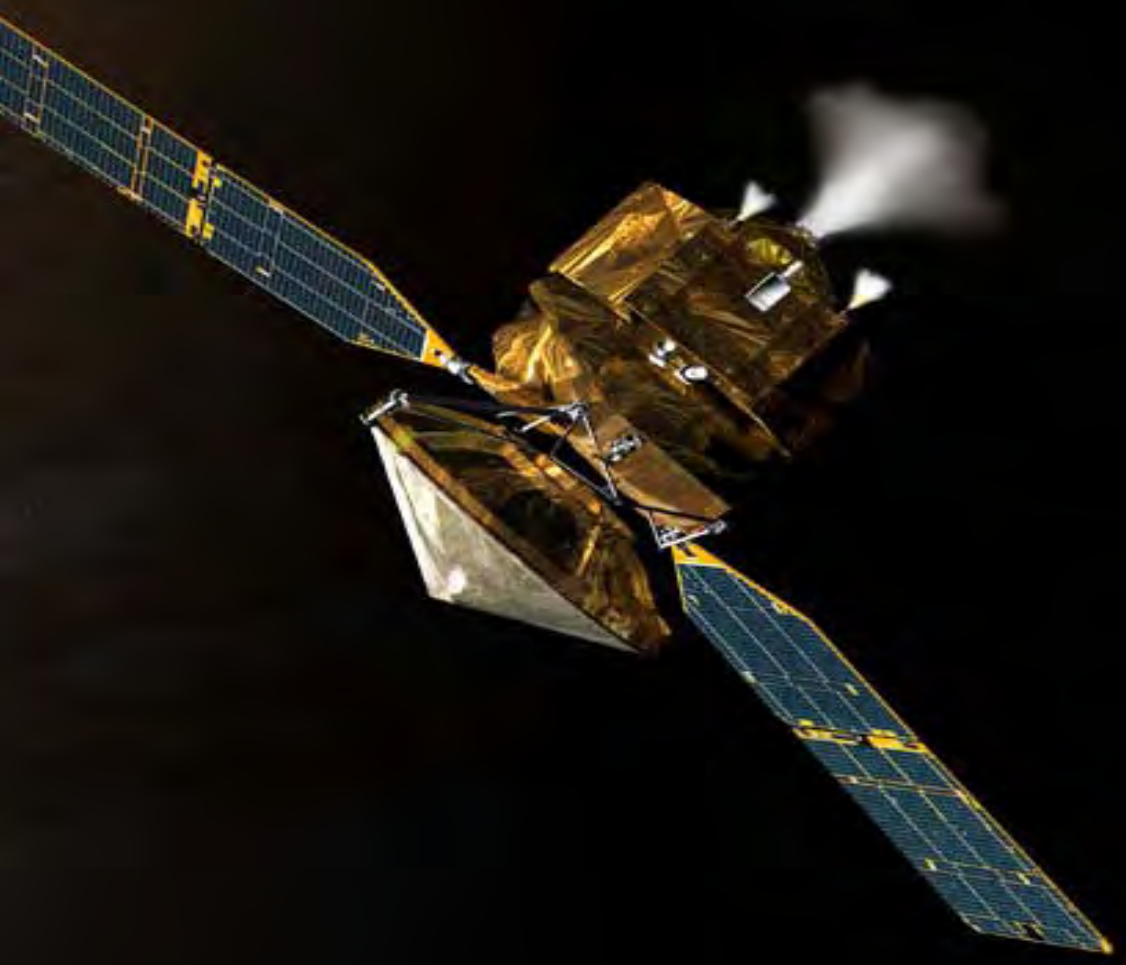
NASA/SP-2014-612
Version 1.0
November 2014

NASA System Safety Handbook

Volume 2: System Safety Concepts, Guidelines, and Implementation Examples



NASA/SP-2014-612
Version 1.0

NASA System Safety Handbook

Volume 2: System Safety Concepts, Guidelines, and Implementation Examples

National Aeronautics and Space Administration
NASA Headquarters
Washington, D.C. 20546

November 2014

NASA STI Program … 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 plays a key part in helping NASA maintain this important role.

The NASA STI program operates under the auspices of the Agency Chief Information Officer. It collects, organizes, provides for archiving, and disseminates NASA's STI. The NASA STI program provides access to the NASA Aeronautics and Space Database and its public interface, the NASA Technical Report Server, thus providing one of the largest collections of aeronautical and space science STI in the world. Results are published in both non-NASA channels and 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 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 co-sponsored by NASA.

SPECIAL PUBLICATION. Scientific, technical, or historical information from NASA programs, projects, and missions, 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 also include creating custom thesauri, building customized databases, and organizing and publishing research results.

For more information about the NASA STI program, 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

Fax your question to the NASA STI Help Desk at 443-757-5803

Phone the NASA STI Help Desk at 443-757-5802

Write to:

NASA STI Help Desk
NASA Center for Aerospace Information
7115 Standard Drive
Hanover, MD 21076-1320

ACKNOWLEDGMENTS

The project manager and the authors express their gratitude to NASA Office of Safety and Mission Assurance (OSMA) management for their support and encouragement in developing this document, the second and final volume of the NASA System Safety Handbook. Building upon the work that resulted in the first volume of this handbook, the development effort leading to this document was conducted in stages, and was supported through reviews and discussions by the NASA System Safety Steering Group (S3G) and by the additional contributors listed below (in alphabetical order).

AUTHORS:

Homayoon Dezfuli (Project Manager)	NASA Headquarters
Allan Benjamin	Consultant
Christopher Everett	Information Systems Laboratories
Martin Feather	Jet Propulsion Laboratory
Peter Rutledge	Quality Assurance & Risk Management Services
Dev Sen	Information Systems Laboratories
Robert Youngblood	Idaho National Laboratory

NASA SYSTEM SAFETY STEERING GROUP MEMBERS:

Michael Blythe	NASA Engineering and Safety Center
Roger Boyer	Johnson Space Center
Steven Broussard	Marshall Space Flight Center
Jonathan Brown	Armstrong Flight Research Center
Alfredo Colon	Headquarters
Anthony Diventi	Goddard Space Flight Center
John Evans	Headquarters
Chester Everline	Jet Propulsion Laboratory
Martin Feather	Jet Propulsion Laboratory
Jennifer Franzo	Stennis Space Center
Robert Gargiulo	Stennis Space Center
Maria Havenhill	Glenn Research Center
Brent Heard	Langley Research Center
K.C. Johnson	Langley Research Center
Crystal L. Jones	Kennedy Space Center
Jeffrey King	Armstrong Flight Research Center
Mark Kowaleski	NASA Safety Center
Jesse Leitner	Goddard Space Flight Center
Burton Lewis	Goddard Space Flight Center
Donovan Mathias	Ames Research Center
Richard Morrison	Ames Research Center
Cami Vongsouthy	Jet Propulsion Laboratory
Clifford Watson	Marshall Space Flight Center
Chi Yeh	Kennedy Space Center

ADDITIONAL CONTRIBUTORS:

Roselynne Strickland	Marshall Space Flight Center
Curtis Smith	Idaho National Laboratory
William Vesely	Headquarters

TABLE OF CONTENTS

CHAPTER/SECTION **PAGE**

Foreword

Preface

Chapter 1: Introduction 1
1.1 Motivation for the Approaches to System Safety Discussed in the Handbook 1
1.2 Principal Themes of the Handbook 4
1.3 Relationship to MIL-STD-882 5
1.4 Relationship to Other Standards and Safety Disciplines 5
1.5 Intended Uses/Users of the Handbook 6
1.6 Coverage and Organization of the Handbook 6

Chapter 2: Key Concepts 9
2.1 System Safety, Risk, and Safety Performance 9
2.2 Acquirers and Providers 10
2.3 Adequate Safety 10
2.3.1 Meeting Minimum Tolerable Levels of Safety Performance 11
2.3.2 Being As Safe As Reasonably Practicable 13
2.3.3 State of the System with Respect to Adequate Safety 14
2.4 Relationship of System Safety to Systems Engineering and Risk Management 14
2.5 The Risk-Informed Safety Case 16

Chapter 3: Overview of the System Safety Framework 19
3.1 Top-Level Objectives and Associated Requirements 19
3.1.1 Top-Level Safety Performance Requirements 20
3.1.2 Lower-Level Safety Performance Requirements 24
3.1.3 Safety-Related Engineering and Process Requirements 25
3.2 System Safety Ensurance Activities to Achieve a Safe System 26
3.2.1 Conducting an Integrated Safety Analysis (ISA) 26
3.2.2 Requirements Development Support 27
3.2.3 System Design Support 27
3.2.4 Program Control and Commitment Support 28
3.2.5 Performance Monitoring Support 29
3.3 Development of the RISC (Argument for Safety) 30
3.4 Evaluation of the RISC 30
3.5 Interactions between the Acquirer and Provider 31
3.6 System Safety throughout the System Life Cycle 34

Chapter 4: Setting Safety Objectives and Associated Requirements: The Acquirer's Role 35
4.1 Setting Minimum Tolerable Levels of Safety Performance at the System Level: the Total Probability of Loss 35
4.1.1 The Acquirer's Responsibilities and Areas to Address 35
4.1.2 Establishing Sub-Cases for Different Key Mission Objectives (KMOs) 36
4.1.3 Developing Safety Performance Risk Margins 37
4.1.4 Developing Thresholds and Goals for the Total Loss Probability 39
4.1.5 The Use of Ratios 40

4.2 Developing Safety Performance Requirements at the System Level: the Probability of Loss from Known Risks 40
4.2.1 The Acquirer's Responsibilities and Areas to Address 40
4.2.2 Developing Performance Requirements for the Probability of Loss from Known Risks 41
4.3 Levying Deterministic Engineering and Process Safety Requirements 41
4.3.1 The Acquirer's Responsibilities and Areas to Address 41
4.3.2 Levying Deterministic Safety Requirements 41
4.3.3 Tailoring Deterministic Safety Requirements 42
4.4 Setting Verification Procedures for the Safety Requirements 43
4.4.1 The Acquirer's Responsibilities and Areas to Address 43
4.4.2 Setting Initial Verification Procedures 43
4.4.3 Negotiating with the Provider to Rebaseline Requirements and Reset Verification Procedures 44
4.5 Example for Chapter 4 – Deriving Safety Risk Margins, Safety Thresholds and Goals, and Probabilistic Safety Requirements 45
4.5.1 Space Shuttle Experience 45
4.5.2 Launch Vehicle Experience 48
4.5.3 Initial Loss Probability Margin 51
4.5.4 Justification for the Use of Ratios 51
4.5.5 Safety Threshold for LEO Missions 53
4.5.6 Safety Goal for LEO Missions 55
4.5.7 Probabilistic Loss Requirement for LEO Missions 55

Chapter 5: Performing System Safety Ensurance Activities: The Provider's Role 57
5.1 Developing and Implementing the System Safety Management Plan (SSMP) 57
5.1.1 The Provider's Responsibilities and Areas to Address 57
5.1.2 Conducting a System Safety Requirements Analysis (SSRA) and Setting Requirements for the SSMP 59
5.1.3 General Contents of the SSMP 60
5.1.4 Specifying Roles and Responsibilities, Controls, Processes, and Protocols 62
5.1.5 Configuration Control and Data Management 64
5.1.6 Management and Organizational Commitment to Safety 65
5.2 Performing an Integrated Safety Analysis (ISA) 67
5.2.1 The Provider's Responsibilities and Areas to Address 67
5.2.2 Rationale for Performing an ISA 69
5.2.3 Characteristics of an ISA 69
5.2.4 Implementing a Graded Analysis Approach 74
5.2.5 Integrating Subsystem Analyses into the ISA 78
5.2.6 Treating Fault Management in an ISA 81
5.2.7 Integrating Software Analyses into the ISA 84
5.2.8 Testing to Support the ISA 85
5.2.9 Adhering to the Modeling and Simulation (M&S) Credibility Assessment Scale (CAS) 86
5.3 Tailored, Derived, and Allocated Requirements 88
5.3.1 The Provider's Responsibilities and Areas to Address 88
5.3.2 Tailoring Levied Requirements 88
5.3.3 Developing Allocated and Derived Requirements 90
5.3.4 Conducting Ongoing Negotiations with the Acquirer 94
5.3.5 Considering Safety Performance Requirements and Levied Engineering Requirements as an Integrated Package 94

5.4 Supporting the System Design 95
5.4.1 The Provider's Responsibilities and Areas to Address 95
5.4.2 Using RIDM and Historical Data 96
5.4.3 Applying ASARP in Combination with the Minimum Tolerable Level 97
5.4.4 Designating and Analyzing Safety Critical Items and Safety Risk Drivers 101
5.4.5 Minimizing Design Complexity 103
5.5 Maintaining Adequate Safety Performance throughout the Life Cycle 105
5.5.1 The Provider's Responsibilities and Areas to Address 105
5.5.2 Using CRM to Manage Emerging Risks 106
5.5.3 Monitoring and Managing Safety Critical Items and Safety Risk Drivers 106
5.5.4 Monitoring and Correcting for Anomalies and Precursors 107
5.5.5 Identifying and Justifying Departures from the Plan 107
5.5.6 Keeping the Design within the Validated Domain 107
5.5.7 Maintaining Realistic Budgets and Schedules 108
5.5.8 Risk-Informing Support Activities 109
5.6 Taking Advantage of Emerging Opportunities to Improve Safety Performance 110
5.6.1 The Provider's Responsibilities and Areas to Address 110
5.6.2 Identifying and Assessing Safety Improvement Opportunities 110
5.6.3 Testing Interactions with the Whole System 111
5.7 Example for Chapter 5 – ASARP Principles Applied to a Proposed Space Shuttle Escape Pod 112

Chapter 6 Developing the Risk-Informed Safety Case (RISC): The Provider's Role 117
6.1 Overall Approach to RISC Development and Documentation 117
6.1.1 The Provider's Responsibilities and Areas to Address 117
6.1.2 Principles for the Overall Approach 118
6.1.3 Deriving Safety Claims from Objectives and Requirements 120
6.1.4 Developing Evidence 122
6.1.5 Optional Use of Goal Structuring Notation in Developing Claims Trees 125
6.1.6 Documenting the RISC 127
6.2 Assigning Responsibilities for RISC Development and Integrating the Parts 128
6.2.1 The Provider's Responsibilities and Areas to Address 128
6.2.2 Assigning Appropriately Qualified Personnel 129
6.2.3 Integrating Subsystem RISCs 130
6.3 Exercising a Graded Approach in RISC Development 131
6.3.1 The Provider's Responsibilities and Areas to Address 131
6.3.2 Exercising a Graded Apprach 131
6.4 Maintaining and Updating the RISC and Addressing Future Life-Cycle Phases 133
6.4.1 The Provider's Responsibilities and Areas to Address 133
6.4.2 Maintaining and Updating the RISC 134
6.4.3 Addressing Future Life-Cycle Phases 135
6.5 Addressing Weaknesses and Unresolved Safety Issues 136
6.5.1 The Provider's Responsibilities and Areas to Address 136
6.5.2 Addressing Weaknesses and Issues 136
6.6 Example for Chapter 6 – RISC Fragment for New Technology on a Robotic System (Electric Ion Thruster) 137
6.6.1 Safety Claims Tree 137
6.6.2 Sources of Evidence 140
6.6.3 Hypothetical Suppositions, Assumptions, and Contexts 140
6.6.4 A Note on Margins 143

Chapter 7: Evaluating the RISC: The Acquirer's Role 145
7.1 Interfacing with the Provider 146
7.1.1 The Acquirer's Responsibilities and Areas to Address 146
7.1.2 Guidance on Interfacing with the Provider to Facilitate Evaluation of the RISC 146
7.2 Reviewing, Evaluating, and Scoring the RISC 146
7.2.1 The Acquirer's Responsibilities and Areas to Address 146
7.2.2 Composition and Independence of the Evaluation Team 147
7.2.3 Sources of Assurance Deficit 148
7.2.4 The Process of Ranking Assurance Deficits and Base Claim Importance 151
7.2.5 Experts' Estimates of Overall Confidence 152
7.2.6 Using Value-of-Information Methods to Analyze Options for Reducing Uncertainty 154
7.3 Documenting the Findings of the RISC Evaluation 155
7.3.1 The Acquirer's Responsibilities and Areas to Address 155
7.3.2 Contents of the Summary and Detailed RISC Evaluation Findings 155
7.3.3 Contents of the RISC Evaluation Report 156
7.4 Example for Chapter 7 – RISC Evaluation for New Technology on a Robotic System (Electric Ion Thruster) 166
7.4.1 Identification and Analysis of Assurance Deficit Sources 166
7.4.2 Experts' Assessment of Overall Confidence 166
7.4.3 Request for Additional Information 166
7.5 Generic Risk Evaluation Tree 169

References 181

Appendix A – Abbreviations and Acronyms 187

Appendix B – Definitions 189

FIGURES

FIGURE **PAGE**

Chapter 2

2-1 Example Set of At-Risk Populations with Respect to Safety 11
2-2 Fundamental Principles of Adequate Safety 11
2-3 Safety Threshold and Safety Goal for a Generic Safety Performance Measure 12
2-4 As Safe As Reasonably Practicable (ASARP) 13
2-5 Relationships among Risk Management, Systems Engineering, and System Safety in a Program/Project Context 15

Chapter 3

3-1 The NASA System Safety Framework 19
3-2 Objectives-Driven Requirements Development (notional) 21
3-3 Interaction/Iteration between the Acquirer and Provider during the Development of System-Specific Safety Requirements 22
3-4 Relationship between the Initial Safety Performance Requirement, Initial Safety Performance Margin, and Safety Threshold 23
3-5 Relationship between the Safety Performance Requirement, Safety Performance Margin, and Minimum Tolerable Level of Safety as the System Matures 24
3-6 A Safety Claim Supported by Two Independent Arguments 30
3-7 Life-Cycle Perspective on the System Safety Framework (Notional) 34

Chapter 4

4-1 Results of a Retrospective Analysis of P(LOC) for the Space Shuttle Compared to Earlier PRA Predictions, from Hamlin, et al 45
4-2 Correlation of Shuttle Risks from Retrospective Analysis with Changes in Design, Fabrication, and Operation 46
4-3 Comparison of Retrospective Analyses of Shuttle Risks Accounting for Versus Not Accounting for Revealed LOC Accidents 48
4-4 Failure History for the Soyuz and Molniya Launch Vehicles 50
4-5 Failure History for the Atlas Launch Vehicle 50
4-6 Failure History for the Delta Launch Vehicle 51
4-7 Loss of Mission and Loss of Crew Probability Contributions by Accident Scenario for an Example Launch System with a Launch Abort Capability, from NASA Ames Study 54
4-8 Correlation of Total Loss Probability with Chronological Flight Number 56
4-9 Loss Probability Requirement as a Function of the Number of Completed Flights 56

Chapter 5

5-1 Relationship between Risk-Informed Safety Case Development and Evaluation and Continuous Risk Management 64
5-2 The Concept of a Scenario 71
5-3 Integration of System Safety Analysis Methods in ISA 73
5-4 The Potential Consequences of an Accident Scenario May Impact Several Safety Objectives or Requirements 73
5-5 Example Directed Graphs of Subsystem-Hardware-Software-Human Interactions to Be Modeled in Integrated Safety Analyses 80
5-6 Transformation of Directed Graph (c) to Event Tree–Fault Tree Representation 81
5-7 Timing Considerations in a Launch Abort Sequence 83
5-8 FAA Reliability Approach 91

5-9 Schematic of Subsystem-Hardware-Software-Human Interactions Leading from Implementation of a Derived Requirement to Effects on the System 93
5-10 Schematic of the Effect of a Design Change on the Probability of Loss from Known Risks at the Time of the First Flight and on the Project Cost 98
5-11 Schematic of the Effect of a Design Change on the Probability of Loss from All Risks and on the Project Cost 98
5-12 Schematic of the Effect of a Design Change on the Probabilities of Exceeding Joint Confidence Levels and Risk Tolerances 99
5-13 An Example Two-Dimensional Utility Function 100
5-14 Example Contours of a Two-Dimensional Utility Function Used for Illustration 100
5-15 Effect of System Complexity on Program/Project Cost and Schedule 104
5-16 Schematic of Proposed Space Shuttle Escape Pod 112
5-17 Effect of Adding an Escape Pod on Loss Probability and Program Cost 114
5-18 Effect of Adding an Escape Pod on Loss Probability and Payload Weight 115

Chapter 6

6-1 Claims Tree Conceptual Diagram 120
6-2 Goal Structuring Network (GSN) 126
6-3 The NASA Systems Engineering Life-Cycle Phases 133
6-4 Coverage of the System Life Cycle in the RISC 135
6-5 Schematic of an Electric Ion Thruster Subsystem and Photo of a Test Model 137
6-6 Claims Tree Pertaining to Wearout Failures for a New Technology Ion Electric Thruster 138

Chapter 7

7-1 Phased Evaluation of the RISC 148
7-2 Illustration of Symbols Used in Goal Structuring Notation 173
7-3 Overall Safety Argument 174
7-4 The System Meets the Minimum Tolerable Level of Safety for Known Risks 174
7-5 ISA is Consistent with Applicable Safety-Related Test Results 175
7-6 Argue that Testing Has Been Performed Properly and Is Trustworthy 175
7-7 Argue that the ISA Has Been Performed Properly and Is Trustworthy 176
7-8 The SSMP Demonstrates that the System Will Continue to Meet all Relevant Safety Performance Requirements in the Future 176
7-9 Argue that the SSMP Has Been Prepared Correctly and Is Trustworthy 177
7-10 The System Is ASARP 177
7-11 The Introduction of Hazards Has Been Minimized 178
7-12 Safety Has Been Prioritized during Decision Making 178
7-13 Safety Performance Requirements Have Been Risk-Informed 179
7-14 Argue that an Adequate Assessment of Allocated Requirements Has Been Performed for Known and UU Risks 179
7-15 Organization Functions Effectively and Prioritizes Safety 180

TABLES

TABLE **PAGE**

Chapter 2
2-1 Possible Combinations of Meeting the Minimum Tolerable Level of Safety and Being ASARP.... 14

Chapter 3
3-1 Interactions between the Acquirer's and Provider's Systems Engineering and System Safety Functions during Formulation of Requirements 32
3-2 Interactions between the Acquirer's and Provider's Systems Engineering and System Safety Functions during Preparation and Evaluation of the Risk-Informed Safety Case..... 33

Chapter 4
4-1 Modification of Assessed Probabilities of the Top Accident Scenarios Leading to LOC at the Time of the First Shuttle Flight Assuming the Challenger and Columbia Accidents Had Not Occurred..... 47
4-2 Suggested Guidelines for Estimating the Ratio of the Initial Probabilistic Safety Performance Margin to the Initial Loss Probability from Known Risks..... 52

Chapter 5
5-1 Cross Reference from Operational Safety Objectives to Provider System Safety Ensurance Activities 58
5-2 Type of Risk Scenario Modeling Used for Different Criticality Conditions 77
5-3 Ranking Criteria for Credibility Assessment Scale (CAS) Factors..... 87

Chapter 6
6-1 Illustration of the Relationship between Base Claims and Operational Objectives for Claims Pertaining to the Acquirer's Roles for Setting Probabilistic Requirements..... 121
6-2 Illustration of the Relationship between Base Claims and Operational Objectives for Claims Pertaining to the Provider's Roles for Managing Unknown and Underappreciated (UU) Hazards..... 122
6-3 Illustration of the Types of Evidence that Pertain to Various Potential Base Claims..... 123
6-4 Amount of Additional Evidence Needed to Satisfy the Criteria for Evidence Completeness for Different Levels of Mission Criticality and Evidence Criticality 132
6-5 Evidence for Selected Base Claims Pertaining to Wearout Failures for the Ion Thruster Subsystem… 141

Chapter 7
7-1 Illustration of Assurance Deficit Sources that Pertain to Evidence for Various Potential Base Claims 149
7-2 Guidelines for Summary Evaluation Findings..... 158
7-3 Guidelines for Detailed Evaluation Findings..... 160
7-4 Example Analysis of Assurance Deficits Affecting Confidence in Long-Term Thruster Operation..... 167

FOREWORD

I am pleased to introduce volume 2 of the NASA System Safety Handbook.

This handbook fits within a set of activities sponsored by the Office of Safety and Mission Assurance aimed at the development of a more objectives-based assurance approach, in which the decomposition of top-level safety and mission success objectives into concrete sub-objectives and associated strategies, form the basis for the planning and review of assurance activities.

This approach does not negate the use of trusted assurance standards, tools, and methods, but treats their application and results as means to substantiate claims that relevant objectives have been addressed. Being focused on a structured and comprehensive set of assurance considerations, rather than a prescribed collection of standards, techniques, and deliverables, is envisioned as a meaningful way to address current and future challenges associated with the safety and mission assurance function for NASA's spaceflight missions by:

- Providing a technical basis for adaptations of or changes to assurance standards and practices driven by new acquisition models, increased use of model-based systems engineering approaches, and other changes to space system development and operations.
- Enhancing consistency and perceived value of assurance models and efforts by documenting how the use of safety and mission assurance standards, methods, and techniques contributes to the confidence in top-level overall safety (and mission success) of a system.
- Further enhancing consistency by allowing considerations associated with various disciplines to be combined in a single assurance framework.
- Explicitly considering quality aspects of the assurance argument, such as the credibility of models and data, and qualification of analysts and reviewers, as part of the claims made by provider and the independent evaluations conducted by assurance organizations.
- Promoting accountability for safety (and mission success) on the part of acquirers, providers, and assurance organizations, by clarifying their roles, and by identifying the range of considerations that (should) underlie claims regarding safety and mission success and associated risks. This is increasingly important in situations where there is a diminished ability to rely on a wide variety of experts to bring concerns forward.

The concepts and guidance in this document promote a better understanding of system safety by defining the term "adequate safety," showing ways in which this abstract term may be broken down into more concrete considerations, thereby providing a model for related ensurance and assurance activities.

I thank the handbook development team and the NASA System Safety Steering Group for their contributions to this document, and encourage the engineering and safety and mission assurance communities to evaluate how the concepts in this document can be adopted and used to improve the safety of NASA missions.

Frank Groen, Ph.D.
Director, Safety and Assurance Requirements Division
Office of Safety and Mission Assurance
June 2015

PREFACE

The NASA system safety framework is in the process of change. A major motivation for this change is the desire to promote an objectives-driven approach to system safety that explicitly focuses system safety efforts on system-level safety performance, and serves to unify, in a purposeful manner, safety-related activities that otherwise might be done in a way that results in gaps, redundancies, or unnecessary work. An objectives-driven approach to system safety affords more flexibility to determine, on a system-specific basis, the means by which adequate safety is achieved and verified. Such flexibility and efficiency is becoming increasingly important in the face of evolving engineering modalities and acquisition models, where, for example, NASA will increasingly rely on commercial providers for transportation services to low-earth-orbit. An objectives-driven approach is also consistent with input from the Aerospace Safety Advisory Panel (ASAP), which in its 2012 Annual Report reiterated its belief that target levels of safety performance must be specified for NASA systems on a mission-specific basis, and that it is necessary to understand the tradeoffs between safety performance and performance in other domains (technical, cost, schedule) when programmatic decisions are made.

System safety has heretofore tended to focus on identifying and controlling individual hazards. This type of focus is evident in existing standards such as MIL-STD-882, which has been a primary reference document for system safety since it was initially released in July 1969. The focus of this NASA handbook is on the framework within which activities such as those prescribed in MIL-STD-882 are conducted, so that such activities are adequate to ensure the achievement of system-level safety performance objectives, and decision-makers are provided with sufficient information, clearly communicated, to enable them to make appropriately informed decisions concerning safety throughout the system life cycle. As such, it is the intent of this NASA handbook to build upon, rather than replace, standards such as MIL-STD-882, by addressing NASA-specific needs that go beyond those addressed by existing documents.

The cost of implementing the framework described in this handbook within a given program or project is a concern to some, and so a word or two on that subject is needed. The use of a graded approach to the implementation of this framework, based on the criticality of the mission and the concerns or scenarios being investigated, should serve to ensure that the overall cost of implementation for a particular system is no more than a small percentage of the overall cost of developing, building, and operating the system. This notion is expected to prove true regardless of the scale or complexity of the system. It is anticipated that the cost of implementing this framework will be on the order of one percent of the system life cycle cost. Such an amount could routinely be included within the initial budget of a program or project as the cost for assuring system safety. Moreover, it is expected that this framework will, in time, result in reduced system life cycle costs through the reduction in unnecessary or duplicative work, more efficient and effective life cycle reviews, and fewer and less consequential mishaps. In practice, it will take some experience, beginning with pilot studies, to determine how the costs compare, and to refine the framework and its implementation to maximize its cost-effectiveness.

The approach to system safety presented in this handbook represents a significant evolution from the traditional approach. It is important to recognize that the transition from today's approach to the new one will not take place at once. Over a period of time, implementation plans will be developed with the broad participation of Agency personnel, and the plans will be implemented gradually but steadily. During this transition, the new concepts will be piloted, lessons will be learned, and the content of this handbook will be updated as necessary to continually reflect a vision of system safety that is optimal for the Agency.

Homayoon Dezfuli, Ph.D.
NASA System Safety Fellow and Chair, NASA System Safety Steering Group
NASA Headquarters
November 2014

1. Introduction

This is the second of two volumes that collectively comprise the NASA System Safety Handbook. Volume 1 (NASA/SP-2010-580) [1] was prepared for the purpose of presenting the overall framework for System Safety and for providing the general concepts needed to implement the framework. Toward this end, Volume 1 addressed the following topics:

- The fundamental principles of adequate safety
- Derivation of operational safety objectives
- System safety activities and their relationships to safety objectives
- Special topics pertaining to integrated safety analysis and to risk-informed allocations of safety thresholds and goals
- General considerations in the collaborative development of controls
- Elements in the development of a risk-informed safety case (RISC), and associated examples
- Elements in the evaluation of a RISC

Volume 2 provides guidance for implementing these concepts as an integral part of systems engineering and risk management. This guidance addresses the following functional areas:

1) The development of objectives that collectively define adequate safety for a system, and the safety requirements derived from these objectives that are levied on the system
2) The conduct of system safety activities, performed to meet the safety requirements, with specific emphasis on the conduct of integrated safety analysis (ISA) as a fundamental means by which systems engineering and risk management decisions are risk-informed
3) The development of a RISC designed to ensure that significant gaps or faults that could lead to safety deficits are identified and corrected
4) The evaluation of the RISC (including supporting evidence) using a defined set of evaluation criteria, to assess the veracity of the claims made therein

1.1 Motivation for the Approaches to System Safety Discussed in the Handbook

The principal motivation for developing Volumes 1 and 2 of this handbook is to prepare the path for transitioning system safety at NASA from present practices to those needed to conduct NASA's mission over the next 10 years and beyond.

Considering the increased complexity of NASA's future missions (e.g., landing and sustaining humans on Mars; capturing and redirecting an asteroid), it was the intent of Volume 1 to develop a system safety framework that fosters the following set of desired attributes:

- A system-level approach to safety that understands that safety is an emergent property that is more than the sum of the parts
- An objectives-driven approach to system safety that focuses system safety efforts on the achievement of an adequately safe system, rather than a product-driven approach that focuses on the production of deliverables
- Explicitly addressing system-level considerations (e.g., aggregate risk, adverse sub-system interactions)

- System safety integrated into systems engineering decision making as an aspect of risk management
- A "probabilistic thinking" mindset that emphasizes effective treatment of uncertainties
- Appreciation for the potential magnitude of unknown and underappreciated risks
- A “graded approach” to system safety, in order to match the resources and depth of safety analysis to the complexity and importance of decisions being addressed, as well as to save on the cost of the analysis
- Model-based system safety as an integrated aspect of model-based systems engineering (MBSE), rather than process-based system safety
- Coherent and compelling presentation of safety-related information at relevant decision forums (e.g., milestone reviews), rather than as a set of prescribed deliverables that are treated as check-the-box questions
- A systematic and principled attempt to identify failure causes and control them

Furthermore, with the increased emphasis on transferring much of the technology for space flight to the private sector and conducting collaborative missions with private enterprises, it appears clear that system safety will also need to foster the following attributes:

- Effective lines of communication between NASA as an acquirer and private companies as providers that lead to mutually shared and accepted agreements regarding the development of safety requirements and verification that they have been met
- Accommodation of a variety of insight/oversight acquisition models such as those related to commercial transportation services without over-constraining provider design, product realization, and operations & sustainment practices

In addition to this list of desired attributes, which formed the original rationale for the handbook, NASA’s System Safety Steering Group (S3G) members were asked in September 2013 to complete a questionnaire seeking to identify areas where improvement in current practices are needed in order to achieve NASA’s future goals. The most commonly cited areas were as follows:

- Adequacy of discussions of the substance of system safety results in project forums
- Integration among system-safety related disciplines; e.g., hazard analysis, reliability analysis, probabilistic risk assessment, and risk management
- Early involvement of system safety in life cycle activities
- Integration of system safety across Centers and projects
- Differentiation between system safety requirements for crewed versus uncrewed missions
- More effective analysis of cross-system interactions
- Adequacy of time allotted to system safety activities
- Better reporting of system safety results to higher levels of the organizational hierarchy
- Better treatment of uncertainties

These needs and desirable attributes form the principal motivators for Volume 2. They can be summed up in the following five statements of rationale:

Rationale 1

One of the foremost motivators for the guidance in Volume 2 is the desire to support the core strategic goals, objectives, and values of the Agency. Specifically, the guidance promotes an ***objectives-driven*** approach to system safety, in accordance with NASA's library of directives, procedural requirements, and handbooks, explicitly focusing system safety efforts on the achievement of systems that are adequately safe. At the same time, it is designed to allow ***flexibility*** in the means by which system safety is achieved, thereby accommodating next-generation engineering modalities and promoting ***innovation*** (a value that is highly emphasized in NASA's 2014 strategic plan). In the process, it also promotes ***technical rigor where needed*** in safety assessments and safety rationale, in order to enhance the credibility of these products, thereby facilitating decision makers' acceptance of the system safety information.

Rationale 2

The guidelines in Volume 2 are also motivated by a desire to promote an approach which explicitly serves to unify safety analysis activities that otherwise might be done in a way that results in gaps, redundancies, or unnecessary work. The guidelines in this handbook promote a means for organizing existing, often disparate, system safety products such as hazard analyses (HAs), failure modes and effects analyses (FMEAs), finite element analyses (FEAs), cross-system fault-failure analyses, and probabilistic risk assessments (PRAs) into a single, integrated, ***system-level*** safety analysis that comprehensively characterizes the hazards and associated accidents that could credibly occur and adheres to the credibility criteria put forth in the NASA Modeling and Simulation (M&S) Standard. This integrated safety analysis should be ***scenario-based*** so that effective controls can be derived, should evolve over the life cycle of the system so that it remains current with accumulating test and operational experience, and should be developed to a level of rigor sufficient to support risk-informed decision making throughout the life of the system. Such areas of decision-making include: design trades; optimization of hazard control strategies; designation and management of ***safety-critical items***; allocation of safety-related performance requirements into sub-systems and components; and determination and maintenance of a ***safe operating envelope*** that is resistant to normalization of deviance. Each of these areas benefits from the ***holistic perspective*** afforded by a comprehensive, system-level safety analysis.

Rationale 3

The guidelines are also motivated by a desire to promote a coherent approach to ***risk acceptance*** decision-making at Key Decision Points (KDPs) through a comprehensive, ***cased-based approach to safety assurance*** that focuses on demonstrating satisfaction of safety objectives to the system (or service) acquirer, rather than on the production of a set of prescribed safety-related deliverables. A case-based approach places the burden on the provider to argue that the safety objectives are met, using system information and system safety products as evidence to substantiate the claims made in the safety argument. It also provides a rational basis for identifying ***assurance deficits*** due to flaws in the safety argument and/or inadequacies in the evidentiary support of the constituent claims.

Rationale 4

Another motivator is a desire to be consistent with existing practices and processes within NASA's systems engineering and safety assurance functions in areas where they have succeeded to date and are likely to succeed in the future. Thus, the guidelines in Volume 2 support the requirements in NPR 7123.1B (NASA Systems Engineering Processes and Requirements) by providing appropriate system safety contributions to the systems engineering process throughout the project life-cycle. This includes providing ***safety ensurance*** within the current systems engineering framework, providing documentation

for the technical review process at key decision points, and adhering to the technical review success criteria. The guidelines also align with processes already exercised by the safety assurance function within NASA. They support the role of the technical authority (TA) in providing ***safety assurance*** and in promoting the integrity of the ***risk acceptance*** function at the highest levels of NASA, and they support the safety goal policy, which introduces probabilistic considerations into the requirements and places emphasis on conducting a broad integrated safety analysis.

