Program Promotion Can Distort Space Systems Engineering and Deny Risk

Harry W. Jones[1]
NASA Ames Research Center, Moffett Field, CA, 94035-0001

NASA's spectacular success in the Apollo moon landings was achieved against the odds by an obsessive dedication to reducing the great risk. However, risk analysis predicted so many astronaut fatalities that it was thought to be unreasonably pessimistic and potentially damaging to the Apollo program. Risk analysis was discontinued, risk was neglected in space shuttle engineering, and so the space shuttle design was unnecessarily dangerous. Since the Apollo era it has been understood that long human space missions would recycle oxygen and water to avoid the very high launch cost of directly supplying them. The development of recycling systems was justified by the need to increase material closure and reduce launch mass. When it was recognized that increasing closure leads to rapidly diminishing returns, the program goal was changed to reducing launch mass and reliability, cost, and risk were considered irrelevant. Systems engineering and especially the discouraging problems of risk and cost have been deliberately ignored because they detract from program promotion, with unfortunate results. Current human launch system design does account for risk and the result strongly resembles Apollo. Current life support design continues to assume recycling, even though the recent great reduction in launch cost now allows direct supply of oxygen and water with significantly better quality, reliability, cost, and risk.

Nomenclature

CAIB = Columbia Accident Investigation Board
ECLSS = Environmentally Controlled Life Support System
ESM = Equivalent System Mass
ISS = International Space Station
LCC = Life Cycle Cost
PRA = Probabilistic Risk Assessment

I. Introduction

THIS paper suggests that the standard systems engineering process should be used more thoroughly and completely in space systems engineering. Two cases are described where harm has resulted from avoiding directly dealing with systems engineering issues, specifically including risk and cost.

Risk analysis predicted that Apollo would suffer many fatalities. This awareness and the Apollo 1 fire focused the program on eliminating risk as much as possible. Unfortunately, risk analysis was discontinued to avoid damaging support for Apollo. The spectacular success of Apollo created over confidence and encouraged disregarding risk in the space shuttle design. The shuttle was unnecessarily dangerous because it placed a fragile spacecraft next to the fuel tanks and did not provide a launch abort system. These design decisions directly caused the Challenger and Columbia tragedies and were reversed in the current rocket and capsule design.

Early missions such as Apollo used tanks supplying oxygen and water, but it seemed obvious that longer missions would need recycling because of the large mass and high launch costs of oxygen and water. Recycling systems were tested in the 1960's and now fly on the International Space Station (ISS). Life support research emphasized increasing closure and reducing system mass, and paid little attention to reliability, cost, and risk. Now that launch cost has been greatly reduced, it seems that more effort should have been spent on reducing cost and risk and less on reducing mass.

[1] Systems Engineer, Bioengineering Branch, Mail Stop N239-8.

While investigating why systems engineering was not fully utilized in life support research and space shuttle design, a surprising and worrying observation was made. Advocating a program requires emphasizing its benefits while not dwelling on its problems, but this can go too far, so that well known problems are not addressed. Program promotion can distort space systems engineering and deny risk.

II. Introductory overview of risk in Apollo and shuttle

1. The risk of failure in Apollo was thought to be very large.
2. The Apollo 1 fire concentrated effort on reducing risk.
3. Risk analysis was thought too negative, damaging to Apollo, and was discontinued.
4. Apollo 11 successfully landed on the Moon.
5. Shuttle was designed assuming risk could be designed out and would be very low.
6. Shuttle was fragile, unsafe, and high risk.
7. After the Challenger loss, renewed risk analysis found the chance of failure was about 1 in 100.
8. Shuttle flights continued at a much reduced rate with the hope that care could beat the odds.
9. After the Columbia loss it was recognized that shuttle could not be made safe and had to be replaced.
10. The replacement has a hardened crew capsule above the rocket boosters with a launch abort system, as in Apollo.
11. Program promotion distorted systems engineering and denied risk.

III. Apollo

The Apollo moon landing was a spectacular success and is a major space milestone. The Apollo program was a success, but not a perfect success. Apollo 1 was a tragedy with an amazingly negligent cause, that the possibility of a fire was simply dismissed. Apollo 13 was a close call that demonstrated the high risk inherent in complex systems. The last three Apollo flights were cancelled and Apollo achieved only six of the ten planned moon landings.

Joseph Shea, the Apollo program manager, chaired the initial Apollo systems architecting team. A participant reported that "calculation was made by its architecting team, assuming all elements from propulsion to rendezvous and life support were done as well or better than ever before, that 30 astronauts would be lost before 3 were returned safely to the Earth. Even to do that well, launch vehicle failure rates would have to be half those ever achieved and with untried propulsion systems."[1]

The awareness of risk led to intense focus on reducing risk. "The only possible explanation for the astonishing success – no losses in space and on time – was that every participant at every level in every area far exceeded the norm of human capabilities."[1]

However, this understanding of the risk was not considered appropriate for the public. During Apollo, NASA conducted a full Probabilistic Risk Assessment (PRA) to assess the likelihood of success in "landing a man on the Moon and returning him safely to Earth." The PRA indicated the chance of success was "less than 5 percent." The NASA Administrator felt that if the results were made public, "the numbers could do irreparable harm." The PRA effort was cancelled and NASA stayed away from numerical risk assessment as a result.[2]

