NASA Glenn Research Center

A Prognostic Launch Vehicle Probability of Failure Assessment Methodology for Conceptual Systems Predicated on Human Causal Factors

Craig H. Williams
NASA GRC

Lawrence J. Ross, J. Joseph Nieberding
Aerospace Engineering Associates LLC

10 Aug 2018

Overview

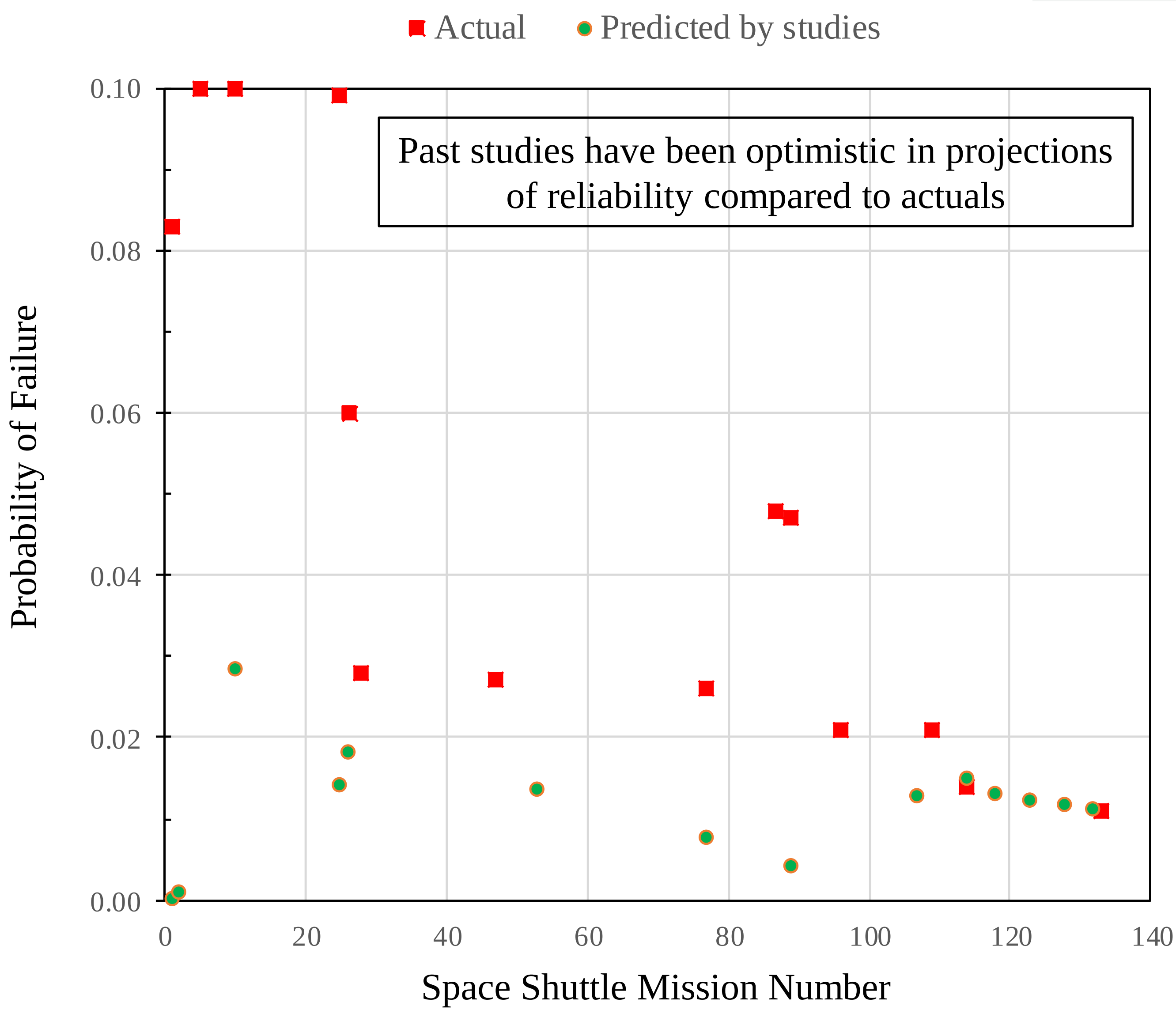
NASA Glenn Research Center

- **Purpose**: Create an improved method to calculate reliability of a conceptual launch vehicle system prior to fabrication by using historic data of actual root causes of failures
  - While failures have unique “proximate causes”, there are typically a finite amount of common “root causes”
  - Heretofore launch vehicle reliability evaluation typically hardware-centric statistical analyses, while most root causes of failures are been shown to be human-centric
  - A method based on human-centric root causes can be used to quantify reliability assessments and focus proposed actions to mitigate problems
  - Existing methods have been optimistic in their projections of launch vehicle reliability compared to actuals

- **Hypothesis:** reliability of a conceptual launch vehicle can be more accurately evaluated based on a rational, probabilistic approach using past failure assessment teams’ findings predicated on human-centric causes

Actual vs. Predicted Probability of Failure for Space Shuttle System

NASA Glenn Research Center



Terminology Regarding Mission Failures

(from NASA NPR 8621.1B – Appendix A ; http://nodis3.gsfc.nasa.gov/)

NASA Glenn Research Center

- ***Proximate Cause****: The event(s) that occurred, including any condition(s) that existed immediately before the undesired outcome, directly resulted in its occurrence and, if eliminated or modified, would have prevented the undesired outcome. Also known as the direct cause(s).*