Rationale 5

The final motivator is a desire to ***streamline*** system safety activities (e.g., safety analysis activities) in order to reduce redundancies and potential inconsistencies, thereby increasing the likelihood of the program/project staying within schedule and budget.

1.2 Principal Themes of the Handbook

The following are some of the main themes underlying the guidance provided in this handbook: (1) that safety is an emergent property of a system that arises when system components interact with each other, and with the environment in which the system is operated, and with the system operators themselves; (2) that engineering, operational, and management activities which affect system safety should be informed by an *integrated* safety analysis (ISA) to help ensure that scenarios that cut across subsystem boundaries are fully addressed; (3) that while a system should meet its specified safety requirements and should be as safe as reasonably practicable (ASARP), it should also be affordable with a high degree of confidence; (4) that the greatest threats to safety, cost containment, schedule adherence, and technical performance are not from the risks that are known and fully appreciated, but from the risks that are unknown and/or known but underappreciated, that are best controlled through organizational and managerial means and through reliance on best engineering practices such as robust margins, adherence to codes and standards, etc.; (5) that the blanket imposition of unnecessary requirements on historical grounds alone may lead to sub-optimal results, (6) that to provide confidence that a system is adequately safe, it is necessary to demonstrate, through a convincing set of arguments backed by evidence, that the system meets its safety objectives; (7) that thorough evaluation of the safety claims and supporting evidence by an assurance entity that possesses expertise in the areas covered by the safety case is essential for the approval authority to make an informed decision; and (8) that because of the inductive nature of safety cases, the evaluation should include a rigorous, interrogative attempt to identify flaws in the safety argument, rather than attempting to prove it in some absolute sense.

Another theme of the handbook is that new opportunities for improving safety should be exploited when the improvement in safety justifies the sacrifice that might be entailed in cost, schedule, or technical performance. New opportunities may arise from various sources, including design improvements, diagnostic improvements, and testing improvements that are enabled by new equipment, new technology, or new applications of an existing technology. The handbook recommends that the management of new safety opportunities be integrated with the management of safety risks. It is suggested that the framework for safety risk and safety opportunity management should be integrated because new opportunities frequently evolve from new risks, and new risks are an expected byproduct of new opportunities.

Taken as a whole, the approach to system safety presented in this handbook represents a significant evolution from the traditional approach that remains in use throughout NASA at the time of publication. It is not expected for the transition from today's traditional or baseline approach to the new one to take place overnight, or for all aspects of that approach to disappear. Rather, this handbook is intended to provide a vision or objective for how system safety should function perhaps ten years from now. Between now and then, implementation plans will be developed with the broad participation of Agency personnel. These plans will be thoughtfully implemented to assure a gradual but steady evolution of system safety practice from today's baseline to the way it is described in this handbook. Some aspects of the new approach will be easy and quick to realize, while others will take more time. During the transition, many

of the new concepts seen here will be piloted, progress will be made, lessons will be learned, and the content of this handbook will be updated as necessary to continually reflect the optimal vision of system safety for the Agency.

1.3 Relationship to MIL-STD-882

System safety has traditionally focused on identifying and controlling individual hazards, as exemplified by standards such as MIL-STD-882 [2], which has been a primary reference document for system safety since its initial release in July 1969. MIL-STD-882 identifies the Department of Defense systems engineering approach to eliminating hazards, where possible, and minimizing risks where those hazards cannot be eliminated. It covers hazards as they apply to systems, products, equipment, and infrastructure (including both hardware and software) throughout design, development, test, production, use, and disposal. The intent of the guidance in this handbook is to build upon, rather than replace, standards such as MIL-STD-882 by addressing NASA-specific needs that go beyond those addressed in existing documents. It is in this context that Volume 2 of the System Safety Handbook, like Volume 1, goes beyond an approach to system safety in which "risks" are individually identified and prioritized for management based on worst-case consequence likelihood. Instead, it advances an approach in which individual "risks" are integrated into an analysis of the system's aggregate safety performance, enabling decision makers to focus on the drivers of aggregate risk when making systems engineering and/or risk management decisions. This approach recognizes that safety performance is intrinsically a system-level attribute whose characterization requires a holistic perspective. For example, this handbook provides guidance to implement a recommendation of the Aerospace Safety Advisory Panel (ASAP) in its 2012 Annual Report, that target levels of safety performance be specified for NASA systems on a mission-specific basis, and that the tradeoffs between safety performance and performance in other domains (technical, cost, schedule) be understood when programmatic decisions are made.

1.4 Relationship to Other Standards and Safety Disciplines

A number of existing mission-type-specific and discipline-specific standards levy requirements deemed necessary to the safety of the particular missions and disciplines to which they apply. These standards tacitly implement a graded approach to safety, and it is not the intention of the guidance in this handbook to supersede them.

Among these safety-related disciplines for which other standards exist, reliability/maintainability (R/M) and quality assurance (QA) are cross-cutting processes that support fulfillment of multiple fundamental objectives (such as those relating to safety and performance). Reliability engineering comprises the set of pursuits aimed at assessment and improvement of reliability performance of systems during their missions. Maintainability engineering consists of the assessments and verification of the system design characteristics and maintenance processes so that downtime is minimized when maintenance action is necessary. Quality assurance consists of the assessments that provide additional independent assurance that risks associated with noncompliance are minimized to an acceptable level.

While this handbook does not suggest comprehensive guidance for them, there are extremely important system safety concerns involving R/M and QA that are discussed herein. One such concern is decision-making about the deployability of new, complex, large-scale systems for which integrated performance experience is lacking, and cannot feasibly be obtained before a decision is made whether to allow crewed flights (i.e., it is infeasible to conduct a large number of full-scale tests to measure system-level reliability empirically). In such a case, the assessment of system-level safety performance is itself a complicated endeavor, for which diverse types of evidence will be needed. The problem is not just one of assessing performance at the system level based on subsystem performance, but additionally examining interactions between subsystems. Moreover, while this problem is difficult enough when only one institution is involved, its difficulty is compounded when the system delivered to NASA is put together by a contractor, assembling subsystems delivered by yet other contractors.

1.5 Intended Uses/Users of the Handbook

Historically, the organizations to which guidance of this type has been applicable have resided within the NASA organizational hierarchy, e.g., programs, projects, and elements. However, NASA is increasingly relying on commercial service providers for cargo and crew transportation, to low Earth orbit (LEO) in particular. This handbook is intended to support such commercial services by providing guidance that supports risk-informed development of safety performance requirements, while allowing flexibility in the particular means by which the requirements are met, and in the means by which safety performance is substantiated via the RISC.

This document is intended to be used by personnel engaged in any systems engineering common technical process, where decisions are made that potentially affect safety, and by decision-makers at all levels of the organization. Common technical processes are enumerated in NPR 7123.1B [3]. It is not the intent of this document to imply that system safety personnel are solely responsible for the safety performance of the system; safety is a team effort. Rather, the view of this handbook is that system safety entails developing an *understanding* of the safety performance of the system and of the safety implications of decisions regarding the system, proposing system or operational *changes* when deemed advisable to improve safety, and *communicating* this understanding and rationale for changes to decision makers as part of a risk-informed decision making process that jointly considers all domains of system performance, including safety along with technical, cost, and schedule, as discussed in NPR 8000.4A [4]. Decisions should prioritize safety, but nevertheless must be informed by their effects over all mission execution domains because measures optimized for only one domain may negatively (or positively) impact performance in other domains.

This handbook is intended for all systems where NASA acts as Acquirer. The guidance set forth in Chapter 4, Safety Objectives and Associated Requirements, is formulated for all such systems. Beyond this minimum set, implementation of the guidance in Chapters 4 through 7 may need to be tailored to the specific circumstances, since it is impractical to promulgate a generic means for implementing detailed guidance that will apply optimally to a broad range of system types. Although the guidance in this volume is intended to suggest what the Acquirer should be seeking, the Acquirer retains the flexibility to decide whether to accept or reject the safety risk for any system regardless of whether or not this guidance is followed. The intent of the guidance, which is essentially system-safety-specific implementation of the general risk management requirements in NPR 8000.4A, is to help the Acquirer make the most appropriate safety risk acceptance decision, recognizing the pitfalls of excessive conservatism on the one hand, and inadequate vigilance on the other. Because implementation of the guidance in this document entail costs to the system, it is essential that they be applied in a cost-effective, graded manner.

1.6 Coverage and Organization of the Handbook

Chapter 2 presents a number of key concepts used throughout this handbook, namely:

- System safety, risk, and safety performance
- Acquirers and Providers
- Adequate safety
- Relationship of system safety to systems engineering and risk management
- The risk-informed safety case

Chapter 3 presents an overview of the System Safety framework and concepts for implementation. It includes a summary description of the following topics:

- Setting objectives and requirements
- Performing system safety analyses
- Developing, evaluating, and approving the RISC
- Managing interactions between the Acquirer and Provider
- Managing system safety throughout the system life cycle

Chapter 4 concerns the Acquirer's role in setting operational objectives and deriving requirements from these objectives. The following pursuits pertaining to this topic are addressed in Chapter 4:

- Setting thresholds and goals for the probabilities of loss or harm
- Developing probabilistic requirements for known hazards
- Developing margins to accommodate unknown and underappreciated hazards
- Levying deterministic engineering and process safety requirements
- Performing a System Safety Requirements Analysis
- Setting procedures for verifying that the requirements are satisfied

An example at the end of Chapter 4 illustrates how realistic safety thresholds, goals, margins, and requirements for new systems may be developed using a combination of historical experience and analyses applicable to other systems. In this example, operational experience attained for the Space Shuttle, the detailed probabilistic risk assessment developed for the Shuttle, failure experience obtained from early launch vehicles, and phenomenological analyses for the Ares 1 launch abort system are used to develop candidate probabilistic safety requirements for a new low-earth-orbit system.

Chapter 5 addresses system safety activities performed by the Provider. The following topics are addressed:

- Developing and implementing a System Safety Management Plan (SSMP)
- Conducting a System Safety Requirements Analysis (SSRA)
- Implementing a graded approach in the Integrated Safety Analysis
- Integrating fault management and software considerations into the analysis
- Adhering to modeling and simulation standards
- Developing tailored, derived, and allocated requirements
- Identifying and protecting safety critical items
- Maintaining adequate safety performance throughout the life cycle
- Taking advantage of emerging opportunities

The last section in Chapter 5 presents an example pertaining to application of the "As-Safe-As-Reasonably-Practicable" principle to decide upon options for improving safety. The example concerns the tradeoffs involved in deciding whether or not to incorporate a crew escape pod into a new system based on a design similar to that considered originally for Apollo missions.

Chapter 6 concerns the development of a RISC by the Provider to demonstrate that the system is adequately safe. Topics addressed include:

- Developing safety claims that align with safety objectives
- Incorporating programmatic, managerial, and organizational considerations into the RISC
- Developing the body of evidence
- Self-assessing assurance deficits in the evidence
- Assigning qualified personnel
- Integrating subsystem RISCs into the system-level RISC
- Applying a graded approach in the development of the RISC
- Life-cycle considerations, including maintaining and updating the RISC
- Addressing weaknesses, limitations, and unresolved safety issues
- Documenting the RISC
- Optional use of Goal Structuring Notation (GSN)

The last section in Chapter 6 provides an example RISC fragment that pertains to the use of a new electric thruster technology on a deep-space science mission. The example illustrates how a safety claims tree might be developed to make the case that the designed spacecraft propulsion system is adequately protected against wearout failure, and how evidence might be accumulated to support the base claims in the tree.

Chapter 7 addresses the evaluation of the RISC by the Acquirer to verify that the RISC is sufficiently robust to make the case for safety. Topics covered in this chapter include:

- Providing the technical basis for decision making on whether and how to proceed to the next phase.
- Different levels of review
- The composition and independence of the evaluation team
- Evaluators' assessment of assurance deficits and their importance
- Rating the evaluators' overall confidence in the RISC
- Using Value-of-Information methods to analyze options for reducing uncertainty
- Documenting the evaluation findings

The next-to-last section in Chapter 7 continues the example of Chapter 6 concerning the use of a new electric thruster technology on a deep-space science mission. The example illustrates how the evaluators might develop their rationale for ranking assurance deficits in the RISC submitted by the Provider and for assessing their overall confidence in the RISC.

The last section of Chapter 7 provides a generic RISC evaluation tree that has been incorporated by NASA's Office of Safety and Mission Assurance (OSMA) into a new NASA software tool called RISC Evaluation Management Tool [5]. The generic evaluation tree provides the evaluator with a means for performing an independent assessment of the Provider's RISC.

2. Key Concepts

2.1 System Safety, Risk, and Safety Performance

System safety is defined within NASA as the application of engineering and management principles,[1] criteria, and techniques to optimize safety[2] and reduce safety risk within the constraints of operational effectiveness, time, and cost throughout all phases of the system life cycle [6]. System safety is more than just the pursuit of safety performance; it also entails making the case to relevant decision makers that the pursuit of safety performance is on track to be successful throughout the system life cycle.

TERMS PERTAINING TO THE DEFINITION OF SYSTEM SAFETY

- **Safety** – Safety is defined as freedom from those hazards that can result in failure to meet one or more safety objectives by causing death, injury, or illness in humans, adversely affecting the environment, and/or causing damage to or loss of equipment or property.
- **Safety Performance** – In this handbook, safety performance is the complement of the probability of harm, i.e., one minus the probability of adverse safety consequences. Thus, a system need not have perfect safety performance in order to eliminate safety risk; it need only achieve the safety performance specified in the levied safety performance requirements.
- **Safety Performance Requirement** – A safety performance requirement is the specification of a minimum acceptable level of safety performance (i.e., a maximum acceptable value for the probability of harm).
- **Risk** – The use of the term risk in this handbook is consistent with its use in NPR 8000.4A [4], namely that risk is the potential for performance shortfalls, which may be realized in the future, with respect to achieving explicitly established and stated performance requirements in any one or more of the mission execution domains of safety, technical, cost, and schedule.
- **Safety Risk** – Safety risk (sometimes referred to as "safety performance risk") is the potential for shortfalls with respect to safety performance requirements. This differs from the common use of the term risk in a safety context as being the probability of some form of harm (e.g., as expressed by measures such as the probability of loss of crew (LOC), the probability of planetary contamination, or the public casualty expectation). Safety risk only arises to the extent that there is uncertainty as to whether the system's safety performance meets requirements.
- **Risk Burndown** – Risk burndown refers to the expectation that as a program/project evolves over time, mitigations are implemented; and as risk concerns are retired and the state of knowledge about the performance measures improves, uncertainty should decrease, with an attendant lowering of risk.

[1] The "application of management principles" includes management's function in both *ensuring* (promoting) safety and *assuring* (verifying) safety. These terms will be defined further in Section 2.2.

[2] The term "optimize safety" is interpreted here to be equivalent to the term "achieve adequate safety," with the implication that both include being as safe as reasonably practicable (ASARP) and meeting minimum tolerable levels of safety.

2.2 Acquirers and Providers

For the purposes of this handbook, the personnel and organizations with responsibilities relating to system safety are generically separated into two basic roles: *Acquirers* and *Providers*.

ACQUIRERS AND PROVIDERS

AND THEIR ROLES RELATIVE TO SAFETY ASSURANCE AND SAFETY ENSURANCE

- **Acquirers** - An Acquirer is a NASA organization that tasks another organization (either within NASA or external to NASA) to produce a product or deliver a service. The Acquirer is responsible for ***safety assurance***, i.e., the development of confidence that safety has been sufficiently ensured by the Provider, such that a decision can be made to accept the safety risk of the system. It is the responsibility of the Acquirer to articulate expectations regarding the performance of the product or service in question, including safety performance, by developing a comprehensive set of performance requirements that the Provider is expected to meet. It is also the responsibility of the Acquirer to evaluate the product or service delivered, or proposed to be delivered, in terms of the degree to which it satisfies those requirements. Correspondingly, by accepting a product or service from a Provider, the Acquirer is also accepting as adequate its assessed safety performance, as well as the risk that its actual safety performance might be less than its assessed performance.
- **Providers** – A Provider is a NASA or contractor organization that is responsible for ***safety ensurance***, i.e., the reduction and elimination of system hazards and the achievement of adequate safety performance through design, procurement, fabrication, construction, and in the case of a service provider, operation. A Provider is tasked by a customer or supervising organization (i.e., the Acquirer) to produce a product or service. It is the responsibility of the Provider to deliver a product or service that is consistent with the stated requirements, including safety performance requirements. To substantiate that the product or service is indeed consistent with these requirements, and to convey to the Acquirer what resources and pursuits are needed in order to achieve and maintain the committed level of safety performance, the Provider develops a RISC to support Acquirer decision making regarding the acceptability of the safety performance of the product or service.

In cases where the Provider organization tasks another organization to produce a product or deliver a service it needs to fulfill its requirements, the Provider organization takes on the role of an Acquirer with respect to the tasked organization, with all the organizational responsibilities that this entails.

Although Acquirers and Providers represent distinct entities, a key message of this handbook is that they work together towards a common goal of achieving a safe system. Correspondingly, it is expected that relationships between Acquirers and Providers will be cooperative, involving timely communication and coordinated action as needed to best achieve the desired result. For both the provider and the acquirer, there is presumed to be a systems engineering function and a system safety function that that are technically independent but work together to provide a product that is timely, cost effective, accomplishes the technical objectives of the mission, and is safe. Agreements negotiated between the acquirer and the provider are sometimes referred to colloquially in this volume as a “handshake.”

2.3 Adequate Safety

For the purposes of this document, safety is defined as freedom from those conditions that can cause death, injury, occupational illness, damage to or loss of equipment or property, or damage to the environment.[3] Although this definition is broad, it focuses exclusively on physical, rather than functional, consequences. For systems such as non-recoverable spacecraft, damage to or loss of equipment may be meaningful only insofar as it translates into degradation or loss of mission objectives. Therefore, it is also

[3] This definition of safety is consistent with NPR 8715.3C and MIL-STD-882E.

reasonable to include, within the definition of safety, freedom from conditions that can cause loss of mission (LOM). In any case, the specific populations included in the definition of safety are context dependent, and it is up to the involved parties, including stakeholders, to unambiguously define what constitutes safety for a given application in a given environment. Figure 2-1 illustrates a decomposition of safety into a specific set of at-risk populations.

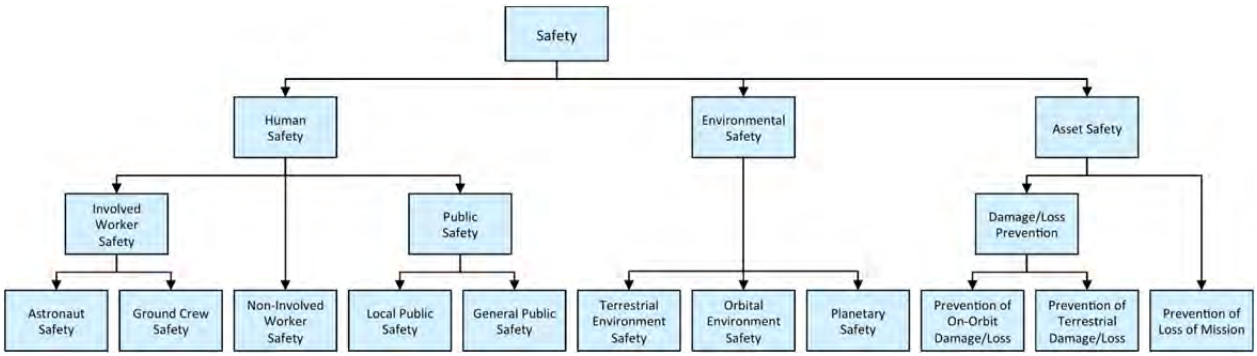


Figure 2-1. Example Set of At-Risk Populations with Respect to Safety

Just as the scope of conditions relevant to safety is application specific, so too is the degree of "safety" that is considered acceptable. NASA does not expect to attain absolute safety, but the Agency strives to attain a degree of safety that fulfills obligations to at-risk populations and addresses Agency priorities. An adequately safe system is not necessarily one that completely precludes all conditions that can lead to undesirable consequences. Rather, an adequately safe system is one that adheres to the following fundamental safety principles: 1) meeting minimum tolerable levels of safety; and 2) being ASARP (see Figure 2-2). These two principles must be maintained throughout all phases of the system life cycle. Opportunities to improve safety exist from concept studies to closeout, and efforts to that end must be operative throughout the life cycle.

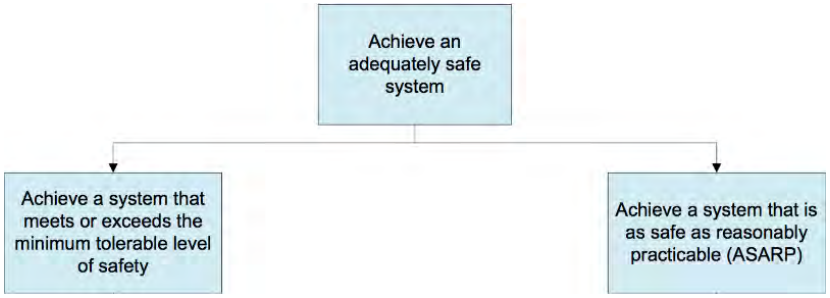


Figure 2-2. Fundamental Principles of Adequate Safety

2.3.1 Meeting Minimum Tolerable Levels of Safety Performance

Minimum tolerable levels of safety performance serve a risk acceptance function. Setting these levels involves consideration not only of societal issues relating to risk tolerance, but also of what is feasible given current capabilities and technological development potential. It involves the conduct of safety studies by the Acquirer during pre-Formulation to better understand the safety risks involved, recognizing that conceptual studies in the early stages of development are often notional due to the large uncertainties involved. The results of such studies should be the starting point for the development of minimum tolerable levels of safety, not as a means of establishing them directly. It is important to set levels of safety that are feasible to achieve and to assure, given technological, cost, and other constraints, but without compromising NASA's "safety-first" core value policy.

These minimum tolerable levels of safety performance are not necessarily constant over the life of a system. As a system is operated and information is gained as to its strengths and weaknesses, design and

operational modifications are typically made, which, over the long run, improve its safety performance. This is typically the case for production line items where operating experience can inform the design and operation of future units, and for reusable systems that can be modified prior to reuse. It is less the case for one-time, non-recoverable systems where the opportunity to modify the system is limited to, e.g., software upgrades from the mission control center.

In particular, an initial level of safety performance may be accepted for a developmental system, with the expectation that it will be improved as failure modes are "wrung out" over time. In such cases the level of tolerable safety can be expressed as a safety threshold against which initial system performance is assessed, and a safety goal against which future performance will be assessed. The application of safety thresholds and goals for this purpose is addressed, for example, in NASA guidance and requirements related to the ISS and to human rating [7, 8]. Figure 2-3 illustrates a safety threshold and safety goal for a generic safety performance measure.

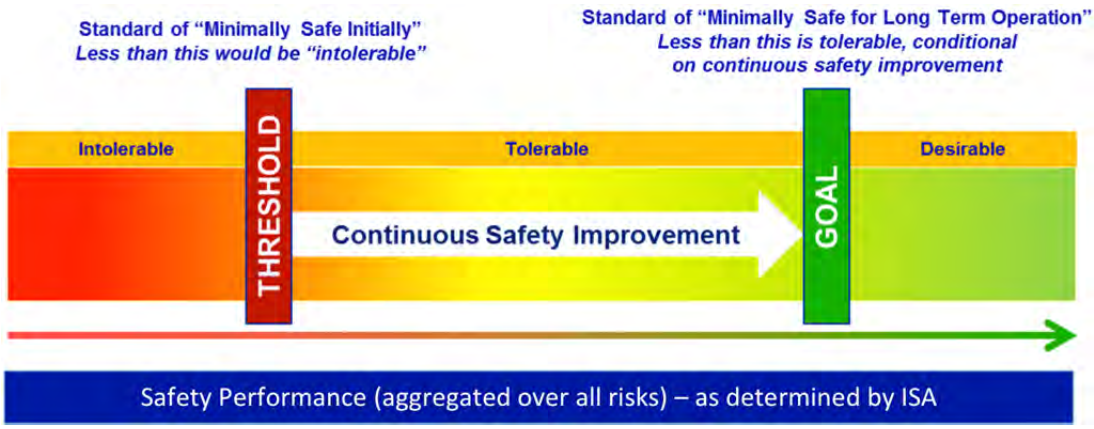


Figure 2-3. Safety Threshold and Safety Goal for a Generic Safety Performance Measure

SAFETY THRESHOLDS AND SAFETY GOALS

Safety Threshold – The level of safety performance against which initial system performance is assessed (i.e., the maximum tolerable probability of harm or loss from all sources of risk when the system is first put into operation)

Safety Goal – A target level of safety performance that is expected from continuous safety upgrades and improvements to the system (i.e., the maximum tolerable probability of harm or loss from all sources of risk when the system has been operational long enough to uncover and correct significant unknown and underappreciated risks)

KNOWN, UNKNOWN, AND UNDERAPPRECIATED RISKS

Known Risk – A scenario affecting safety performance that has been correctly identified and accurately assessed with respect to its likelihood of occurrence and potential severity of harm or loss

Underappreciated Risk – A scenario affecting safety performance that has been correctly identified but for which the likelihood of occurrence and/or potential severity of harm or loss are underestimated

Unknown Risk – A scenario affecting safety performance that has not been identified and is therefore unknown at the time of analysis

Minimum tolerable levels of safety need not be defined for every safety performance measure defined for the system. They are most appropriate for safety consequences considered catastrophic by the relevant stakeholders. Such consequences may include human death, planetary contamination, or loss of vital

assets. For safety consequences of lesser severity, such as loss of a replaceable asset, ASARP implementation may be sufficient for adequate safety.

2.3.2 Being As Safe As Reasonably Practicable

An adequately safe system is as safe as reasonably practicable (ASARP).[4] A determination that a system is ASARP entails weighing its safety performance against the commitments and tradeoffs needed to further improve it. The system is ASARP if an incremental improvement in safety would require an intolerable or disproportionate deterioration of system performance in other areas. Thus, a system that is ASARP is one where safety is given the highest priority within the constraints of operational effectiveness, time, and cost, throughout all phases of the system life cycle. Figure 2-4, a close adaptation of Figure 3-4 in Volume 1, illustrates the ASARP region for a generic set of alternatives.

Being ASARP is a separate and distinct consideration from meeting a minimum tolerable level of safety. ASARP makes no explicit reference to the absolute value of a system's safety performance or the tolerability of that performance. It is strictly concerned with the system's safety performance relative to a comprehensively identified set of other alternatives. ASARP reflects a mindset that values safety improvement regardless of the current level of safety. It is an integral aspect of a good systems engineering process that guides risk-informed decision making throughout the system life cycle. Correspondingly, the condition of ASARP is demonstrated to some extent through process considerations.

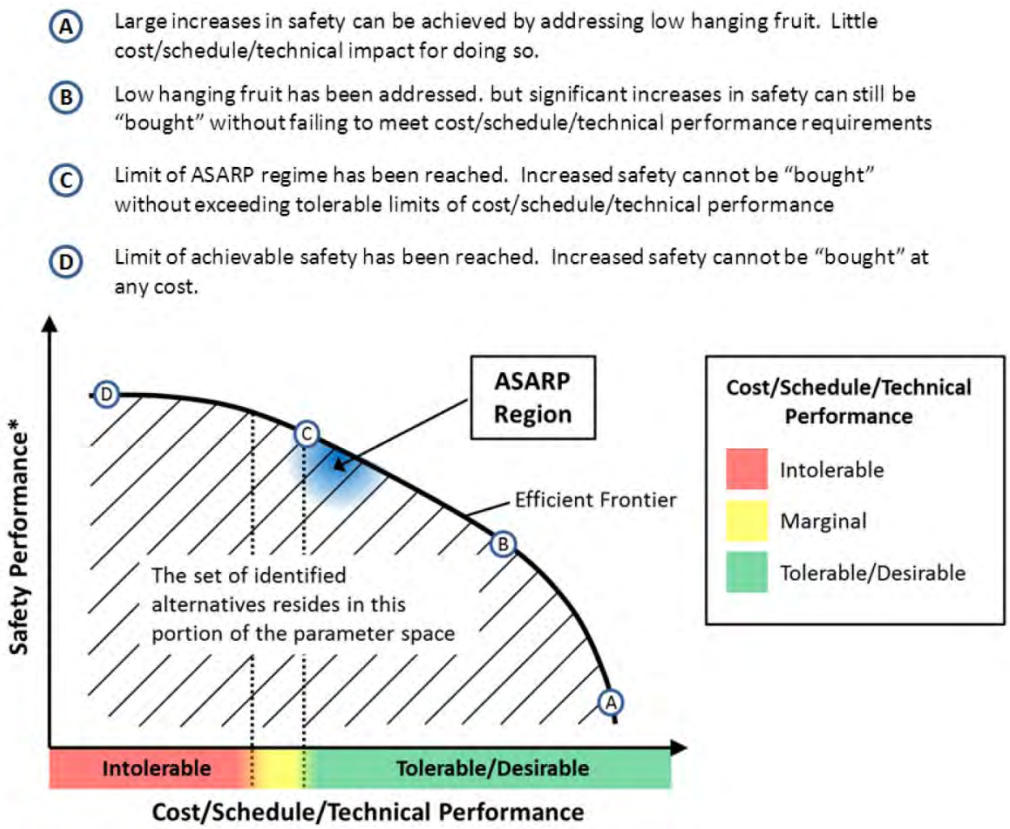


Figure 2-4. As Safe As Reasonably Practicable (ASARP)

[4] The ASARP concept is closely related to the "as low as reasonably achievable" (ALARA) and "as low as reasonably practicable" (ALARP) concepts that are common in U.S. nuclear applications and U.K. Health and Safety law, respectively [9, 10].

2.3.3 State of the System with Respect to Adequate Safety

Table 2-1 illustrates the four possible states of a system with respect to adequate safety, based on whether or not it meets the minimum tolerable level of safety and whether or not it is ASARP.

Table 2-1. Possible Combinations of Meeting the Minimum Tolerable Level of Safety and Being ASARP

System is ASARP?	System meets minimum tolerable level of safety? Yes	System meets minimum tolerable level of safety? No
Yes	System is adequately safe	System is inherently unsafe
No	System is sub-optimally safe as designed	System is unsafe as designed

Table 2-1 illustrates that:

- A system that meets or exceeds the minimum tolerable level of safety, and is also ASARP, is deemed adequately safe (upper left quadrant). Such a system not only meets existing expectations of safety performance (for a system of its type), but to the extent that it exceeds them, it raises safety performance expectations going forward. Few, if any, technologies have inherent limits on their safety performance. The ASARP principle is the driver of safety growth over the long-term, and in its wake, minimum tolerable thresholds of safety rise as stakeholders revise their understandings about what levels of safety performance are possible and, consequently, expected.
- A system that meets or exceeds the minimum tolerable level of safety, but is not ASARP, is sub-optimally safe as designed (lower left quadrant). Opportunities to improve safety are still available within the existing program constraints and should be pursued.
- A system that does not meet the minimum tolerable level of safety, despite being ASARP, represents an inherently unsafe system type in that minimum safety is not achievable without intolerable sacrifice of performance in other domains (upper right quadrant).
- A system that neither meets minimum tolerable level of safety, nor is ASARP, is unsafe as designed (lower right quadrant). Modifications that make the system ASARP may or may not also result in meeting the minimum tolerable level of safety.

2.4 Relationship of System Safety to Systems Engineering and Risk Management

In NPR 7123.1B, the systems engineering approach is defined as: “the application of a systematic, disciplined engineering approach that is quantifiable, recursive, iterative, and repeatable for the development, operation, and maintenance of systems integrated into a whole throughout the life cycle of a project or program.”

The approach to system safety articulated in this handbook recognizes the substantial overlap between systems engineering, risk management, and system safety. Risk management and systems engineering are both concerned with the achievement of defined objectives. Broadly speaking, systems engineering is the means by which the objectives are met, and the role of risk management is to provide a control function

for systems engineering to assure that the development is, and will remain, on track to meet the objectives, across all mission execution domains.[5] This relationship is illustrated in Figure 2-5.

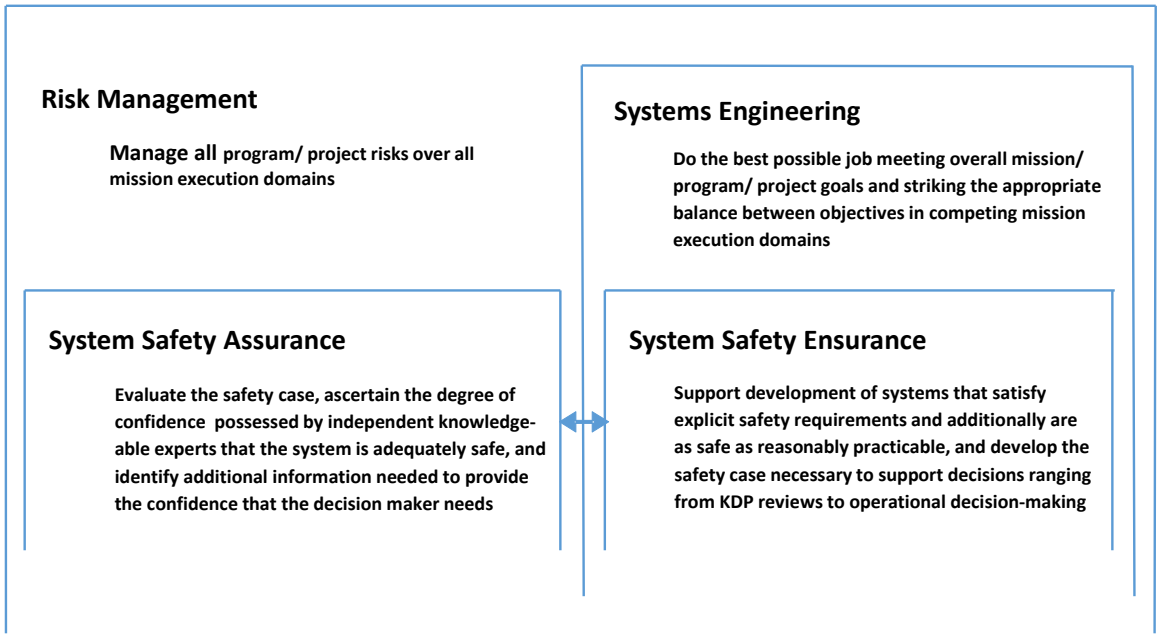


Figure 2-5. Relationships among Risk Management, Systems Engineering, and System Safety in a Program/Project Context

System safety is an input to both systems engineering and risk management. Considering that the safety performance of a system is a stakeholder concern in much the same way that technical performance capabilities such as payload mass to orbit are stakeholder concerns, system safety *ensurance* is an integral part of systems engineering efforts to develop a system that satisfies stakeholder objectives across all domains of performance, including safety, technical, cost, and schedule. System safety *assurance* is an evaluation activity that maintains functional independence from systems engineering in order to be most effective in providing confidence that the system is indeed adequately safe. Transfer of information between the ensurance and assurance parts of system safety occurs through the natural interactions between provider and acquirer, as will be discussed in Section 3.6, and both parts contribute to risk management efforts to identify and respond to conditions that may arise and threaten the achievement of the system's baselined safety performance.

Although each of the mission execution domains of safety, technical, cost, and schedule have unique characteristics, system performance in the safety domain is of particular concern to NASA, as reflected in safety being a NASA core value according to NPD 1000.0A, Governance and Strategic Management Handbook [11]. Moreover, safety performance is typically probabilistic, in that adverse safety consequences occur only sporadically, which means that 1) safety performance is not directly observable, but must be inferred; and 2) assessments of safety performance are inherently uncertain, since they are subject to the same margin of error considerations that apply to stochastic phenomena generally. Accordingly, while system safety is carried out as an integral part of systems engineering and risk management, Chapters 4 through 7 of this volume provide guidance that incorporate the risk management requirements in NPR 8000.4 and the systems engineering processes in NPR 7123.1 into system safety activities.

[5] Technical risk management, which is a subset of risk management, is a systems engineering technical control process as discussed in NPR 7123.1B.