A. Apollo 1

During a simulated flight conducted in the Apollo 1 capsule on the launch pad, the astronauts reported a fire. The three astronauts died from smoke and flames before escape or rescue was possible. Joseph Shea recalled a fire discussion a few months before, "I got a little annoyed, and I said, 'Look, there's no way there's going to be a fire in that spacecraft unless there's a spark or the astronauts bring cigarettes aboard.'"[3] "Deeply involved in the investigation of the 1967 Apollo 1 fire, Shea suffered a nervous breakdown as a result of the stress that he suffered. He was removed from his position and left NASA shortly afterwards."[4]

The NASA administrator established a review board which found that:

> "The fire in Apollo 204 was most probably brought about by some minor malfunction or failure of equipment or wire insulation. This failure, which most likely will never be positively identified, initiated a sequence of events that culminated in the conflagration. Those organizations responsible for the planning, conduct and safety of this test failed to identify it as being hazardous. … The Command Module contained many types and classes of combustible material in areas contiguous to possible ignition sources. … The Command Module Environmental Control System design provides a pure oxygen atmosphere. ... This atmosphere presents severe fire hazards."[5]

The review board recommended that NASA continue the program to the reach the moon by 1969, but make safety more important than schedule.

The cause of the Apollo 1 failure was a failure to anticipate a known hazard. Astronaut Frank Borman, on the NASA review board, stated "none of us gave any serious consideration to a fire in the spacecraft."[5] The Apollo 1 fire was unexpected, unpredicted even though several fires in other pure oxygen atmospheres had caused deaths.

Later spacecraft designs used Earth normal atmosphere, considered the combustibility of materials, and developed capabilities and procedures for escape and rescue.

After the tragedy of the Apollo 1 fire, the reliability of Apollo was made central by an engineering culture that encouraged an environment of open communications, attention to detail, and ability to challenge technical assumptions. "Anyone could challenge a design at any time. … Reliability was a concern at all levels."[6]

B. The Apollo 11 success

Apollo 11 successfully landed on the moon in 1969. A major factor in the success of Apollo was the extreme attention paid to reliability and crew safety. The policy was to speak and to listen, to always bring up issues that were not fully understood. The Apollo success showed that by intense effort to reduce risk, a dedicated organization can achieve results far beyond reasonable expectation.

Unfortunately, this amazing success led to extreme overconfidence. The head of Apollo reliability and safety decided, "Statistics don't count for anything," and that risk is reduced by "attention taken in design." This attitude was carried forward from Apollo to shuttle. A NASA safety analysis for Galileo explained that shuttle "relies on engineering judgment using rigid and well-documented design, configuration, safety, reliability, and quality assurance controls." It was also thought that, with the attention given to safety and reliability, "standard failure rate data are pessimistic."[2]

IV. Shuttle

The space shuttle transported cargo and crew to orbit from 1981 to 2011. There were 133 successful missions and two tragic failures. Things were different than Apollo.

The Apollo program had much stronger congressional support than later manned space programs. "Administrator Webb did not have to 'sell' the program to his political overseers with exaggerated claims or by acquiescing to unrealistic budget compromises." Later programs came much closer to being cancelled.[7, 8]

Now NASA shuttle program advocacy over-optimistically promised rapid turnaround, frequent flights, and lower launch costs. It would pay for itself launching satellites and carrying out science experiments. The most disastrous error was a blind unreasonable belief that the shuttle was safe.

A. Denying risk in shuttle

The initial design of the shuttle emphasized capability and cost without serious consideration of risk. "During the early shuttle studies, there was a debate over the optimal shuttle design that best balanced capability, development cost, and operational cost."[9] A retired NASA official stated, "some NASA people began to confuse desire with reality. … One result was to assess risk in terms of what was thought acceptable without regard for verifying the assessment. … Note that under such circumstances real risk management is shut out."[2]

Shuttle risk was generally neglected. "Although every knowledgeable observer recognized that there was some potential for a major shuttle failure, the press and the broader public in the early 1980s paid little attention to the risks of human spaceflight. Even those close to the shuttle system let down their guard."[10]

Quantitative mission and system wide PRA was avoided. System level space shuttle risk assessments were all qualitative. They included preliminary hazards analysis, failure modes and effects analysis with critical items lists, and various safety assessments. There was some quantitative analysis conducted for specific subsystems.[2] The space shuttle requirements for safety were too simplistic. Subsystems were to "fail-operational after the failure of (the) most critical component" and to "fail-safe for crew survival the failure of (the) two most critical components."[11]

Safety was simply assumed rather than designed into the shuttle.

> "The final shortcoming was that the Shuttle was designed as if it had the inherent operating safety of an airliner. It was not equipped with any provision for crew rescue in case of booster failure during ascent to orbit, or being stranded in orbit, or structural failure during re-entry. The crew was not even provided with spacesuits, despite the lessons of the Soviet space program. This seemed an extraordinary act of engineering hubris, given that contemporary military aircraft were equipped with pressure suits and ejection seats. But the weight problem also meant that there was no margin for crew safety measures without (to NASA) unacceptable impact to the net payload. … Following the Columbia disaster, NASA finally realized it could not make the Shuttle safe. The only way to continue American manned spaceflight would be to develop a replacement manned spacecraft with an escape system, and meanwhile fly the Shuttle as little as possible."[12]

B. Over promising cost performance in shuttle

NASA overpromised what the shuttle would achieve. Shuttle would perform all NASA, military and commercial launches and the high number of launches would make it cost-effective.