- ***Root Cause****: An event or condition that is an organizational factor that existed before the intermediate cause and directly resulted in its occurrence (thus indirectly it caused or contributed to the proximate cause and subsequent undesired outcome) and; if eliminated or modified, would have prevented the intermediate cause from occurring, and the undesired outcome. Typically, multiple root causes contribute to an undesired outcome.*

Examples of Relationship between Proximate Cause of Failure vs. Root Causes

NASA Glenn Research Center

Aerospace Engineering Associates LLC

- Example #1: Titan IVB/Centaur-32 failure
  - Proximate cause of failure
    - Loss of upper stage roll control due to software error
  - Root causes
    - Erroneously (human) entered flight constant
    - Human software checks failed to detect the error due to lack of understanding by staff
    - Software testing lacked formality, performed with default values (not the entered flight values)
    - Cape personnel did not diligently follow-up when they noticed something atypical

- Example #2: Atlas/Centaur-62 failure
  - Proximate cause of failure (sequence of events)
    - 1st: Minor LOX tank leak escaped build, test, and inspection procedures
    - 2nd : SOX accumulated in interstage adapter during ascent
    - 3rd : SOX amplified shape charge firing shock, exceeding tank design, caused crack
    - 4th : LOX escape through crack exceeded control authority of attitude control system
  - Root causes
    - More effective Systems Engineering (test/inspection technologies insertion, noting missing analysis tasks)
    - Test program in synch with design

Proximate causes tend to be unique and are difficult to anticipate in future programs.
Root causes, however, tend to exhibit commonalities among failures.

Assessments of Existing Methods

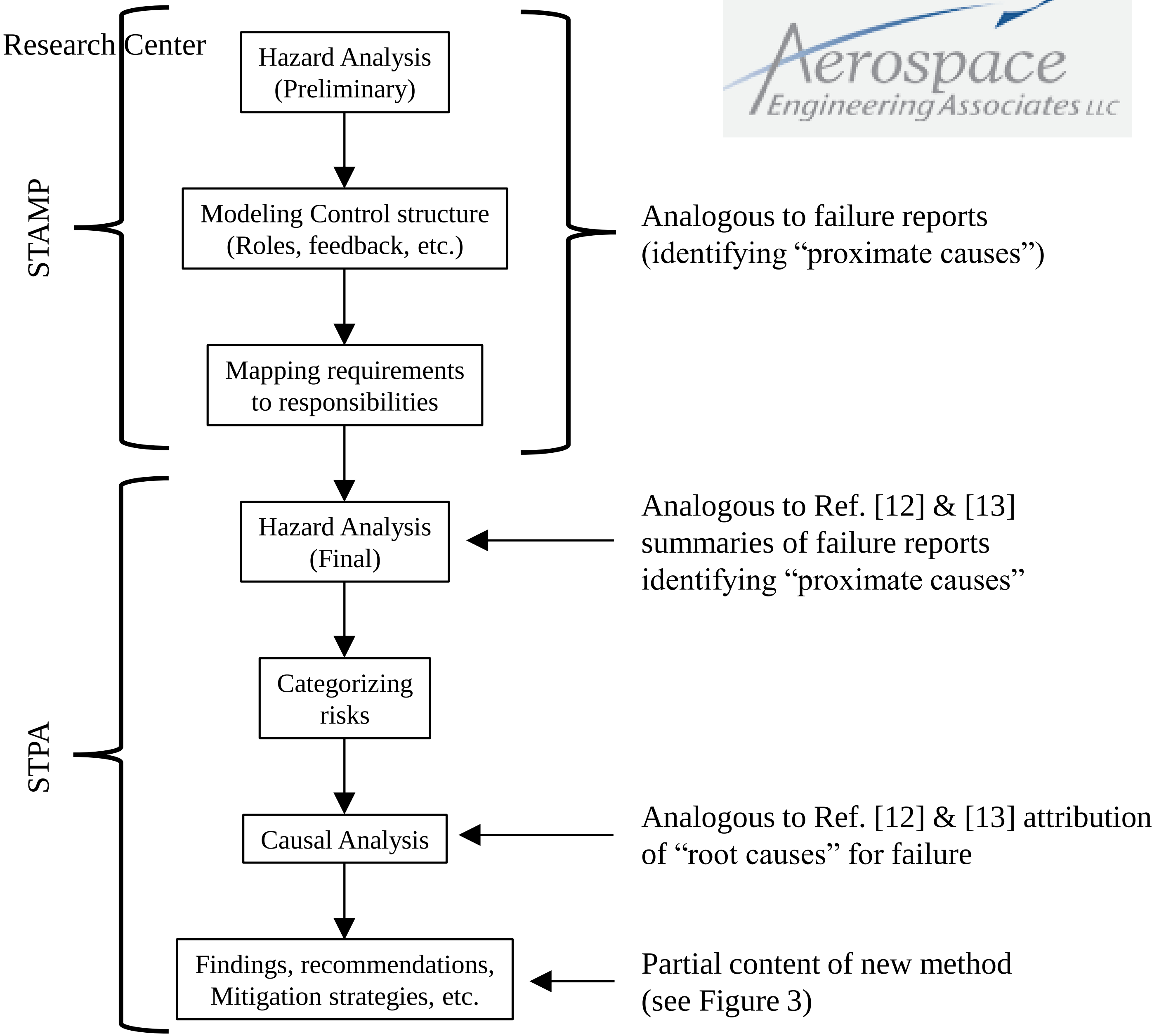
NASA Glenn Research Center

- "Human Reliability Analysis Methods Selection Guidance for NASA"
  - Chandler F.T., et al., NASA HQ/OSMA study group, July 2006
  - Outside HRA experts from academia, other federal labs, and the private sector
  - 50 system reliability methods considered, fourteen selected for further study, four selected as best suited for human flight
  - Probabilistic Risk Analysis (PRA) + Human Reliability Analysis (HRA) enabled incorporating effects, probabilities of human errors
  - While four down-selected methods deemed appropriate for failure assessment, it did not appear that these methods could be concisely applied to perform major system-wide assessment of probability of failure of a conceptual design without becoming unwieldy