2.5 The Risk-Informed Safety Case

A key element of this objectives-driven approach is the use of the risk-informed safety case (RISC): a structured argument, supported by a body of evidence that provides a compelling, comprehensible, and valid case that a system is or will be adequately safe for a given application in a given environment. The RISC addresses each of the objectives defined for the system, providing a rational basis for making informed risk acceptance decisions at relevant decision points in the system life cycle. The RISC is not an add-on to today's system safety practices; it is a means of organizing existing, often disparate, system safety products such as HAs and PRAs into a unified, coherent, evidence-based argument that the required level of safety has been attained. In addition to furnishing the risk acceptance rationale, the RISC can serve as a roadmap for risk management during deployment, including pursuits such as precursor analysis, a process for learning from flight experience that is now considered a necessity for risk management of human space flight. RISCs are prepared for the purpose of making the case for safety at Key Decision Points (KDPs).

The term "risk-informed," within the acronym "RISC," is intended to convey the idea of a safety case that not only attempts to argue inductively that the system is adequately safe, but also attempts to argue deductively that all credible scenarios that could lead to a significant risk of system failure have been identified, rigorously analyzed, and conscientiously responded to. One of the principal intents of a RISC is to prevent or mitigate "confirmation bias," a term used to denote the tendency to favor information that confirms one's prior beliefs or hypotheses, thereby preventing analysts from using the safety case as "simply a paper exercise that repeats what the engineers are most likely to have already considered" [12]. Confirmation bias is addressed by approaching safety ensurance from two complementary directions: first, by attempting to discover the ways in which the system may be unsafe (by looking at possible failure scenarios), and second, by attempting to show that the system is safe nonetheless (by formulating the overall safety argument).

SAFETY CASES AND RISK-INFORMED SAFETY CASES

- **Safety Case** - In general, a "safety case" is a structured argument, supported by a body of evidence that provides a compelling, comprehensible, and valid case that a system is or will be adequately safe for a given application in a given environment.
- **Risk-Informed Safety Case** - The term "risk-informed" is used here to emphasize that a determination of adequate safety is the result of a deliberative decision making process that necessarily entails an assessment of risks and tries to achieve a balance between the system's safety performance and its performance in other areas. The RISC, which evolves over the course of the system life cycle, supports decision making at system life-cycle reviews and other major decision points.

Because one of the functions of the RISC is to argue that a comprehensive effort has been made to identify all credible and significant accident scenarios, the RISC provides a means for helping to ensure that all significant hazards that could lead to system failure have been accounted for, properly analyzed, capably managed, and adequately controlled. In addition, because the RISC emphasizes a holistic system-wide approach, it fosters attention to hazards whose causes and/or effects may cut across subsystems and emerge at system level rather than subsystem level. In this sense, the development of the RISC serves as a self-assessment aid for the providers of the RISC, helping them to ensure that all important knowable scenarios have been identified and that all important knowable interactions between different parts of the system have been analyzed.

Equally important, the RISC submitted by a provider undergoes thorough evaluation by an independent evaluation team, acting for the acquirer, consisting of both subject matter experts and people who are highly knowledgeable about the system. This evaluation of the RISC by a qualified, independent team helps provide the assurance needed by decision makers who are called upon to either accept or reject the safety risks.

A RISC consolidates and organizes the applicable body of evidence into a valid case that argues that a system is (or will be) adequately safe for a given application in a given environment. The applicable body of evidence may include (but not necessarily be limited to):

- Safety analyses
- Safety-critical item designations
- Test results
- Safety management program elements
- Qualifications of the workforce
- Verification and validation procedures
- Adherence to norms and standards
- Attention to best practices and lessons learned
- Risk-informed justification for seeking waivers
- Adequacy of margins

Much, if not most, of this evidence derives from safety ensurance activities that need to be conducted regardless of whether or not a RISC is produced. The job of the RISC is to use such evidence to maximum effect in communicating the safety of the system.

In the context of NASA systems engineering, *RISC* refers to the totality of safety-related documentation submitted to a given technical review. As such, the documentation requirements of the RISC are consistent with the entrance criteria for the relevant review, as itemized in NPR 7123.1B, NASA Systems Engineering Processes and Requirements. Similarly, the criteria for evaluating the adequacy of the RISC are consistent with the corresponding technical review success criteria in the same NPR.

The RISC supports the informed acceptance of safety risk by the various parties who need to accept the risk. It enables a structured, critical, and skeptical evaluation by the system acquirer, facilitating the identification of specific assurance deficits whose remediation may be designated as a condition of further system development or use. It organizes the relevant evidence into a coherent case for safety, supporting the needs of decision makers as the system moves through life-cycle reviews and other major decision points. Properly formulated, the RISC meets the needs of foreseeable “certification” processes, e.g., for human rating or for autonomous systems.

A RISC is not a “one-shot deal” that sits on a bookshelf or hard drive once it has been completed. When significant programmatic changes occur (e.g., budget reductions or unanticipated technological challenges), the existing safety case may no longer be valid and may need to be changed to reflect evolving realities. An updated RISC is presented at each major milestone, making it a living document that evolves over the system life cycle in tandem with system evolution and operational experience. This use of a RISC is compatible with NASA's established practices of critical review and evaluation by experts tasked with assessing the adequacy of the system's safety on behalf of the decision makers and at-risk communities.

Within a planning context, a RISC may also be used to facilitate evaluation of the effects of *proposed* programmatic changes on the case for safety. This use of a RISC is compatible with NASA’s desire to be prepared for new challenges or new opportunities before they manifest.

3. Overview of the System Safety Framework

The NASA system safety framework, introduced in Volume 1, provides a structured model for planning, conducting, and documenting system safety activities in a manner that meets stakeholder objectives and provides the technical basis for decision making, including risk acceptance.

The framework consists of the following system safety elements as illustrated in Figure 3-1 (which is a modified version of Figure 2-4 in Volume 1): requirement setting, safety ensurance, safety assurance, and risk acceptance (see color-coded Key to Figure 3-1). Objectives and requirements, RISC evaluation, and risk acceptance decisions are the within the purview of the Acquirer, whereas ensurance activities and RISC preparation are within the purview of the Provider. The handbook provides guidance for all elements of the system safety framework except for risk acceptance. The acceptance of risk is a decision process that is intended to be informed by the RISC and by the evaluation of the RISC (i.e., RISC-informed) but not constrained by the RISC or by its evaluation (i.e., not RISC-based).

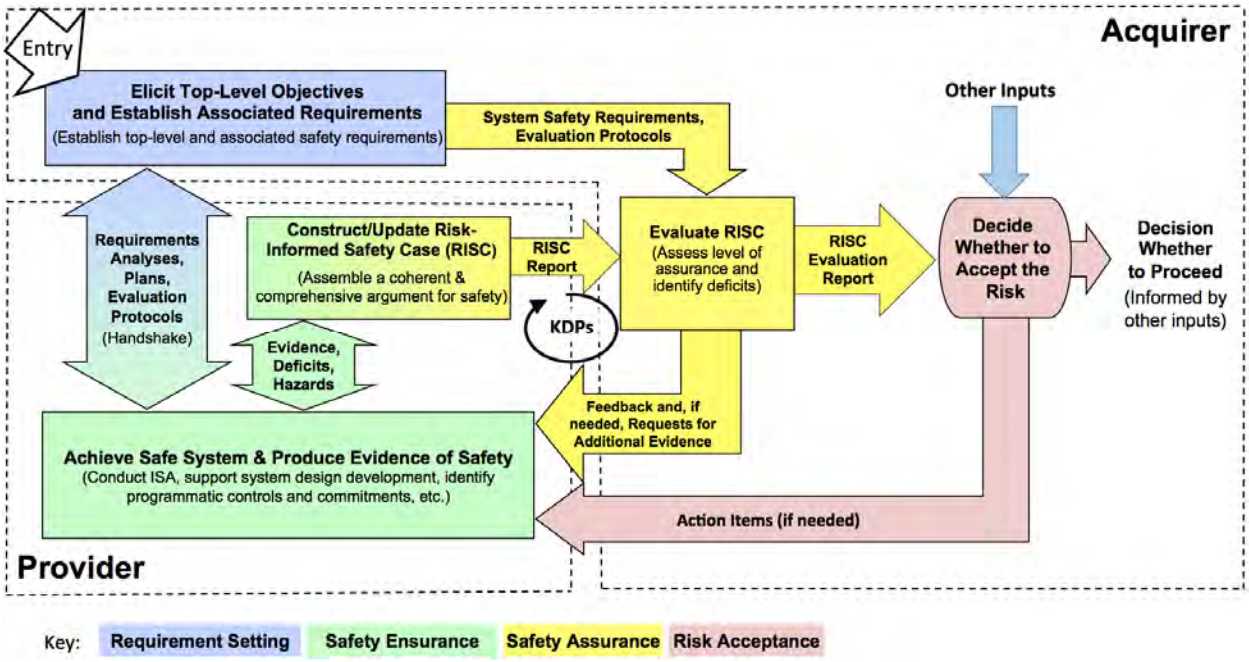


Figure 3-1. The NASA System Safety Framework

3.1 Top-Level Objectives and Associated Requirements

As discussed in NPR 8000.4A, at the outset of a program or project, the objectives, deliverables, performance measures, baseline performance requirements, resources, and schedules that define the task to be performed are negotiated between the organizational unit performing the task (Provider) and the organizational unit sponsoring the task and responsible for oversight (Acquirer). As part of this process, a set of safety objectives is negotiated consistent with the two fundamental safety principles discussed above, namely meeting minimum tolerable levels of safety and being ASARP.

These two principles are decomposed into specific safety objectives to be met by the system. By specifying safety objectives down to a level where they can be clearly addressed by systems engineering processes, an operational definition of safety is created that enables the processes to be developed and evaluated in terms of the safety objectives. By adequately meeting these so-called *operational safety objectives*, then by virtue of their derivation from fundamental safety principles, the system can be said to be adequately safe.

Figure 3-2 illustrates the derivation of generic operational safety objectives, including the application of safety performance margin to minimum tolerable levels of safety, so that safety requirements can be developed that maintain a reserve for the expected presence of unknown and/or underappreciated (UU) scenarios. The figure shows that even in the absence of minimum tolerable levels of safety, the ASARP principle is still operative. Traditional, deterministic safety practices such as requiring redundancy where practical, have implicitly recognized the ASARP principle as fundamental to system safety.

Figure 3-2 also implies that organizations provide the framework in which safety is achieved, since objectives such as "Prioritize safety during design/realization/operation decision making" and "Be responsive to new safety-related information" imply that the organization puts safety first and that it is structured to respond to safety-related opportunities. The reader will find many organizational factors throughout the document that have the ability to affect the achievement of safety objectives. Organizational factors are crosscutting and have the ability to affect the achievement of safety objectives across the objectives hierarchy.

The development of safety requirements is the purview of the Technical Requirements Definition Process of the NASA Systems Engineering Engine [13]. It is carried out collaboratively between the Provider and the Acquirer, but the evaluation regarding the appropriateness of the collection of safety requirements resides with the Acquirer. Early in the system life cycle, the Provider conducts a System Safety Requirements Analysis (SSRA).[6] The SSRA serves to clarify the detailed requirements (including, but not limited to, engineering requirements) that the Provider expects to address in the ensuing development, and which form the basis of the Provider's System Safety Management Plan (SSMP). In addition, the SSMP focuses on the collection of process requirements.

Figure 3-3 illustrates the interaction/iteration between the Acquirer and Provider during the development of system-specific safety requirements.

The transition from safety objectives, which are defined by the Acquirer, to system safety ensurance activities, which are conducted by the Provider, is accomplished by translating the objectives into safety requirements that the Acquirer then levies on the Provider. Broadly speaking, these requirements fall into the following categories: *top-level safety performance requirements*, *lower-level safety performance requirements*, *safety-related engineering requirements*, and *safety-related process requirements*.

3.1.1 Top-Level Safety Performance Requirements

Requirements associated with minimum tolerable levels of safety performance are typically probabilistic, because the safety performance measures that they explicitly constrain are typically probabilistic or statistical, such as the probability of loss of crew, the probability of loss of mission, or casualty expectation.[7] Such probabilistic requirements are often verified by *synthetic* analysis methods such as PRA, which quantify the system's safety performance based on explicit identification of scenarios leading to adverse safety consequences. This is particularly true in aerospace applications, where the relatively low numbers of flights are insufficient in and of themselves to provide a statistically sound basis for claims of satisfaction of safety performance requirements.

Verification of probabilistic requirements is particularly challenging, because synthetic methods[8] are vulnerable to incompleteness of scenario identification as well as an incomplete understanding of the

[6] The SSRA is similar in function to the System Requirements Hazard Analysis (SRHA) of MIL-STD-882E, Task 203.

[7] In addition to probabilistic requirements, top-level safety performance requirements generally include top-level functional requirements, which describe the capabilities of the system (what the system must do).

[8] By "synthetic analysis methods" we mean methods that produce system-level risk estimates by aggregating the effects of explicitly identified individual contributors to that risk.

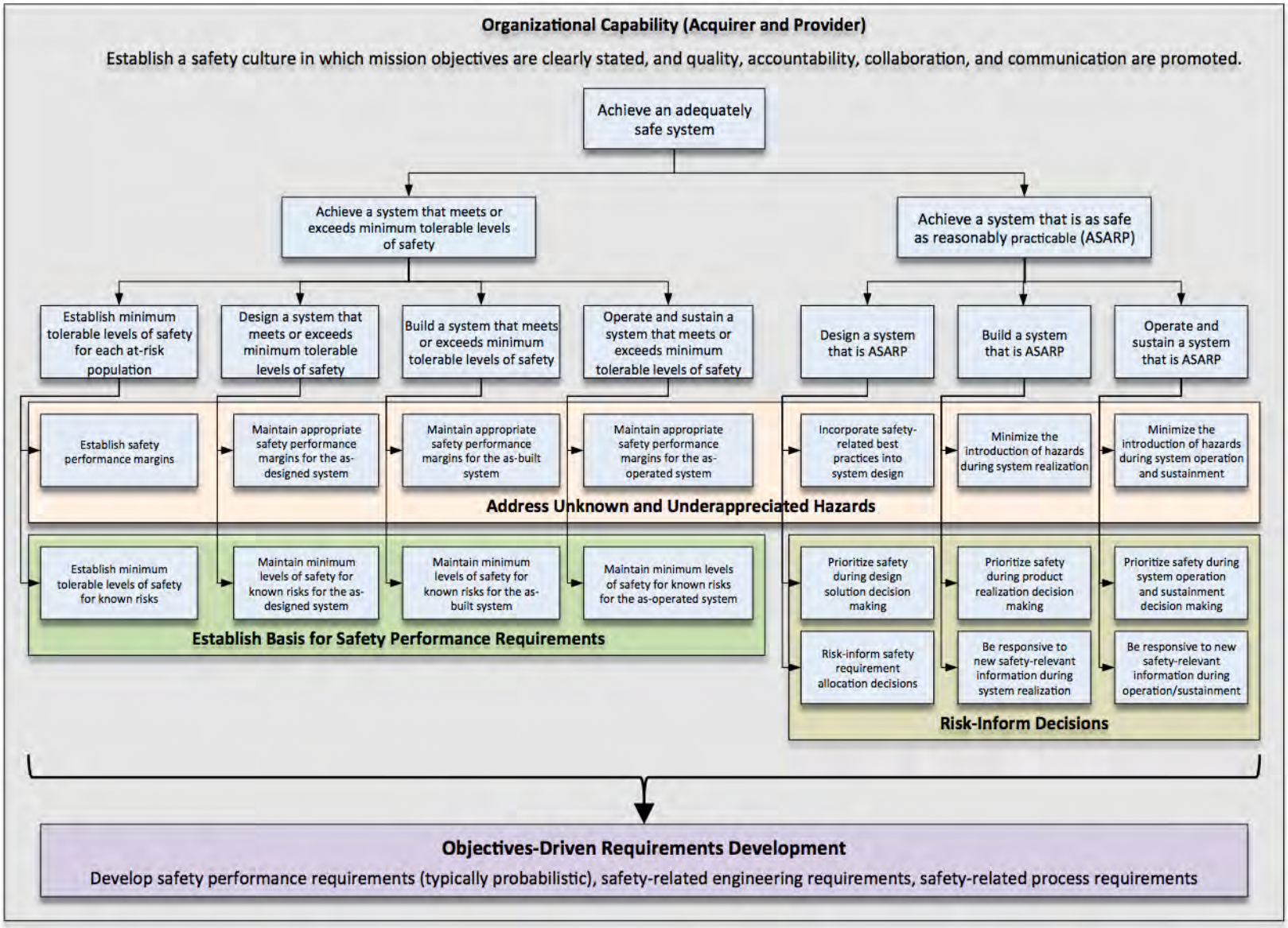


Figure 3-2. Objectives-Driven Requirements Development (notional)

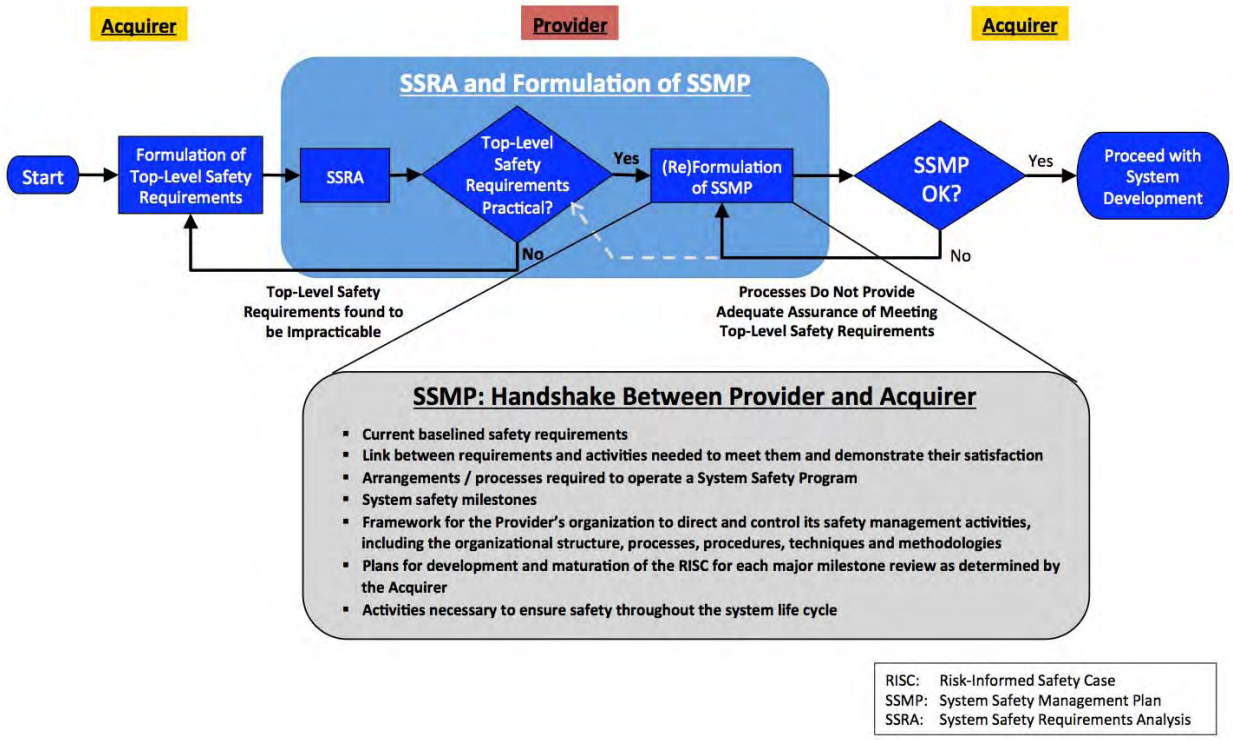


Figure 3-3. Interaction/Iteration between the Acquirer and Provider during the Development of System-Specific Safety Requirements

probability of occurrence, leaving a potentially substantial portion of the actual safety risk unaccounted for by the verification protocols.

The factors that tend to increase the likelihood of UU scenarios have been extensively studied in various contexts, and it has been found that the principal risk factors that affect safety also affect cost, schedule, and technical performance risks. For example, the following factors are often cited as having a strong influence on the ability to meet requirements pertaining to all of these mission execution domains [14-27]:

- Over-optimism, complacency, and discounting of risk
- Diffusion of responsibility and authority with inadequate oversight
- Limited communication channels and poor information flow
- Lack of testing and modeling at as-flown or as-claimed conditions
- Poor or missing specifications
- Unnecessary complexity
- Reuse of technologies without appreciation for the differences in how they are applied
- Inadequate review
- Pressures to meet budget and schedule constraints that are not only tight but also changing

Therefore, safety management cannot be accomplished in isolation, and is affected in a positive way by managing risk to costs, schedule, and technical performance.

To account for the contribution of UU scenarios to the system's actual (though unknown) safety performance, safety performance margins are introduced between the minimum tolerable levels of safety (initially the safety threshold), which refer to actual safety performance, and the (more stringent) safety performance requirements against which the results of synthetic analyses are compared (see Figure 3-4).

Therefore, if the system meets the safety performance requirements levied on it, and if the safety performance margins used to develop the requirements have a sound basis (e.g., consistency with historical experience), then there is a rational basis for the claim that the system's actual safety performance is within minimum tolerable levels.

SAFETY PERFORMANCE MARGIN

Safety Performance Margin – An incremental margin subtracted from the safety threshold or goal to account for the estimated total effects of unknown, un-quantified, and under-evaluated hazards. It is estimated from analysis of historical experience with similar technologies taking into account the complexity of the system, the degree to which new technology is being used, and the degree to which new operating environments are being introduced. The size of the margin decreases with time in operation as unknown and underappreciated risk scenarios are uncovered and corrected.

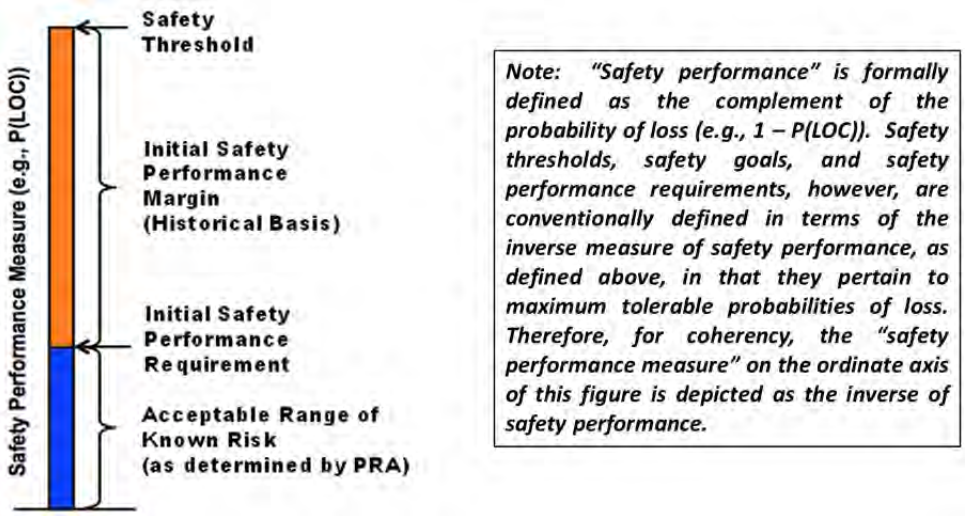


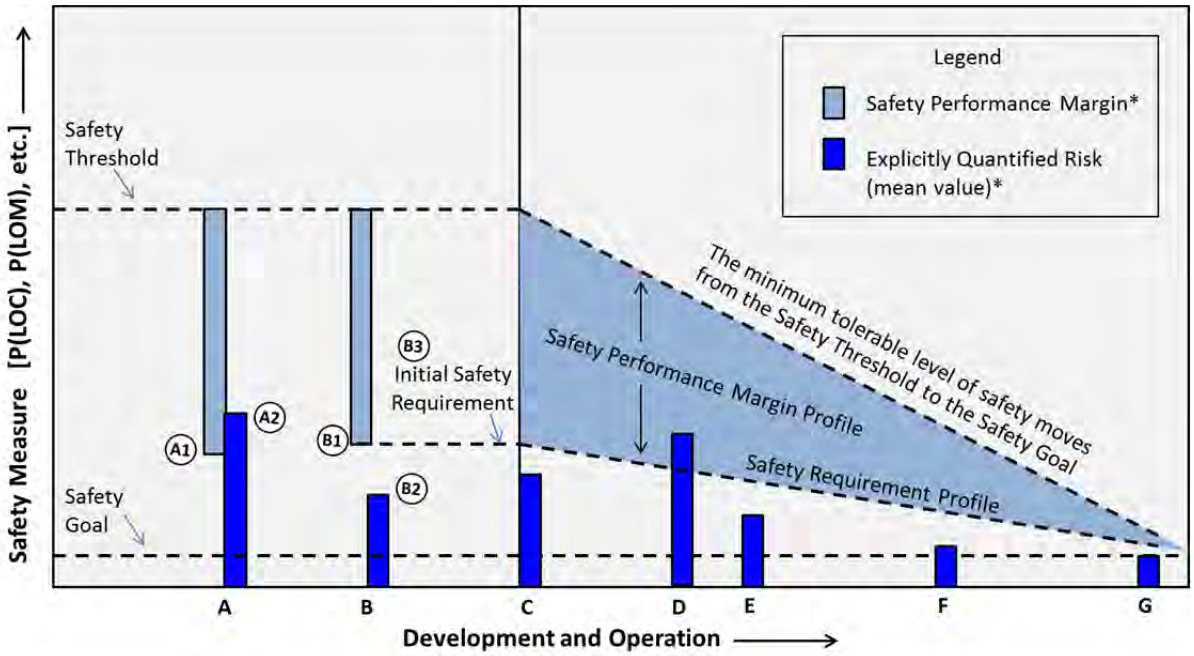
Figure 3-4. Relationship between the Initial Safety Performance Requirement, Initial Safety Performance Margin, and Safety Threshold

Both the safety threshold and the safety performance margin normally diminish to lower values as safety information is gained through both testing and actual flight experience, and as both preventive and corrective actions are undertaken to make the system increasingly safe. For example, Figure 3-5 illustrates how the safety performance requirement and margin might vary within a defined profile for the minimum tolerable level of safety that moves over time from the safety threshold (for initial flights) to the safety goal (for long-term operation). Initially, given a concept of operations, an analysis is performed to determine a reasonable value for the safety performance margin (the height of the light blue bar at time A). Subtraction of this margin from the safety threshold leads to the limit on the explicitly quantified risk (the point marked A1). In addition, a "first-order" risk analysis is performed on the preliminary system design to scope out the mean value of risk due to known scenarios (the point marked A2). Since the value of the safety measure at A2 exceeds the value at A1, the system does not satisfy the minimum tolerable level of safety.

To remedy the situation, the following actions may be taken: 1) the safety performance margin may be reduced by making provisions to reduce the UU risks (B1); and/or 2) the system design details may be refined and controls may be added to further mitigate the explicitly quantified risk and improve safety (B2).[9] At this point an initial safety requirement for explicitly quantified risks is specified (B3) as being

[9] The two are not independent. Organizational, programmatic, and design philosophy changes can impact the quantitative assessment of risk, and design refinements can impact the factors that affect safety performance margin.

equal to the value of the safety measure at point B1. Additionally, a decreasing safety performance margin profile can be derived from historical experience and subtracted from the profile for the minimum tolerable level of safety, to obtain a safety requirement profile against which the explicitly quantified risk will be assessed over the operational life of the system. At time "C" the known risk of the as-built system satisfies the safety requirement. At time "D," newly discovered scenarios are added to the risk analysis, increasing the explicitly quantified risk beyond the safety requirement. Mitigations are introduced and the system is re-analyzed (E) to demonstrate that the known risk has been brought back within the requirement.[10] During system operation, proactive safety upgrade and improvement programs reduce the risk in increments (F) until the safety goal is met (G).



* The bars represent the value of an appropriate statistic that reflects an acceptable degree of certainty that the safety requirement is met.

Figure 3-5: Relationship between the Safety Performance Requirement, Safety Performance Margin, and Minimum Tolerable Level of Safety as the System Matures

3.1.2 Lower-Level Safety Performance Requirements

To facilitate the achievement of the high-level probabilistic requirements without requiring multiple design iterations, it is common practice for probabilistic allocations to be made from the top level down to the lower elements of the system. In this way, the ownership of the safety requirements is similarly allocated from program/project level down to the lower organizational units supporting the program/project.

Approaches for allocating top-level safety performance requirements to lower levels were discussed in Section 4.4 of Volume 1.

[10] If the risk could not be brought within the requirement, the issue would be elevated according to the program/project risk management process, and could potentially be resolved by rebaselining the safety requirement, if organizational and/or programmatic changes provide an adequate basis for reducing the safety performance margin.

3.1.3 Safety-Related Engineering and Process Requirements

In addition to setting probabilistic requirements at various levels of the system, there must be assurance that the assumptions used in the analysis of potential accident scenarios are valid, e.g., that controls have the capabilities, reliabilities, and availabilities that they are credited with. Therefore, verifiable, system-specific engineering and process requirements are levied on those attributes of the system that demonstrably contribute, via safety analysis, to that performance. It is the combination of satisfactory safety analysis (per agreed-upon analysis protocols) and compliance with safety-related engineering and process requirements that counts as compliance with probabilistic safety performance requirements.

Safety-related engineering requirements are those that are related to observable system attributes that the Acquirer levies on the Provider either as derived requirements or as a matter of best practice (e.g., a particular level of failure tolerance or a particular design safety factor).

ALLOCATED REQUIREMENTS

- Allocated requirements are quantitative requirements that are apportioned from system or subsystem level to lower levels, where the units of measure remain the same as at the higher level. For example, the overall maximum allowable probability of LOC during abort may be apportioned to the failure probabilities of the abort motor, the jettison motor, the attitude control motor, other subsystems, and their interfaces, since it is important not only that they work correctly individually, but also that their composition works correctly when the individual subsystems work correctly. The maximum failure probabilities at the lower level are allocated to stay within the maximum failure probability at the higher level.

DERIVED REQUIREMENTS

- Derived requirements may be quantitative or qualitative and are developed at a lower level of a system to implement a higher-level requirement. Derived requirements arise from constraints, consideration of issues implied but not explicitly stated in the Acquirer's requirements, or factors introduced by the selected architecture or the design. For example, it may be determined that in order for the overall probability of LOC during abort to be less than X, the maximum acceleration during abort must be less than Y. The limit on acceleration is a derived requirement.

Safety-related engineering requirements need not have an explicit basis in the ISA. They may also be levied on the basis that not doing so would be inconsistent with best practices and lessons learned. The levying of such requirements is considered part of ASARP in the objectives decomposition of Figure 3-2. They typically reflect the application of consensus engineering codes and standards where these are deemed to promote good performance in the system. However, the blanket imposition of requirements that are historical in origin has the potential to lead to sub-optimal results, or even to over-constrain the system, particularly if the system is novel in some respect. Therefore, the ASARP principle also requires the ability to tailor the set of safety requirements to modify or remove those that are counterproductive in the particular application in question.

Safety-related process requirements are those that specify particular processes intended to support system safety. They specify *how* the system is designed, realized, or operated, rather than the system attributes themselves. Process requirements may relate to risk management, quality assurance, accident precursor, testing, training, or other processes and programs that affect safety.

3.2 System Safety Ensurance Activities to Achieve a Safe System

System safety ensurance activities are conducted by the Provider as part of overall systems engineering technical process activities and are focused on the achievement of the stated safety objectives. System safety ensurance activities not only ensure the safety of the system, but also produce the evidence of safety that will be used to support claims in the RISCs provided to the Acquirer at key decision points in the system life cycle.

Because of the diversity of Providers, each with its own particular set of systems engineering, system safety, and risk management processes, NASA recognizes the need for flexibility in the nature and composition of system safety activities, as long as they are able to achieve their operational safety objectives.

Within the current framework, system safety activities typically fall into the following categories: *conducting an integrated safety analysis*; *requirements development support*; *system design support*; *program control and commitments support*; and *performance monitoring support*.

3.2.1 Conducting an Integrated Safety Analysis (ISA)[11]

ISA refers to the development and analysis of potential accident scenarios that can credibly affect the safety performance of the system. ISA adheres to two key principles:

1. *ISA is Scenario Based* – ISA entails the development of a comprehensive set of accident scenarios that may lead to adverse consequences with respect to safety. ISA includes methods for identifying and analyzing these scenarios. This includes accident causes, contributing factors, effectiveness of controls (both existing and proposed), subsystem interactions, analysis of physical responses of the system to the environments it encounters, and analysis of the probability that the undesirable consequences will be realized. Conducting a scenario-based ISA is essential, because it is only by developing an understanding of how adverse safety consequences can be produced that effective measures can be taken to prevent (or mitigate) them.
2. *ISA is Conducted at the System Level* – ISA is by definition a system-level analysis that aims at comprehensively identifying all credible safety-related accidents associated with the system in the context of its intended operation. The scope of ISA is somewhat broader in this context than is sometimes invoked within NASA by the term integrated hazard analysis (IHA), which is typically used to describe the coordinated analyses between projects or elements that address those hazards or causes that are controlled by a project or element other than the one who is producing the analysis. The term "ISA" in this handbook defines an evolution of this more narrow definition, wherein it becomes an analysis of the whole system rather than a set of separate analyses to fill in gaps that remain after analyses of the subsystems.

ISA integrates different types of safety analyses (e.g., HA, PRA, phenomenological modeling) to the greatest extent possible. The ISA consolidates these separate analyses to produce a single comprehensive set of quantified safety performance measures that can be used to assess the standing of the system with respect to the levied safety performance requirements. The ISA is then used to risk-inform system design and operational decision making.

The ISA must be tailored to the particular phase in the life cycle at which it is conducted. As the system design evolves, the ISA is kept current, typically through the use of progressively more rigorous analysis

[11] The term "integrated safety analysis" and acronym ISA used in this handbook should not be confused with the term "integrated design and safety analysis," which has been used elsewhere with reference to the determination of failure tolerance requirements and the required amount of redundancy in design. IDSA in this sense refers to only a part of ISA.

techniques that model the system at progressively finer levels of detail. The ISA is maintained during system realization so that it can be used to inform decisions related to safety, such as test protocols. During system operation, the ISA is updated to reflect such things as design modifications and accumulating operational experience, including anomalies.

A principal use to which the ISA is put is demonstration of satisfaction of requirements, including safety requirements such as those derived from NASA safety thresholds and goals, and additional requirements derived or allocated from top-level requirements or levied independently by the Acquirer. As such, the ISA must conform to any analysis protocols that have been established for quantifying the safety performance measures used to assess compliance.

The focus of the ISA is on safety; however, in order to risk-inform trade studies and other decisions ensuring that the system is ASARP, the ISA must be integrable with other performance models in the mission execution domains of cost, schedule, and technical performance, as discussed in the NASA Risk Management Handbook [28].

3.2.2 Requirements Development Support

System safety activities support requirements development in a number of distinct ways. Through the early conduct of an SSRA, the Provider identifies applicable requirements by reviewing NASA, military, and industry standards and specifications, historical documentation on similar and legacy systems, etc. Additionally, using the ISA, the Provider translates any system-level probabilistic safety performance requirements levied by the Acquirer into objectively verifiable system-specific derived requirements (and associated verification procedures) that protect the assumptions underlying the system's assessed safety performance. Such derived requirements may involve, for example, prescribing specific levels of component reliability, specifying limits on environments produced by components such as particulate emissions or vibrations, or requiring a certain level of failure tolerance in a subsystem. In cases where the flowdown of requirements crosses organizational boundaries, the ISA provides a rational basis for allocating requirements from higher to lower levels of the system, and thereby from program/project level to subordinate organizations.

It may be the case that a levied requirement proves to be overly burdensome (such as by adding too much mass to the system) or sub-optimal (e.g., where alternate means are available to meet the intent of the requirement). System safety plays a role in these cases by assessing the potential consequences of tailoring the requirement, both through explicit modeling of safety performance using the ISA, and by qualitative consideration of the potential erosion of protection against unknown and underappreciated scenarios. This provides a technical basis for tailoring the requirement.

REQUIREMENT TAILORING

Tailoring refers to the process of adjusting or seeking relief from a levied requirement. Tailoring may be thought of as a process for winnowing down the sum total of best practices and lessons learned to those that are specifically relevant to the mission and/or system being investigated. The tailoring process results in the generation of deviations and waivers depending on the timing of the request.