> "In keeping with agency survival instincts Fletcher and others engaged in political hype to sell their program. … Fletcher initially quoted sixty flights a year (with full payload), an utterly unrealistic figure but politically essential if human space flight was to survive and compete with expendable launch vehicles for cost savings, and even turn a profit. Although the flight rates were reduced to fifty, agency personnel were still being asked to support a mythical figure. In the words of one subordinate, 'We had to argue that [the shuttle] was cheaper. It would be cheaper than all the expendable launch vehicles. It would be better than all the expendable launch vehicles. Well, there was a feeling that we were on the razor's edge. That if we said the wrong thing, or anything like that, the shuttle would be killed.' "[13]

NASA never came close to achieving such launch numbers but before Challenger, NASA "would not admit to problems matching resources to launch rates."[13]

C. Shuttle PRA's were distorted or disregarded

Contrary to NASA's general approach, two PRA's were externally compelled for shuttle. PRA was required because shuttle was to be used to launch Galileo to Jupiter, and Galileo contained plutonium in a thermonuclear generator which could be dispersed by an accident. The first contractor study done for NASA very misleadingly reported that the risk of losing a shuttle during launch was between 1 chance in 1,000 and 1 in 10,000. In fact, it found a much greater risk in the solid-fuel rocket boosters, which had a failure rate of about 1 in 40. However, rather than use this historical data, the NASA sponsor made an "engineering judgment" and "decided to assume a failure probability of 1 in 1,000" or even 1 in 10,000.[2]

A second study for the Air Force noted that the earlier study involved both gathering failure data "and the disregarding of that data and arbitrary assignment of risk levels apparently per sponsor direction" with "no quantitative justification at all." After reanalyzing the data, the study found that the boosters' track record "suggest[s] a failure rate of around one-in-a-hundred."[2]

NASA's internal analysis also minimized safety concerns. NASA Johnson Space Center conducted its own internal safety analysis for Galileo in 1985. The Johnson authors went through failure mode worksheets assigning probability levels. A failure in the solid rocket booster (the failure that destroyed Challenger) was assigned a probability of 1 in 100,000.[2]

Even after the Challenger accident, the NASA chief engineer in a hearing on the Galileo thermonuclear generator said, "We think that using a number like 10 to the minus 3, as suggested, is probably a little pessimistic." He thought the actual risk "would be 10 to the minus 5." The number was derived "based on engineering judgment."[2]

D. Challenger

The Challenger broke up at 73 seconds into flight when an O-ring in the right solid rocket booster failed and allowed a flare to reach the external fuel tank, which separated causing aerodynamic forces that disintegrated the shuttle. The crew cabin hit the ocean at unsurvivable speed at 2 minutes and 45 seconds after the breakup.

NASA's internal investigation was initially conducted in secrecy and was suspected of not presenting relevant information. The presidentially appointed Rogers Commission identified failure causes in NASA's management culture and decision-making processes.

> "testimony reveals failures in communication that resulted in a decision to launch (Challenger) based on incomplete and sometimes misleading information, a conflict between engineering data and management judgments, and a NASA management structure that permitted internal flight safety problems to bypass key shuttle managers."[14]

The flaw in the O-ring design and the potential for flare blow-by had been known for many years but had been ignored and the risk improperly minimized. This has been labeled "normalization of deviance."[15] Before the flight, engineers had warned about the danger of launching in much colder than previously experienced temperatures.

The Nobel physicist Richard Feynman provided "Personal Observations on Reliability of Shuttle" as an appendix to the Rogers Commission report on the Challenger accident.

> "It appears that there are enormous differences of opinion as to the probability of a failure with loss of vehicle and of human life. The estimates range from roughly 1 in 100 to 1 in 100,000. The higher figures come from the working engineers, and the very low figures from management. …An estimate of the reliability of solid rockets was made by the range safety officer, by studying the experience of all previous rocket flights. Out of a total of nearly 2,900 flights, 121 failed (1 in 25). ... NASA officials argue that the figure is much lower. They point out that these figures are for unmanned rockets but since the Shuttle is a manned vehicle 'the probability of mission success is necessarily very close to 1.0.' … It would appear that, for whatever purpose, be it for internal or external consumption, the management of NASA exaggerates the reliability of its product, to the point of fantasy."[14]

The neglect of O-ring and tile damage, the normalization of deviance, and the fantastic exaggeration of shuttle reliability are well documented.[14, 15, 16] There is an even more obvious but usually unmentioned reason for the failures of shuttle. The space shuttle simply was not designed to minimize risk. Unlike the hardened Apollo capsule, the shuttle crew compartment was fragile, unlike the Apollo command module, the shuttle crew compartment was

next to rather than above the dangerous rockets, and unlike Apollo, the shuttle had no launch abort system. These design errors can be considered the fundamental causes of the Challenger and Columbia accidents. These early basic design errors have been deemphasized in favor of blaming operational people who by extraordinary capability might have beaten the bad odds. These design errors are implicitly acknowledged by the fact that NASA's post shuttle rocket and crew vehicle designs replicate the Apollo approach. In response to the Rogers Commission's recommendations, NASA redesigned the solid rocket boosters and created a new Office of Safety, Reliability and Quality Assurance reporting directly to the administrator.