- "Engineering a Safer World",
  - Detailed, comprehensive study external to NASA
  - Leveson N. G., MIT, 2011.
  - Systems-Theoretic Accident Model and Processes (STAMP)
    - All-encompassing accident model based on systems theory
    - Analyzed accidents after they occurred and created approaches to prevent occurrence in developing systems
    - Not focused on failure prevention per se, but rather reducing hazards by influencing human behavior through use of constraints, hierarchical control structures, and process models to improve system safety
  - System Theoretic Process Analysis (STPA)
    - Addresses predictive part of problem (a "hazard analysis")
    - Includes all causal factors identified in STAMP: "……design errors, software flaws, component interaction accidents, cognitively complex human decision-making errors, and social organizational and management factors contributing to accidents"
    - Can guide design process rather than require it to exist before hand
    - Did not appear capable of concise application for system-wide assessment of probability of failure of a conceptual design without becoming unwieldy

Comparison of STAMP/STPA to New Method

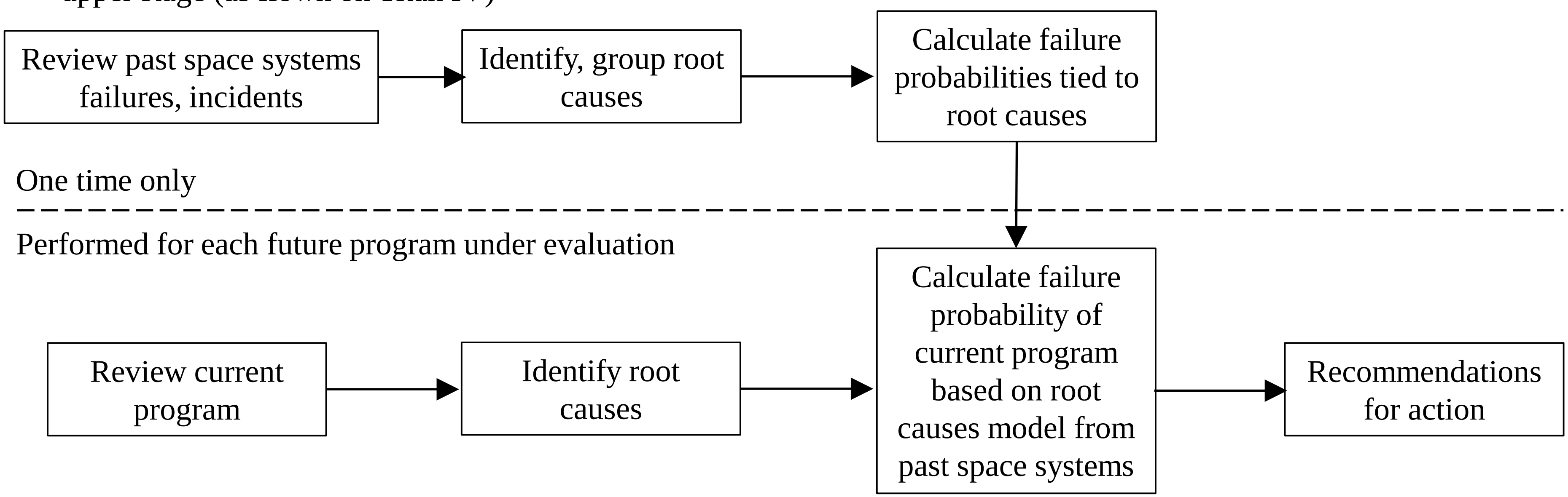
NASA Glenn Research Center



Approach of New Method to Assess Probability of Failure

NASA Glenn Research Center

- Establishing new method's basis
  - Review of past proximate causes of launch vehicle failures
  - Establishing root causes of past launch vehicle failures based on expert judgment
  - Categorizing, consolidating similar root causes into finite categories
  - Establishing baseline model using root causes of past launch vehicle failures
    - Selection of cases to be used
    - Scoring of root and sub-root causes
    - Plotting resultant data
  - Derivation of function for probability of failure of launch system
- Applying the new method to conceptual designs (example): NASA/USAF Shuttle/Centaur G-Prime upper stage (as flown on Titan IV)