3.2.3 System Design Support

It is generally recognized that an effective way that system safety activities promote safety is through the influence they can have over system design when properly integrated into the systems engineering process. System design support is of two broad types: best-practice-informed and ISA-informed. Best-practice-informed design support promotes safety by identifying applicable historically-applied safety-related engineering requirements and by assuring that proven strategies for optimizing safety are considered during system design decisions. ISA-informed support promotes safety by risk-informing

design decisions with an assessment of the safety performance of each contending alternative. These two types of design support work synergistically to achieve a design that is ASARP.

One traditional approach to design support is to apply a system safety design order of precedence. For example, MIL-STD-882E specifies that safety risk should be reduced by (in order): elimination of hazards through design selection; risk reduction through design alteration; incorporation of engineered safety features; provision of warning devices; and/or incorporation of signage, procedures, training, and personal protective equipment (PPE). However, this approach presumes that the strategies are ordered in terms of decreasing effectiveness, which might not be true in a specific application. This handbook takes the position that the strategies enumerated in such orders of precedence represent potentially fruitful design approaches, but that their relative impacts on safety performance should be analyzed as part of a risk-informed decision making process.

3.2.4 Program Control and Commitment Support

System safety promotes the development of program controls and commitments needed to ensure that the framework for safety is backed by sound administrative and management practices. Of particular importance to the maintenance of the system's safety performance is the identification of *safety-critical items* (SCIs) that are explicitly relied on for safety. A major vehicle for SCI identification is the ISA, which is used to identify the hardware, software, human, operational, and managerial system features upon which safe system operation depends. Such items can be explicit in the ISA (e.g., redundancies, backup systems) or they can be implicit (e.g., assumptions regarding component structural integrity). In either case, designating these items as safety-critical protects their safety functions by imposing rigorous and highly visible safety management provisions on them.

In the context of system safety, critical items have a broader meaning than in the context of reliability and maintainability. SCIs can include any element or attribute of the system that is important to safety, including hardware, software, interfaces between hardware and software, the human interface, operating procedures, and management practices. Therefore, SCIs have to evolve from a top-down approach that starts from an integrated system model. This is somewhat orthogonal to the more standard approach for deriving a critical items list, where a bottom up approach such as failure modes and effects analysis is used exclusively.

SCIs are, by definition, those items that have to function to ensure safety. They include any item whose failure could cause critical safety consequences, regardless of the likelihood of failure. For cases where the threat to safety involves harm to humans or to the environment, SCIs are the items that have to be managed to high standards in order to ensure safety. On the other hand, for cases where the threat to safety involves loss of equipment, property, or mission objectives without involving harm to humans or to the environment, this handbook (within the discretion of the decision maker) recommends that items to be managed to high standards be defined in terms of risk drivers rather than SCIs. A risk driver is any significant contributor to safety performance risk. Because the definition of risk drivers is based on the combination of probability and consequence (i.e., risk), rather than just consequence, there are fewer risk drivers than SCIs. Thus, the focus on risk drivers when the measure of safety does not involve harm to humans or to the environment results in fewer items to be managed with associated cost savings. In should be emphasized, however, that even SCIs that are not risk drivers have to be continually monitored to ensure that the basis for their not being risk drivers remains intact.

The ISA is a necessary, though not necessarily exclusive, basis for safety-critical item designation. The ISA-based SCIs may be thought of as the items that have to be made to come true so that the ISA results will, in turn, also come true. Functioning of SCIs at properly allocated levels of capability, reliability, and availability assures that the likelihood of adverse safety consequences is reduced to the required level.

As the program or project evolves, the designation of SCIs may evolve to include elements that are identified in the RISC as being critical to the safety case even though they may not be explicitly considered in the ISA. Some of these additional SCIs might then be incorporated into the next iteration of the ISA if they have an impact on the analysis of safety performance. In this way, the formulation of models for the ISA and the development of the RISC work in tandem.

The adequacy of safety-critical item designation is ultimately at the discretion of the Acquirer, and may include (for example) items associated with safety-related engineering requirements levied for reasons such as defense-in-depth or margin preservation, independent of the ISA. Other aspects of program controls and commitments covered within the system safety framework, but not necessarily as part of an ISA, include configuration management, quality assurance, training and certification of personnel, use of best practices and lessons learned, and assurance that safety requirements are being complied with. Any specific levied requirements, controls, and commitments that are considered critical to safety could be designated as SCIs.

SAFETY CRITICAL ITEMS

Safety critical items (SCIs) are elements or attributes of the system that are important to safety, including hardware, software, interfaces between hardware and software, the human interface, operating procedures, and management practices. This definition is somewhat broader than the definition in MIL-STD-882E, which states that an SCI is "a hardware or software item that has been determined through analysis to potentially contribute to a hazard with Catastrophic or Critical mishap potential, or that may be implemented to mitigate a hazard with Catastrophic or Critical mishap potential." In this handbook, the set of SCIs is defined to include all things that need to be assured if the RISC is to be valid, and in particular, if the Acquirer's safety requirements are to be met with the needed assurance.

SCIs pertains to safety in the context of freedom from harm to humans or to the environment, but it may not always be necessary to apply SCIs in the context of safety that pertains to freedom from loss of equipment, property, or mission objectives. The decision maker may determine that safety in the latter context may be assured by attending to risk drivers, a subset of SCIs that, in addition to being critical to safety, have to have a high enough failure probability or probability of occurrence to be significant contributors to the safety performance risk. SCIs that are not risk drivers may not have to be managed to the same level of certification when the threat to safety does not involve harm to humans or the environment, but still have to be monitored to ensure that the basis for their not being risk drivers remains intact.

The Administrator's letter to the NASA staff dated April 19, 2013, provides a rationale for accepting higher risk tolerance for loss of mission or loss of equipment than for loss of crew or harm to workers. Quoting from part of that letter: "… As long as we ensure that our people are protected we can manage and tolerate failures as part of the price of progress. … As we prepare to undertake the many challenges offered in the President's 2014 budget for our agency, I ask you to continue to think about how we can identify and seize opportunities to make progress quickly and affordably, identify and manage risks, learn fast and adapt our plans to take the next steps. While we do this, we must constantly balance our risks and rewards and always, always put the lives and safety of our people first."

3.2.5 Performance Monitoring Support

System safety supports effective performance monitoring, both in the development of monitoring protocols and in responding to performance data. The ISA is used to risk-inform the selection of system attributes that will be monitored, to ensure both that significant uncertainties are reduced as experience accumulates and that important performance-related assumptions in the ISA remain valid over the system life cycle. Anomalous performance data are scrutinized for their potential impact on safety (e.g., via accident precursor analysis [29]) and managed accordingly.

3.3 Development of the RISC (Argument for Safety)

A RISC is a structured argument, supported by a body of evidence that provides a compelling, comprehensible and valid case that a system is or will be adequately safe for a given application in a given environment. It addresses each of the operational safety objectives of the system, including plans for achieving safety objectives that are applicable to later phases of the system life cycle. RISCs are prepared for the purpose of making the case for safety at KDPs.

The elements of the RISC are [30]:

- An explicit set of safety claims about the system(s), for example, the probability of an accident or a group of accidents is low
- Evidence justifying the claims, for example, representative operating history, redundancy in design, or results of analysis
- Structured safety arguments that link claims to evidence using logically valid rules of inference

The interaction of these elements is illustrated in Figure 3-6 for a safety claim supported by two independent arguments.

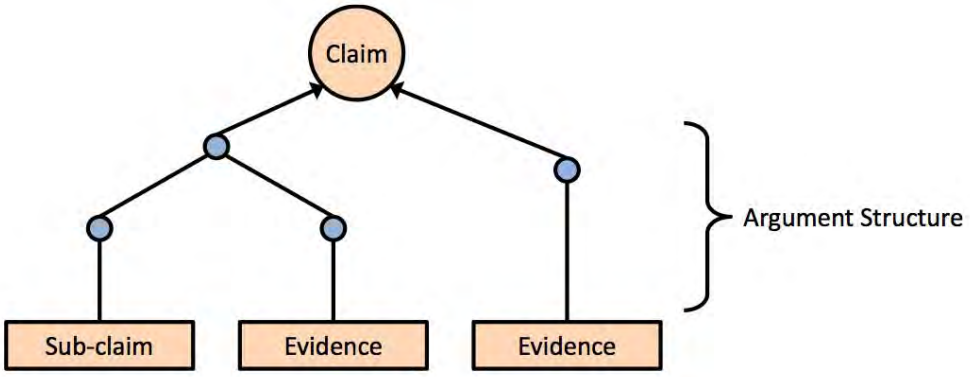


Figure 3-6. A Safety Claim Supported by Two Independent Arguments

3.4 Evaluation of the RISC

Upon submittal of the RISC to the Acquirer, the Acquirer, usually through a designated Evaluation Team (referred to as the Evaluator), conducts an evaluation of the RISC to determine the technical adequacy of its safety claims. RISC evaluation is carried out based on defined evaluation protocols for the system at a particular point in its life cycle. For each claim in the RISC, it is the task of the Acquirer to:

- Understand the evidence behind the claims.
- Evaluate the evidence to determine its validity.
- Provide judgment as to validity of the claims.

In other words, to evaluate the RISC, the claims in the RISC are critically reviewed, thereby making use of the collected evidence related to the safety of the system. The Evaluator ultimately rates the RISC overall as being Acceptable or Unacceptable. In order to provide rationale that the RISC is acceptable, the Evaluator must be able to infer from the evidence in its totality that the truth of the top claim (e.g., that the system is adequately safe) has been demonstrated with high confidence. The ability to make this inference is based on the Evaluator's knowledge of the system as a whole and of the various combinations of requirements that have to be satisfied in order for the system to be demonstrably safe.

In the evaluation process, it is important for the reviewer(s) to evaluate the RISC from a critical viewpoint, examining the supporting evidence as necessary to develop confidence in the claims at all levels. The output of the RISC evaluation is a set of evaluation findings summarizing the review and indicating potential areas of weakness in the RISC.

Safety case evaluation is part of the Acquirer's safety assurance and risk management processes. It is a particular instance of supporting a risk acceptance decision.

RISC development an iterative process that is complete when the Acquirer is satisfied that the technical basis of the RISC is sound. The completed RISC, communicated in the form of a RISC Report, and the RISC evaluation, communicated in the form of a RISC Evaluation Report, become the safety-specific technical bases supporting the decision for which the RISC was developed.

RISC submittal entails a commitment to maintaining its validity, upon which the Acquirer's decision will be predicated. As such, the commitments and understandings captured in the RISC become part of the performance baseline to be managed subsequently under the Continuous Risk Management (CRM) elements of NPR 8000.4A. The risk of a shortfall in safety performance relative to this baseline is managed under the same risk management process within which all other performance risks are managed. If thresholds have been established for safety performance measures, then risk elevation will be required should threshold satisfaction be threatened by emergent conditions.

As mentioned earlier, the RISC is not intended to be the sole basis for the decision-maker's risk acceptance decision. The acceptance of risk is a decision process that is intended to be informed by the RISC and by the evaluation of the RISC (i.e., RISC-informed) but not constrained by the RISC or by its evaluation (i.e., not RISC-based).

3.5 Interactions between the Acquirer and Provider

In one sense, the Acquirer-Provider paradigm defines a division of responsibilities, where the Provider is responsible for developing the safety case and the Acquirer for evaluating it. In another sense, it also defines a collaborative environment wherein the setting of requirements at a program or project level and the means for verifying them depend upon agreements forged between the Provider and the Acquirer. The collaboration is consistent with the fact that the missions for space exploration are becoming more complex and the risks are increasingly cross-cutting.

Toward this end, the Provider of a system is asked to conduct an SSRA early in the system life cycle, which serves as a basis for subsequent communication with the Acquirer about the necessity and sufficiency of the requirement set that is agreed to. This collaboration continues throughout the life cycle in the form of periodic reviews, requests for additional information (RAIs), and compliance with such requests. At each KDP, the objective is to provide the decision-making-authority within the acquiring organization with sufficient confidence to justify a decision to proceed to the next KDP.

In describing the interactions between the two, there is presumed to be a systems engineering (SE) function or its equivalent and a system safety (SS) function or its equivalent for both the Acquirer and the Provider, that together provide an integrated product. Their interactions with respect to system safety are summarized in Tables 3-1 and 3-2 and described below.

As shown in Table 3-1, the interactive process starts with the Acquirer's SE function, in consultation with the Acquirer's SS function, specifying a set of requirements including safety requirements that are deemed to be achievable within the technical, cost, and schedule constraints of the program/project. These requirements and their rationale are documented in the system's requirements documents as specified in the Systems Engineering Handbook [13].

The Provider's SE function, in consultation with the Provider's SS function, develops the framework for the design, realization, and operation of the system and determines whether or not the safety requirements

specified by the Acquirer's SE function can be achieved. The Provider's framework for design, realization, and operation and the justification for claiming that they will meet the Acquirer's requirements are documented in the Provider's System Safety Management Plan. If the Provider believes that it is not possible to satisfy all the Acquirer's requirements, there may be a need for the Acquirer and Provider to negotiate and reach an agreement on how to resolve this problem. The results of the negotiation should be promptly documented in a Memorandum of Understanding, and if the negotiation results in a rebaselining of the requirements, the appropriate system requirements documents should be updated to reflect this rebaselining.

Table 3-1. Interactions between the Acquirer's and Provider's Systems Engineering and System Safety Functions during Formulation of Requirements

	Acquirer		Provider	
Mission Phase	**Systems Engineering Function**	**System Safety Function**	**Systems Engineering Function**	**System Safety Function**
Pre-Design Concept	Specify system safety objectives and requirements	Consult on achievability of system safety requirements		
Prelim. Design			Develop framework for design, realization, and operation	Perform a System Safety Requirements Analysis (SSRA); evaluate whether framework meets Acquirer's requirements
	Negotiate changes in order to meet requirements	Corroborate whether framework meets Acquirer's requirements		
Design, Build, Test, Operate			Develop the system	Evaluate whether system continues to meet Acquirer's requirements
	Negotiate changes in order to meet requirements	Corroborate whether framework meets Acquirer's requirements		

The SSRA facilitates this process. It serves to clarify the detailed requirements (including allocated and derived requirements) that the Provider expects to address in the ensuing development of the system, and evaluates whether the satisfaction of these requirements will provide satisfaction of the top-level requirements. It also serves to facilitate negotiations between the Acquirer and Provider concerning the Acquirer-levied safety requirements.

Upon further development and maturation of the system design concept by the Provider's SE function, the Provider's SS function develops a RISC, as shown in Table 3-2. The RISC demonstrates how the following aspects of system safety have been and/or will be treated during the program/project development and operation:

- Objectives-based safety claims
- Integrated analysis and testing, including implementation of a graded analysis approach
- Modeling and simulation credibility assessment
- Safety risk margin assessment

- Incorporation of relevant best practices and lessons learned from present and previous programs/ projects
- Processes to evaluate departures from the plan
- Design, programmatic, and organizational provisions to reduce UU risks
- Risk-informing of key system safety decisions through a risk-informed decision making (RIDM) process, and continuous management of the implementation of those decisions through a CRM process
- Risk-informing of important system safety support activities including testing, training, quality control, and maintenance
- Processes for tailoring and allocating requirements
- Implementation of the ASARP philosophy
- Implementation of levied requirements

The Provider documents the RISC in a RISC report.

Table 3-2. Interactions between the Acquirer's and Provider's Systems Engineering and System Safety Functions during Preparation and Evaluation of the Risk-Informed Safety Case

	Acquirer		Provider	
KDP	Decision Making Entity	Evaluation Group	Systems Engineering Function	System Safety Function
At Each KDP:	*After requirements have been set and agreed to:*			
KDP n ↓			Incorporate the RISC into the program/ project review package	← Develop/ Update the RISC
	Accept or reject the Provider's input and issue a RAI if needed	← Evaluate the RISC		
			Respond to the RAI and update the SSMP	← Provide additional information requested by Acquirer
	Accept or reject the additional information and approve/ disapprove decision to proceed	← Evaluate the additional information		

At the behest of the Acquirer's SE function, the Acquirer's SS function evaluates the RISC and determines whether the system is adequately safe for its intended mission(s). The results of this evaluation are documented in the RISC Evaluation Report. If the Acquirer's SE and SS functions believe that the system may be adequately safe but the Provider has not sufficiently demonstrated it, then the Acquirer and Provider will negotiate the remaining evidence to be developed by the Provider in order to demonstrate adequate safety. The Provider's SE function develops the additional evidence, and the

Provider's SS function incorporates it into the RISC and updates the documentation in the RISC report accordingly.

This negotiation between the Acquirer and Provider after the evaluation of the RISC may include discussions of possible architectural and design changes, additional controls, relaxation of technical, cost, and/or schedule constraints, and/or relaxation of the safety requirements that are not satisfied. The results of such negotiations should be promptly documented in a Memorandum of Understanding and should later be included in the final documentation of the RISC.

The Acquirer's SE function, in consultation with the Acquirer's SS function, provides their judgment of whether the system has been demonstrated to be adequately safe, and the Acquirer's decision making authority makes the decision whether or not to authorize the Provider to proceed.

These processes and the associated interactions are repeated between each major program/project review and the documentation is updated accordingly.

3.6 System Safety throughout the System Life Cycle

NASA programs and projects are managed to life cycles, the division of the program's and project's pursuits over the full lifetime of the program or project, based on the expected maturity of information and products as they move through defined phases in the life cycle. Figure 3-7 shows a simplified version of a project's life cycle to illustrate the relationship between the phases, the key decision points and application of the system safety framework. (Program and project life-cycle phases are described in NPR 7123.1B.) The vertical thickness of each shape in the figure is intended to notionally indicate the level of effort and/or rigor of each activity. In general, it is expected that adequate safety performance is best assured when system safety activities are conducted beginning early in the system life cycle.

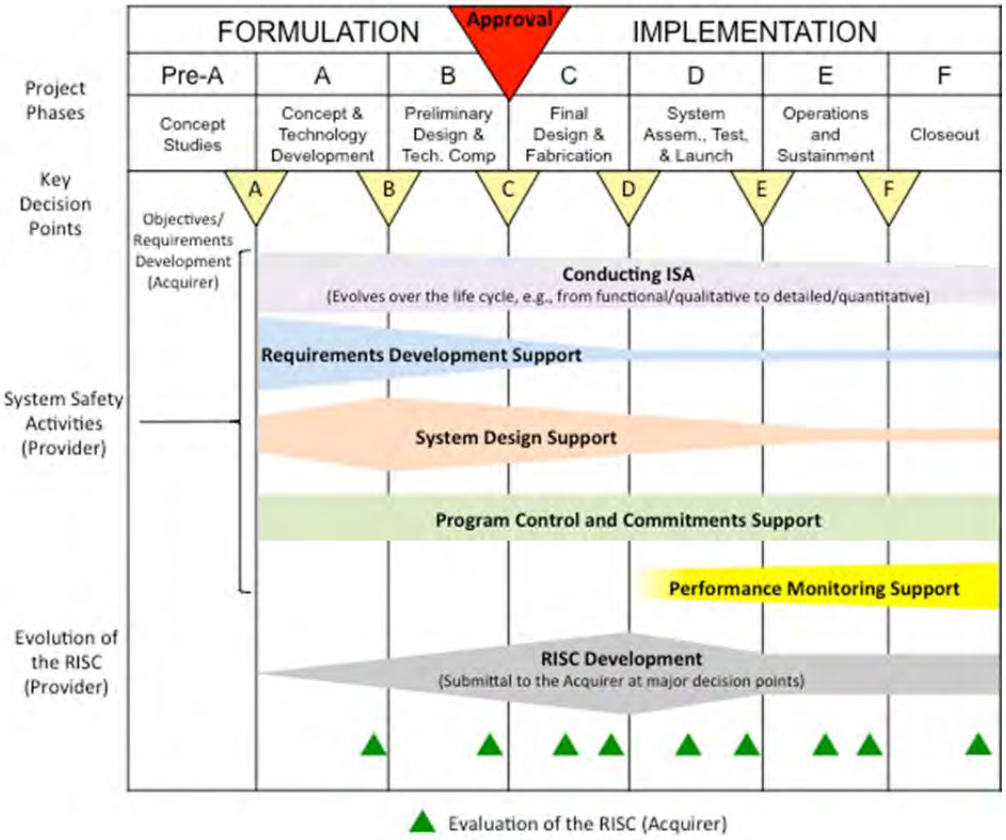


Figure 3-7. Life-Cycle Perspective on the System Safety Framework (Notional)

4. Setting Safety Objectives and Associated Requirements: The Acquirer's Role

This chapter provides guidance and examples that expand upon the overview for specifying safety objectives and associated requirements presented in Section 3.1. The following subjects from that overview are discussed sequentially in Sections 4.1 through 4.4:

- Setting minimum tolerable levels of safety performance at the system level: the total probability of loss
- Developing safety performance requirements at the system level: the probability of loss from known risks
- Levying deterministic engineering and process safety requirements
- Setting verification procedures for each requirement

4.1 Setting Minimum Tolerable Levels of Safety Performance at the System Level: the Total Probability of Loss

4.1.1 The Acquirer's Responsibilities and Areas to Address

The areas to be addressed in setting minimum tolerable levels of safety at the system level, and the depth to which they should be addressed, depend upon the value of the system to NASA. For a system that is of critical value, minimum tolerable levels of safety should be developed with the following considerations:

- They should address the total probability of loss, including loss from known risks and loss from UU risks.
- They should reflect appropriate stakeholder input.
- They should span the set of adverse safety consequence types, including death, injury, occupational illness, damage to or loss of equipment or property, degradation or loss of mission objectives, or damage to the environment.
- They should be informed by the safety performance of applicable baselines (e.g., Shuttle).
- They should be informed by an assessment of what is feasible (e.g., via ISA).
- They should be balanced against the potential rewards of system use.
- They should include limits imposed by applicable requirements and standards.
- For crewed systems, they include limits on the probability of loss of crew (P(LOC)) [12].
- They should be decomposed into safety thresholds that specify a minimum tolerable initial safety performance, safety goals that specify a minimum tolerable level of long-term safety performance, and timeframes or safety growth profiles that define *long-term* for each safety goal (e.g., for systems that experience design maturation over the long term).
- For systems without a defined mission, they should be defined on a suitable alternate basis (e.g., the safety associated with executing a specific system function or capability).

[12] NASA has established minimum tolerable levels of safety for crew transportation missions to the International Space Station (ISS) [7]. Minimum tolerable levels of safety for crew transportation to other destinations may be developed as matters of policy as mission concepts develop.

For systems of lower (noncritical) value, the Decision Maker in combination with the Technical Authority should select which of these areas of consideration apply.

This next subsections address the following topics pertaining to the setting of minimum tolerable levels of safety performance:

1. Establishing sub-cases for different key mission objectives (KMOs)
2. Developing system safety risk margins for each KMO
3. Developing thresholds and goals for the total loss probability for each KMO
4. Justification of the use of ratios

4.1.2 Establishing Sub-Cases for Different Key Mission Objectives (KMOs)

In general terms, the missions that NASA is concerned with include single-objective missions and multiple-objective missions. A single-objective mission may involve a single launch with a single purpose (e.g., the launch of the Hubble satellite) or could include multiple launches with a single purpose (e.g., multiple launches of the same system to deliver supplies to the ISS). The distinguishing feature of a single-objective mission from the point of view of this handbook is that a single safety case suffices.

On the other hand, there are missions with multiple objectives, where the decision maker needs to be presented with safety sub-cases for different mission objectives in order to make decisions affecting the program/project as a whole. When there is a need for the safety case to be divided into sub-cases, it is usually not necessary for the various sub-cases to stand alone. The same set of arguments may cover more than one sub-case as long as the differences between the different sub-cases are recognized and accounted for. For example, separate safety sub-cases may appear in the form of separate claims within the same RISC or in the form of separate evidence within the same claim. The key point is that the differences in the safety arguments and evidence required for different mission objectives are highlighted and explained.

Two examples of multiple objective missions are noted below:

- For a planetary exploration mission that lasts for years and has many rendezvous events, a safety sub-case would be needed for each major rendezvous event. For example, in the case of the Galileo mission, a safety case might be provided for the Jupiter probe shock wave crossing, for rendezvous with Jupiter's moons (Ganymede, Callisto, Europa, and Io), and for final impact with Jupiter. Failure to rendezvous with Io might not be considered as critical as failure to rendezvous with Europa, because the Europa rendezvous occurs prior to the Io rendezvous and is considered more important from a scientific perspective. Thus, there might be a decision to proceed even if the safety sub-case for Io rendezvous were weak, as long as the sub-case for Europa rendezvous was strong.
- For a mission consisting of many flights with each flight having one destination, a safety sub-case should be provided for any flight that marks a significant change from previous flights. For example, in the case of the Space Shuttle, separate safety sub-cases would be expected for the Hubble repair mission as opposed to ISS missions. On the basis of these safety sub-cases, the decision maker might opt to proceed with one but not the other.

For purposes of this handbook, the objectives for which separate safety sub-cases are required are referred to as key mission objectives (KMOs) and are specified by the Acquiring organization. They are not the same as KDPs as defined in the NASA Systems Engineering Handbook [13], as most of the KDPs occur prior to initial operation, whereas most KMOs refer to different operational states.

Consistent with the notion that there may be a need for separate safety sub-cases for different KMOs, the Acquirer may also wish to specify different requirements for different KMOs. Generally speaking, the

requirements will be presented in terms of a set of requirements that are applicable to all KMOs augmented by sets of requirements that are applicable only to specific KMOs.

The safety cases for different KMOs are initially prepared during preliminary design and then updated at key reviews during the design, realization, and operation of the system. For example, updates might be required at Preliminary Design Review (PDR), Critical Design Review (CDR), System Integration Review (SIR), Critical Events Readiness Reviews (CERRs), and Decommissioning Review (DR).

4.1.3 Developing Safety Performance Risk Margins

Risk model completeness has long been recognized as a challenge for synthetic methods of risk analysis such as PRA as traditionally practiced [8]. These methods are generally effective at identifying system failures that result from combinations of component failures that propagate through the system due to the functional dependencies of the system that are represented in the risk model. However, they are understandably deficient at identifying system failures that result from unexpected or underappreciated scenarios, frequently involving complex intra-system interactions that may have little to do with the intentionally engineered functional relationships of the system. Such underappreciated interactions (along with other factors) were operative in both the Challenger and Columbia disasters [31, 32].

Because of the nature of UU risks, involving interactions between subsystems, they tend not to be revealed by subsystem testing [33]. Full-up testing has the potential to reveal them, but the cost of full-up testing in as-flown environments is generally too high to allow a quantity of tests that would demonstrate low probabilities of occurrence.

It is possible, however, to gain major insights into the historical importance of UU risks, as compared to known and fully analyzed risks, by examining programs for which there is a history of catastrophic accidents and/or near misses[13].

NASA's probabilistic safety thresholds do not explicitly address the question of how to account for UU risks. Yet, the accepted expectation is that the demonstration of safety threshold satisfaction is to be based on the actual risk, which includes both known, adequately modeled sources, as well as UU risks.

Volume 1 introduces the concept of safety risk reserve as an approach for compensating for UU risks, and states that there are methods for estimating how large the reserve should be. In this section of Volume 2, we are concerned with how such reserves should be estimated. Also in this section, we use the term "margin on the loss probability" in place of the term "safety risk reserve" to emphasize the fact that the thresholds and goals are expressed in terms of maximum allowable loss probabilities, which are concrete measure of safety performance. Aside from this semantic distinction, there is no real difference between the meanings of the two terms.

Safety performance margins were defined in Section 3.1.1 and their relationship to safety thresholds and safety requirements were depicted in Figure 3.4. A safety performance risk margin is a resource (programmatic) margin similar to mass margins and management budgetary reserves, not a desired design (physical) margin on the end product (such as a margin on load carrying capacity of a structural element). Likewise, applying a margin on loss probability to account for the risk from UU scenarios is analogous to the systems engineering approach of using margins to protect against exceeding established limits in the technical, cost, and schedule mission execution domains [13, 34-38]. In all cases, the main purpose of the margin is to accommodate unknown or underappreciated conditions or events that tend to arise during design, development, fielding, and operation. This principle applies to any metric that has a numerical constraint: i.e., technical metrics (launch mass, payload mass, power, thrust, throughput, etc.), safety metrics (probability of LOC, LOV, LOM, or equipment loss), cost, and schedule.

[13] Sections 4.5.1 and 4.5.2 will provide examples of how such assessments might be conducted.

Comparison of System Safety Risk Margins with Mass Margins and with Cost and Schedule Reserves

Mass margins are adopted for use within NASA Systems Engineering during the design phase of a project in order to insure that the total vehicle mass stays within its allowable limit as the design progresses. The policy for specifying and adhering to mass margins is expressed in [35], wherein the following terms are used:

- Mass Limit: The maximum mass that is physically consistent with the system's operating constraints (e.g., the maximum mass that the launch vehicle can accommodate)
- Mass Margin: A contingency or reserve intended to mitigate potential mass increases during the design process emanating from omissions or refinement of existing design requirements
- Mass Growth Allowance (MGA): An allowance for the expected mass growth during the design process resulting from lack of maturity in the current design data
- Customer Reserve: A mass contingency reserved by NASA
- Basic Mass: The current calculated mass based on an assessment of the most recent baseline design. The designers are required to ensure that the basic mass of the design is enough below the mass limit to accommodate the margin, the MGA, and the customer reserve.

According to [35], the recommended amount of mass margin varies depending upon the complexity of the design and the degree to which it emanates from a heritage design (i.e., whether the space system is a new system, a modified system, or a production system). It also varies according to the risk that the procuring authority and the contractor are willing to accept. Tables of recommended values for MGA and margins based on historical mass growth due to both anticipated and unexpected sources are provided in [33]. The recommended values are intended to cover the upper limit of the uncertainty in the predicted mass. Thus, the margin is expected to provide high confidence that the total mass of the design prior to the first launch will not be greater than the mass limit, although there is no specific numerical confidence level cited in the reference.

Cost and schedule margins, or reserves, are governed by the Joint Confidence Level (JCL) rule. The NASA Cost Estimating Handbook [39] states that at Confirmation Reviews and Authority to Proceed decision points, the cost estimate must include an appropriately chosen level of unallocated future expense (UFE)/reserves. The term "reserves," according to [39], includes "funding, performance, manpower, and services allocated to and managed by the program/project manager for the resolution of problems normally encountered to mitigate risks while ensuring compliance to the specified program/project scope."

There are six activities associated with developing UFE/reserves:

- Determine the project's cost drivers with input from the project manager and staff.
- Develop probability distributions for the cost model uncertainty.
- Develop probability distributions for the technical and schedule cost drivers.
- Run the risk model.
- Identify the probability that the actual cost is less than or equal to the point estimate.
- Recommend sufficient UFE/reserves to achieve a 70% confidence level for both cost and schedule.

Unlike the mass margin or the reserves for cost and schedule, the magnitude of the margin for loss probability can be interpreted as a best estimate of the effect of UU scenarios on the actual, or total, risk. The use of a best estimate rather than a high-confidence estimate is consistent with the fact that the inclusion of UU risks together with known risks in defining the threshold introduces an element of high uncertainty and corresponding arbitrariness that would make a high-confidence estimate difficult to justify. Furthermore, whereas the possibility of over-mass, over-budget, or behind-schedule conditions can potentially be showstoppers for a mission, the implications of over-threshold for P(LOC), P(LOV), or P(LOM) are not so clear-cut.

4.1.4 Developing Thresholds and Goals for the Total Loss Probability

As mentioned in Section 4.1.3, the safety threshold (i.e., threshold value for the loss probability) should reflect an achievable expectation for the actual probability including both known risks and UU risks. The known risks are generally quantified using PRA methods, which may be more or less detailed depending on the criticality of the mission[14]. Realistic expectations for the UU risks are inferred from historically verifiable margins. Separate considerations apply for low earth orbit missions compared to other missions.

Safety Thresholds for Low Earth Orbit (LEO) Missions

In the absence of a launch abort system (LAS), a reasonable threshold value for LEO missions may be inferred from a combination of the Space Shuttle known risk at the time of the first launch and the appropriate multiplier to account for UU risks[15]. An argument can be made that the same threshold would also be appropriate for loss of vehicle for robotic missions involving LEO, since vehicle failure during LEO might result in threats to public safety on Earth.

Because all future crewed missions will have to have LAS capability by virtue of the human-rating requirements for space systems [40], the threshold for P(LOC) for future crewed missions will be more stringent than the value that is achievable without LAS capability. An estimate for the likelihood that a LAS would save the crew in the event of an accident that results in LOV can be inferred from existing analyses available in the public literature [41-46][16].

Safety Thresholds for Other Types of Missions

Known risks for new systems intended for missions other than LEO can be estimated by extending risk models for existing systems intended for LEO to include additional design elements and mission profiles that characterize the new system in the new mission. As a reasonable approximation, it could be assumed that the safety performance factors[17] to account for UU risks are the same as those derived for LEO systems.

Safety Goals

The safety goal is the goal value for the loss probability and should reflect an achievable expectation for the total loss probability after the system has matured. The goal is relevant for missions that involve many individual flights and presumes that after a sufficient number of them, the loss probability will have decreased to a steady-state value. By this time, significant contributions from UU risks will have been uncovered and corrected for (i.e., infant mortality effects will have been wrung out). In general, based on

[14] The level of detail based on mission criticality will be discussed further in Section 5.2.4.
[15] An example will be provided in Section 4.5.5.
[16] Use of these analyses to derive a realistic threshold for P(LOC) including LAS capability will be explored further by way of example in Section 4.5.5.
[17] The safety performance factor is the ratio of the loss probability from all risks to the loss probability from known risks

experience in the Space Shuttle program and for launch vehicles, steady state can be assumed to have been achieved after 100 to 150 flights[18].

Wringing out the significant UU risks is a safety growth process. Safety growth is analogous to reliability growth, but the former focuses on reducing system level failures, whereas the latter has tended to focus on reducing subsystem and component level failures. Historical evidence for Shuttle and for some of the earlier launch vehicles (Soyuz/Molniya and Delta) indicates that during the period of safety growth, the loss probability from known risks decreases at the same time as the loss probability from UU risks, though not as rapidly. A factor-of-2 reduction in loss probability from known risks during this period is typical[19]. Thus, it is reasonable for the safety goal to be more-or-less a factor of 2 less than the initial loss probability from known risks. This reduction in the known risks over 100 to 150 flights is attributable to improvements in hardware, software, and processes that occur normally during the operational phase as part of ASARP implementation.

4.1.5 The Use of Ratios

In estimating the margin to account for UU risks, it is helpful to think that the ratio of the loss probability from all risks (known plus UU) to the loss probability from just known risks correlates, at least qualitatively, with various design, organizational, and programmatic factors. Although it is not necessary to assume that the correlation applies only when using a ratio, it simplifies the process. The example in Section 4.5 and its subsections (particularly Section 4.5.4) will explore the degree to which the ratio appears to correlate with knowable design, organizational, and programmatic factors and will provide a rationale for why this may be a reasonable supposition.

4.2 Developing Safety Performance Requirements at the System Level: the Probability of Loss from Known Risks

4.2.1 The Acquirer's Responsibilities and Areas to Address

For a system that is of critical value, safety performance requirements should be developed in concert with the following considerations:

- They should address the probability of loss from known risks.
- They should consider each consequence type, including death, injury, occupational illness, damage to or loss of equipment or property, degradation or loss of mission objectives, or damage to the environment.
- They should reflect the application of safety performance margins relative to minimum tolerable levels of safety, to account for unknown and underappreciated sources of safety performance risk.
- For crewed systems, they should include a limit on the probability of loss of crew (i.e., P(LOC)).
- They should be consistent with any safety growth or degradation profiles that have been specified.
- They should be reevaluated and possibly rebaselined if changes to the system or its management invalidate the bases upon which the safety performance margins were developed.
- They should apply to all milestone reviews for which a RISC is required.

[18] This inference will be demonstrated in Section 4.5.2.
[19] This inference will also be demonstrated in Section 4.5.2.

- They should be accompanied by specification of safety verification methods (e.g., analysis, demonstration, inspection, or test) according to which the Provider may argue compliance, and according to which the Acquirer may deem the requirement to have been satisfied.

For systems of lower (noncritical) value, the Decision Maker in combination with the Technical Authority should select which of these areas of consideration apply.