In her investigation of the Challenger disaster, Diane Vaughan found that, because of difficult goals and limited resources, NASA's Apollo safety culture became a "culture of production" that emphasized productivity, efficiency, obeying orders and following rules rather than problem solving or concern about safety. The result was "the normalization of deviance," the acceptance of what should have been alarming indications of incipient failure. Blocked communications, Vaughan's "structural secrecy," prevented effective action.[15]

Initial qualitative assessments of shuttle reliability were based on expert judgment rather than reliability analysis. After Challenger, PRA was adopted and applied to the space shuttle, space station, and some unmanned space missions. NASA then developed realistic estimates of the probability of space shuttle failure, roughly 1 in 100.[17]

E. Columbia

The Columbia astronauts perished when the shuttle heat shield failed on reentry. The Columbia Accident Investigation Board (CAIB) reported:

> "The physical cause of the loss of Columbia and its crew was a breach in the Thermal Protection System on the leading edge of the left wing, caused by a piece of insulating foam which separated from the left bipod ramp … and struck the wing ... During re-entry this breach in the Thermal Protection System allowed superheated air to penetrate through the leading edge insulation and progressively melt the aluminum structure of the left wing, resulting in … break-up of the Orbiter. This breakup occurred in a flight regime in which, given the current design of the Orbiter, there was no possibility for the crew to survive. … The organizational causes of this accident are rooted in the Space Shuttle Program's history and culture, including the original compromises that were required to gain approval for the Shuttle, subsequent years of resource constraints, fluctuating priorities, schedule pressures, mischaracterization of the Shuttle as operational rather than developmental, and lack of an agreed national vision for human space flight. Cultural traits and organizational practices detrimental to safety were allowed to develop, including: reliance on past success as a substitute for sound engineering practices (such as testing to understand why systems were not performing in accordance with requirements); organizational barriers that prevented effective communication of critical safety information and stifled professional differences of opinion; lack of integrated management across program elements; and the evolution of an informal chain of command and decision-making processes that operated outside the organization's rules."[16]

The physical cause of the Columbia tragedy was identified and it was noted that the shuttle design provided no crew escape and no possibility for the crew to survive. The CAIB's emphasis was on the organizational practices detrimental to safety, the barriers that prevent communication of critical safety information, the lack of integrated management, and the informal chain of command that were immediate contributors to the failure. The goal of the prescribed independent program technical authority, the independent safety assurance organization, and the learning organization culture is to "more safely and reliably operate the inherently risky Space Shuttle."

The CAIB found that the post-Challenger changes in NASA management and culture were ineffective.

> "(T)he Rogers Commission … recommendations centered on an underlying theme: the lack of independent safety oversight at NASA. … NASA's response to the Rogers Commission recommendation did not meet the Commission's intent: the Associate Administrator did not have direct authority, and safety, reliability, and mission assurance activities across the agency remained dependent on other programs and Centers for funding."[16]

The CAIB believed that Columbia and Challenger were both lost because of similar failures in NASA's organizational system. "(T)he causes of the institutional failure responsible for Challenger have not been fixed."[16] NASA during Apollo had a good safety culture but lost it before Shuttle. NASA had lost the ability to recognize and repair threats that were obvious in hindsight.[18] The CAIB concluded, "The Shuttle is now an aging system but still developmental in character. It is in the nation's interest to replace the Shuttle as soon as possible."[16]

F. Neglect of risk in Apollo and shuttle

Apollo was a fantastic success despite high risk because of it made a great effort to reduce risk. However, risk analysis was then discontinued to avoid the potential political damage caused by awareness of its high risk. Shuttle was designed assuming good engineering could eliminate risk, which allowed `an intrinsically unsafe approach. The cause of the shuttle tragedies was that misguided program promotion prevented risk analysis and led to the unconscious acceptance of risk.

The Challenger tragedy is an often used example of management failure, but the key top level management error is never discussed. The short term focus is on the Challenger launch decision, the last minute failure of

communication, the inability of engineers to have their O-ring concerns heard and acted on. The longer term organizational cause is the gradual "normalization of deviance," when safety issues gradually became neglected due to a production culture and short launch schedules.[15, 16] Management readily accepts the need to emphasize safety and to lead a safety culture change, but seems unaware that misleading program promotion can impair systems engineering and lead to pervasive unawareness of risk.

The risk-negligent design of the shuttle produced a system that was too dangerous. Unlike the hardened Apollo capsule head shield, the shuttle crew compartment used fragile tiles, unlike the Apollo crew module, the shuttle crew compartment was next to rather than above the dangerous rockets, and unlike Apollo, the shuttle had no launch abort system. These design errors directly led to the Challenger and Columbia accidents. The current rocket and crew vehicle designs are similar to the safer design configuration of Apollo, with a hardened crew capsule, the crew capsule above the rocket and fuel, and a launch abort system.

The space shuttle was not designed for minimum risk because risk was misunderstood during its design. The ultimate cause of the shuttle tragedies was the choice by the Apollo-era NASA administrator to discontinue risk analysis to avoid damaging the Apollo program.

V. Introductory overview of space life support

1. Early brief human space missions used Earth resupplied and stored oxygen and water.
2. Recycling would be needed for long missions because of large resupply masses and high launch costs.
3. The classic recycling system was tested with humans in closed chambers in the 1960's.
4. A similar system was developed in the 1990's, flown to ISS in the 2000's, and continues in use.
5. Research focused on increasing closure and reducing system mass, emphasizing recycling's mass savings.
6. Management deemphasized reliability, cost, and risk - major recycling weaknesses compared to resupply.
7. Great reductions in launch cost remove the need to increase closure and reduce mass, favoring resupply.
8. Program promotion distorted systems engineering and caused disregard of reliability, cost, and risk.