Source data from

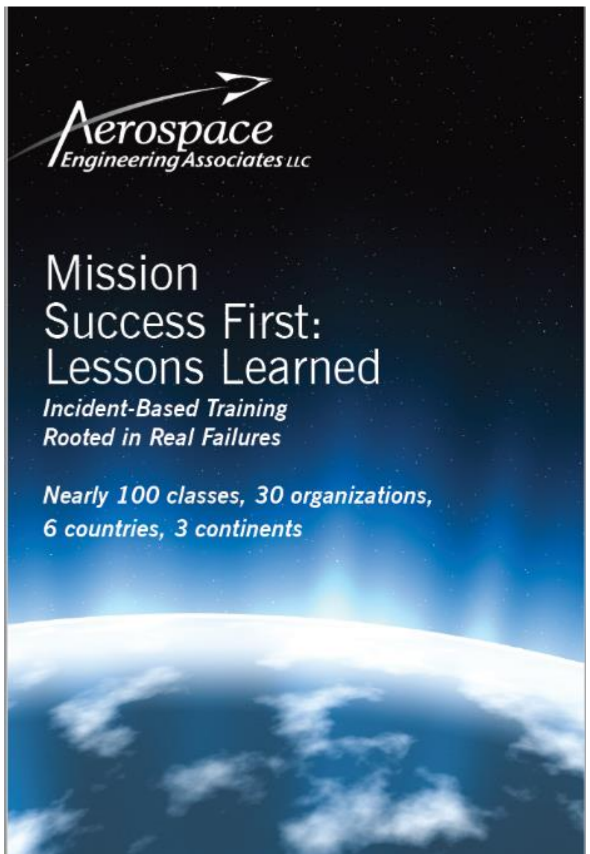
"*Lessons Learned Applied to Space Systems Developments*"

J. Nieberding & L. Ross, Aerospace Engineering Associates LLC

NASA Glenn Research Center

Former NASA GRC executives:
successfully led several launch vehicle development programs, 60+ launch teams

Case studies of 40+ NASA and international case failures, major incidents, and shortfalls, where proximate causes given from failure review boards and root causes proposed

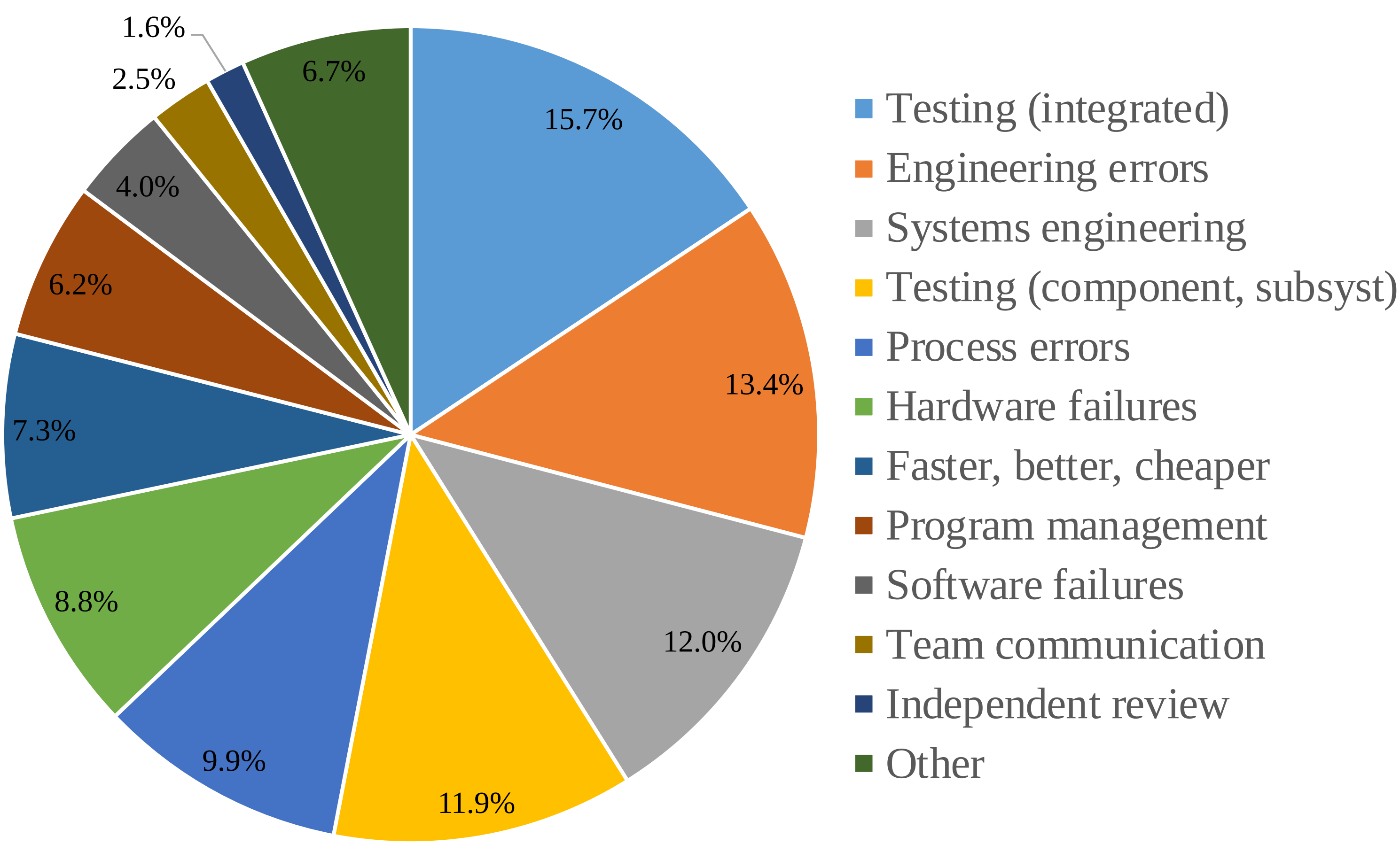


Lessons from Past Missions

- Screening Out Design Errors
- Screening Out Procedural Errors
- Impact of Weak Testing Practices
- Systems Engineering Lapses
- Software Mishaps
- Flawed Processes
- Information Flow Breakdown
- Component Failure
- Experienced Teams make Mistakes
- Normalizing Deviance
- Missed Advanced Warnings
- Perils of Heritage Systems
- Sabotage
- Management Factors Have Lost Missions