4.2.2 Developing Performance Requirements for the Probability of Loss from Known Risks

Continuing the thread from Section 4.1 and its subsections, the requirement for the probability of loss from known risks at the time of the first flight is related to the safety threshold and the safety performance factor as follows: Initial Requirement = Safety Threshold Divided by Safety Performance Factor. For example, if the safety threshold for total loss probability is 0.01 and the safety performance factor is 5 (reflecting a belief that the loss probability from unknown and underappreciated risks is four times the loss probability from known risks), then the initial requirement for the known loss probability is 0.01 / 5 = 0.002.

If the mission consists of many flights so that all significant UU risks are wrung out by the time of the last flight, then at the time of the last flight the value of the requirement should be reduced to the value of the safety goal. If, as is typical, the expected loss probability for the matured system is assumed to be half the initial known loss probability, then the mature-system requirement for the known loss probability would be 0.002 / 2 = 0.001.

Between the first and last flights, the requirement decreases at a rate that is consistent with the burndown of known risks that has been observed for other space missions. It typically takes about 125 flights for the system to reach full maturity and the burndown relationship is exponential[20].

4.3 Levying Deterministic Engineering and Process Safety Requirements

4.3.1 The Acquirer's Responsibilities and Areas to Address

The Acquirer's responsibility in levying deterministic and process safety requirements can be summarized as follows:

- Based on the System Safety Requirements Analysis (SSRA) that the Acquirer develops as per Section 3.1 and on current best practices, identify and baseline safety-related engineering requirements and process requirements that comport with fulfillment of the needed level of safety performance and reflect application of the ASARP principle.

The following two topics pertaining to this responsibility are discussed in the next subsections:

1. Levying deterministic safety requirements
2. Tailoring deterministic safety requirements

4.3.2 Levying Deterministic Safety Requirements

Deterministic safety requirements (which include both engineering and process requirements) are based on the Acquirer's understanding of best practices and lessons learned from relevant experience. There are many sources of information pertaining to best practices and lessons learned from NASA's experiences with spaceflight.

[20] These observations will be demonstrated in Section 4.5 (Figures 4-4 through 4-6, and 4-8).

Of these, the following provide a suitable starting point:

- Design best practices: GSFC-STD-1000F (The Goddard Open Learning Design (GOLD) Rules), "Rules for the Design, Development, Verification, and Operation of Flight Systems," Feb 2013 [47].
- Design and test best practices: "NASA Preferred Practices for Design and Test of Robust Systems," Jet Propulsion Laboratory, http://oce.jpl.nasa.gov/preferred_practices.html#test [48].
- Program and project management best practices, "NASA Space Flight Program and Project Management Handbook," May 2013 [49].
- Reliability and maintainability best practices, NASA Technical Memorandum 4322, "NASA Reliability Preferred Practices for Design and Test," and NASA Technical Memorandum 4628, "Recommended Techniques for Effective Maintainability," see website http://www.hq.nasa.gov/office/codeq/rm/prefprac.htm [50, 51].
- NASA Lessons Learned System, NASA Engineering Network, see website http://llis.nasa.gov [52].

As described in the NASA Systems Engineering Handbook [13], deterministic engineering requirements may be of several types or categories. Some of these are summarized below:

- Functional requirements define what the system (element/subsystem/component) must do. For example, a requirement that the system must provide crew abort capability from ground to low earth orbit (LEO) is a functional requirement.
- Constraint requirements limit or restrict the function or performance of the system. For example, a requirement that the acceleration on astronauts during an abort not exceed 7g's is a constraint requirement.
- Verification requirements establish how functional, performance, and constraint requirements will be verified. For example, a requirement for performance and constraint requirements on the Launch Abort System to be demonstrated through modeling and simulation is a verification requirement.
- Interface requirements concern constraints involving interconnections between components or subsystems, interfaces with support and test equipment, and interfaces between the system and the external world. For example, a requirement concerning the signals that a Launch Abort System receives from a Flight Control Computer is an interface requirement.

4.3.3 Tailoring Deterministic Safety Requirements

A tailoring process (to be discussed here and in Section 5.3.2) is used to winnow down the sum total of best practices and lessons learned to those that are specifically relevant to the mission and/or system being investigated. Because the list of safety requirements levied on a program or project can be very long (e.g., see Constellation Architectural Requirements Document [53]), there can be considerable savings in time and cost if some of the requirements levied by the Acquirer can be excluded based on their not being relevant or not being practicable for the system being developed. The Acquirer should therefore attempt to avoid levying requirements that are not relevant and/or not capable of being applied to the program or project or to a specific system or component. The Acquirer may also opt not to levy requirements that would normally be supported by best practices/lessons learned for either or both of the following reasons:

- They do not provide an evident or discernible net increase in safety for the present mission.

- They lead to a disproportionate or impracticable penalty in cost, schedule, and/or technical performance.

When a relevant best practice or lesson learned is not translated into a levied requirement, the Acquirer should state the associated rationale for not doing so. The process for reporting the tailoring of requirements should follow the procedures in NPR 7120.5E [34].

Section 5.3.2 will provide additional information and examples relating to the conditions under which levied requirements corresponding to best practices and/or lessons learned may be opted out based on ASARP considerations.

4.4 Setting Verification Procedures for the Safety Requirements

4.4.1 The Acquirer's Responsibilities and Areas to Address

The Acquirer's responsibility in setting verification procedures for the safety requirements can be summarized as follows:

- For every safety requirement levied, establish a safety verification method (e.g., analysis, demonstration, inspection, or test), negotiated with the Provider as part of the SSRA process.

The following two topics pertaining to this responsibility are discussed in the next subsections:

1. Setting initial verification procedures for the safety requirements
2. Negotiating with the Provider to rebaseline requirements and reset verification procedures

4.4.2 Setting Initial Verification Procedures

Because safety is an emergent property of a system, which involves discovery on the part of both the Provider and the Acquirer as the system is designed, built, and operated, there is no catch-all set of verification procedures that can be prescribed during the initial requirement development phase that will assure ultimate achievement of the top operational safety objective: namely, that the system is adequately safe throughout all phases of the program/project. For example, if during the substantiation of the safety case, the results of testing and analysis are different from what is expected, future testing and analysis plans may need to be modified or reformulated. With this understanding, however, it is possible and desirable to set preliminary verification procedures with the expectation that as the program/project progresses, these verification procedures will evolve. Just as the requirements may need to be rebaselined as new information surfaces, the verification procedures may need to be rebaselined.

The Acquirer initially specifies a set of tests, demonstrations, analyses, reporting procedures, or other approaches that the Acquirer posits as being sufficient if there are no surprises. The general process is for specific safety tests and analyses to be integrated into appropriate system Test and Evaluation (T&E) plans, including verification and validation plans. Where system-level integrated safety tests are not feasible, the Acquirer may specify that verification of compliance will be demonstrated using engineering analyses, analogies, laboratory tests, functional mockups, or models and simulations. The Acquirer will also specify what is expected in the way of review plans and the documentation of safety verification results.

As part of the verification of requirements, per NPR 7123.1B [3], the Acquirer will specify a set of measures of performance (MOPs) and technical performance measures (TPMs) by which the safety of the overall system will be judged. Typical MOPs for probabilistic requirements might include the computed probability of loss of the system and the mean failure rates of major subsystems or components for specified conditions. For deterministic requirements they might include proof that design specifications

have been met (e.g., that a certain subsystem is two-failure tolerant) or values of induced environmental parameters (e.g., accelerations, temperature, pressure, radiation).

TPMs are used for progress measurement and generally meet the following criteria: (1) be a significant qualifier of the system to be monitored at critical events (e.g., inspections, planned tests); (2) be measurable both in terms of present values and in terms of projected progress profiles; and (3) include (as subsets) both leading indicators and margins.

4.4.3 Negotiating with the Provider to Rebaseline Requirements and Reset Verification Procedures

As mentioned earlier, there may be a need for the Acquirer and Provider to renegotiate on safety requirements and on verification procedures. The need centers around the fact that safety is an emergent property of a system that involves a discovery process as the system is developed and later operated.

Guidelines for rebaselining performance requirements are provided in Section 2.3.2 of the RM Handbook [28]. The decision to request a rebaselining occurs within the Provider's CRM activity, where disparities between requirements and the ability to satisfy them are first perceived.[21] The concern is elevated within the Provider's organization to the appropriate management authority. If the adjustments to requirements are straightforward and easily resolved, they may be agreed to by the appropriate level of management within the Acquirer's organization and the agreement may be recorded in the SSMP and in the appropriate program's/project's requirements document. If the amount of change in the requirements is sufficiently large, however, the RIDM process should be invoked by the Provider to produce a proposed set of rebaselined requirements to present to the Acquirer. The Acquirer will then initiate a formal decision making process in accord with NASA procedural requirements.

A decision to rebaseline certain safety requirements carries with it a need to reset the verification procedures that accompany those requirements. In addition, a need to reset verification procedures may occur even when there has been no change in the requirements, if it is discovered that the current verification procedures will not provide the desired confidence that the system is adequately safe. The process for resetting the verification procedures starts out similarly to the process for rebaselining requirements. The decision to request new or modified verification procedures initiates in the Provider's CRM activity where the need is first recognized, and the concern is elevated within the Provider's organization to the appropriate management authority. The negotiated agreement between the Provider and the Acquirer may involve little more than a handshake with an accompanying documented entry in the SSMP. Changes in verification procedures would not require a formal decision making process.

[21] The term 'Provider' here refers to the organization responsible for providing the system. The system provider can also be an acquirer with respect to contracted subsystem providers and may determine that a system level requirement needs to be rebaselined if a contractor's allocated requirement cannot be met.

4.5 Example for Chapter 4 – Deriving Safety Risk Margins, Safety Thresholds and Goals, and Probabilistic Safety Requirements

In this example, the Acquirer uses knowledge gained from the historical record of successes and failures for the Space Shuttle, Atlas, Delta, Soyuz, and Molniya launch systems, together with results from the comprehensive Shuttle probabilistic risk assessment and general observations about organizational and programmatic failure causes from the literature to assist in the development of realistic safety performance risk margins, safety thresholds and goals, and probabilistic safety requirements.

Although this example uses real data to infer plausible safety performance margins, **we wish to emphasize** that it is intended to illustrate a method and manner of thinking rather than argue for particular values for the margins. Alternative views on how to select and interpret the data may be equally valid. Different ways of looking at the data contribute to our perspective on the uncertainty, and these different perspectives are a fundamental consideration to be addressed in the development and evaluation of the RISC.

A summary version of the analyses presented below may also be found in [54].

4.5.1 Space Shuttle Experience

A recent JSC study by Hamlin, et al., [55] provides a basis for comparing the actual risk of LOC for the Space Shuttle prior to each flight with the risk of LOC that would have been calculated using known risks only. The calculations utilize the most recent Space Shuttle full-scope PRA model [56] in a retrospective, or backward-looking, mode. The risk model, since it was created after the Columbia accident, includes the knowledge gained from the Challenger and Columbia accidents, as well as from all the other flights that occurred during the Shuttle lifetime. Accordingly, the JSC authors were able to use the risk model to estimate, in hindsight, what the total risk of LOC was at the time of each launch up to the very last one. The result is shown in Figure 4-1.

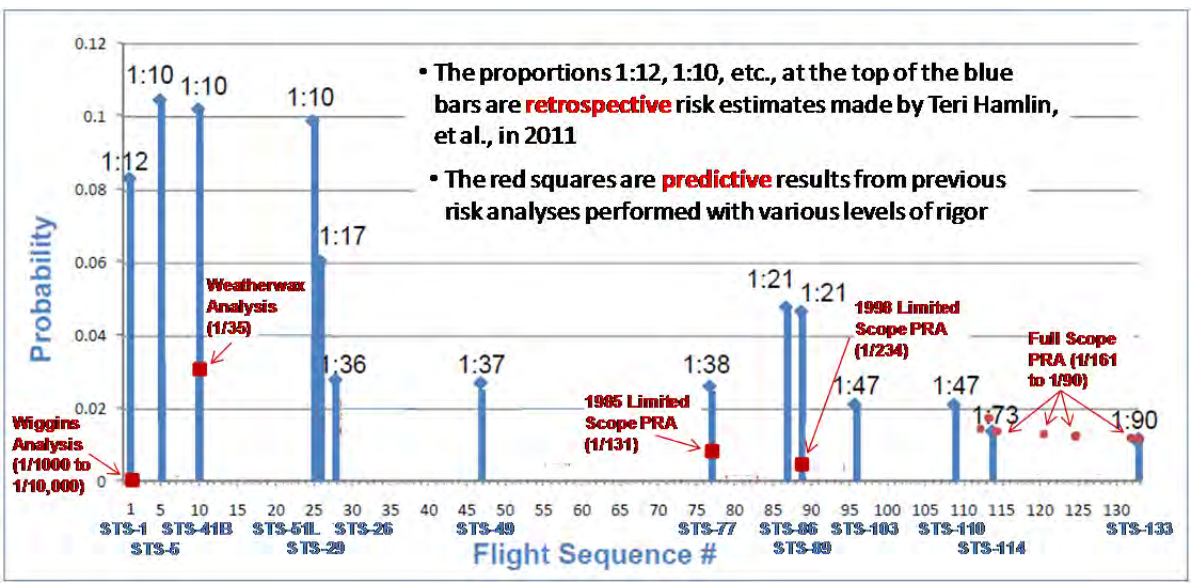


Figure 4-1. Results of a Retrospective Analysis of P(LOC) for the Space Shuttle Compared to Earlier PRA Predictions, from Hamlin, et al

Also shown in Figure 4-1 are results for P(LOC) obtained from various risk assessments exercised in a predictive mode. These include the following results:

- In 1982, J. H. Wiggins Co. estimated P(LOC) for the Space Shuttle to be between 1/1000 and 1/10000 based on engineering judgment [57].
- In 1983, R. K. Weatherwax of SERA Inc. applied more of a database analysis to the Wiggins approach to estimate P(LOC) at ~1/35 [58].
- The first in house limited-scope PRA for the Shuttle in 1995 included ascent and entry/ landing and covered 3 Orbiter systems and the propulsion elements. It resulted in P(LOC) = 1/131.
- An unpublished analysis in 1998 using QRAS was similar to the 1995 analysis but had no integration of elements. It resulted in P(LOC) = 1/234.
- The full-scope PRA models developed and applied post-Columbia between 2003 and 2010 have resulted in P(LOC) values between 1/61 and 1/90.

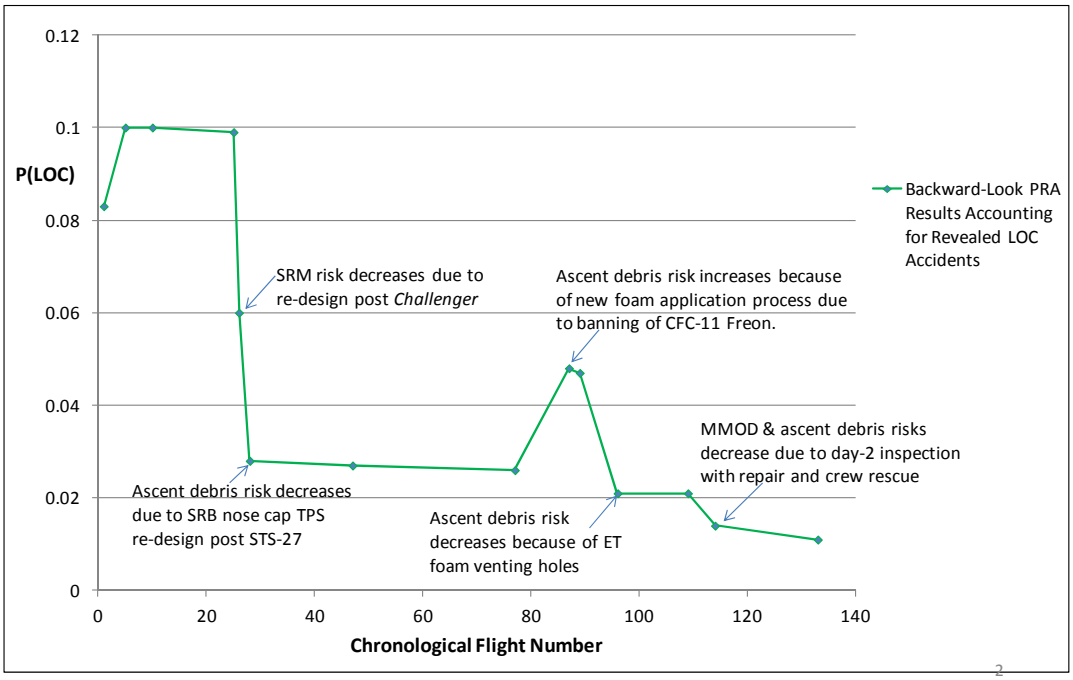


Figure 4-2. Correlation of Shuttle Risks from Retrospective Analysis with Changes in Design, Fabrication, and Operation

The jagged nature of the retrospectively estimated total risks in Figure 4-1 is caused by responses to unexpected events that resulted in changes to the design, fabrication, or operation of the system. These are shown on Figure 4-2. The first major change was the re-design after the Challenger accident, which resulted in a reduction of the total risk of LOC by about 40%. Following soon after was a modification to the SRB nose cap to provide protection against debris impacts, resulting in a further reduction in the total risk by about 53%. Little change occurred thereafter until STS-86, when NASA's compliance with an OSHA directive to discontinue the use of Freon in applying foam to the external tank unexpectedly caused a significant increase in the number of debris strikes on the Orbiter and raised the total risk of LOC by about 80%. The inclusion of vent holes in the foam to alleviate this problem led to a risk decrease of about a factor of 2.2. Return-to-flight changes after the Columbia accident during STS-114 resulted in a further risk decrease of about 35%.

The model used in the JSC analysis provided probabilities for all modeled accident scenarios that could lead to LOC. A list of the top accident scenarios and their probabilities prior to the first flight, STS-1, is reproduced in Table 4-1. Original values were calculated by Hamlin, et al., using the full-scale Shuttle PRA model modified to account for the design features at the time. Also shown in red are edited values based on assuming the Challenger and Columbia accidents had not occurred. The difference between the original and edited values is the effect of underappreciated risks based on the knowledge available at the time of STS-1.

Using the process illustrated in Table 4-1 for all of the Shuttle flights, it is possible to construct a second curve on top of the one in Figure 4-2 that represents the historical variation of known risks for the Shuttle. The result is shown in Figure 4-3. As a point of reference, the actual risk before the 25th flight (STS-51L) was about a factor of 5 times the risk that would have been predicted if a detailed PRA had been conducted based on information known before the Challenger accident. Similarly, the actual risk before the 87th flight (STS-86) was about a factor of 3 times the risk that would have been predicted.

Table 4-1. Modification of Assessed Probabilities of the Top Accident Scenarios Leading to LOC at the Time of the First Shuttle Flight Assuming the Challenger and Columbia Accidents Had Not Occurred

Rank	% of Total	Cumulative Total	Probability (1:n)	Description
1	53.5	53.5	**1.1E-03 (1:940)** ~~4.5E-02 (1:22)~~	Ascent debris strikes Orbiter TPS leading to LOCV on orbit or entry
2	19.2	72.8	**6.5E-04 (1:1500)** ~~1.6E-02 (1:63)~~	SRM-induced SRM catastrophic failure and ejection seats fail to save the crew
3	6.4	79.2	5.3E-03 (1:190)	MMOD strikes Orbiter on orbit leading to LOCV on orbit or entry
4	5.0	84.2	4.2E-03 (1:240)	SSME-induced SSME catastrophic failure and ejection seats fail to save the crew
5	3.7	87.9	3.1E-03 (1:320)	Orbiter APU Shaft Seal Fracture Entry and ejection seats fail to save the crew
6	2.9	90.8	2.4E-03 (1:420)	APU external leak on entry and ejection seats fail to save the crew
7	2.0	92.8	1.7E-03 (1:600)	Orbiter flight software error results in catastrophic failure during ascent and ejection seats fail to save the crew
8	1.1	93.9	9.0E-04 (1:1100)	APU external leak on ascent and ejection seats fail to save the crew
9	1.1	95.0	8.8E-04 (1:1100)	Orbiter APU Shaft Seal Fracture Ascent and ejection seats fail to save the crew
10	0.8	95.7	6.3E-04 (1:1600)	SSME-induced safe shutdown of the SSME and ejection seats fail to save the crew
Total		**100.0**	**2.4E-02 (1:42)** **8.3E-02 (1:12)**	

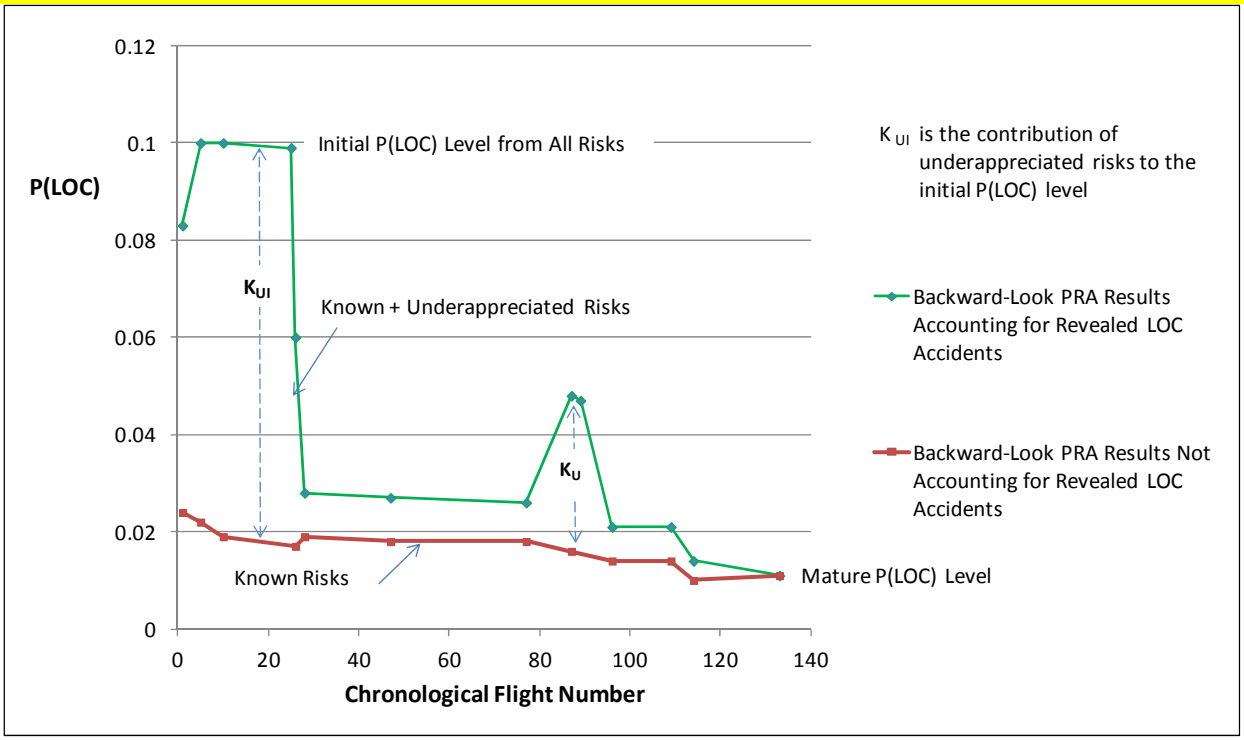


Figure 4-3. Comparison of Retrospective Analyses of Shuttle Risks Accounting for Versus Not Accounting for Revealed LOC Accidents

Unknowns had been fairly well wrung out prior to STS-1 because the technology used in the Space Shuttle program was fairly mature. First, the Shuttle launch vehicle was based to a large extent on the design of the Titan launch vehicle and its successors, comprising a large liquid fuel booster with attached solid rocket motors. Features on the Shuttle that were essentially new (e.g., a payload that included the use of heat protective ceramic tiles) were extensively tested prior to the first crewed flight through both ground testing and unmanned flight testing.

4.5.2 Launch Vehicle Experience

For programs where there are large numbers of catastrophic accidents, it is possible to compare observed system failure rates early in the program with observed failure rates much later in the program to infer estimates of the magnitude of UU risks initially and how they burn down with time. By examining how these estimates vary across different launch systems, it is possible to draw insights about the attributes of the system and the management of the program/project that contribute to higher ratios of UU risks to known risks.

There has been a long history of launch vehicle successes and failures since the 1950s. Between 1957 and 1999, for example, there were 390 launch vehicle failures out of 4378 attempts throughout the world [59]. With such a large sampling of successes and failures, it is possible to perform meaningful statistical analyses of how the system failure rate has varied with time for a number of launch systems.

At least four analyses of launch vehicle failure data across various systems have been performed previously. Two used a Bayesian methodology [60, 61] and the other two a frequentist methodology [62, 63]. In the example that follows, we utilize a simple frequentist approach to illustrate, for each of three launch systems, the observed failure rate near the beginning of the operational phase, the observed failure

rate after the UU risks were wrung down, and the number of launches that it took to reach the latter.* The purpose of using a simple approach in this example when other more rigorous statistical methods are available is to increase the tutorial benefit; i.e., to illustrate the thought process with a minimum amount of mathematical complexity.

It should be recognized that because there are different ways of looking at the data (e.g., Bayesian versus frequentist), significant differences in results might be obtained depending on the method used. Rather than viewing these differences as representing a flaw in one analysis approach versus another, they should be viewed as contributing to our perspective on the uncertainty associated with the analysis of the data. Thus, such differences provide a contribution to the evaluation of the confidence of the RISC

Soyuz and Molniya

The data for Soyuz and Molniya are shown in Figure 4-4. They are grouped together both here and in [59] because they are of the same family and are very similar in design. Molniya was the predecessor for Soyuz. Between the two systems, as of 1998, there were 1458 launches with 69 failures. The Soyuz/Molniya system is an example of a launch system that was developed under very high time constraints during the early phase of the Cold War by an entity that did not possess a strong safety culture (i.e., the Soviet Government in the late 1950s and early 1960s), and thus it is not surprising that the initial total risk divided by the initial known risk appears in Figure 4-4 to be considerably larger than for the Shuttle in Figure 4-3.

Atlas

Atlas is an example of a launch system that was developed under significant time constraints during the early phase of the Cold War by an entity that placed more-or-less equal emphasis on safety and schedule (i.e., the U.S. Government in the late 1950s and early 1960s). Based on the data for Atlas shown in Figure 4-5, the ratio of the initial total risk to the initial known risk appears to be about the same as that for the Shuttle in Figure 4-3. However, both the numerator and the denominator in this ratio were about a factor of 5 higher for Atlas than for the Shuttle.

Delta

Unlike Soyuz/Molniya and Atlas, Delta is an example of a launch vehicle that was based on heritage technology. It was developed starting from the Thor vehicle with the objective of being more reliable. To accomplish this objective, components found to be unreliable in Thor were replaced by more reliable ones in Delta. The ratio of total initial risk to known initial risk for Delta in Figure 4-6 appears to be considerably less than for the Shuttle in Figure 4-3, as would be expected from the fact that Delta was more of a heritage system than the Shuttle.**

---

* The specific approach we used to develop the illustration in this example was to take a running snapshot of the number of launches required to produce 10 failures, using launch data provided in the International Reference Guide to Space Launch Systems [59]. For example, the Atlas launch vehicle had the first ten failures occurring within the first 19 flights, so we took a failure rate of 10/19 = 0.53 as being representative of the first 19 flights.

** The first five Thor flights were failures but the next five were successes. By comparison, the first Delta flight was a failure but the next 22 were successes. The failure on the first Delta flight was due to an avionics problem, not a propulsion system problem.

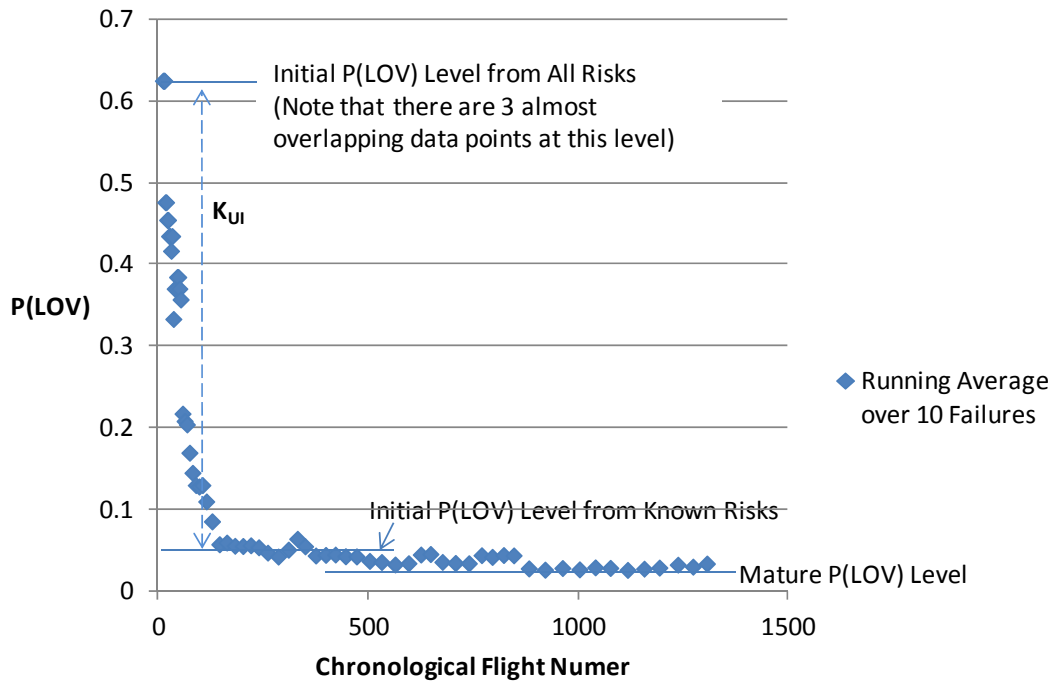


Figure 4-4. Failure History for the Soyuz and Molniya Launch Vehicles

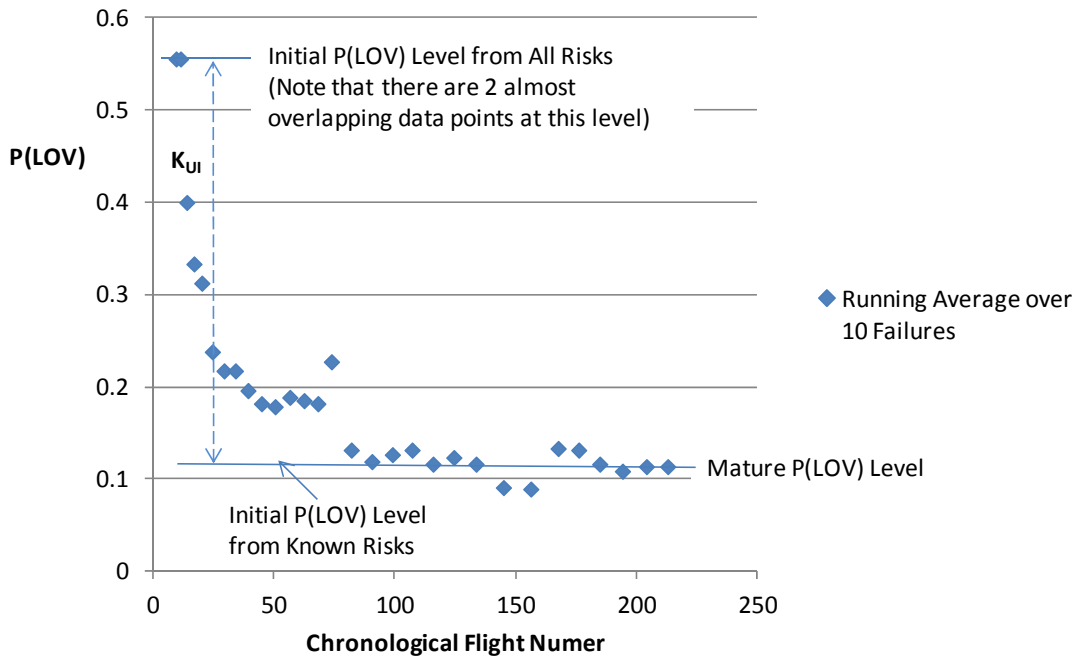


Figure 4-5. Failure History for the Atlas Launch Vehicle

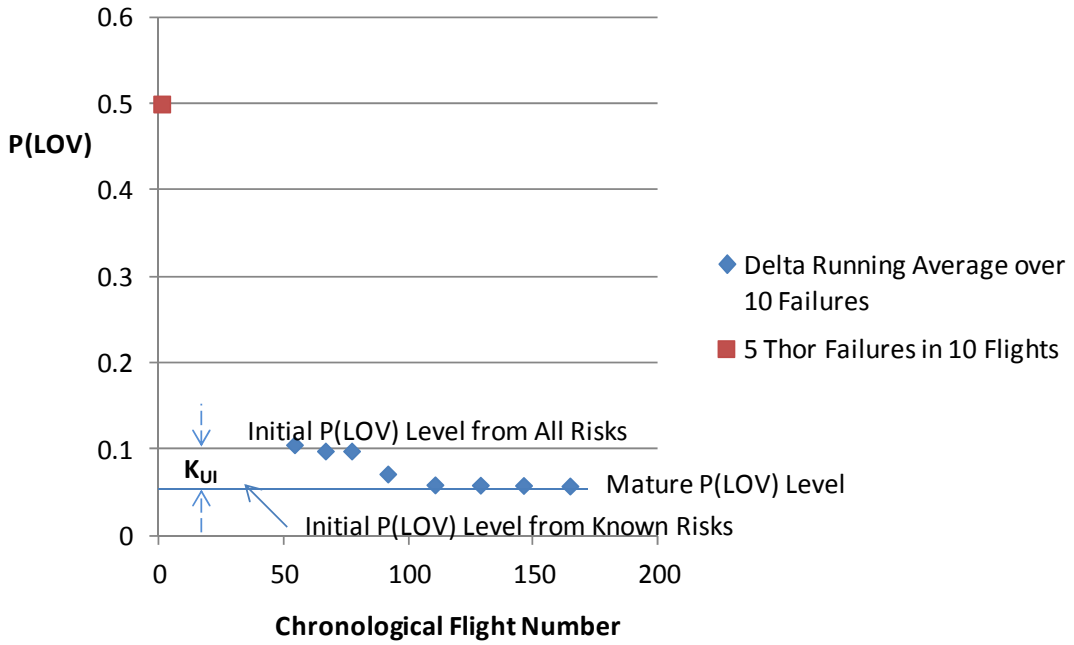


Figure 4-6. Failure History for the Delta Launch Vehicle

4.5.3 Initial Loss Probability Margin

Table 4-2, adapted from [54], provides suggested guidelines for specifying safety performance margins based on attributes of the system design, the project priorities, and the management culture. These guidelines were developed using a set of evidence that included but was not limited to the analyses in Sections 4.5.1 and 4.5.2 of this handbook. It also included operational experience with commercial nuclear reactors and military systems, as well as human reliability experience as encoded within performance shaping factors. The suggested guidelines were not meant to be prescriptive for all applications but rather to give an indication of the magnitudes of safety performance margins that are typical based on a wide variety of experience. As mentioned earlier, alternative estimates based on other data sources and other analysis methods may be used when considered appropriate.

As indicated in the parenthetical note under the title in Table 4-2, the applicability of the factors in the table is based on the assumption that the known probability of loss evaluated for the present system (i.e., the denominator in each factor) is consistent with results from analogous systems that have substantial operating experience accompanied by full-scope PRAs. Clearly the Space Shuttle and the ISS are examples of potential analogous systems. The factors in the table do not apply if the known probability of loss is evaluated only from a limited-scope PRA or other analysis method that consciously neglects potentially important sources of risk.

4.5.4 Justification for the Use of Ratios

As mentioned earlier in Section 4.1.4, it is convenient to suppose that it is the ratio of the total loss probability (from known and UU risks) to the loss probability from just known risks that correlates with the qualitative factors cited in Table 4-2. The results from the preceding sections in this example appear to corroborate this supposition. The discussion below provides a rationale for why it makes sense.

Table 4-2. Suggested Guidelines for Estimating the Ratio of the Initial Probabilistic Safety Performance Margin to the Initial Loss Probability from Known Risks

(Assumes that the known risk evaluated for the present system is consistent in likelihood with results from analogous systems that have substantial operating experience accompanied by full-scope PRAs.)