VI. Recycling life support justifications

The promotion of physical/chemical recycling life support emphasizes the need for high closure and reducing the mass of recycling equipment while deemphasizing reliability, cost, and risk.

A. Setting high closure as the main goal of life support

The system closure metric for life support is the percentage of all life support material - oxygen, water, food, and other supplies for the crew - that is provided by recycling rather than supplied from Earth. Because the mass of oxygen, water, and other supplies increases for longer missions, more recycling and increasing closure are needed for long duration missions. However, limitlessly increasing closure becomes less cost effective because of diminishing returns.

After development of the current space station recycling systems, life support research was justified as needed to increase closure. The percent closure was used as a planning and progress metric. The goal of life support research and development was to approach a totally closed human ecosystem, independent from Earth. The future life support system for the Moon and Mars was expected to include food production, waste recycling, and ultimately to be "totally closed except for losses due to leaks, EVA's, etc.," and to approach "complete closure of the food and solid waste loops."[19, 20] "The goal for these (Moon and Mars) missions is a higher level of mass recovery, perhaps achieving 95% closure."[19]

However, diminishing returns means that achieving higher closure requires using technologies with lower cost-benefit. The most abundant and easiest to treat wastes, such as condensed atmosphere humidity or hygiene water, are recovered first. Methane produced from exhaled carbon dioxide or human solid waste are more difficult to recycle and constitute a much smaller part of the total spacecraft material circulation. Growing food is obviously needed for full closure but requires a large greenhouse, high power for lights, and massive supporting equipment. Since the easiest resources are exploited first, increasing closure by recycling more material using more equipment always has diminishing returns.

The closure metric can be used to justify developing uneconomic technology for recovering difficult minor wastes and for growing food plants in space. Percent closure was replaced as the major metric for life support research and development by Equivalent System Mass (ESM) in the late 1990's. Although ESM is still used as a guiding metric, increasing closure remains the main goal and is used to justify difficult and costly recycling efforts.[21,

It is important to recognize that life support analysis and projects in addition to considering closure usually include standard systems engineering goals such as reducing system and logistics mass, improving reliability, maintainability, and supportability, and reducing risk. Some technology development projects or system research areas may have no relation to closure.[21, 42] Closure is a very attractive goal for space life support. Many people are concerned about unsustainability in Earth's ecology and the need for much more recycling. Closing the loop in space habitats creates ecological awareness and can be expected to produce technologies useful on Earth. Standard practical systems engineering is focused on reducing cost, launch mass, system complexity, and risk. Increasing closure has increasing costs and diminishing returns, so that careful optimizing of closure is a practical way to incorporate this goal.[42]

B. Setting Equivalent System Mass (ESM) as the metric for life support

Equivalent System Mass (ESM) was established as the guiding life support metric in the late 1990's. Closure was criticized as impractical due to diminishing returns and being unable to meet the need for a metric to guide cost saving research. ESM is the total launch mass needed to provide and support a system, including its mass, volume, power, cooling, and materials and spares logistics. The mass equivalent of the crew time to operate and maintain the hardware is added.[22] ESM was selected as the basis of the life support progress metric.[23] ESM was used in life support technology selection.[24]

Life support research and development previously aimed at reducing resupply mass for long missions by increasing closure. However, increased closure reduces cost only if it reduces the launch mass and launch cost. Reducing ESM directly reduces the launch cost. Reducing ESM tends to counteract the diminishing returns of increasing closure, since increasing closure tends to add hardware mass and so increase ESM.

ESM reflects only launch cost and neglects hardware development cost and operations cost. Life support research and development project planning should consider the entire Life Cycle Cost (LCC). LCC includes development, launch, and operations costs.[25] ESM emphasizes the launch mass reduction benefit of recycling systems. Supplying all materials has very high launch mass and ESM but very low development cost for the materials and their containers. Recycling systems have much lower launch mass and ESM, but they have very high development cost for the recycling systems themselves. The LCC of life support recycling systems can be much higher than the LCC of storage systems.[26, 27] Recently much lower launch cost has reduced the portion of LCC due to ESM, and has made resupply relatively more attractive than recycling.[28, 29]

C. Suggestion that higher reliability is not the highest priority in life support

It is obvious that that Mars life support needs much higher reliability than the current ISS system since a Mars mission lacks the options of rapid resupply or crew return if needed. Increased focus on reliability was opposed by three different groups. Within life support research, the extensive analysis effort was largely restricted to ESM, which alone formally guided decisions. Reliability work was not included in ESM or life support. The Constellation moon base planners paid detailed attention to reliability, since it affects the number, mass, and cost of spare parts. However, they realized that the moon base was similar to ISS since it had the options of rapid resupply and crew return. The moon base could be kept operating safely with systems having the same lesser reliability as those on ISS. ISS life support design engineers explained that they followed the prescribed reliability procedures, did as well as possible given the resources available, and have provided a working system.

Not having higher reliability is acceptable when it is not needed, but it is needed for Mars. It is understood that life support reliability needs more attention. "(N)o consensus has been reached on what is meant by improving on reliability, or on how to assess reliability within the AES (Advanced Exploration Systems) projects."[30] A paper assessing the potential of ISS life support for Mars observes that, "With several readily apparent exceptions, …. equipment has been shown to be capable of achieving operational lifetimes on the order of those needed to support such missions."[31] It is helpful that most equipment could last through A Mars mission, but the "several readily apparent exceptions" are a serious problem. The ISS life support engineers have argued that ISS life support can be "refined" for use for Mars[31] but this seems excessively optimistic.[32] Some reliability analysis mistakenly assumes that all failures can be repaired using spares.[33, 34] Recently the goal of increasing life support reliability has become prominent. Our goal is often cited as "increasing life support reliability and closure." While closure has been, off and on, a goal for decades, reliability is more recent and has less accumulated experience. However, there is a growing body of useful documented research. The recent focus on the moon before Mars may reduce the urgency of the higher reliability needed for Mars, and moon and Mars life support reliability development should be coordinated.