Distribution of Root Causes In Launch Vehicle Development/Operation

NASA Glenn Research Center



Distribution of distinct root causes is fairly even,
leading to decision to merely sum all causes to calculate a total

Selected Case Studies of Launch Vehicle Development/Operation Failures

NASA Glenn Research Center

Aerospace Engineering Associates LLC

		Mission	Problem	Result	Number in Series	Description of Total Number in Series
Research & Development						
	1	Atlas/Centaur F-1	Premature sheild seperation	Loss of mission	8	Test flights: 7 LeRC led + F-1
	2	Atlas/Centaur AC-5	Premature booster engine shutdown	Loss of mission, pad		See A/C F-1
	3	N-1 #1 (Russian)	Stage 1 failure	Loss of mission	4	Four N-1's in series
	4	N-1 #2 (Russian)	T - 0 explosion	Loss of mission, pad		See N-1 #1
	5	N-1 #3 (Russian)	Uncontrolled roll	Loss of mission		See N-1 #1
	6	N-1 #4 (Russian)	POGO	Program termination		See N-1 #1
	7	Titan IIIC/Centaur TC-1	Centaur engine start failure	Loss of mission	1	Test flight only
	8	X-43A	Loss of control	Loss of mission	3	Three (expendable) vehicles; one failure
Operational						
	1	Apollo 13	LOX tank explosion	Loss of mission	20	Total Service Module flights
	2	Apollo 13 Stage II	POGO	Potential loss of mission	13	Total Saturn V flights
	3	Ariane 5 (501)	Loss of control	Loss of mission	92	Total up through May 2017
	4	Atlas/Centaur AC-21	Fairing seperation failure	Loss of mission	61	Total non-test flight A/C up to 1990 (AC-69)
	5	Atlas/Centaur AC-24	Avionics hardwear failure	Loss of mission		See A/C-21
	6	Atlas/Centaur AC-33	Loss of control	Loss of mission		See A/C-21
	7	Atlas/Centaur AC-43	Booster engine failure	Loss of mission		See A/C-21
	8	Atlas/Centaur AC-62	Loss of control during coast	Compromised mission		See A/C-21
	9	Atlas/Centaur AC-67	Lightining strike	Loss of mission		See A/C-21
	10	Space Shuttle Challenger	SRM failure	Loss of mission	135	Total Space Shuttle flights
	11	Space Shuttle Columbia	Launch-induced wing damage	Loss of mission		See Space Shuttle Challenger
	12	Titan IIIC/Centaur TC-6	Stage 2 LOX tank problem	Potential loss of mission	6	Post TC-1
	13	Titan IVB/Centaur -32	Loss of control	Loss of mission	16	Total Titan IV/Centaur flights
					359	

Color and Numerical Scoring of Root Causes of Past Launch Vehicle Failures

NASA Glenn Research Center

0.0 | 0.2 | 0.4 | 0.6 | 0.8 | 1.0

No or minimal problems

Correctable problems within existing program definition

Prominent problems requiring prompt resolution

Serious problems Threatening program viability

Scoring of Root Causes of Titan IVB/Centaur– 32 Failure

NASA Glenn Research Center

Aerospace Engineering Associates LLC

		Sub-root cause Qualitative Scores	Root Cause Scores	Total Score
				3.55
Insufficient testing (integrated system)			0.70	
	Lack of prudent integrated system testing			
	Not pursuing "test as you fly; fly as you test"			
	Insufficient understanding of interactions within entire system			
	Lack of test data of functioning system while in relevant environment			
Engineering errors			0.60	
	Faulty hardware design, fabrication			
	Incorrect analytical modeling or computational errors			
Ineffective Systems Engineering			0.00	
	Inadequate SE / engr judgment / understanding, resolving crit problems			
	Insufficient meaningful reviews			
	Failure to challenge analyses, heritage, assumptions			
	Analytic models uncorrelated w/ actuals, ill- scaled, or questionable validity			
Insufficient testing (components, sub-systems)			0.00	
	Lack of prudent component, sub-system testing			
	Verification by analysis or comparison with requirements only			
	Heritage hardware/software: not validating for new application			
	Not establishing instrumentation needs			
Process errors			0.80	
	Fabrication, test, integration, or launch process not followed			
	Non-standard events, work-arounds not incorporated into process			