Safety Performance Factor*	Margin Ratio**	Applicable Conditions	Justification
1	0	Systems that can take credit for at least 125 actual cycles of operation of the same or equivalent systems with positive indication that the risk has leveled off to a mature system value	Results for Shuttle, Atlas, Delta, Molniya/ Soyuz after 125 flights
~2	~1	New systems that are developed and operated under at most mild time pressure, with reliability and safety having a higher priority than cost and schedule, with an inclusive management structure, and with a design philosophy that does not involve significantly new technology or new integration of an existing technology or scaling of an existing technology beyond the domain of knowledge or tight functional coupling	Results for Delta, first 75 flights***
~3	~2	New systems that are developed or operated under at least moderate time pressure, with cost and schedule having at least an equal priority with reliability and safety, and with a tendency for the management structure to be hierarchical, but with a design philosophy that does involve significantly new technology or new integration of an existing technology or scaling of an existing technology beyond the domain of knowledge or tight functional coupling	Results for Atlas, first 75 flights***
		New systems that are developed or operated under significant time pressure, and with a design philosophy that involves either new technology or new integration of an existing technology or scaling of an existing technology beyond the domain of knowledge or tight coupling, but with reliability and safety having a higher priority than cost and schedule, and with an inclusive management structure,	Results for Shuttle retrospectively, first 75 flights, if post-Columbia return-to-flight improvements had been in place***
~5	~4	New systems that are developed or operated under significant time pressure, with cost and/or schedule having at least an equal priority with reliability and safety, with a tendency for the management structure to be hierarchical, and with a design philosophy that involves either new technology or new integration of an existing technology or scaling of an existing technology beyond the domain of knowledge or tight coupling.	Results for Shuttle, first 75 flights. Anecdotally nuclear reactor experience and human reliability experience***
~10	~9	New systems that are developed or operated under extreme time pressure, with cost and/or schedule having significantly higher priority than reliability and safety, with a highly hierarchical management structure, and involving either new technology or new integration of an existing technology or scaling of an existing technology well beyond the domain of knowledge	Results for Molniya/ Soyuz first 75 flights. Factors of this magnitude and larger are also suggested in [63].

* Safety Performance Factor is the loss probability from all risks divided by the loss probability from known risks before the first flight

** Margin Ratio is the loss probability from UU risks divided by the loss probability from known risks before the first flight. Margin Ratio = Safety Performance Factor – 1.

*** Ratios of 2 to 5 are also consistent with historical reliability growth estimates cited in Table I of MIL-HDBK-189A for commercial and military systems.

When an accident occurs, the pursuits undertaken to prevent further accidents of that type involve identifying the causes of the accident and instituting design changes, operational changes, and/or administrative controls to prevent them from happening again. Most of the time, these changes and controls are formulated to affect a broader spectrum of accidents than just the one that is promulgating the action. For example, after the Columbia accident, one of the main corrective actions was to photographically scan the surface of the Shuttle while in orbit to detect damage caused by foam debris so as to be able to initiate astronaut extra-vehicular activities to repair any damage that might be significant enough to endanger re-entry. This corrective action had the effect of protecting not only against foam debris impacts but also against damage caused by micrometeoroids and orbital debris (MMOD), which is considered to be one of the main sources of risk for orbiting space vehicles. In addition, the return-to-flight activities associated with Columbia included a restructuring of the management within NASA to address generic shortcomings identified in the CAIB report [32]. These types of corrective action have a generic character that provides protection against many potential accident scenarios.

The implication is that the reduction of known risks also reduces UU risks. Clearly, however, that reduction is only possible when the protection against the known risks has a generic character as was the case for Columbia. It would not be the case if the reduction of known risks was focused very narrowly on the specific events contained in a known scenario.

4.5.5 Safety Threshold for LEO Missions

For LEO missions without LAS capability, an achievable threshold for P(LOC) for crewed missions or for P(LOV) for high-cost robotic missions might be inferred by combining the value of P(LOC) for known risks for the Shuttle in Figure 4-3 with the appropriate multiplier for UU risks in Table 4-2. For example, if it can be argued that the system is developed under moderate time pressure, with reliability and safety having the top priority over cost and schedule, and involving new integration and significantly new technology, then an initial loss probability of about 2.0 times the early known loss probability for the Shuttle (0.02) would be considered achievable based on Table 4-2. This would suggest a threshold value for initial operation of around 2.0 x 0.02 = 0.04.

For LEO missions with LAS capability, results from the study performed by NASA Ames [41] can be used. In the Ames study, the analysts calculated that if a LAS were integrated with the Ares 1 launch vehicle, the probability of LOC given a scenario that produced LOM during ascent would be reduced by another 80% ; i.e., the probability of LOC given LOM is about 0.2 (see Figure 4-7). The study considered a long list of launch system scenarios that could lead to the need to abort and their probabilities of occurrence. It also considered the effects of the harsh environment on the LAS and its passengers (blast wave, heat, fragments) which could kill the crew even if the abort initiation was successful. More information about this study will be provided in Section 5.2.6.

All in all, the evaluation of abort effectiveness for a LAS should include several factors:

- The probability of failure of the LAS to initiate abort (as considered, for example, in [45])
- The possibility of false positives (abort being initiated when not needed, see [46])
- The possibility of unintended interactions between the LAS and the launch vehicle (e.g., interactions between the control software for each vehicle, see [64])
- The effects of the harsh environment on the LAS and on the crew inside the LAS as considered, for example, in the cited Ames study)
- The probability of the crew failing to survive during or after touchdown or splashdown
- UU scenarios from other sources

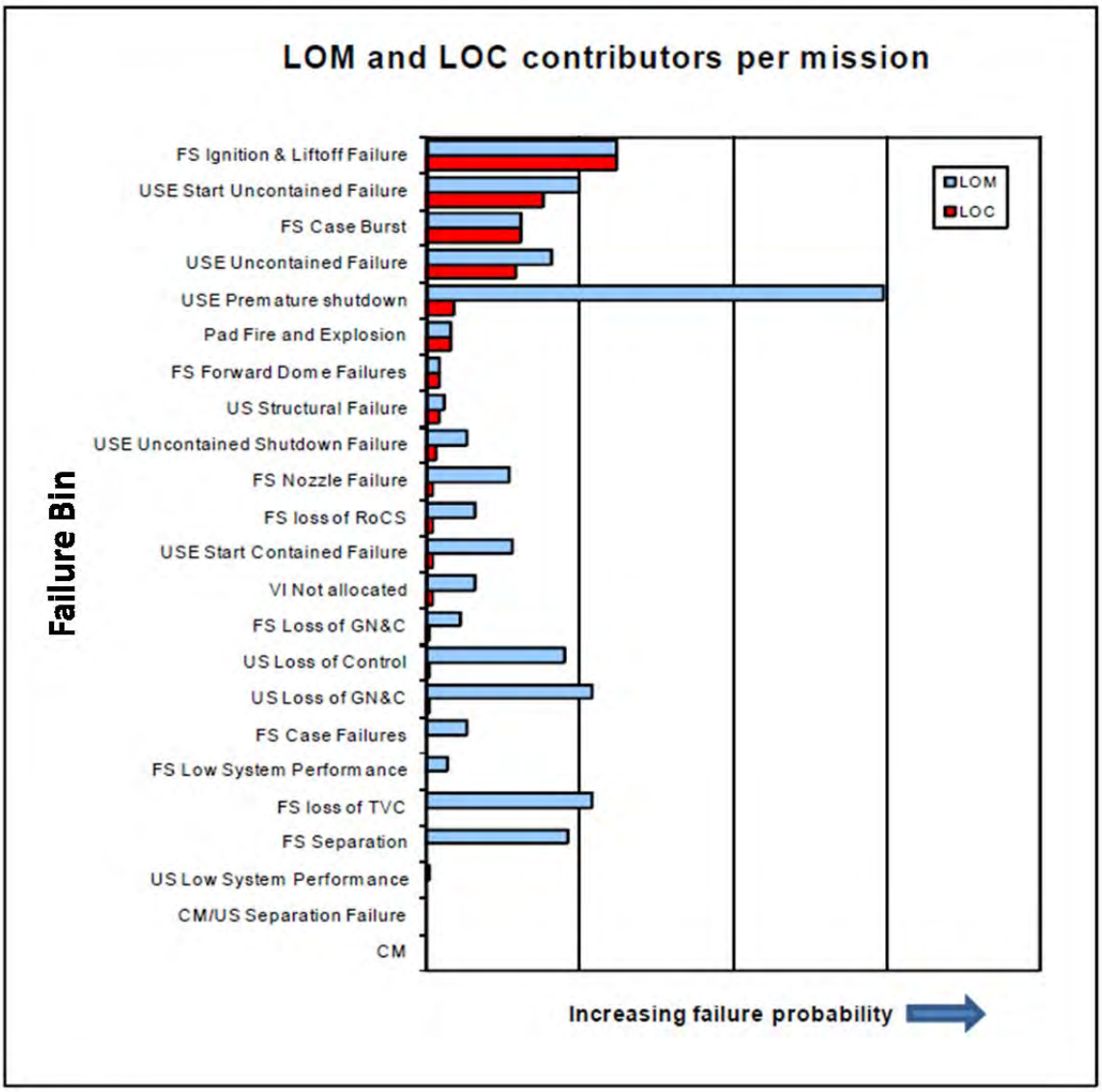


Figure 4-7. Loss of Mission and Loss of Crew Probability Contributions by Accident Scenario for an Example Launch System with a Launch Abort Capability, from NASA Ames Study

Given the harshness of a significant fraction of the abort environments and the fact that the analysis of these environments employs conservative assumptions, it is possible to argue that in most cases the probability of loss of crew given an accident that causes LOM during ascent will be dominated by the effects of the harsh environment on the LAS and the crew rather than by the other factors listed above. Thus, the assessment that the LAS reduces the probability of LOC from accidents during ascent by around 80% is likely valid despite the fact that some potential sources of accidents such as false positives and system-wide interactions were not considered.

In addition to the abort effectiveness for accidents during ascent, the overall effectiveness of the LAS for LEO missions must consider accidents occurring during other phases of the mission, such as orbit, docking with another craft such as the ISS, and entry. In the case of the Space Shuttle, about 23% the probability of LOC from known risks at the time of the first flight was assessed to be due to accidents initiated during

orbit and during entry. (This figure is calculated using the data in Table 4–1 by adding 5.3E-3 + 3.1E-3 + 2.4E-3 and dividing the total by 2.4E-2.) To be specific, about 11.5% of the total P(LOC) from known risks was associated with impacts of micrometeoroids and orbital debris (MMOD) occurring during orbit, and roughly an equal amount was attributable to APU failure occurring during entry (see Table 4-1). Increasing the total probability of LOC by 23% causes the conditional probability of LOC given LOM to change from 0.2 to 0.24, a small effect that is within the noise of the calculation. (The value of 0.24 is calculated by dividing 0.2 x 1.23 by 0.2 x 1.23 + 0.8.)

Given this information, an achievable threshold for crewed missions with launch abort capability might be estimated by multiplying the Shuttle loss probability from known risks (0.02) by the appropriate factor in Table 4-2 and by the conditional probability of LOC given LAS initiation obtained from the Ames study (around 0.2). For cases where the system is developed under moderate time pressure, with reliability and safety having the top priority over cost and schedule, and involving new integration and significantly new technology, an achievable threshold for initial operation using this logic would be around 0.02 x 2.0 x 0.2, or 0.008.

The achievable threshold value of 0.008 is not necessarily a limit that applies to all LEO systems. A lower safety threshold could be justified for cases where the LAS could be designed to handle a larger fraction of the accident scenarios initiated during ascent, or for cases where a capability for crew escape during orbit and entry could be designed into the system.

4.5.6 Safety Goal for LEO Missions

The example in Section 4.5.5 indicated that typical values for known risk for an LEO mission prior to the first flight would be around 0.02 for systems without a LAS (including both crewed and high-cost robotic systems) and 0.004 for systems with a LAS (crewed missions only). The value for systems without a LAS was based on the retrospectively derived known risk for the Shuttle prior to the first flight (Figure 4-3), and the value for systems without a LAS was 80% lower based on results for abort effectiveness from the NASA Ames study. This implies that an achievable safety goal for systems without a LAS would be around 0.5 x 0.02 = 0.01, and for systems with a LAS would be around 0.5 x 0.004 = 0.002. These goals compare to achievable threshold values of around 0.04 and 0.008 respectively (as quoted in Section 4.5.5), assuming the system is developed under moderate time pressure, with reliability and safety having the top priority over cost and schedule, and involving new integration and significantly new technology.

4.5.7 Probabilistic Loss Requirement for LEO Missions

Between the first and last flights, the requirement decreases at a rate that is consistent with the burndown of total risks that has been observed for other space missions. The burndown rate for the total loss probability for the Shuttle and for several launch vehicles has been displayed earlier in Sections 4.5.1 and 4.5.2 (see Figures 4-3 through 4-6). The relevant data are collected in a single chart in Figure 4-8, where it may be seen that the burndown rate tends to follow an exponential relationship. As noted earlier, the total loss probability tends to reach its steady state value after about 125 flights. Thus, it is reasonable to specify that the burndown rate for the requirement follows an exponential decay starting at the initial value prior to the first flight and ending at the final steady state value after the 125th flight. A graph of this relationship in log-linear form is shown in Figure 4-9.

Per the example presented in Sections 4.5.4 and 4.5.5 for an LEO mission with launch abort capability, suppose the safety threshold value has been specified at 0.008, the safety performance factor at 2.0, and the safety goal at 0.002. Then the safety requirement varies from a value of 0.008 / 2.0 = 0.004 prior to the first flight to a value of 0.002 after the 125th flight according to the following relationship: Requirement = 0.004 exp (− 0.0055 N). If the safety performance factor were 5, consistent with a system that is developed under

significant time pressure, with cost and/or schedule having an equal priority with reliability and safety, and involving new integration and significantly new technology, and if the threshold value and goal were specified at 0.03 and 0.001, respectively, then the pertinent relationship would be: Requirement = 0.006 exp (− 0.014 N).

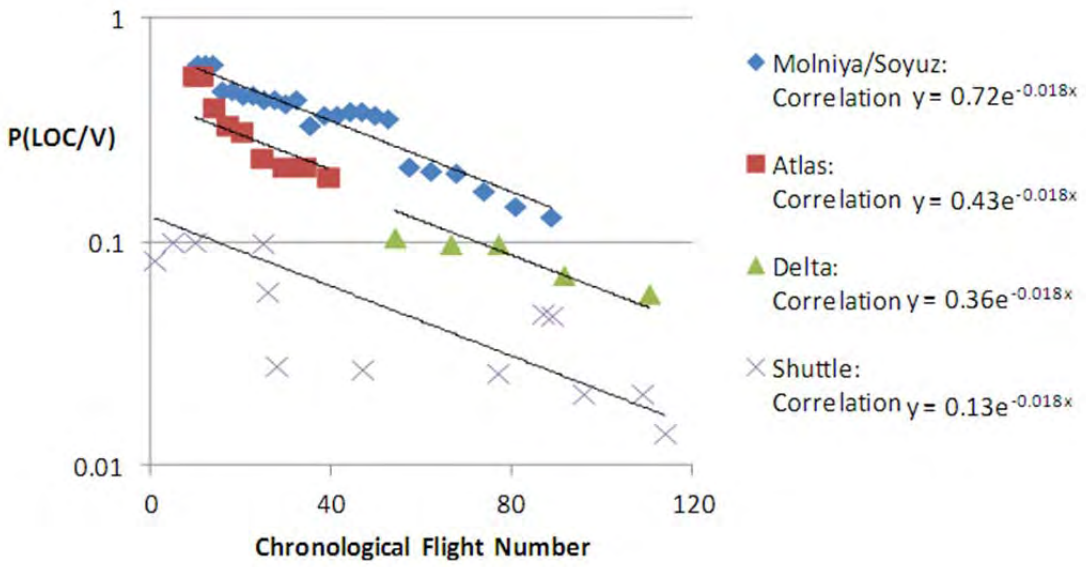


Figure 4-8. Correlation of Total Loss Probability with Chronological Flight Number

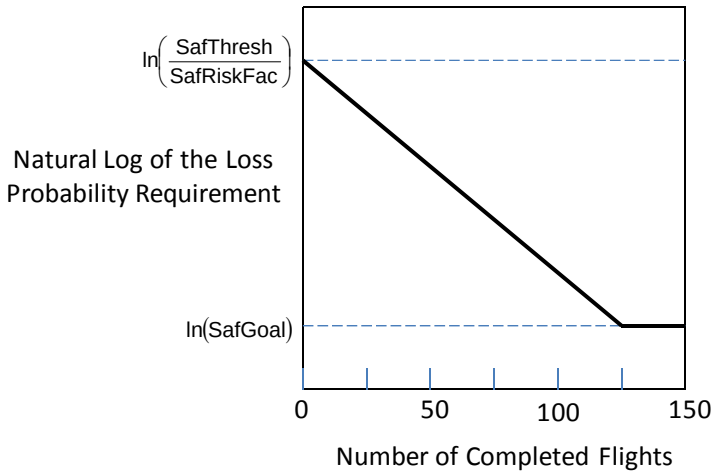


Figure 4-9. Loss Probability Requirement as a Function of the Number of Completed Flights

5. Performing System Safety Ensurance Activities: The Provider's Role

System safety ensurance activities are conducted by the Provider and support the design, development, and operation of a safe system that meets the safety performance requirements levied on it. System safety *ensurance* differs from system safety *assurance*, which is a responsibility of the Acquirer and evolves from the review by the Acquirer's evaluation team of the Provider's RISC submittal. System safety assurance is covered in Chapter 7.

Each system safety ensurance activity may address or influence one or more operational safety objectives, and conversely, each operational safety objective may be addressed by one or more system safety ensurance activities. Table 5-1 indicates the relationship between system safety ensurance activities and the operational safety objectives they address. The set of system safety ensurance activities is the same as in Section 3.2, with an additional system safety ensurance management activity to provide a management function over the other activities.

The following subjects from the overview in Section 3.2 are discussed sequentially in Sections 5.1 through 5.6:

- Developing a System Safety Management Plan (SSMP) and maintaining an auditable system safety program
- Conducting an Integrated Safety Analysis (ISA) and assessing the performance of Safety Critical Items (SCIs)
- Support of requirements development including tailoring of requirements and development of derived and allocated requirements
- System design support, including hazard elimination and control, implementation of failure tolerance, safing, emergency operations, and design phase testing
- Support of program control and commitments and performance monitoring throughout the life cycle
- Exploiting opportunities to improve safety

5.1 Developing and Implementing the System Safety Management Plan (SSMP)

5.1.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in developing and implementing the SSMP can be summarized as follows:

- Prepare the SSMP content as established by the Acquirer, and agreed to by the Provider, to accomplish some or all of the following objectives:
  - Provide the link between baselined safety requirements (including associated corollary process requirements) and system safety activities.
  - Detail the specific actions and arrangements required to operate a System Safety Program and define system safety milestones for the project.
  - Delineate the framework for the Provider's organization to direct and control its safety management activities, including the organizational structure, processes, procedures, techniques and methodologies.

Table 5-1. Cross Reference from Operational Safety Objectives to Provider System Safety Ensurance Activities[22]

Operational Safety Objectives	Provider System Safety Ensurance Activities					
	System Safety Ensurance Management	Integrated Safety Analysis	Requirements Development Support	System Design Support	Program Control and Commitments Support	Performance Monitoring Support
Establish safety performance margins	X	X				
Establish minimum tolerable levels of safety for known risks	X	X				
Maintain appropriate safety performance margins for the as-designed system	X				X	X
Maintain minimum levels of safety for known risks for the as-designed system	X	X	X	X	X	X
Maintain appropriate safety performance margins for the as-built system	X				X	X
Maintain minimum levels of safety for known risks for the as-built system	X	X	X	X	X	X
Maintain appropriate safety performance margins for the as-operated system	X				X	X
Maintain minimum levels of safety for known risks for the as-operated system	X	X	X	X	X	X
Prioritize safety during design solution decision making	X	X		X		
Prioritize safety during product realization decision making	X	X		X		
Prioritize safety during system operation and sustainment decision making	X	X		X		
Risk-inform safety requirements allocation decisions	X	X	X	X		
Be responsive to new safety-relevant information during system realization	X	X		X	X	
Be responsive to new safety-relevant information during operation/sustainment	X	X		X	X	X
Incorporate safety-related best practices into system design	X			X		
Minimize the introduction of hazards during system realization	X			X	X	
Minimize the introduction of hazards during system operation/sustainment	X			X	X	X

[22] For more information, see Chapter 4 of [1].

- Describe, in appropriate detail, plans for development of the RISC for each major milestone review as determined by the Acquirer.
- Address the activities necessary to ensure safety throughout the system life cycle.

- Incorporate a Product Assurance Plan into the SSMP.
- Review and update the SSMP regularly throughout the life of the system, especially as safety requirements are rebaselined.
- Address those aspects of the systems engineering approach that affect safety, and those aspects of the risk management approach related to safety.
- Ensure that the SSMP is coordinated with the Systems Engineering Management Plan, if one exists, and with the governing Risk Management Plan.

The following topics pertaining to these responsibilities and areas of consideration are addressed in Sections 5.1.2 through 5.1.6:

- Conducting an SSRA and setting requirements for the SSMP
- The general contents of the plan
- Specifying roles and responsibilities within the plan
- Configuration control and data management planning
- Addressing management buy-in within the plan

5.1.2 Conducting a System Safety Requirements Analysis (SSRA) and Setting Requirements for the SSMP

If the development work being undertaken has an inter-organizational character, then externally imposed requirements on planning processes may be appropriate.[23] Organizational units providing a system or service to other organizations as part of a large-scale development effort are subject to numerous requirements imposed by the acquiring organizational unit, or deriving from higher-level agency requirements (e.g., safety thresholds) or institutional requirements (e.g., NPR 8000.4), or external requirements (e.g., environmental regulations). Satisfaction of all these requirements will need to be planned in a coordinated fashion, and isolated plans for each cannot therefore serve.

For those reasons, the conduct and evaluation of a System Safety Requirements Analysis (SSRA) is essential. Subsequent to that process, the set of safety requirements can be baselined. Once that occurs, the Provider finalizes a System Safety Management Plan (SSMP) or its equivalent that lays out the process for satisfying the baselined safety requirements. The SSMP is submitted to the Acquirer for approval. Just as the SSRA serves as the safety requirements handshake between Acquirer and Provider, an SSMP developed according to the following guidelines can serve as the *process* handshake between the Acquirer and the Provider. The SSRA may be included within the SSMP or may be a stand-alone document, in which case it is referenced by the SSMP.

Unless the safety risks are trivial, the case for risk acceptance needs to be based on a combination of demonstrated product attributes (e.g., test results) and development process attributes (QA, validation and verification, peer review, etc.) which, essentially by definition, need to have been planned, and whose execution needs to have been documented. It may be possible to establish product performance after the fact in a small number of tests, but an emergent property such as safety cannot be established in that way.

[23] This would not usually be the case for a single organization carrying out an activity in isolation for its own reasons. A written plan would be strictly that organization's concern, just as its project management practices would be its own concern.

System safety process cannot, by itself, assure safety either, but many acquisition decisions need to be supported, at least in part, by evidence of process. Planned and documented process is evidence in the RISC.

It is not strictly necessary to address the above content within a single stand-alone document; the Acquirer may prefer to see portions of it incorporated into the Systems Engineering Management Plan, and other portions incorporated into the Risk Management Plan.

5.1.3 General Contents of the SSMP

Following is a listing of topics that may be addressed in the SSMP. This list is consistent with the discussions in the preceding subsections, and is also informed by the Johnson Safety Center Safety & Health Handbook [65]. Each program/project may have different requirements for an SSMP. The list below generically summarizes the expectations of an SSMP for a program/project that has high criticality[24]. The SSMP for a high criticality program/project should:

1. Describe the mission and the scope of the program/ project.
2. Describe the system being analyzed and the general safety philosophy of the design and operation of the system.
3. List any documents and specifications that the system safety effort will use either as directives or as guidance.
4. Describe the operational safety objectives of the program/project.
5. Describe the system safety organization or function, including charts to show the organizational and functional relationships and lines of communication.
6. Describe the responsibility, authority, and accountability of system safety personnel and other organizations (including contractors and subcontractors) involved in the system safety effort.
7. Assign an organizational unit for each task and an authority for resolving each unresolved safety issue.
8. Describe how the system safety organization is staffed for the length of the program/project, including labor loading.
9. Describe the qualifications of key personnel including both technical and managerial personnel.
10. Describe the interfaces between the system safety organization and other related disciplines such as systems engineering, reliability, and quality assurance at all levels of the program/ project (NASA, contractor, and subcontractor).
11. Identify safety milestones and reviews of the effectiveness of the system safety effort at critical safety checkpoints (e.g., design reviews, self-evaluations, operational readiness reviews, audits, etc.).
12. Schedule safety tasks, show start and finish dates, report dates, review dates, and labor loading, as they relate to other program/project milestones.
13. Identify other engineering tasks such as design analyses, tests, or demonstrations that also apply to the system safety program, and identify the system safety personnel who will participate in these tasks.

[24] Expectations associated with programs/projects of different levels of criticality will be defined further in Section 5.2.4

14. List the safety standards, system specifications, and levied requirements the program/project either must follow or will adopt.
15. Describe the procedures for assessing risk, including the acceptable risk levels for the program/project.
16. Describe the management controls to make sure the program/project follows safety requirements, including the process for making management decisions, the level of management required to accept different levels of risk, and methods to make management aware of and take action on risks.
17. Describe the past experience of the organization in managing large programs/projects similar to the one being addressed herein.
18. Describe the analysis techniques and format to be used to identify risks, their causes, their effects, and recommended responses.
19. Identify when each analysis technique will be used.
20. Describe how system safety analyses from different organizations such as contractors and subcontractors will be integrated.
21. Describe how cross-system interactions and interfaces between hardware, software, and the human element will be factored into the system safety analyses.
22. Describe how fault management will be included in the design and operation of the system and how these provisions will be factored into the system safety analyses.
23. Describe how software safety concerns will be factored into the system safety analyses.
24. Describe how system safety analyses will be graded according to the criticality of the mission and the importance of the risk scenarios being investigated.
25. Describe how the system safety analyses will be updated throughout the life cycle as relevant new information emerges.
26. Describe the approach for researching, distributing, and analyzing historical hazard or mishap data, best practices, and lessons learned.
27. Describe the approach for identifying and tracking precursors, anomalies, and near misses during the course of the program/project and for incorporating knowledge gained from these events into the system safety analyses.
28. Describe how safety critical items (SCIs) and risk drivers will be derived from the results of the integrated safety analyses and how these SCIs and risk drivers will be managed throughout the program/project life cycle.
29. Describe how a risk-informed safety case (RISC) will be developed and how the arguments will be structured to provide confidence that the system is adequately safe for each key mission objective.
30. Describe how the principles of the minimum tolerable level of safety and as-safe-as-reasonably practicable (ASARP) will be implemented for the system.
31. Describe how the potential contributions from unknown and underappreciated risks will be incorporated into the RISC and how these will be minimized by incorporating relevant best practices and lessons learned.

32. Describe how evidence will be gathered to support the RISC and how the evidence will be evaluated.
33. Describe how the credibility assessment scale (CAS) factors in the NASA Modeling and Simulation (M&S) Standard and Handbook will be used to assess validity of the evidence.
34. Describe how emerging opportunities to improve safety will be considered within the ASARP context and how they will be implemented into the system design and operation when shown to be effective and practicable.
35. Identify the data management needs and methods for making risk-informed safety decisions.
36. Describe the verification and audit requirements and procedures to make sure that the system safety program has been implemented and that the levied requirements have been satisfied,
37. Describe the procedures to make sure that safety information is available for management and engineering review and analysis.
38. Describe the review procedures to make sure that hazardous tests, and especially tests involving human test subjects, are conducted safely.
39. Describe training, techniques, and procedures to make sure that engineers, test subjects, technicians, operators, and support (including maintenance) personnel understand the objectives and requirements of the system safety program.
40. Identify and describe any reviews by experts outside the program/project.
41. Describe how the SSMP will be updated whenever new information emerges that would affect the conduct of the safety program or change the results of the risk-informed safety case, and how comments from program/project reviews will be incorporated and resolved.
42. Describe how the content of the SSMP has been integrated with the program/project system engineering and risk management plans and how they collectively address the needs of system safety, systems engineering, and risk management.

5.1.4 Specifying Roles and Responsibilities, Controls, Processes, and Protocols

As discussed in [33], information to be provided in the SSMP should include, for each organizational unit in the structure:

- Roles and responsibilities
- Controls
- Process model requirements
- Coordination and communication protocols
- Contextual (environmental and behavior-shaping) factors that might bear on the unit's ability to fulfill its responsibilities
- Inputs and outputs to other units in the control structure

Roles and responsibilities for safety management is part of the safety control structure for the organization. The SSMP should provide information about where the responsibility for implementing each safety requirement rests. If there is a current safety organization with roles and responsibilities assigned at different levels of the organization, a gap analysis should be performed to identify where requirements are not being implemented (enforced) anywhere. Thereafter, the safety control structure

needs to be evaluated to determine whether it is potentially effective in enforcing the system safety requirements and constraints.

The SSMP should be able to show that the control structure is capable of allocating responsibility to continuously demonstrate, as the program/project progresses, that the safety claims made in the RISC are satisfied. Providing assurance that the safety claims are satisfied is synonymous with ensuring that the assurance deficits that are identified for each claim are maintained within acceptable limits[25]. Thus, an effective safety control structure is an organizational structure within which there is an assignment of responsibilities to designated owners for tracking, evaluating, and managing each source of assurance deficit and for documenting/communicating any effects on the RISC.

DEFINITION OF ASSURANCE DEFICIT

An assurance deficit is "any knowledge gap that prohibits perfect (total) confidence" [66]. Assurance deficits are caused by variability or lack of knowledge concerning the data or models being used to produce the evidence, the parameter inputs to the models, and the interpretation of model outputs. Examples of assurance deficits include high statistical uncertainty associated with an insufficient amount of data, the possibility of externally imposed major programmatic changes, and incompleteness in the identification of hazards. Although "assurance deficit" is more-or-less synonymous with "confidence deficit," there is a slight but significant difference in that "assurance" implies an active process of gathering and assessing evidence, whereas "confidence" implies the result attained by providing adequate assurance.

As an example, following is a partial list of potential unplanned events and conditions that, if they should emerge, could negatively affect the level of confidence in being able to meet the claims in the RISC:

- Externally imposed decisions (such as Congressional funding decisions) could result in a need to alter the key mission objectives, possibly resulting in a reduction of planned launches and in-flight events.
- New information from ongoing programs/projects within NASA could result in changes in the perception of reasonableness for previously determined safety thresholds, goals, margins, and requirements.
- New analyses using newer models and data sources could improve the accuracy and completeness of the ISA but also cast doubt on previous results.
- New developments affecting the ratings for the M&S credibility assessment scale (CAS) factors for various analyses that have been performed could negatively affect confidence in the ISA.
- Changes in management personnel could lead to changes in the confidence that safety is being managed effectively.

The SSMP should contain the following to ensure that such changes are adequately addressed:

- Delineation of how responsibilities for tracking, evaluating, and managing each assurance deficit source have been or will be assigned to individuals and teams that have the proper qualifications and authority to act
- Explanation of how the information from their activities will be communicated, documented, and used for updating the RISC

[25] This topic will be addressed further in Section 6.1.2.

The process for managing safety claim assurance deficits should be similar to and commensurate with the organization's process for managing program/project risks. In fact, management of assurance deficits can and should be part of the overall risk management process. In this context, the sources for assurance deficit identified in the RISC convert to "risk statements" and become part of the organization's "risk list." The tracking and control of these risks is assigned to "risk owners" who, as discussed in NPR 8000.4A [4], have the lead for overseeing the implementation of the agreed disposition of the risks that are assigned to them.

The organizational aspects of risk management within NASA are discussed in NPR 8000.4A. The relationship between RISC development/evaluation and risk management is illustrated in Figure 5-1.

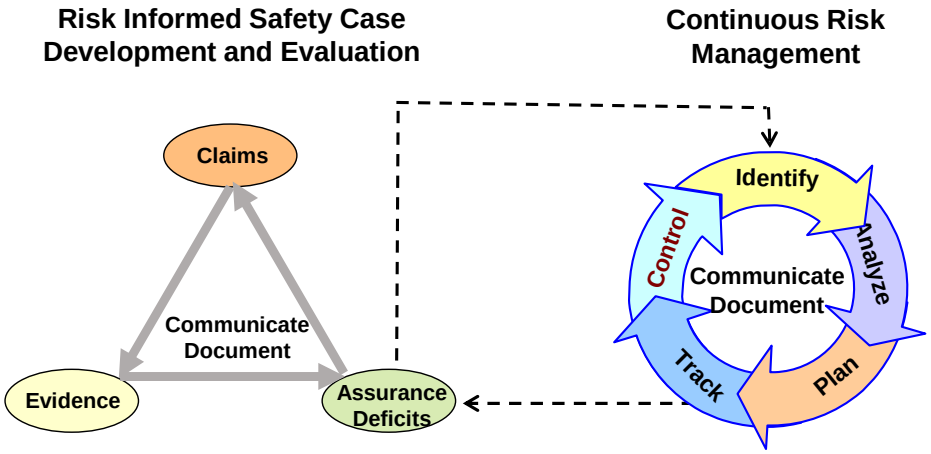


Figure 5-1. Relationship between Risk-Informed Safety Case Development and Evaluation and Continuous Risk Management

5.1.5 Configuration Control and Data Management

Improper configuration control and data management have been among the root causes of several mission failures, including the Mars Climate Orbiter mishap [67] and several software-related spacecraft accidents [64]. Therefore, sound configuration control and data management practices should be part of the RISC and should be addressed in the SSMP. NASA's preferred practices for configuration control and data management are explained in the NASA Configuration Management Standard [68] and in the NASA Systems Engineering Handbook [13].

Following are some of the characteristics of a good configuration control program (adapted from the NASA Systems Engineering Handbook):

- Enough time is allocated for engineering, SMA, or other Configuration Control Board (CCB) members to meet to evaluate changes.
- Dissent in the record for CCB documentation is allowed for.
- Redlines and other informal practices to keep track of changes are avoided.
- There are clear procedures for the integrating contractor to approve all subcontractors/vendors changes.
- There is a designated process for preparing and coordinating change requests.
- Verification data can be traced to the appropriate hardware/software and specification.

- Supporting change procedures adequately involve approval by the originating organization.
- The contractor has a Configuration Management web site that NASA is able to access to verify the latest released changes.

There should also be a data management plan that covers the following subjects (from the NASA Systems Engineering Handbook):

- Identification/definition of data requirements for all aspects of the product life cycle
- Control procedures—receipt, modification, review, and approval
- Guidance on how to access/search for data for users
- Data exchange formats that promote data reuse and help to ensure that data can be used consistently throughout the system, family of systems, or system of systems
- Data rights and distribution limitations such as export-control Sensitive But Unclassified (SBU)
- Storage and maintenance of data, including master lists of where documents and records are maintained

For purposes of the SSMP, the Provider should describe their plans for configuration control and data management, demonstrate through past performance how their management and their teams have handled configuration control and data management in related programs/projects, and explain how configuration control and data management will be handled for proprietary or restricted information and how such information will be made available to members of the team who have a need to know it

5.1.6 Management and Organizational Commitment to Safety

Admiral Hyman G. Rickover famously stated that the principal characteristics of high-reliability organizations are: (1) top management being commitment to safety as an organizational goal, (2) implementation of personnel redundancy as well as engineering redundancy, (3) development of a culture of reliability, and (4) the valuation of organizational learning [25]. When these principles are implemented, they have the effect of countering design characteristics that can produce UU risk such as interactive complexity and tight coupling in the design[26]. On the other hand, not implementing them makes it possible for design deficiencies to inordinately affect the risk.

It is pointed out as well [26] that avoidance of failures requires a nonhierarchical and consultative relationship, at least in the planning stages and general operational processes, and that two-way flows of information are especially essential in technological systems to maximize the sharing of information among all personnel regardless of position in the organizational hierarchy. Bureaucratic barriers to cooperation are particularly dysfunctional, given our limited understanding of technological systems and our limited ability to control them. However, when a crisis arises in the operations of a technological system, the command model – namely, a hierarchical and single-directional mode of communication – should supersede the nonhierarchical consultative model in an effort to contain the crisis and limit the damage.

As evidenced by the Mars Climate Orbiter mishap (launched in 1998; lost in 1999), interfaces between different elements of the system provided by different suppliers require stringent oversight by the managing agency (NASA). In that case, inadequate oversight resulted in a catastrophic failure because one organization had written the flight system software to calculate thruster performance using metric units, while another was entering course correction and thruster data using Imperial units.