D. Suggestion that cost analysis would be damaging to life support

Cost like reliability is an important consideration in systems engineering. Examining cost has been strongly discouraged in life support planning and only very limited work has been done. The use of ESM favors recycling life support over resupply because of its emphasis on launch mass and its omission of recycling's higher development and operational costs. For recycling life support systems, the development and operations cost is usually much larger than the launch cost. The generally accepted NASA cost metric is Life Cycle Cost (LCC), which includes development, launch, and operations costs.[25]

One of the reasons for establishing ESM was to institutionalize a method for avoiding cost analysis in life support. To quote the NASA ESM guidelines, "ESM is typically used as a transportation cost measure in ALS (Advanced Life Support) trade studies, to avoid the complications, both technical and political, of using dollar costs for comparisons."[35] ESM is the favored alternative to cost. "Cost would be a superior metric if we had the data, but flight hardware cost estimates are typically derived from mass during early development. Furthermore, dollar cost is overly dependent on timeframe and fiscal issues, such as the cost of money, and rapidly becomes politicized."[36]

The cost of recycling life support development and operations and its continuing support can be very large. "The total cost of the ISS life support seems to be about 2 billion current dollars, one billion for hardware development, and one billion for launch and operations."[37] Neglecting cost, like neglecting reliability, helps promote recycling but seriously distorts life support engineering.

E. Suggestion that a system similar to ISS life support can be used for Mars

The idea that space station life support can be used for Mars has often been suggested, with the reservation that improvements are needed in closure and reliability. "The current ISS regenerative air and water systems form the basis for these future architectures. Key improvements such as increased reliability and additional loop closure will be needed for future missions."[38] "(E)volving systems used successfully aboard the International Space Station (ISS) to realize gains in reliability, logistics reduction, and resource recovery is a leading technical approach."[39]

It has been unconvincingly suggested that ISS life support can become sufficiently reliable for long duration missions.[31] The hope for improvement assumes the usual reliability growth process of testing, operating, investigating failures, redesigning to fix problems, and continuing until a satisfactory design is obtained. This often occurs in systems development, but it has not happened and cannot be expected for the ISS Environmentally Controlled Life Support System (ECLSS). The ISS ECLSS did not have significant reliability growth its first decade.[40] The ISS ECLSS exists in only one proto-flight version in orbit. Trouble shooting, devising improved designs, and then testing them on ISS is very difficult. There is a more fundamental problem. The ISS ECLSS was not designed to have high reliability, which is needed for Mars but not Earth orbit or the moon.

Another obvious difficulty in using space station life support for Mars is that the mass payback of recycling systems will be much less for Mars than for space station, because Mars is a much shorter mission. Space station had a planned ten-year life and will operate longer. The recycling systems can be quite massive and yet pay back 10 or 20 times their mass in water and oxygen over the ten-year mission. However, Mars transit round trip and Mars surface missions are only about a year and a quarter long. A system that pays back 10 or 20 times its mass in ten years will pay back only 1 or 2 times its mass in one year. Recycling systems similar to those on the ISS will not save significant mass for Mars.[26, 28]

Space station life support is designed for space station requirements, which are significantly different from Mars transit and Mars surface requirements.[41] Evolving space station life support for Mars seems implausible.

F. The life support objective is a reliable closed loop system based on ISS

NASA's life support program goal is to develop a highly reliable closed loop life support system based on ISS.

> "NASA … strives to develop … (h)ighly reliable, closed-loop life support systems … (s)tarting with the International Space Station (ISS) LSS systems." "Starting with ISS systems as a point of departure, where applicable, the project is evolving these systems and developing new systems to be smaller, lighter and more reliable and to further close the water and air loops to reduce the consumable mass needed."[21]

The current approach is upgrading the ISS life support and testing improved or new systems on ISS. Future work will be needed develop the actual deep space life support.

> "The NASA ECLSS community currently anticipates the development of the final exploration ECLSS for deep space missions through a "Two Build" development effort. The first build would include the demonstration of new systems as well as multiple upgrades and enhancements to current systems on ISS. These new systems and enhancements would be designed to enhance closure and improve maintainability. … The second build would be the actual exploration ECLSS used to support deep space missions and would be installed on the first long duration human habitat vehicle."[42]

But ISS systems will be used "where applicable" and "new systems" and a different "actual" deep space life support system may be needed. This suggests that "highly reliable closed loop life support systems" based on ISS

may not be sufficient or even appropriate. High closure was a reasonable goal when launch costs were much higher and were significant in life cycle cost, but this is clearly no longer true.[29] Earth-supplied material is less costly and more reliable than recycling systems.[28] In many cases there is no need for "highly reliable" life support. The crew can be kept safe with a less reliable system if there are redundant systems, resupply fall backs, safe havens, and escape options. The focus on the moon rather than Mars has reduced the need for higher reliability. A better life support engineering goal would be to support the current mission safely and efficiently without constraining the approach. Systems engineering would then start top-down, from requirements, through trade-offs, to a cost-effective design that considered reliability, cost, risk, and more. If instead the design is predefined, systems engineering can be considered distracting and damaging and so is discouraged, distorted, and disregarded.