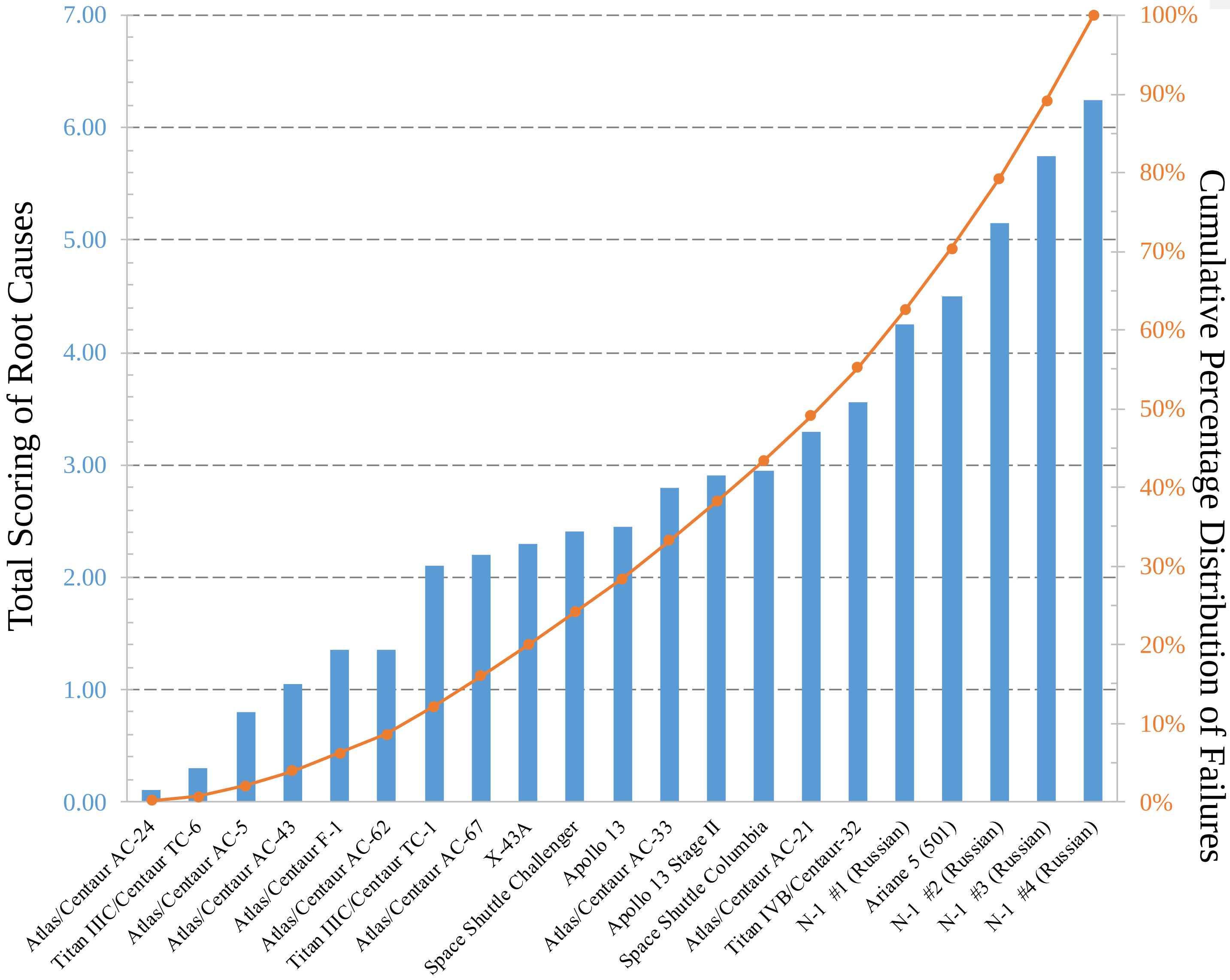
Hardware failure (flight or ground)			0.00	
	Poor quality or statistically out of tolerance component			
	Multiple unforeseen program/environment changes, or secondary effects			
Faster, Better, Cheaper			0.00	
	Overworked staff due to imprudently short schedule			
	Imprudently low funding			
Poor program management			0.00	
	Lack of leadership integrity			
	Inattentiveness to (or ineffectiveness in) managing problems			
	Normalization of Deviance (unexpected deviation, revised expectation)			
Software failure (flight or ground)			0.80	
	Differences between functional specifications and true requirements			
	Insufficient (or no) IV&V			
Poor team communication			0.65	
	Organization-to-organization differences			
	Insufficient formality between working groups			
Insufficient use of independent review team guidance			0.00	
	Absence of independent assessment			
	Failure to heed or fully implement recommendations			
Others			0.00	
	International pressures			
	Loss of key leader without comparable replacement			
	Others			

Root Causes Totals per Failure and their Cumulative Percentage Distribution

(all other 338 cases were successful (scores ≡ 0))

NASA Glenn Research Center

Aerospace Engineering Associates LLC



Cumulative Distribution Function to Calculate the Probability of Failure

NASA Glenn Research Center

- Since any non-zero scores of root causes can result in case failure, the probability (P) of failure event (E) ***should be a cumulative distribution F***
  - $F_x(b) = P\{E^X_b\} = P\{\omega \mid X(\omega) \leq b\}$
  - Where ω are the possible cases
  - Where b is limiting score of root causes
  - X is random variable of interest (the score of root causes for any case)

- The probability of a successful case (i.e. score = 0) ≡ $F_x(a) = F_x(0) = P\{\omega \mid X(\omega) \leq 0\}$
  - Number of case studies considered (the sample space Ω) = 359
  - Number of failures: 21
  - Thus, probability of success ***of entire sample space*** Ω = (359-21)/359 = 0.9415
  - Chance of failure = (1 - 0.9415) = 0.0585 or one chance of failure out of ~ 17 attempts