[26] This point will be discussed further in Section 5.4.5.

Two sources of information are pertinent for providing evidence of management's dedication to safety and effective oversight: (1) documentation in the SSMP and the Program/Project Management Plan, and (2) history of the effectiveness of the management team in present and past programs/projects.

Indications of management and organizational commitment to safety that need to be discussed in the SSMP and demonstrated in the RISC are summarized below (excerpted from a list provided by the International Association of Oil and Gas Producers [69]):

Demonstration of management commitment

- Safety Policy Statement is endorsed by Chief Executive of the Company (or if the Providing organization resides within NASA, by the Administrator). It is prominently displayed at worksites and in company documentation, and is regularly discussed.
- Senior Management is visibly committed to safety, as evidenced by involvement in the appropriate company safety committees; carrying out safety 'walkabouts' and site visits; making personal statements on safety; holding regular, routine safety performance reviews; active involvement in incident reviews and follow-up (incidents include injuries, emergencies and near misses); inclusion of safety as a high priority item during management and staff meetings; and appropriate resourcing is committed to safety.
- Senior management compares safety performance on an inter- and intra-company basis.
- Management participates in setting safety standards and measurable targets, and in establishing performance measurement systems and procedures.
- Management personally practices safety as a top priority and ensuring that their organization does not compromise safety in pursuit of other goals.
- Management and supervision demonstrate a deep belief that safety is part of their job, and behaving accordingly.

Demonstration of organizational motivation and communication

- There is a clearly defined and implemented communication structure which enables open and frank two-way communication.
- There is clear communication of collective and individual safety performance.
- There is consultation and communication of safety objectives, tasks, and targets.
- There is a program for recognizing safety performance.
- There is a system in place for the reporting of hazards, follow-up by manager, and feedback to the initiator.
- Regular meetings have defined purpose, structure, and records.
- There are ad hoc meetings for special tasks/reviews.
- There is a system in place to encourage safety suggestions and ideas.
- There is an effective safety awareness program which includes promotion, publicity, off-the job safety, etc.

Demonstration of contractor involvement and commitment

- Safety is an integral part of the contractor selection procedure.
- Contracts contain clear provisions requiring the contractor to meet safety conditions and specifications consistent with the company's standards for its own employees.
- There is an audit of contractor performance and facilities prior to the start-up of work.
- There is a system for approving contractor working procedures for use on company operations.
- There is a system for verification of contractors' training programs, individual competence/ qualifications, skill achievements, and where appropriate, team experience.
- The contractor has to demonstrate safety commitment from their senior management, clearly defined safety responsibilities, an active safety program, a system for reporting safety performance and investigation of incidents, an effective communication and consultation system, and effective control of sub-contractors.
- Company supervision monitors contractors' performance.

5.2 Performing an Integrated Safety Analysis (ISA)

5.2.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in developing and implementing the SSMP can be summarized as follows:

- Beginning at Formulation, conduct an ISA of the system that has the following characteristics:
  - Is scenario-based, accounting for credible departures from intended system operation that have the potential to lead to adverse safety consequences, accident propagation pathways through the system including interactions among subsystems, the potential for dependent effects, and the potential for latent faults/failures
  - For critical systems and systems of high value, is scoped to encompass all credible risks affecting safety, and uses tools and techniques as necessary to achieve the highest practicable level of completeness given the current maturity of the system, recognizing that completeness of such risk analyses at the root cause level cannot be guaranteed *a priori*
  - For noncritical or lower value systems, is scoped to encompass the most important risks according to a graded analysis approach
  - Addresses uncertainty and sensitivities
  - Is consistent with relevant verification procedures
  - Adheres to analysis protocols emerging from SSRA
- From the results of the ISA, designate as SCIs a preliminary set of system features whose performance (e.g., capability, reliability, and availability) at levels documented in the ISA assures the satisfaction of system-level safety performance requirements. Note: the list of SCIs will be updated later to include other items that are found during the RISC development to be important to safety.
- As part of the ISA, assess the ability of each SCI to perform its safety function(s) in the ISA accident scenario(s) that rely on that function for the prevention/mitigation of adverse safety consequences. For each SCI, this responsibility involves the following objectives:
  - Identify the accident scenario(s) in the ISA that involve the SCI and the function performed by the SCI in each accident scenario.

    - Evaluate the capability of the SCI to carry out its safety function in the context of the accident scenario(s) for which that safety function is needed, including uncertainty.
    - Evaluate the reliability of the SCI in performing its safety function in the context of the accident scenario(s), including uncertainty.
    - Evaluate the sensitivity of the safety performance of the system, as a whole, to the capability, reliability, and availability of the SCI.
- Designate as risk drivers those SCIs whose failure probabilities are sufficiently high to make them significant contributors to the safety performance risk
- Keep the ISA current through timely incorporation of the following:
    - System design and operational changes and/or refinements
    - Corrections to, additions to, and/or refinements of existing ISA models
    - Safety-related information gained through operational experience with the system, including the occurrence of anomalies deemed as precursors to adverse safety consequences
    - Safety-related information gained through testing or other data gathering activities
    - Safety-related concerns that have been accepted into the risk management process
- Document the ISA in a report, to include:
    - The system description, which provides the physical and functional characteristics of the system and its subsystem interfaces, and refers to more detailed system and subsystem descriptions, including specifications and detailed review documentation, when such documentation is available
    - ISA methods and techniques, providing a description of each analysis method and technique used, where in the system they are applied, and how these techniques are integrated into the ISA. Include a description of assumptions made for each analysis and the qualitative or quantitative data used.
    - ISA analysis results. Contents and formats may vary according to the individual requirements of the program and methods and techniques used.
    - ISA findings and recommendations

The following topics pertaining to these responsibilities and areas of consideration are addressed in Sections 5.2.2 through 5.2.9:

- Rationale for performing an ISA
- Characteristics of an ISA
- Implementing a graded analysis approach in the formulation of the ISA
- Integrating subsystem analyses into the ISA
- Integrating software analyses into the ISA
- Implementing special considerations for fault management into the ISA
- Conducting tests to support the ISA
- Adhering to the Modeling and Simulation (M&S) Credibility Assessment Scale (CAS) in conducting the ISA

Integration of human interactions with hardware and software is an area that is also discussed in this section,[27] albeit qualitatively. Furthermore, the importance of various sources of human error that tend to increase the probability of UU risks will be discussed later,[28] and safety performance margins to account for the effects of various human-error-producing factors were discussed earlier.[29] However, specific means for modeling human errors of commission and omission in an ISA are not developed herein.

5.2.2 Rationale for Performing an ISA

Of central importance to the design, development, and operation of a safe system is the early conduct and subsequent maintenance of a comprehensive ISA by which the Provider identifies and evaluates safety risks and documents the resulting state of knowledge regarding the safety performance of the system. The ISA serves as a technical basis for imparting system-specific safety performance information into systems engineering and risk management decision making. The scope of system safety involvement in systems engineering is necessarily broad, reflecting the reality that practically all systems engineering decisions have the potential to affect safety.

A scenario-based ISA is considered essential to achieving adequate safety, because it is only by developing an understanding of how adverse safety consequences can potentially be produced (i.e., what the accident scenarios are) that 1) effective measures can be taken to prevent or mitigate them (e.g., by managing safety risk drivers), and 2) the measures upon which safe operation depend can be identified and maintained at an acceptable level of functional effectiveness over the life of the system.

A scenario-based understanding of accident potential is considered essential regardless of the quantity of operating experience that a system has. For a system with little or no operating experience, ISA can be thought of as an elicitation process involving postulated accidents whose cumulative safety risk is then managed to acceptable levels. For an operational system, the ISA reflects whatever mishaps or anomalies have occurred over its operational lifetime, in addition to those postulated by the ISA analysts. The main difference is the degree of completeness of scenario identification. As a system gains operational experience in the form of successes as well as failures that are subsequently fixed, there is increasing confidence that the system is free from higher probability/frequency unknown accident scenarios.

5.2.3 Characteristics of an ISA

An ISA a system-level identification and analysis of scenarios that may lead to undesirable safety consequences such as 1) death, injury, or occupational illness; 2) damage to or loss of equipment or property, loss of mission; or 3) damage to the environment (e.g., Earth or planetary). The term "integrated" in ISA refers to four different aspects:

- *Analysis Methodologies* – ISA is methodologically non-prescriptive, allowing for the application of different methodologies as appropriate for the level of system definition (which evolves over the life cycle), the specifics of the part of the system/mission being addressed (e.g., software, launch abort), the safety significance of the issue being analyzed, etc. The ISA integrates safety performance information obtained from these separate/complementary methodologies into a single, comprehensive, system-level accident scenario set that represents the Provider's state of knowledge regarding the safety of the system.

- *Subsystem Analyses* – ISA is a system-level safety analysis that utilizes analyses performed at lower levels of the product breakdown structure in the service of a single overarching system-level analysis. All scenarios that affect the safety of the system are within the scope of ISA, regardless of whether or not they propagate across subsystem boundaries.

[27] See Section 5.2.5, Figure 5-5.
[28] For example, see Section 5.5.7.
[29] For example, see Section 4.5.3, Table 4-2.

- *Safety Performance Measures* – ISA integrates the analysis of different safety-related performance measures to the greatest extent possible, recognizing the extent to which multiple at-risk entities (crew, public, assets, etc.) are affected by the same scenarios.

- *Risk Analysis* – ISA is integrable with analyses in other performance domains (technical, cost, schedule) to support the ASARP principle.

ISA adheres to two key principles:

1. *ISA is Scenario Based* – The purpose of an ISA is to comprehensively identify and analyze what can go wrong in the system, in terms of a comprehensive set of accident scenarios that connect accident causes to safety consequences of concern. Accident scenarios identify and analyze the response of the system to the accident as it progresses, including the potential for success and failure of preventive and mitigative features. Depending on the level of rigor of the ISA (as discussed in section 5.2.4 below), the set of scenarios may be defined qualitatively in terms of the credible sequences of events that can lead to adverse consequences, or also quantitatively in terms of the probabilities/frequencies of occurrence of the sequences and the magnitudes of the consequences.

2. *ISA is Conducted at the System Level* – ISA is by definition a system-level analysis that aims at comprehensively identifying all credible safety-related accidents associated with the system in the context of its intended operation. This differs from integrated hazard analysis (IHA) as traditionally conducted at NASA, which is a coordinated analyses between elements or projects that addresses only those hazards or causes that are controlled by another project or element than the one who is producing the analysis. Instead, ISA is concerned with the totality of the system and its potential accidents, regardless of whether or not they cross subsystem boundaries as they progress from cause(s) to consequences. The comprehensive, system-level perspective provided by ISA enables:

    - <u>The holistic development of controls</u>: By understanding the accident potential of the system as a whole, it becomes possible to optimize the development of control sets. Controls can be implemented to provide broad coverage over multiple scenarios, as opposed to implementing controls on a scenario-by-scenario or subsystem-by-subsystem basis. This helps to minimize the cost and complexity of the control set, and to free up resources (e.g., mass) for use in the service of other performance attributes (e.g., payload).

    - <u>Determination of aggregate safety performance</u>: Although not necessarily done (e.g., Priority 3 projects), quantification of system safety performance (e.g., P(LOC), P(LOM)) requires aggregation of the scenario-based contributors to safety risk over the entirety of the system. This is particularly true when uncertainties are correlated across subsystems.

    - <u>System-level risk acceptance</u>: Whether the safety performance of the system is quantified or not, it is necessary to have a single organizational point of responsibility for system risk acceptance. In cases where the ISA is qualitative, this entails the acceptance of individual scenarios in terms of the adequacy and acceptability of measures taken to eliminate or control the underlying hazards. This acceptance cannot properly be done outside the context of a system-level perspective, because without such a perspective there is a risk that separate points of responsibility will accept scenarios/controls without a proper appreciation, however qualitative, of the scenarios/controls that have been accepted or rejected by others in other parts of the system. In cases where there are system-level safety performance requirements, such as on P(LOC), system-level risk acceptance is necessary by definition.

This is not to say that subsystem safety analyses should not be done. Typically, the bulk of the engineering expertise required for a credible safety analysis resides in subsystem level organizations, and

it is reasonable for these organizations to develop subsystem safety analyses as part of their SE processes. However, at the level of the system, these analyses should be taken as technical input to the ISA, rather than as collectively comprising the ISA (augmented by IHA). Section 5.2.5 provides guidance on integrating subsystem analyses into the ISA.

SCENARIO ORIENTATION OF ISA

In realistic engineering situations, scenario-based modeling within ISA is central to building a strong understanding of the system and a strong RISC, because it is necessary to understand what scenario elements need to be prevented or mitigated, in order to formulate, justify, implement, and (for purposes of the RISC) defend the strategies needed to prevent or mitigate those events. Moreover, in typical system safety applications, besides identifying scenarios, it is necessary to quantify scenario likelihoods, and to address uncertainty. This is true in the context of safety prioritization and safety tradeoff exercises, or as part of addressing safety requirements, goals, or thresholds.

As indicated in Figure 5-2, a scenario begins with an initiating event that perturbs the system away from its nominal condition. Subsequent pivotal events that are relevant to the evolution of the scenario may (or may not) occur, and may have either a mitigating or exacerbating effect on the accident progression. The successful functioning of controls will in general have a mitigating effect, whereas the failure of controls to function, the defeating of controls due to overwhelming stresses, or the involvement of hazardous material will tend to exacerbate the scenario. The spectrum of possibilities for the evolution of the accident is represented by the multiple pathways that can be followed and the multiple end states that can be produced.

Hazard analysis as traditionally practiced by NASA has focused specifically on the worst-case credible consequences of identified scenarios, which generally occurs under bounding stresses and/or significant control set failure. Although this scope of analysis is valuable, it is insufficient to support the calculation of probabilistic safety metrics such as P(LOC), P(LOV), and P(LOM). ISA needs to go beyond the examination of worst-case end states also to address systematically less severe end states. Scenario development requires systematic analysis of complex interactions, dependencies, and combinatorial effects. NASA/SP-2011-3421, Probabilistic Risk assessment Procedures Guide for NASA Managers and Practitioners [8], contains additional guidance on scenario development.

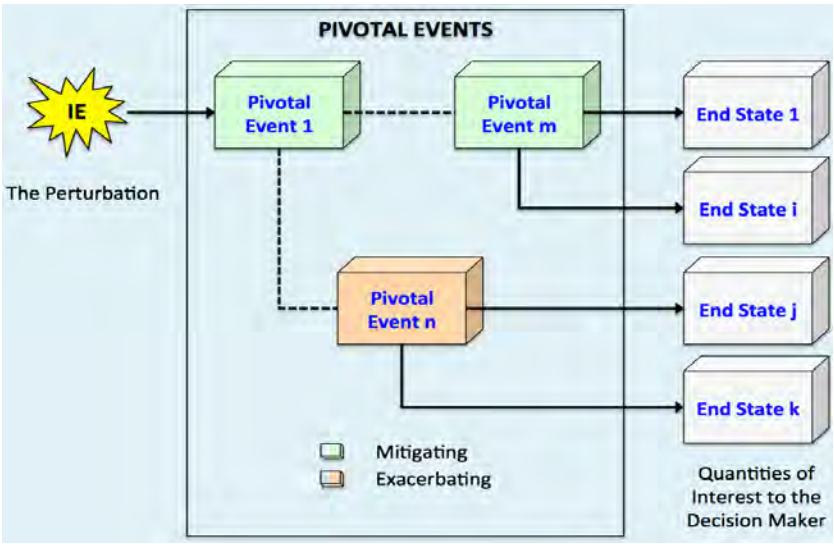


Figure 5-2. The Concept of a Scenario

The primary task of an ISA is the development and analysis of accident scenarios. These scenarios include consideration of accident causes, contributing factors, effectiveness of controls, subsystem interactions, physical responses of the system to the environments it encounters, and the means by which undesirable consequences may be realized. An ISA is a system-level analysis that integrates potentially differing types of safety information and produces results such as the probability of occurrence of each accident scenario and the severity of the consequences. The results may be qualitative, quantitative, or both, depending on the context and degree of ensurance needed.

The comprehensiveness of the ISA depends upon the criticality of the program/project, as will be discussed in Section 5.2.4. For any program/project that is considered to have high criticality, the ISA is designed to be comprehensive with respect to accident scenario types. Potential sources of or contributing factors to accidents include component failures, software errors/failures, human errors, unintended functional behaviors (including unintended performance/absence/magnitude/timing of a function), environmental stresses, intra-system stresses, unintended or improper communication (including electronic, optical, magnetic, and/or mechanical communication), absence of an expected communication, the possible presence of latent failures, and the possible occurrence of common-cause failures.

ISA models also consider uncertainties, dynamics, and dependencies. The analysis of uncertainties leads to probabilistic models and the propagation of probabilistic information. The consideration of dynamics leads to an understanding of changes in boundary conditions, information, and states over time. The consideration of dependencies leads to an understanding of the correlations and causal mechanisms that cause risks to be coupled.

ISAs can, and typically will, use a variety of safety analysis methodologies, such as failure modes and effects analysis (FMEA), hazard and operability (HAZOP) analysis, fault and event tree analysis, probabilistic risk assessment (PRA), bounding analysis, and physics-based simulation, to assess safety performance. Each method has its own domains of application such that the methods should generally be selected with regard to their intended purpose and applied in a coordinated fashion to ensure that all relevant accident types are addressed at an appropriate level of detail. The key characteristic in this regard is that the choice of methodologies used depends mainly upon the characteristics of the system or sub-system being evaluated, including the level of design detail that has been achieved at a particular point in the system life cycle; the type of event being analyzed (e.g., component failures, human errors, environmental stresses, software errors), as well as the skills and preferences of the analyst. For example, FMEA is typically applied to a detailed system design in order to assess the possible effects of individual component failures. (It can also be applied at a functional level to systems of less mature design to assess the effects of functional failures.) Physics-based simulation, on the other hand, can be used to determine the capability of the system to withstand applied loads or to determine whether a particular subsystem can function successfully under accident conditions.

Individual analysis methods are not intended to be conducted as separate, stove-piped, stand-alone system analyses. Rather, the set of selected methods should work synergistically in service of the development of a single, comprehensive accident scenario set that represents the core output of ISA. To that end, the methods should be executed using consistent system data and assumptions, and should be managed in a coordinated fashion to appropriately address the scope of hazards and system interactions at issue. Figure 5-3 notionally illustrates how these different types of analysis methods integrate into a comprehensive analysis of accident scenarios.

The ISA scenario set is quantified to address system-level safety performance measures that pertain to the objectives and requirements of the program/project. The ISA integrates the analysis of each performance measure to the greatest extent possible, recognizing that many accident scenarios impact multiple performance measures, as shown in Figure 5-4.

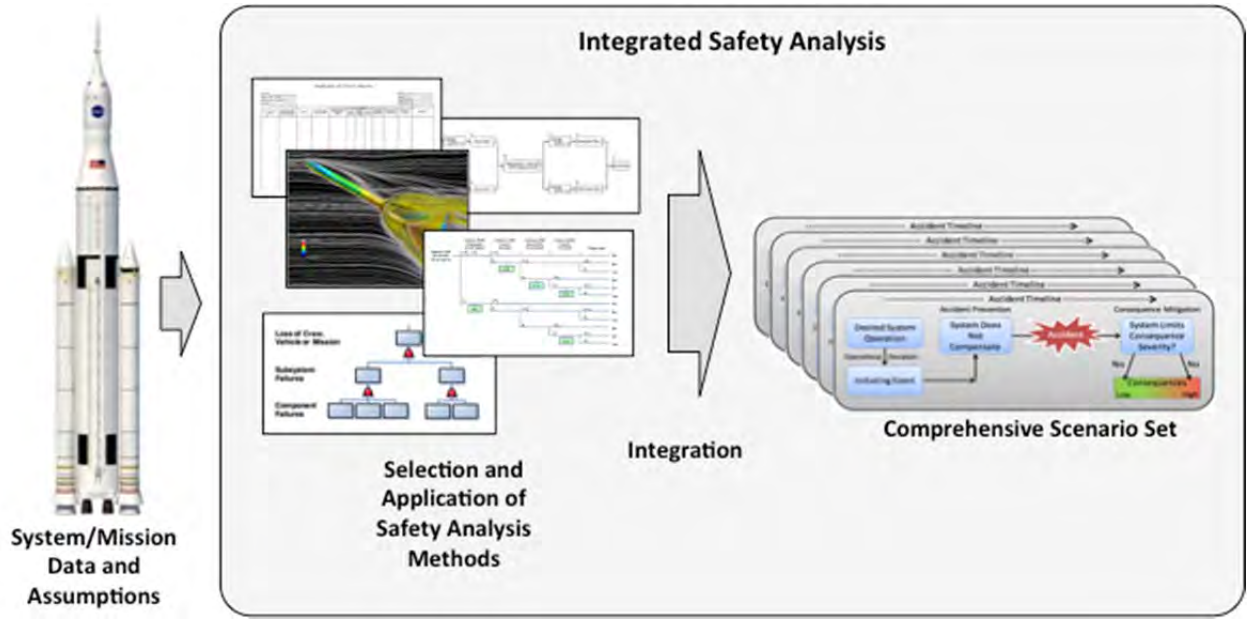


Figure 5-3. Integration of System Safety Analysis Methods in ISA

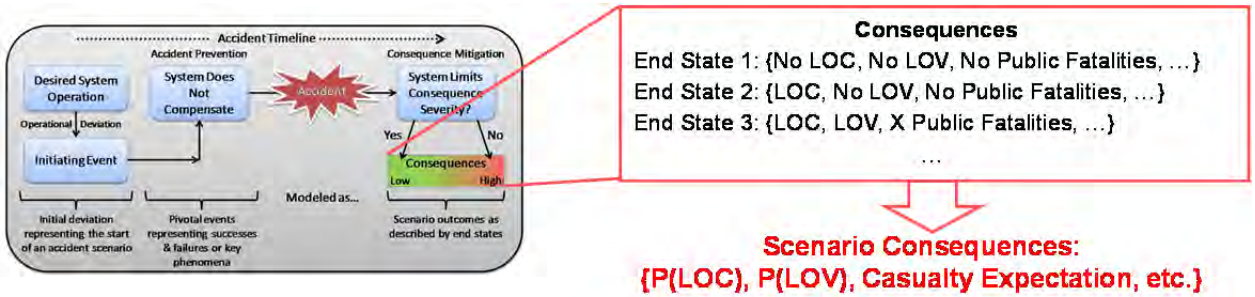


Figure 5-4. The Potential Consequences of an Accident Scenario May Impact Several Safety Objectives or Requirements

An ISA is a part of Risk Management. Therefore, an ISA is integrated into the analyses of performance in other mission execution domains, including technical performance, cost, and schedule. The integration across mission execution domains assures that the system configuration and the analysis assumptions are treated in a common and consistent manner and that the dependencies among the consequences are appropriately correlated (e.g., lower safety risk may be correlated with higher cost).

ISA should be led by a single, system-level, organizational entity that has ownership of and responsibility for the ISA, and that is a customer for supporting subsystem analyses.

An ISA is tailored to support the RISC at the particular phase in the life cycle at which it is conducted. Thus, as the system design evolves, the ISA is likewise evolved and kept current, typically through the use of progressively more rigorous analysis techniques that model the system at progressively finer levels of detail. The ISA is maintained during system realization so that it can be used to inform decisions related to the as-built system, and during system operation in order to reflect design and operational modifications and accumulating knowledge/experience (e.g., precursors and anomalies).

Examples of ISAs, or analyses that illustrate ISA principles, that are applicable during different phases during the life cycle include the Exploration Systems Architecture Study (ESAS), which was developed for concept studies; the Simulation Assisted Risk Assessment (SARA) study for Launch Vehicle Design,

which was developed for conceptual design; the Crew Exploration Vehicle (CEV) Smart Buyer study, which was developed for preliminary design; and the PRA for the International Space Station, which was developed for the operations and sustainment phase.

Various ways in which an ISA may support the RISC can be found in Table 7-3 in Section 7.3.2. In particular, the column titled "Typical Evidence from Provider" makes various references to analyses that are a part of the ISA, especially with respect to the rows designated as Index 2.1.1, 2.2.1, 2.2.2, 3.1.1, 3.1.2, 3.2.1, and 3.3.1.

5.2.4 Implementing a Graded Analysis Approach

While Section 1.1 states that System Safety includes "specific emphasis on the conduct of integrated safety analysis (ISA) as a fundamental means by which systems engineering and risk management decisions are risk-informed," a key to the successful performance of an ISA within budgetary and time constraints is the use of a graded approach. For example, the expectation of rigor in modeling and analysis would be much higher for missions that are critical to NASA than for missions that are not. In fact, routine missions for which there is already a storehouse of system failure data may not require more than a simple analysis of the failure rate from the data.

An ISA consists of an integration of various individual models. Within an integrated safety analysis, some individual models may require more rigor than others. Generally speaking, the amount of rigor to be applied within the individual models and in the integration process depends on three factors:

- The criticality of the mission and its importance to NASA
- The importance of the concern or scenario being investigated by the individual model relative to the ability to demonstrate that the system meets the minimum level of safety
- The level of detail that has been achieved in the design being investigated

Criticality of the Mission

The criticality of a mission and its importance to NASA are addressed in several NASA NPRs, including (in order of date of latest update):

- NPR 8705.5A, Technical Probabilistic Risk Assessment (PRA) Procedures for Safety and Mission Success for NASA Programs and Projects, dated June 7, 2010 [70].
- NPR 7120.5E, NASA Space Flight Program and Project Management Requirements, with Changes 1-10 dated August 24, 2012 [34].
- NPR 8715.3C, NASA General Safety Program Requirements, with Change 9 dated February 8, 2013 [6].

Specific requirements pertaining to graded analysis as affected by mission criticality are found in the last of these, wherein project priority rankings are defined as follows:

- Project Priority Ranking 1: Any project that involves human spaceflight, or requires White House approval per PD/NSC-25, or is subject to planetary protection requirements, or has high strategic importance to the Agency, or has life-cycle cost exceeding $1 billion. Priority 1 projects require probabilistic risk assessment (per NPR 8705.5) supported by qualitative system safety analysis.
- Project Priority Ranking 2: Any project that does not have Priority Ranking 1 and has a life-cycle cost between $250 million and $1 billion. Priority 2 projects require qualitative system safety analysis supplemented by probabilistic risk assessment where appropriate.

- Project Priority Ranking 3: Any project that does not have Priority Ranking 1 or 2 and has a life-cycle cost less than $250 million. Priority 3 projects require qualitative system safety analysis only.

The term "high criticality mission" is taken here to be synonymous with a project that has a project priority ranking of 1. Likewise, medium and low criticality missions correspond to project priority rankings of 2 and 3, respectively.

It is important to make two caveats with respect to NPR 8715.3C. First, in the NPR, a Priority 1 ranking is also recommended for projects that are conducted within a limited launch window, thereby producing a limited time frame for development. A short time frame does increase the importance of acting expeditiously, but it does not necessarily imply that a project has high strategic priority (i.e., that it is important with respect to the achievement of NASA's strategic goals), and it does not necessarily mean that greater rigor should be exercised in the integrated safety analysis. On the contrary, even if the mission is critical in other respects (e.g., human safety and cost), if it has a short time frame the recommended approach is to conduct a preliminary simplified analysis that can be demonstrated to be bounding, as discussed in the Risk Management Handbook [28]. This enables decisions to be made without having to wait for detailed analysis. To compensate for the shortage of rigor, a large safety risk margin is applied consistent with the fact that high time pressures lead to a higher contribution from UU risks[30]. Thus, until such time as a more detailed analysis can be performed befitting the criticality of the mission, a high criticality mission with a short time frame is treated similar to a low criticality mission with respect to analysis rigor, but similar to a high criticality mission with respect to margin.

Second, some relatively low-cost missions (such as transport of supplies and equipment to the ISS) may technically be ranked at Priority 1 according to the project priority rankings in NPR 8715.3C because they involve interactions with onboard crew members during or after docking, but not before. In such cases, risk scenarios that can affect the safety of the onboard crew would be considered critical from the viewpoint of the degree of rigor to be exercised in a graded approach, but those that do not affect the safety of the crew would be considered noncritical from that viewpoint. Differences between the treatment of high criticality scenarios and low criticality scenarios will be discussed later in this section.

Integrated Safety Modeling Characteristics for Missions of Different Criticality

For missions with high criticality and sufficient time window, the integrative qualities of an ISA include:

- Comprehensiveness in the identification of known hazards and risk scenarios that could be significant actors, with particular attention to scenarios that cut across subsystem boundaries
- Common treatment of elements that cut across different models (e.g., risk scenarios, assumptions, inputs, outputs, and uncertainties)
- Treatment of correlations and dependencies between different risk scenarios, inputs, and outputs

For medium or low criticality missions, or for missions with high criticality on short timeframes that require rapid decisions, the degree of completeness, commonality of treatment, and treatment of correlations and dependencies can be relaxed as long as the results obtained can be shown to be conservative relative to what would be obtained if a more rigorous integrated analysis were performed.

With regard to high criticality missions in particular, the biggest challenge for an integrated analysis is to ensure that all potentially important emergent concerns or scenarios (i.e., those that cut across subsystem boundaries and reveal themselves at a system level) have been identified and are included within the integrated models [33]. It is generally recommended that system safety analyses start with models that are based on functional representations of the system to ensure that cross-system interactions are accounted

[30] This point will be demonstrated further in Section 5.5.7.

for. Later, when cross-system interactions are fully understood and the system has reached a sufficient level of maturity, subsystem models can be developed and integrated together in a meaningful way.

In a functional representation, as described in [71-73], each component is modeled as a functional unit that operates on certain energy, material, and signal flows. The components perform certain functions on these flows to transform them from an input to a desired output state. The logic models capture the pathways and timing of these flows. Functional modeling differs from subsystem modeling in that accidents emerge when component malfunctions produce unexpected pathways for flows of energy, mass, and electrical signals at times when the system is vulnerable to departures from the planned pathways. This is different from traditional subsystem models wherein accidents emerge from a propagation of failures.

In addition to facilitating the examination of subsystem interactions, functional models help ensure that inputs to and outputs from different subsystems are treated consistently[31].

Criticality of Individual Safety Risk Scenarios

An ISA starts with analyses designed to identify safety risk scenarios. The process for identifying risk scenarios in general, formulating risk statements, and developing risk scenario diagrams is described in Sections 4.2 and 4.3 of the Risk Management Handbook [28]. The risk statement consists of a condition, a departure, an asset, and a consequence. An example risk statement from the RM Handbook for a particular safety risk scenario is as follows: "Given that [CONDITION] the state of knowledge of Planet X's atmosphere is limited, there is a possibility of [DEPARTURE] unanticipated higher than expected stresses during the aerocapture maneuver at Planet X, impacting the structural integrity of [ASSET] the spacecraft, thereby resulting in [CONSEQUENCE] radioactive contamination of Planet X."

A process for developing a criticality ranking for each risk scenario is described in the RM Handbook (Section 4.3). Two attributes are considered in making this ranking on the basis of strategic importance:

- Safety risk attribute (also known as likelihood-severity attribute): An estimation of the effect of the risk scenario on the ability to demonstrate that the probability of loss is less than the maximum allowable
- Uncertainty attribute: A set of qualitative factors that are correlated with uncertainty in the evaluation of the probability of loss

The safety risk attribute is a measure of the contribution of an individual safety risk scenario to the overall probability of not meeting the quantitative safety requirement for known risks. If, on the basis of a simplified bounding analysis, the inclusion of a particular risk scenario causes the overall loss probability from known risks to change from a value clearly below the requirement to a value above it, the safety risk attribute is ranked as unacceptable (red) for that scenario. Otherwise, it is ranked as marginal (yellow) or acceptable (green). The safety risk attribute can also be applied to collections of risk scenarios. For example, if a minimal set of risk scenarios collectively causes the overall loss probability from known risks to rise above the requirement, the safety risk attribute for the minimal set is ranked as red.[32]

Because the safety risk attribute is evaluated using a simplified bounding analysis, there is a need to account for uncertainties as part of a separate scenario criticality attribute. This attribute is also ranked as red, yellow or green. The qualitative factors used to develop the ranking include the uniqueness of the risk scenario, its cross-cutting character, its complexity, its propagation potential, and its detectability.

[31] More guidance on functional modeling as a means for investigating cross-system interactions will be provided in Section 5.2.6.

[32] The concept of minimal sets of risk issues that cause a safety requirement to be exceeded in safety criticality analysis is similar to the concept of minimal cut sets that cause a system failure to occur in probabilistic risk assessment, see [8].

Details of the analysis process to determine an overall red, yellow, or green ranking for the criticality of each risk scenario are found in the RM Handbook.

Modeling Characteristics for Individual Risk Scenarios of Different Criticality Within Missions of Different Criticality

Although high criticality missions require more rigorous modeling than low criticality missions in general, the level of rigor applied to the modeling of individual risk scenarios within the integrated model should vary according to the criticality of each scenario, the time frame, and the level of design maturity. Table 5-2 and the paragraphs below provide a summary of the type of modeling that should be developed for different variants of these factors.

Table 5-2. Type of Risk Scenario Modeling Used for Different Criticality Conditions

(a) For Different Levels of Mission Criticality and Risk Scenario Criticality

		Mission Criticality		
		Low	**Med.**	**High**
Risk Scenario Criticality	**High**	Bounding	See Table Below	See Table Below
	Med.	Bounding	Bounding	See Table Below
	Low	None	Bounding	Bounding

(b) For Different Levels of Time Frame and Design Maturity when the Mission and Risk Scenario Criticalities are Significant

		Time Frame	
		Short	**Long**
Design Maturity	**High**	Bounding with High Margin	Probabilistic and High Resolution Deterministic
	Low	Bounding with High Margin	Probabilistic and Low Resolution Deterministic

For risk scenarios near the high end of the scale for both medium and high risk scenario criticality, and with a long enough time frame to permit rigorous analyses to be performed, probabilistic models along with deterministic phenomenological models, such as structural and thermal models, may be needed as shown in the Table. When the design is mature, the modeling should be both broad and deep. That is, it should account for all important aspects of the phenomenology and also include a level of resolution that is consistent with the matured level of design detail. When the design is immature, on the other hand, the modeling should be broad but not deep. It should still account for all important aspects of the phenomenology but include a reduced level of resolution commensurate with the level of design detail. Models that are appropriate for different levels of design maturity are summarized in Section 3.2.1.3 of the RM Handbook and in Section 4.3.3.1 of Volume 1 of the SS Handbook.

Quantitative analysis of uncertainty through probabilistic modeling is a key element of the analysis for risk scenarios near the high end of the mission and risk scenario criticality scales with long enough timeframe. In this context, the uncertainties being considered pertain to the characterization of risks that are known and for which there is a means for constructing meaningful uncertainty distributions. Unknown

risks and risks that are underappreciated are also sources of uncertainty, but these sources of uncertainty are not amenable to quantification. They are handled separately by applying margins to the loss probability from known risks.

As described in Section 3.2.2.5 of the RM Handbook, both modeling uncertainty and uncertainty in the inputs to models should be accounted for when analyzing known risk scenarios that have significant mission and risk scenario criticality rankings. The uncertainties in the modeling of the critical risk scenarios are propagated through the aggregate risk model, e.g., via a Monte Carlo process, to obtain the resultant uncertainty distribution for the loss probability from known risks.

For risk scenarios that have low criticality rankings within a high criticality mission, it is not necessary to perform analyses that are broad or deep. To obtain estimates of the loss probability contributed by these risk scenarios, it is usually sufficient to refer to system safety analyses performed for other missions and adapt them as needed to account for differences relative to the present mission. Furthermore, uncertainties do not have to be accounted when analyzing known risk scenarios with low criticality. It is sufficient to use bounding analyses. The same is true of missions with short time frames, although as discussed above, it is necessary to include extra margin in such cases to compensate for the increased time pressures in general and the lack of time available to perform detailed probabilistic analyses in particular.

5.2.5 Integrating Subsystem Analyses into the ISA

The preceding section noted that for high criticality missions, the biggest challenge is to ensure that all knowable, potentially important, emergent risk scenarios (i.e., those that cut across subsystem boundaries and reveal themselves at the system level) have been identified and are included within the integrated models. In an ISA, this means that the analysis must start from a system level model that accounts for potential interactions between subsystems and between the hardware, software, and human elements of the system. Once a system model has been formulated, it becomes apparent where subsystem level models are needed.