VII. Conclusion

The motivating problem for this work was that life support does not usually rely on standard systems engineering methods. The systems engineering processes attempt to manage systems problems such as requirements, reliability, cost, and risk. Ignoring them can easily do harm.

Challenger was examined for insights. The cause of the shuttle tragedies was the disregard of risk ultimately producing an unsafe design. Similarly, the cause of life support misdirection was disregard of cost in favor of high closure and low launch mass. In a mirroring way, shuttle misrepresented cost savings while life support has sometimes downplayed reliability and risk.

Positive project promotion prevents the open discussion of potentially negative systems engineering issues. In an open agency like NASA, the entire program must support the same point of view, so that disregarding problems means that engineers are blocked from realistically dealing with issues such as cost and risk. The resulting systems can be surprisingly costly and risky.

Shuttle was unable to acknowledge performance, cost, and risk problems because of threats of cancellation. But even Apollo feared to publish its risk estimates and made the fateful decision to drop risk analysis, which led to the unsafe shuttle design.

What can be done? If all manned space was placed under the military, the cost, risk, and other systems issues could be accepted and worked on without inhibiting public scrutiny. As far as manned space remains civilian, good systems engineering should be enforced with public analysis and critical reviews. If a program cannot be accepted based on a truthful description, it should not be supported.

References

[1] Rechtin, E., *Systems Architecting of Organizations*, CRC Press, Boca Raton, 2000.

[2] Bell, T. E., and Esch, K., "The Challenger Disaster: A Case of Subjective Engineering," Jan. 28, 2016 (June 1989), https://spectrum.ieee.org/tech-history/heroic-failures/the-space-Shuttle-a-case-of-subjective-engineering, accessed July 24, 2018.

[3] Shea, J. F., Edited Oral History Transcript, NASA Johnson Space Center Oral History Project, 1998, https://www.jsc.nasa.gov/history/oralhistories/SheaJF/SheaJF11-23-98.htm, accessed Feb. 2, 2018.

[4] nasa.wikia, Joseph Francis Shea, NASA Johnson Space Center Oral History Project Biographical Data Sheet, 2006, http://nasa.wikia.com/wiki/JosephFrancisShea, accessed Feb. 6, 2018.

[5] Benson, C. D., and Faherty, W. B., "Moonport: A History of Apollo Launch Facilities and Operations," NASA Special Publication-4204 in the NASA History Series, 1978. http://www.hq.nasa.gov/office/pao/History/SP-4204/contents.html

[6] Oberhettinger, D., NASA Public Lessons Learned Entry: 1806, Capture of Apollo Lunar Module Reliability Lessons Learned: Program/Engineering Management, 9/25/2007.

[7] McCurdy, H. E., *Faster, Better, Cheaper: Low-Cost Innovation in the U.S. Space Program*, The Johns Hopkins University Press, 2001.

[8] Mahler, J. G., *Organizational Learning at NASA: The Challenger and Columbia Accidents*, Georgetown University Press, Washington, DC, 2009.

[9] Wikipedia, Space Shuttle design process, https://en.wikipedia.org/wiki/Space_Shuttle_design_process, accessed July 25, 2018.

[10] Williamson, R. A., "Developing the Space Shuttle: Early Concepts of a Reusable Launch Vehicle," in Logsdon, J. M., editor, Exploring the Unknown: Selected Documents in the History of the U.S. Civil Space Program, NASA SP-4407, 1995.

[11] Camarda, C, Space Shuttle Design and Lessons Learned Presentation, March 2014, https://www.researchgate.net/publication/296652080SpaceShuttleDesignandLessonsLearned, accessed July 18, 2018.

[12] Encyclopedia Astronautica, Space Shuttle, www.astronautix.com/s/spaceShuttle.html, accessed July 19, 2018.

[13] Trento, J. J., *Prescription for Disaster: From the glory of Apollo to the betrayal of the Shuttle*, Crown, New York, 1987. Quoted in Schwartz, H. S., *Narcissistic Process and Corporate Decay: The Theory of the Organization Ideal*, New York University Press, New York, 1990.

[14] Rogers Commission, *Report of the Presidential Commission on the Space Shuttle Challenger Accident*, 1986. http://history.nasa.gov/rogersrep/genindex.htm

[15] Vaughan,, D., *The Challenger Launch Decision: Risky Technology, Culture, and Deviance at NASA*, University of Chicago, 1997.

[16] CAIB, Columbia Accident Investigation Board, Vol. I, August 2003.

[17] Paté-Cornell, E., and Dillon, R., "Probabilistic risk analysis for the NASA space Shuttle: a brief history and current work," Reliability Engineering & System Safety, V. 74, 3, December 2001.

[18] Boin, A., and Schulman, P., "Assessing NASA's Safety Culture: The Limits and Possibilities of High-Reliability Theory," Public Administration Review November-December, 2008.

[19] Wieland, P., "ECLSS Development for Future Space Missions," AlAA 90-3728, AIAA Space Programs and Technologies Conference, September 25-28, 1990, Huntsville, AL.

[20] Bilardo, V. J., Jr., "The Physical/Chemical Closed-Loop Life Support Research Project," AIAA-90-3729, AIAA Space Programs and Technologies Conference, September 25-28, 1990, Huntsville, AL.