- Example: probability that a case <u>is</u> a failure <u>and</u> its score is ≤ 3.60 is given by:
  - $P\{\omega \mid a < X(\omega) \leq b\} = F_x(b) - F_x(a)$ = ((359-21) + 16)/359 - 0.9415 = 0.9861 - 0.9415 = 0.0446
  - Chance of failure: one out of ~ 22 attempts
  - Corresponding reliability = 95.5 %

Shuttle/Centaur G-Prime Upper Stage and Titan IV Launch Vehicle

NASA Glenn Research Center

Scoring of Shuttle/Centaur G-Prime Upper Stage

NASA Glenn Research Center

	Sub-root cause Qualitative Scores	Root Cause Scores	Total Score	
			4.20	
Insufficient testing (integrated system)		0.50		
Lack of prudent integrated system testing				No altitude propulsive stage test at 109%; PLIS mount failures; CISS prop valves erratic ops p. 206
Not pursuing "test as you fly; fly as you test"				Structural dynamic test campaign, system integration facility (for avionics, S/W, others) System Level III/IV Program PDR (March 1983) and CDR (
Insufficient understanding of interactions within entire system				Most of Centaur adopted/leveraged from existing, long heritage Atlas/Centaur program
Lack of test data of functioning system while in relevant environment				Most of Centaur adopted/leveraged from existing, long heritage Atlas/Centaur program
Engineering errors		0.00		
Faulty hardware design, fabrication				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) reports
Incorrect analytical modeling or computational errors				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Ineffective Systems Engineering		0.70		
Inadequate SE / engr judgment / understanding, resolving crit problems				Repeated JSC safety-driven changes in critical fluid dump system interface
Insufficient meaningful reviews				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Failure to challenge analyses, heritage, assumptions				Repeated LeRC challenging of astronauts' LH2 concern with Centaur vs. I
Analytic models uncorrelated w/ actuals, ill- scaled, or questionable validity				Modal survey performed on test article, trajectory design code based on p
Insufficient testing (components, sub-systems)		0.00		
Lack of prudent component, sub-system testing				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Verification by analysis or comparison with requirements only				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Heritage hardware/software: not validating for new application				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Not establishing instrumentation needs				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Process errors		0.30		
Fabrication, test, integration, or launch process not followed				Observed lower quality manufacturing, transport, and contractor staff acti
Non-standard events, work-arounds not incorporated into process				None identified
Hardware failure (flight or ground)		0.20		
Poor quality or statistically out of tolerance component				n/a
Multiple unforeseen program/environment changes, or secondary effects				Change from "Element" to "Payload" designation drove critical hardware c
Faster, Better, Cheaper		0.50		
Overworked staff due to imprudently short schedule				Contractor, LeRC leadership 50 to 70 hr weeks year after year p. 196-19
Imprudently low funding				~$2B current year funding over 4.5 years; Joint NASA & USAF funding
Poor program management		0.60		
Lack of leadership integrity				LeRC securing 109% SSME throttle baseline (TLH p. 205, 208, 209)
Inattentiveness to (or ineffectiveness in) managing problems				JSC intergration staff rather than JSC engineering staff; delayed tech respo
Normalization of Deviance (unexpected deviation, revised expectation)				JSC Shuttle lift capability/commitment
Software failure (flight or ground)		0.00		
Differences between functional specifications and true requirements				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Insufficient (or no) IV&V				System Level III/IV Program PDR (March 1983) and CDR (Dec 1983) r
Poor team communication		0.90		
Organization-to-organization differences				JSC unresponsive to LeRC technical data requests; difference in Center cultures, JSC Integration vs. Technical staff experise (TLH; p196 to 219)
Insufficient formality between working groups				Sufficient technical working groups between LeRC and GDSSD
Insufficient use of independent review team guidance		0.50		
Absence of independent assessment				No NAR convened ; continued Safety concerns by astronauts p.197-199 and 206-207
Failure to heed or fully implement recommendations				n/a
Others		0.00		
International pressures				n/a
Loss of key leader without comparable replacement				n/a
Others				n/a

- Probability of failure:
  - Projected: 4.46% (score: 4.20)
  - Actual: 6.67 %
- Observation of sub-root causes: almost complete reverse scoring between Titan IVB/Centaur – 32 and Shuttle/Centaur G-Prime
  - change in program leadership
  - change in manufacturer
  - 15 year gap (1985 to 1999)
- Implies necessity of scoring program development and launch within reasonable time periods and similar staff

Caveats

NASA Glenn Research Center

- While generally acknowledging shortcomings of accepted methods and need for improvement, concerns were raised by
  - NASA Headquarters Safety Center
  - NASA GRC Safety, Reliability and Quality Assurance Office
- "Non-zero score successes" should be incorporated into cumulative distribution function: requires reviewing of 338 successful mission post flight reports (considerable effort)
- Positive actions (adaptations to new information or feedback loops in decision making) not incorporated
  - Widely acknowledged as essential for successful outcomes
  - Omission represents meaningful modeling deficiency in assessments of probability of failure
- Sample set incomplete, lacking launch scrubs/delays: rejected due to seemingly infinite amount of "what if" speculation
- "Color coded" results helpful, but the numerical scoring might imply precision which does not exist
- Existing methods (Failure Modes Effects Analysis, Fault Tree Analysis, Human Reliability Analysis, etc.) already accommodate human factors: rejected due to anticipated resource-intensive needs if applied system-wide
- Small sample size of 21 launches implies significant statistical error
- Scoring was greatly influenced by sample space definition:
  - Greatest probability of failure of any case considered was 5.85 % (corresponding to a score of > 6.25)
  - Consideration of more failure cases could increase range of potential scores (and more representative of history)
- Potential major weaknesses if there is a significant change in
  - Organization which leads development or performs launch operations (or both)
  - Time between application method and launch operations
- Greatest vulnerability to criticism : "20-20 hindsight bias"
  - Comprehension of circumstances more important than judging past actions as imprudent, insufficient
  - Failure/mishap reports frequently do not describe in great detail various options available
  - Obvious poor decision in hindsight frequently appears to be correct decision in heat of the moment
  - Thus, reliance on (even) complete accident investigation board reports and experts with impressive comprehensive experience can still be subject

