.

The challenge that remains for future programs is how to balance the quantifiable cost savings achieved by trimming the T&E strategy with the unquantifiable accrual of the intangible risks of things you won't know because the system was not given sufficient opportunity to reveal an underlying problem. We assert that NASA's current human-rating requirements do not adequately address the unexpected things the hardware can reveal to the designer during test, nor does it provide sufficient rationale to sustain a human-rating certification test plan in the face of budget challenges.

T&E objectives associated with safety and survival must be identified and protected in the plan in its earliest conceptual form rather than waiting to identify and then protect them in the completed

strategic-level plan delivered at PDR. A cogent, integrated T&E plan cannot withstand budget challenges if the underlying requirements and rationale (certification criterion, residual risk) are left open to broad interpretation.

NASA must continue to explore this topic more deeply. This must include Constellation's attempt to understand historic failure trends in addition to individual historic failure causes. The tendency is to only address known design risks in test plan formulation and to focus only on verification to a specification. Query of hardware for unknown-unknown risks must be a part of the certification criteria.

The Agency should interpret its current certification requirements to a level such that a program's plan for design, analysis, and test can be clearly seen to either meet or not meet those requirements, particularly in the face of budget pressures. Not doing so raises the real risk that unforeseeable problems will not be discovered, and test data needed to understand residual risk will not be collected for decision makers charged with certifying a new system for human flight.

## References

---

- <sup>i</sup> Rhatigan, J.L., Neubek, D.J., Thomas, L.D., and Stegemoeller, C. *Constellation Program's Lessons Learned, Vol. I: Executive Summary*, NASA SP-2011-6127-VOL-1, June 2011 and –VOL-2, Sept 2011.
- <sup>ii</sup> NASA 2011, “Readiness for First Crewed Flight” NESC-RP-10-00619, April 12, 2011.
- <sup>iii</sup> NASA, NPR 8705.2B, Human-Rating Requirements for Space Systems.
- <sup>iv</sup> O’Conner, B., NASA Associate Administrator for Safety and Mission Assurance, personal communication 2011, and NASA memos to file from 1973, ‘74, and ‘77 documenting discussions and decisions related to the Space Shuttle first flight: manned vs. unmanned.
- <sup>v</sup> Rhatigan, J.L., Hanley, J.M., and Geyer, M.S., *Formulation of NASA’s Constellation Program*, IAC-07-B3.1.06, NASA/SP-2007-563, October 2007.
- <sup>vi</sup> NASA 2003. National Aeronautics and Space Administration *Columbia Accident Investigation Board—Report Volumes I-VI*. August 2003.
- <sup>vii</sup> Tomei, E.J., *Assessment of First 3 vs First 7 Launches, 2008 Update*, presentation by The Aerospace Corporation, 2008.
- <sup>viii</sup> Rhatigan, J.L., Thomas, L.D., and Hanley, J.M., *NASA’s Constellation Program: From Concept to Reality*, IAC-08-B3.1.7, July 2008.
- <sup>ix</sup> Rhatigan, J.L., Hanley, J.M., Thomas, L.D., Boatright, T., “NASA’s Constellation Program: Milestones Toward the Frontier” *International Astronautics Congress*, Daejeon, Korea, October 2009.
- <sup>x</sup> Tomei, E.J. and Chang, I-Shih. *51 years of Space Launches and Failures*, IAC-09-D1.5.1, Oct. 2009.
- <sup>xi</sup> Tomei E.J. and Chang, I-Shih. *U.S Homan Space Transportation Failures*, Trans, JSASS Space Tech. Japan, Vol. 7, No. ISTS 26, 2009.
- <sup>xii</sup> Hamlin, T., et al. “Shuttle Risk Progression by Flight” presentation by the Space Shuttle Safety and Mission Assurance Office, NASA JSC, Feb. 7, 2011.



REPORT DOCUMENTATION PAGE			Form Approved OMB No. 0704-0188	
Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden, to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.				
1. AGENCY USE ONLY (Leave Blank)	2. REPORT DATE October 2011	3. REPORT TYPE AND DATES COVERED NASA Technical Memorandum		
4. TITLE AND SUBTITLE First Crewed Flight: Rationale, Considerations & Challenges from the Constellation Experience			5. FUNDING NUMBERS	
6. AUTHOR(S) Carlos Noriega; William Arceneaux; Jeffrey A. Williams; Jennifer L. Rhatigan Formerly of the Constellation Program Office NASA Johnson Space Center, Houston, TX				
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) Lyndon B. Johnson Space Center Houston, Texas 77058			8. PERFORMING ORGANIZATION REPORT NUMBERS S-1108	
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES) National Aeronautics and Space Administration Washington, DC 20546-0001			10. SPONSORING/MONITORING AGENCY REPORT NUMBER TM-2011-216161	
11. SUPPLEMENTARY NOTES				
12a. DISTRIBUTION/AVAILABILITY STATEMENT Available from the NASA Center for AeroSpace Information (CASI) 7121 Standard Hanover, MD 21076-1320                      Category: 16			12b. DISTRIBUTION CODE	
13. ABSTRACT (Maximum 200 words) NASA's Constellation Program has made the most progress in a generation toward building an integrated human-rated spacecraft and launch vehicle. During that development, it became clear that NASA's human-rating requirements lacked the specificity necessary to defend a program plan, particularly human-rating test flight plans, from severe budget challenges. This paper addresses the progress Constellation achieved, and problems encountered, in clarifying and defending a human-rating certification plan, and discusses key considerations for those who find themselves in similar straits with future human-rated spacecraft and vehicles. We assert, and support with space flight data, that NASA's current human-rating requirements do not adequately address "unknown-unknowns," or the unexpected things the hardware can reveal to the designer during test.				
14. SUBJECT TERMS spacecraft, launch vehicles, flight tests, safety factors, risk, risk assessment, Constellation Program			15. NUMBER OF PAGES 30	16. PRICE CODE
17. SECURITY CLASSIFICATION OF REPORT Unclassified	18. SECURITY CLASSIFICATION OF THIS PAGE Unclassified	19. SECURITY CLASSIFICATION OF ABSTRACT Unclassified	20. LIMITATION OF ABSTRACT Unlimited	



---