[21] Schneider, W. F., and Shull, S. A., "NASA Advanced Explorations Systems: 2017 Advancements in Life Support Systems," AIAA 2017-5152, AIAA SPACE and Astronautics Forum and Exposition, 12 - 14 Sep 2017, Orlando, FL.

[22] Levri, J. A., Vaccari, D. A., and Drysdale, A. E., "Theory and Application of the Equivalent System Mass Metric," 2000-01-2395, 30th International Conference on Environmental Systems, Toulouse, France. July 10-13, 2000.

[23] Drysdale, A.E., and Hanford, A.J., (1999) Advanced Life Support Research and Technology Development Metric. Baseline. CTSD-ADV, JSC 39503.

[24] Maxwell, S., and Drysdale, A., "Assessment of Waste Processing Technologies for 3 Missions," SAE 2001-01-2365, 31st International Conference on Environmental Systems, 2001.

[25] Jones, H. W., "Equivalent Mass versus Life Cycle Cost for Life Support Technology Selection," 2003-01-2635, 33rd International Conference on Environmental Systems, 2003.

[26] Jones, H. W., "Storage or Recycling Life Support for Mars?" AIAA 2013-3407, 43rd International Conference on Environmental Systems, July 14- 14-18, 2013, Vail, CO.

[27] Jones, H. W., "Design and Analysis of a Flexible, Reliable Deep Space Life Support System," AIAA 2012-3418, 42nd International Conference on Environmental Systems, 15 - 19 July 2012, San Diego, California.

[28] Jones, H. W., "Much Lower Launch Costs Make Resupply Cheaper Than Recycling for Space Life Support," ICES-2017-87, 47th International Conference on Environmental Systems, 16-20 July 2017, Charleston, South Carolina.

[29] Jones, H. W., "The Recent Large Reduction in Space Launch Cost," ICES-2018-81, 48th International Conference on Environmental Systems, 8-12 July 2018, Albuquerque, New Mexico.

[30] Sargusingh, M. J., and Nelson, J. R., "Environmental Control and Life Support System Reliability for Long-Duration Missions Beyond Lower Earth Orbit," ICES 2014-180, 44th International Conference on Environmental Systems, 13-17 July 2014, Tucson, Arizona.

[31] Bagdigian, R. M., Dake, J., Gentry, G., and Gault, M., "International Space Station Environmental Control and Life Support System Mass and Crewtime Utilization In Comparison to a Long Duration Human Space Exploration Mission," ICES 2015-094, 45th International Conference on Environmental Systems 12-16 July 2015, Bellevue, Washington.

[32] Jones, H. W., "The International Space Station (ISS) Oxygen Generation Assembly (OGA) Is Not Feasible for Mars Transit," ICES-2016-103, 46th International Conference on Environmental Systems, 10-14 July 2016, Vienna, Austria.

[33] Jones, H. W., "We Can't Count on Repairing All Failures Going to Mars," ICES-2016-113, 46th International Conference on Environmental Systems, 10-14 July 2016, Vienna, Austria.

[34] Jones, H. W., "High Reliability Requires More Than Providing Spares," 2019-14, submitted to 49th International Conference on Environmental Systems, 7-11 July 2019, Boston, Massachusetts.

[35] Levri, J. A., Drysdale, A. E., Ewert, M. K., Fisher, J. W., Hanford, A. J., Hogan, J. A., Jones, H. W., Joshi, J. A., and Vaccari, D. A. (2003) *Advanced Life Support Equivalent System Mass Guidelines Document*, NASA/TM-2003-212278, National Aeronautics and Space Administration, Ames Research Center, Moffett Field, California.

[36] Drysdale, A., "Metrics and System Analysis," 981746, Society of Automotive Engineers, Warrendale, PA, 28th International Conference on Environmental Systems, 1998.

[37] Jones, H., "Humans to Mars Will Cost About "Half a Trillion Dollars" and Life Support Roughly Two Billion Dollars," ICES-2016-111, 46th International Conference on Environmental Systems, 10-14 July 2016, Vienna, Austria.

[38] Bagdigian, R., Gatens, R., Metcalf, J., Stephan, R., Broyan, J., Shull, S., and Macatangay, A., "National Aeronautics and Space Administration Environmental Control and Life Support Technology Development and Maturation for Exploration," ICES-2014-19, 44th International Conference on Environmental Systems, Tucson, Arizona, 2014.

[39] Howard, D., Perry, J., Sargusingh, M., and Toomarian, N., "Notional Environmental Control and Life Support System Architectures for Human Exploration beyond Low-Earth Orbit," AIAA-2015-4456, AIAA SPACE 2015, Pasadena, California, 2015.

[40] Jones, H. W., "Reliability Growth in Space Life Support Systems," ICES-2014-075, 44th International Conference on Environmental Systems, 13-17 July 2014, Tucson, Arizona.

[41] Jones, H. W., Hodgson, E. W., and Kliss, M. H., "Life Support for Deep Space and Mars," ICES-2014-074, 44th International Conference on Environmental Systems, 13-17 July 2014, Tucson, Arizona.

[42] Stromgren, C., Escobar, F., Anderson, M., Stambaugh, I., Sargusingh, M., and Goodliff, K., "Assessment of Desired ECLSS Closure Rates for Human Mars Missions," AIAA 2017-5123, AIAA SPACE and Astronautics Forum and Exposition, 12 - 14 Sep 2017, Orlando, FL.