Summary and Conclusions

NASA Glenn Research Center

- Considerable number of methods to evaluate reliability of systems already exist
  - Most from nuclear power industry, a few from NASA
  - Both reassessing and prognosticating
  - Many incorporate human-causal factors
  - Most are best suited for detailed analysis of focused sub-systems, components
- Reliability estimates from existing methods
  - Create optimistic failure probability estimates when compared to actuals
  - Create complex, resource-intensive efforts if applied launch vehicle system-wide
  - Predicated on component hardware reliability and statistical analysis --- minor historic root cause of failures
  - Typically do not focus on human–centric root causes
- While proximate causes are failure case-unique, root causes tend to aggregate into finite, common categories
- Proposed new method to assess reliability of conceptual launch vehicle system based on historic data of human-centric root causes
- Single example
  - Totals agree well with actuals
  - Sub-root causes had almost complete reverse scoring (between Titan IVB/Centaur – 32 and Shuttle/Centaur G-Prime), attributed to change in program leadership, change in manufacturer, and 15 year gap (1985 to 1999)
  - More testing warranted
    While lacking in precision and accuracy, it is based on comprehensive, known root causes of launch vehicle failures.

Recommendation: apply new method to access probability of failure to currently in-development launch vehicle programs

References

NASA Glenn Research Center

- [1] Hamlin, T., et al., "Shuttle Risk Progression: Use of the Shuttle Probabilistic Risk Assessment (PRA) to Show Reliability Growth", AIAA paper 2011-7353, September 2011.
- [2] "NASA Procedural Requirements for Mishap and Close Call Reporting, Investigating, and Recordkeeping", NPR 8621.1C, 19 May 2016.
- [3] "Titan IVB-32/Milstar-3 Accident Investigation Board Report", USAF Form 711, USAF Mishap Report, date unknown.
- [4] Chandler F.T., et al., "Human Reliability Analysis Methods Selection Guidance for NASA", NASA HQ/OSMA Technical Report, July 2006.
- [5] "Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners", NASA/SP-2011-3421 Second Edition, December 2011.
- [6] Leveson N. G., "Engineering a Safer World", MIT Press, Cambridge, MA, 2011, Chapters 6, 7, and 8, pp. 169 – 249, and Appendix B pp. 469 – 493.
- [7] Huang, Z., et al, "Key Reliability Drivers of Liquid Propulsion Engines and a Reliability Model for Sensitivity Analysis", AIAA paper 2005-4436, July 2005.
- [8] Gernand, J. L., et al., "Constellation Ground Systems Launch Availability Analysis: Enhancing Highly Reliable Launch Systems Design", AIAA paper 2010-2180, April 2010.
- [9] Morse, E., L., et al, "Modeling Launch Vehicle Reliability Growth as Defect Elimination", AIAA paper 2010-8836, Sept 2010.
- [10] Guikeme, S., D., et al, "Bayesian Analysis of Launch Vehicle Reliability", AIAA paper 2003-1175, January 2003.
- [11] "NASA Reliability and Maintainability (R&M) Standard for Spaceflight and Support Systems", NASA Technical Standard 8729.1A, June 13, 2017.
- [12] Nieberding, J. J.; Ross, L. J., "Lessons Learned Applied to Space Systems Developments", presentation, Aerospace Engineering Associates LLC, Bay Village, OH, Vol 1, Ver. 1.0, 2006.
- [13] Nieberding, J. J.; Ross, L. J., "Mission Success First: Lessons Learned", Aerospace Engineering Associates LLC, Bay Village, OH, Class #100 presentation, 9-10 November 2016.
- [14] ASTRO-E-2 Mishap Report (Appendix), Table 7-1, NASA Safety Center, Mishap Investigation Board, Type A Mishap, IRIS No. 2005-273-00003, 2005.
- [15] Shuttle/Centaur Level III/IV Program PDR at LeRC, General Dynamics Convair Division, San Diego, CA, March 1983.
- [16] Shuttle/Centaur Level III/IV Critical Design Review at LeRC, General Dynamics Convair Division, San Diego, CA, December 1983.
- [17] Dawson, V., Bowles, M., "Taming Liquid Hydrogen: the Centaur: Upper Stage Rocket, 1958-2002", NASA History Series, NASA SP-2004-4230, Washington, DC, 2004, Chapter 7, pp. 189 – 219.
- [18] "NASA's Plans for Human Exploration Beyond Low Earth Orbit", NASA Office of Inspector General, Report No. IG-17-01713, Washington, DC, April 2017.
- [19] "NASA Human Space Exploration", U.S. Government Accountability Office, GAO-17-414, Washington, DC, April 2017.
- [20] Lilley, S., NASA email, NASA HQ Safety Center, Cleveland, OH, personal communication 21 December 2017.