There is a rational process for identifying subsystem level models and analyses that are needed to support system level safety assessment. The main steps of that process are as follows:

1. Identify all of the planned functions of the system in terms of intended movements or flows of mass and energy (thermal, electrical, mechanical).

2. Identify all of the potentially significant unplanned functions of the system, in terms of unintended movements or flows of mass and energy.

3. Identify the subsystems that may be involved with each planned and unplanned function of the system, including effects that may be caused by propagation of mass and energy between subsystems.

4. To an extent consistent with the level of design, identify the hardware, software, and human elements that are associated with each planned and unplanned function, including consideration of the potential propagation of mass and energy across the system.

5. Construct system level logic models that capture the cross-system interactions between subsystems and between hardware, software, and human elements.

6. From the system level logic models, construct the framework for system level phenomenological modeling needed to evaluate consequences at the system level.

7. From the system logic and phenomenological modeling framework, construct the framework for subsystem level probabilistic and deterministic modeling and analyses that are needed for the evaluation of probabilities and consequences within the system level models.

Steps 1, 2, 3, and 4 are encapsulated in Figure 5-5, which illustrates three examples of how different initiating events may lead to cross-system interactions that need to be modeled in an ISA:

- In Case (a), an external event initiates a possible sequence of events that transcends subsystems. For example, an MMOD impact or a strong solar flare might cause damage to hardware in multiple subsystems, demanding responses from software to unanticipated environments and, if the responses are harmful, demanding the need for ad-hoc human actions.

- In Case (b), a human error initiates a transcendent event sequence. For example, an uploading of incorrect software from the Mission Control Center may lead to unanticipated demands on software in other subsystems causing hardware failures across subsystems leading to the need for ad-hoc human actions from the crew.

- In Case (c), a software anomaly initiates a transcendent event sequence. For example, an erroneous command that initiates a spurious signal in one subsystem may lead to spurious signals in other subsystems causing multiple effects on hardware in various subsystems.

Directed graphs, as illustrated in Figure 5-5, are a recommended approach to help construct the framework for integrated safety analyses. Directed graphs and event sequence diagrams (a form of directed graph) also facilitate the derivation of risk models (Step 4 in the process). For example, the directed graph for Case (c) is converted to a risk model in Figure 5-6 that involves event trees and fault trees. In this example, the spurious signal generated by a software error in Subsystem A can lead to several propagating effects that cross subsystems, such as the following:

- The spurious signal from A may cause a spurious signal in Subsystem B, which causes outright failure of hardware in Subsystem B.

- The spurious signal from A may cause spurious emission of matter from Subsystem A (such as liquid or foam) which impacts on Subsystem B and causes structural failures in that subsystem.

- The spurious signal from A may cause outright failure of hardware in Subsystem A that propagates to failure of hardware in Subsystem B.

And so forth. As noted, many of the events that describe cross-system interactions are expressed as functional failures rather than component failures. For example, the generation of spurious signals and spurious emissions of matter or energy are functional events that do not require precise descriptions of the hardware or software involved[33].

Subsystem level probabilistic and deterministic models and analyses needed to support the system level model and analysis (Steps 6 and 7 in the process) are implied in Figures 5-5 and 5-6, and in similar figures generated for other initiating events, by the references to subsystems in the figures. For example, each circle in Figure 5-5 that refers to a subsystem implies a subsystem level analysis, as does each fault tree that is identified by a green box in Figure 5-6. The cross-system interactions are identified by the arrows connecting the circles in the directed graph and by the logic in the risk model.

[33] Section 5.2.6 will discuss further the subject of functional analysis and its usefulness for depicting cross-system interactions.

- **Functional Representation: Arrows Represent Flows of Mass or Energy**
- **Operational Representation: Arrows Represent Cause-Effect Relationships**

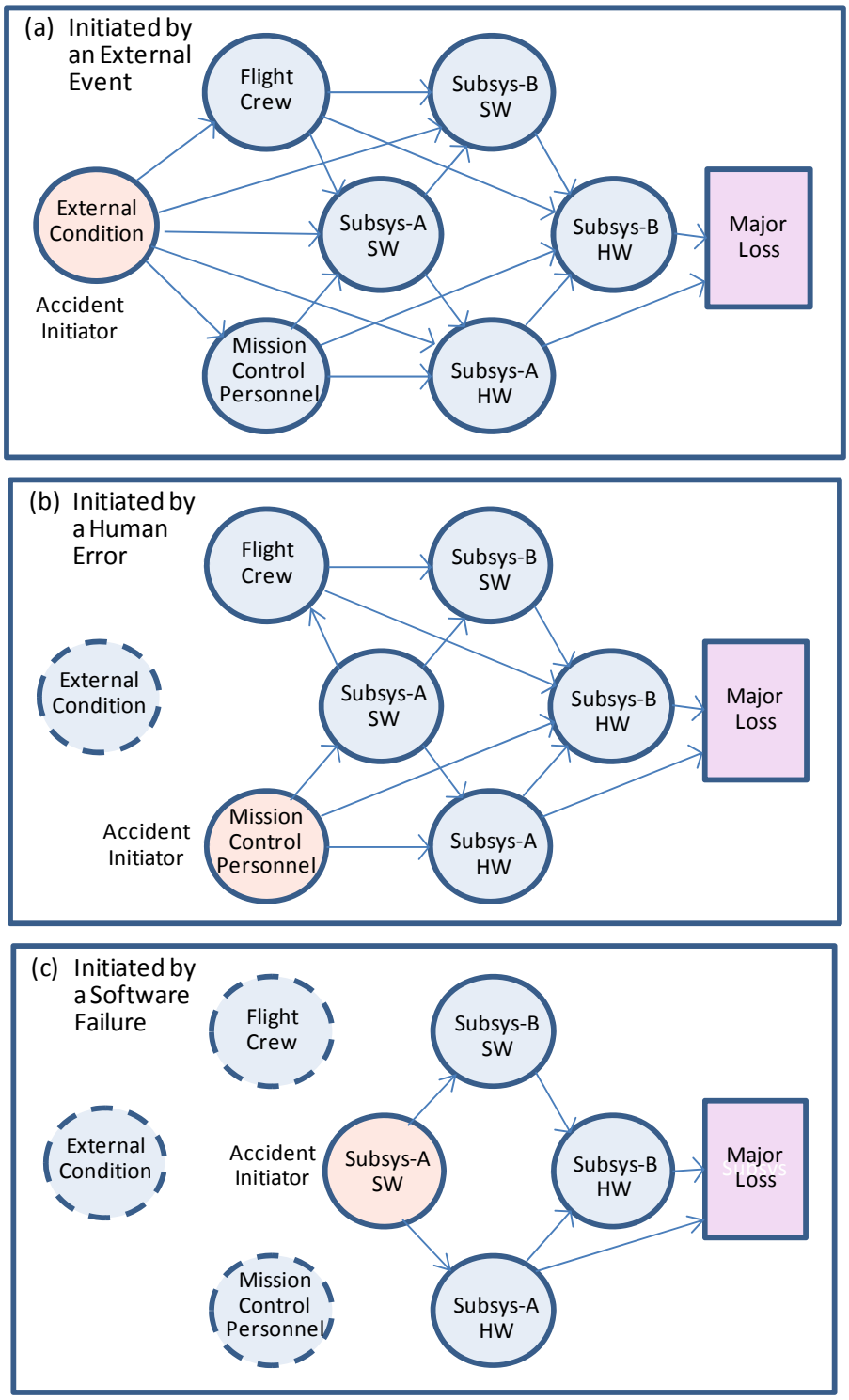


Figure 5-5. Example Directed Graphs of Subsystem-Hardware-Software-Human Interactions to Be Modeled in Integrated Safety Analyses

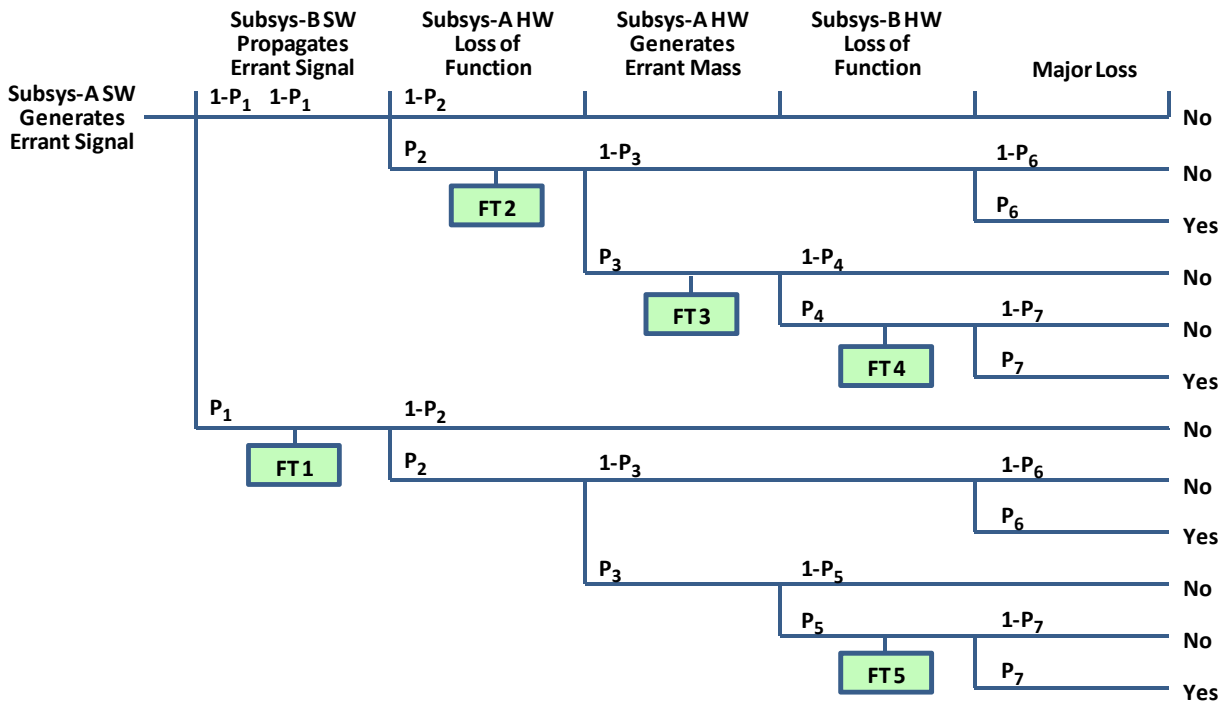


Figure 5-6. Transformation of Directed Graph (c) to Event Tree–Fault Tree Representation

5.2.6 Treating Fault Management in an ISA

Additional Risks Produced by Fault Management Subsystems

Fault management provisions introduce new functionalities within a system. There are signals that are generated and received to detect faults and isolate them; algorithms that are used to process signals, identify causes of faults, predict future states, and determine corrective actions; and additional signals generated and received to initiate corrective actions and track progress. These new functions may introduce new risks. For example, the signals may not occur when intended; they may occur when not intended; they may follow spurious pathways leading to unwanted effects; or the design of the logic in the algorithms may be deficient or incorrect for the situations encountered. For accidents leading to abort, there are still additional risks. Loss of crew may be caused by the inability of the launch abort system (LAS) to withstand the environments produced by a launch vehicle failure (blast wave, high-speed fragments, fireball); by the inability of the crew to survive the environment produced by the LAS (accelerations, heat, depressurization from punctures); or by failure of the crew to survive touchdown/ splashdown.

Methods for Evaluating the Risks

For the reasons cited above, fault management subsystems are often best evaluated using functional analysis techniques. In analysis approaches exercised by Kurtoglu [71], Tumer [72], and others, each component is modeled as a functional unit that operates on certain energy, material, and signal flows. For example, a “combustion chamber” component will have input flows for propellant liquids (material), command signals (signal), output flows for exhaust gases (material), thermal and mechanical energy (energy), and pressure and temperature readings (signal). The components perform certain functions on these flows to transform them from an input to a desired output state. The logic models capture the pathways and timing of these flows. The flow models are frequently dynamic in nature and are integrated with the static risk models that characterize the remainder of the system.

The additional risks associated with crew abort require modeling of the phenomenology of the accidents. Such modeling is concerned with the magnitude and characteristics of the environments produced, including timing and spatial distributions, and their effects on equipment and crew

Integrated Risk Modeling for Fault Management Systems

Various analyses performed within or for NASA illustrate different aspects of the risk modeling.

Mathias et al. [41] describe an integrated method for evaluating the probability of loss of crew given the availability of a launch abort system. Their method has the following features:

- Specific emphasis is placed on the use of physics-based models to characterize the failure environments that pose the greatest threat to the crew: blast overpressure, fragmentation, and thermal radiation environments.
- The entire failure development is modeled using combinations of empirical data, engineering models, and detailed first principle physical models.
- Appropriate analytic techniques are selected through consideration of the failure scenario's overall impact on the integrated system design through risk contribution, sensitivity of the results, uncertainty in the existing knowledge, and complexity of the physics.
- Abort success depends on the warning time, severity of the failure environment, launch abort system, and robustness of the crew module.
- A dynamic PRA model is used to model the time dependence of the abort initiation and execution process.

In their method, many simplifying analysis assumptions are used early in the design process when there is less detail about the design of the system. Once the primary risk drivers are identified via the PRA, they are screened to determine if the results are artificially impacted by the conservatism of the assumptions. If this appears to be the case, the analysis inputs are refined through further analysis of the failure propagation, failure detection, or by decomposing the initiator bins into subsets more representative of actual failures. The process is repeated until the modeling is adequate and the risk representation stems from the physics of the failure and abort process.

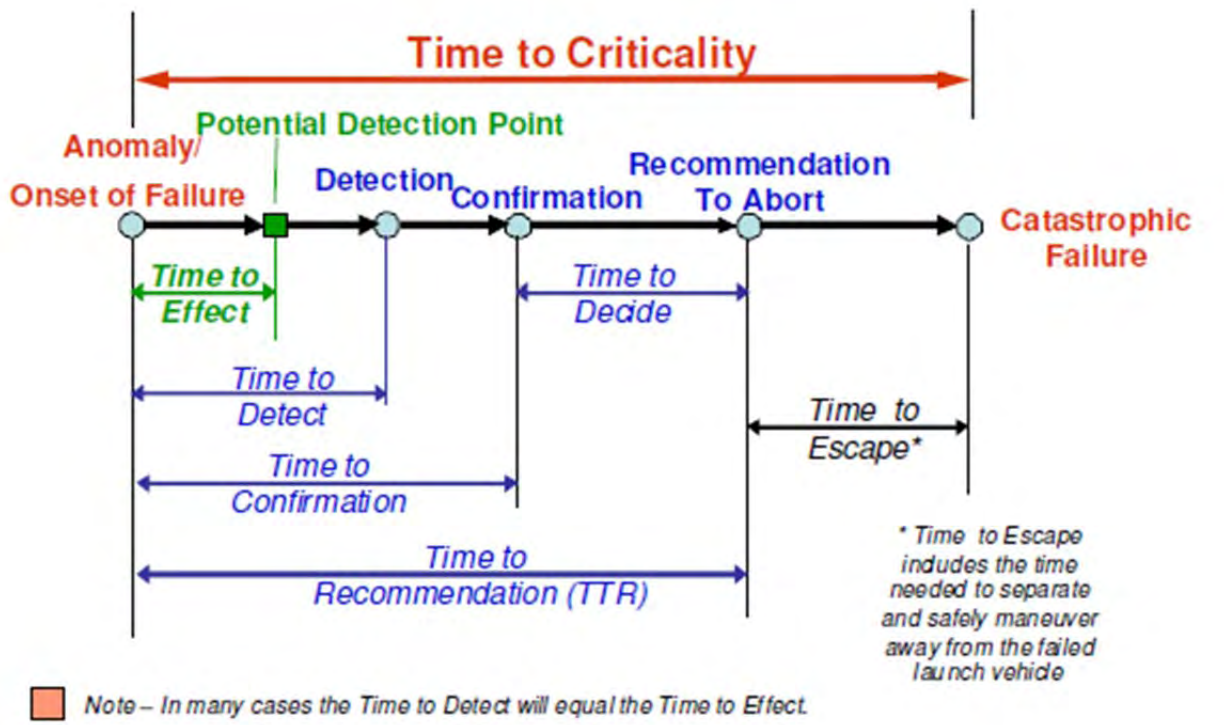
Kurtoglu [71], Tumer [72], and Johnson [73] show how a process called functional fault analysis (FFA) can provide a systems modeling and integration framework that combines information from functional modeling, FMEA, FTA, and failure modes of sub-system components into a single framework. The big advantage of the FFA modeling approach is that it facilitates cross-subsystem thinking by expressing scenarios in terms of flows that cross subsystem boundaries. For example, the following scenario description states how a spurious signal could affect three subsystems, involving both hardware and software, and lead to loss of vehicle and loss of mission:

> ***"A false signal in the landing gear (subsystem 1, hardware) is received by the flight control computers (subsystem 2, hardware) and is incorrectly interpreted by the flight control software (subsystem 2, software) to imply that the spacecraft has landed, thereby resulting in premature termination of the braking engines (subsystem 3, hardware) and crashing of the spacecraft into the planet surface (LOV and LOM)."***

Many will recognize that this scenario is thought to have caused the crash of the Mars Polar Lander into the Martian surface in 1999. An effective fault management system would be one that is capable of recognizing when a signal is spurious and overriding the system's programmed response to the signal.

In addition, the FFA modeling approach provides a capability that accounts for timing information as it relates to functional failures and their propagation. An illustration of the timing elements that affect the

success or failure of a launch abort system is shown in Figure 5-7. Based on the failure mode, the mission phase, and other assumptions relevant to the failure mode (such as size of the leak to be modeled), these times may be based on more or less sophisticated simulations and analyses.



- **Time to Effect** – The time from the initiation of a fault until its effects (symptoms) become potentially detectable.
- **Time to Detect** – The time from initiation of a fault to fault detection. Fault detection means that the system has decided that a sensed anomalous behavior is actually a fault. This includes both the actual sensing of anomalous measurements, and the decision that this anomalous behavior is a real fault and not merely a transient noise reading. This will usually be longer than time-to-effect, as designers may decide that it is better to detect the fault later in the fault propagation path (which generally requires fewer sensors).
- **Time to Criticality** – The time between the initiation of a fault that can cause loss of vehicle or crew until the system loses containment; the time at which a critical fault propagation cannot be stopped.
- **Component Fault Propagation Time** – The time required for anomalous behavior at the input of a component to cause anomalous behavior at the output of that component. The component itself was not the cause of the anomalous behavior, but rather responds to upstream anomalous behavior.

Figure 5-7. Timing Considerations in a Launch Abort Sequence (Source: [71])

In related work, Prassinos, et al. [45], show how the probability of failure of the launch abort system to operate can be modeled, assuming a defined architecture, by using historical failure data. Hansen et al. [46] provide guidelines on how to estimate distributions for abnormal accelerations produced by an out-of-control launch vehicle and for the likelihood of false positives causing the LAS to abort in error.

Graded Analysis Approach

The analysts' decisions on which models to use in evaluating the effect of fault management provisions on the probability of loss is subject to the same considerations as those described in Section 5.2.4 concerning the implementation of a graded analysis approach. As discussed in that section, these decisions are based on the criticality of the mission, the criticality of each scenario, and the timeframe available for analysis.

5.2.7 Integrating Software Analyses into the ISA

Risks Associated with Software

Attention to software risks has strongly increased in recent years as reliance on software has increased [74-76]. For example, the fraction of incident, surprise, and anomaly reports from the JPL planetary missions database (post-Mars-Observer) attributable at least in part to software increased from about 10% at the end of 1996 to about 35% at the end of 2006, three-quarters of which were attributable to ground software. In addition, according to [76], "about half of the losses suffered by NASA in high visibility missions carried out in that time span [the last two decades] have been traced to software faults as a root-cause or as a critically contributing factor." On the other hand, there are ongoing arguments within NASA circles about whether software failures in themselves were a critical contributor to many of the mission-ending failures, or whether most of the failures specifically caused by software were more benign. Either way, it is agreed that the incidence of software failures resulting in some effect on mission success has increased as reliance on software has increased, and so the risk contributed by software failures needs to be evaluated.

The majority of software failures have been related to design or specifications. That is, "the software behaved according to its design, but the design rationale was inappropriate for a particular off-nominal mission condition that had been encountered, which had not been anticipated or had not been sufficiently well understood by the system and software designers." [76]

Other factors affecting the importance of software risks are basically the same as those that affect UU hardware risks, e.g., time and budget pressures, organizational issues, and design complexity [33]. Because software failures usually emanate from faulty design rationale rather than random failures, the correction of software-induced problems has tended to occur on a one-by-one basis by attempting to find and correct each individual cause. Simply adding redundancy has not generally worked and often complicates the problem by making it more difficult to locate the individual causes.

Integrated Risk Modeling for Software and Hardware

One integrated approach for risk modeling of software and its interfaces with hardware has been developed for NASA and is known as the Context-based Software Risk Model (CSRM) [76, 33]. The process is designed for application within the framework of a traditional PRA. It starts from a set of Boolean event-tree /fault-tree models that have been developed (or are being developed) for the non-software portions of a space system. The analyst proceeds to identify the mission-critical software functions in the PRA framework and incorporates them into the risk models.

In the current CSRM approach, the hardware models are component based, but the software models are functional. Typical software functional failure events include the following:

- Redundancy management software error
- Over-correction in a control function
- Under-correction in a control function
- Control set-point set too high
- Control set-point set too low

The software-governed events are analyzed in various levels of detail using static and dynamic models.

The mixing of component-based hardware models with function-based software models is not necessarily a desirable aspect of the integrated modeling. Rather it is a convenience based on the choice to utilize existing hardware models instead of deriving new ones. While this mixing of the modeling may save some time, a more holistic approach would be to utilize the FFA framework discussed in the preceding section for both the hardware and software portions of the problem. The FFA approach is the preferred approach under most circumstances because it facilitates the treatment of cross-cutting scenarios (i.e., that

involve hardware and software contained within various subsystems). Thus, it is more suitable for dealing with the emergent nature of safety concerns.

Data Sources

Several types of software reliability and failure data have been used successfully with these methods. In order of preference, these include:

- Risk-Informed Test Data: Data generated via tests specifically formulated and prioritized in accordance with risk information produced in the course of a software PRA and CSRM application
- System-specific Operational Data: Data collected during software test or operation after the end of any planned and systematic fault-correction process for the same system and software modules that are of interest
- System-specific Debugging Data: Data collected during a process of software test and fault-correction for the same system and software modules that are of interest
- Surrogate Parametric Data: Surrogate data from other systems and operations that are no longer accessible in their original raw state, but have been processed and collapsed into some type of parametric correlation model
- Surrogate Raw Data: Data collected from systems and mission operations judged to be similar to the ones being assessed

Surrogate data should only be used when system specific data are not available, which is generally the case during the concept development and early design phases of a program/project.

Graded Analysis Approach

The amount of rigor applied to the software portion of the model should be graded according to the criticality of the mission, the maturity of the design, and the criticality of the risk scenario being evaluated, much the same as for the hardware portion of the models[34].

Margins

Unknown and underappreciated (UU) risks are as important for software as for hardware, so the need for sufficient safety risk reserves applies to the software side as well as the hardware side

5.2.8 Testing to Support the ISA

Testing is an essential part of model development for the following reasons:

- It establishes much of the basis for validating the models (CAS Factor 2 in Section 5.2.9, Table 5-3).
- It helps establish the input pedigree (CAS Factor 3).
- It serves as a basis for uncertainty quantification (e.g., through a likelihood function) (CAS Factor 4).
- It can help to identify concerns that were heretofore unknown or underappreciated, thus improving results robustness (CAS Factor 5).

[34] Graded analysis not specific to software was discussed in Section 5.2.4.

Testing should be used to provide confidence that the assumptions in the models of the ISA are true and that the overall robustness of the ISA is acceptable. To accomplish this, the tests should be designed to demonstrate the following:

- That each assumption or parameter value used in each model is applicable for each environment where it is assumed to apply
- That each model is applicable for each environment where it is used
- That the integration of models that have been individually verified is applicable for each environment where it is used
- That there is a reasonable margin before failure modes begin to manifest when the environment parameters are allowed to exceed the ranges specified in the environmental requirements document

It is advisable to utilize a validated test planning tool that applies "design of experiments" [77], or an equivalent process, as a means to ensure that the value of testing is optimized for the least cost.

5.2.9 Adhering to the Modeling and Simulation (M&S) Credibility Assessment Scale (CAS)

The credibility assessment scale (CAS) factors, as defined in the Modeling and Simulation (M&S) Standard [78] and Handbook [79], include the following eight attributes:

- Model verification (numerical errors small for all important features)
- Model validation (results agree with real-world data)
- Input pedigree (input data agree with real-world data)
- Results uncertainty (nondeterministic and numerical analysis)
- Results robustness (sensitivity known for most parameters)
- Use history (model is the *de facto* standard for the application in question)
- M&S management (continual process improvement)
- People qualifications (extensive experience in and use of recommended practices for this particular M&S)

For high criticality risk scenarios within high criticality missions, high scores are sought for each of these factors, both for the individual models and for the integrated model. The ranking of the factors is performed on a scale of 0 to 4 using Table 1 of Appendix B of NASA-STD-7009 [78]. The table is reproduced below as Table 5-3, slightly changed to utilize a scale of 1 to 5, in accordance with the ranking scale for assurance deficits to be introduced later[35].

For the individual CAS factors that are underlined in the headings of Table 5-3, the 7009 Standard [78] and Handbook [79] further refine the rankings to include separate rankings for two separate sub-factors: one for evidence and the other for technical review. The evidence sub-factor considers the quality of the evidence, whereas the technical review sub-factor considers the quality of the peer review process. The two sub-factor rankings are each multiplied by a weighting factor, which is selected by the responsible party designated by program and project management and is subject to approval by program/ project management and by the Technical Authority. The weighted sub-factors are summed to yield a single ranking for each CAS factor. To evolve an overall ranking for the M&S credibility assessment, the 7009 Standard [78] and this handbook suggest using the minimum value of the rankings for each CAS factor.

[35] A suggested scale for ranking assurance deficits will be presented in Section 7.2.4.

So, for example, if "Validation" had an evidence ranking of 4 with a weighting factor of 0.6, and a technical review ranking of 3 with a weighting factor of 0.4, the overall ranking for "Validation" would be 4 x 0.6 + 3 x 0.4 = 3.6. If "Results Uncertainty" had an overall ranking of 2.7 and none of the other CAS factors had a lower ranking, the overall ranking for the M&S credibility would be 2.7. We round this value up to 3 for present purposes, to be consistent with the assurance deficit ranking process.

Table 5-3. Ranking Criteria for Credibility Assessment Scale (CAS) Factors

(Factors with a Technical Review subfactor are underlined)

Level	Verification	Validation	Input Pedigree	Results Uncertainty	Results Robustness	Use History	M&S Management	People Qualifications
5	Numerical errors small for all important features.	Results agree with real-world data.	Input data agree with real-world data.	Non-deterministic & numerical analysis.	Sensitivity known for most parameters; key sensitivities identified.	De facto standard.	Continual process improvement.	Extensive experience in and use of recommended practices for this particular M&S.
4	Formal numerical error estimation.	Results agree with experimental data for problems of interest.	Input data agree with experimental data for problems of interest.	Non-deterministic analysis.	Sensitivity known for many parameters.	Previous predictions were later validated by mission data.	Predictable process.	Advanced degree or extensive M&S experience, and recommended practice knowledge.
3	Unit and regression testing of key features.	Results agree with experimental data or other M&S on unit problems.	Input data traceable to formal documentation.	Deterministic analysis or expert opinion.	Sensitivity known for a few parameters.	Used before for critical decisions.	Established process.	Formal M&S training and experience, and recommended practice training.
2	Conceptual and mathematical models verified.	Conceptual and mathematical models agree with simple referents.	Input data traceable to informal documentation.	Qualitative estimates.	Qualitative estimates.	Passes simple tests.	Managed process.	Engineering or science degree.
1	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.	Insufficient evidence.
	M&S Development		**M&S Operations**			**Supporting Evidence**		

In practice, there will be a separate M&S ranking for each model that is used to perform calculations for a critical risk scenario within a critical mission. For purposes of obtaining a single assurance deficit ranking for the base claim that the models used satisfy the M&S Credibility Assessment Scale (CAS) criteria, we recommend adopting the minimum of all the M&S CAS rankings for all the models. The justification is that all the models that undergo CAS examination apply to critical risk scenarios within critical missions, and so all of the CAS rankings apply to essential models.

5.3 Tailored, Derived, and Allocated Requirements

5.3.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities for tailoring, deriving, and allocating requirements are summarized below:

- Along with any proposed tailoring of safety requirements, include an assessment of the effects of the proposed tailoring on the safety performance of the system, using the ISA as a baseline, and a qualitative assessment of the potential effects of the proposed tailoring on the prevalence of unknown and underappreciated sources of safety risk.
- Specify requirements and associated verification procedures on each SCI, sufficient to assure its performance at levels of capability, reliability, and availability that are documented or implicitly assumed in the ISA. Examples of such derived requirements include safety limits, limiting control settings, limiting conditions for operation, administrative controls, safety factors, surveillance and inspection requirements, maintenance requirements, and quality assurance requirements.
- Use the ISA to risk-inform the allocation of safety requirements to assure that they are achievable at a level of confidence that is consistent with the risk tolerance of the Acquirer relative to system-level safety requirements.

Note that in addition to items that are identified through analysis as being critical to safety, SCIs may include support items or items in series whose proper functioning is necessary for the analyzed items to function properly.

The following topics pertaining to these responsibilities and areas of consideration are discussed in Sections 5.3.2 through 5.3.5:

- Tailoring levied requirements
- Developing allocated and derived requirements
- Conducting ongoing negotiation with the Acquirer pertaining to tailored, allocated, and derived requirements
- Considering safety performance requirements and levied engineering requirements as an integrated package

5.3.2 Tailoring Levied Requirements

When the Acquirer levies requirements on the Provider, the Acquirer has already executed a tailoring process to eliminate potential requirements that are deemed to be inapplicable, unneeded, or impracticable from the Acquirer's point of views[36]. Having received those requirements, the Provider will generally perform his/her own tailoring process to argue for waivers on levied requirements that the Provider deems to be inapplicable, unneeded, or impracticable based on the more detailed specification of the design, realization, and operation of the system that the Provider develops. The ultimately agreed-upon set of levied requirements results from negotiations between the Acquirer and the Provider, and the Acquirer exercises approval authority[37]. The process of negotiation, adjustment, and approval can occur at any time during the program/project, whenever the emergence of new conditions warrants a rebaselining of the levied requirements.

[36] This role of the Acquirer in tailoring requirements was discussed in Section 4.3.3.
[37] This negotiation was discussed in Section 4.4.3.

Since the levied requirements emanate from best practices and lessons learned from past experience, it is the responsibility of the Provider in requesting a waiver to substantiate the argument that the best practice or lesson learned is inapplicable, unneeded, or impracticable. Consider, for example, the following best practices cited in the NASA Goddard GOLD Rules [47]:

- ***Gold Rule #1.25:*** *When redundant systems or functions are implemented for risk mitigation, the redundant components, or functional command paths, shall be independent, such that the failure of one component or command path does not affect the other component or command path.*

  Reason for waiving: Both systems or functions might be allowed to have a common component to save on weight and cost if failure of the component were extremely unlikely. For example, a common tank or manifold in a propulsion system might be desirable because separate tanks or manifolds might increase the total number of valves, resulting in a net decrease in safety due to the increased probability of valve leakage.

- ***Gold Rule #1.30:*** *The Attitude Control System (ACS) shall have stability margins of at least 6db for rigid body stability with 30 degrees phase margin, and 12db of gain margin for flexible modes.*

  Reason for waiving: The same bus is being used with fully measured masses and inertias, along with the same algorithm and ACS hardware, flying in the same environment, so that there is much less uncertainty in the system dynamics.

- ***Gold Rule #4.11:*** *Mechanical environmental testing (sine, random, & acoustic, shock, etc.) of flight hardware shall be performed with the test article in the flight like configuration. Mechanisms are configured for flight, and the flight or flight-like blankets and harness shall be present for test.*

  Reason for waiving: Suppose the need to perform both mechanical environmental testing of flight hardware in a flight-like configuration and system end-to-end testing using actual flight hardware caused the schedule for launch to slip beyond the decision maker's tolerance because of fabrication time or acquisition time for the necessary hardware. The mechanical testing could be curtailed or eliminated based on the argument that its purpose is adequately served by the system testing.

- ***Gold Rule #4.23:*** *A life test shall be conducted, within representative operational environments, to at least 2x expected life for all repetitive motion devices with a goal of completing 1x expected life by CDR*

  Reason for waiving: Life testing for 2x expected life might be impossible for devices that have to last for a decade or more (e.g., the ion thrusters for the Dawn mission)

The NASA waiver process is described in NPR 7120.5E [34], which states:

- "The request for relief from a requirement includes the rationale, a risk evaluation, and reference to all material that provides the justification supporting acceptance." (NPR Section 3.5)
- "Provide a rationale consistent with program characteristics such as scope, complexity, visibility, cost, safety, and acceptable risk." (NPR Appendix D)

NPR 7120.5E also states that for levied requirements that have been retained and not waived, the Provider should provide evidence of such retention in the form of references to design drawings, test plans, program/project management plans, etc.

The request for a waiver by the Provider should be backed up by an assessment of the effect of the waiver on the ability to meet the overall top-level objectives of maintaining the minimum required level of safety and being as safe as reasonably practicable. For practical purposes, this assessment can be based on

quantitative bounding estimates or on qualitative arguments. The principal arguments should provide confidence that the following statements would remain true if the waiver request were accepted:

- The resultant value of the loss probability from known risks for each key mission objective would remain within the appropriate probabilistic requirement established in Section 4.2.
- The resultant value of the total loss probability (known risks plus margin) for each key mission objective would remain within the appropriate threshold commitment established in Section 4.1.4.
- The relative increase in the loss probability from known risks for each key mission objective would be small enough compared to savings in cost, schedule, and/or improvements in technical performance, to justify the waiver or modification from an ASARP point of view.
- The relative increase in the total loss probability (known risks plus margin) for each key mission objective would be small enough compared to savings in cost, schedule, and/or improvements in technical performance, to justify the waiver or modification from an ASARP point of view.

If these statements can be justified, then the Acquirer can be confident that the waiver does not cause an inordinate concern about either known or UU risks.

5.3.3 Developing Allocated and Derived Requirements

This section concerns lower level requirements that are allocated or derived by the Provider starting from higher level requirements that have been negotiated between the Acquirer and the Provider. The distinction between allocated and derived requirements was discussed in Sections 3.1.2 and 3.1.3 and is repeated below:

- Allocated requirements are quantitative requirements that are apportioned from system or subsystem level to lower levels, where the units of measure remain the same as at the higher level. For example, the overall maximum allowable probability of LOC during abort may be apportioned to the failure probabilities of the abort motor, the jettison motor, the attitude control motor, and other subsystems. The maximum failure probabilities at the lower level are allocated to stay within the maximum failure probability at the higher level.
- Derived requirements may be quantitative or qualitative and are developed at a lower level of a system to implement a higher level requirement. Derived requirements arise from constraints, consideration of issues implied but not explicitly stated in the Acquirer's requirements, or factors introduced by the selected architecture or the design. For example, it may be determined that in order for the overall probability of LOC during abort to be less than X, the maximum acceleration during abort must be less than Y. The limit on acceleration is a derived requirement.

As described in the NASA Systems Engineering Handbook [13], the requirements at higher and lower levels may be of several types or categories[38]. The process of allocating or deriving lower level requirements from higher level requirements can apply to many quantitative deterministic requirements. For example, a requirement that the weight of an external tank be less than X may be allocated to lower level requirements on the weights of the H_2 tank, the LOX tank, the external structure, and the instrumentation. Allocation is also customarily applied to probabilistic requirements. From a safety viewpoint, the most important types of probabilistic requirements are as follows:

- Probabilistic safety requirements define maximum allowable probabilities of loss. For example, a requirement that P(LOM) be no greater than 1 in 500 is a quantitative safety requirement (since loss of mission is considered a safety performance metric according to NASA's procedural requirements).

[38] Some of the different types of deterministic engineering requirements were summarized in Section 4.3.

- Reliability requirements define minimum required probabilities of success. For example, a requirement that the main engine fire successfully at least 999 times out of 1000 is a reliability requirement.
- Probabilistic technical performance requirements define how well the system must execute its functions. For example, a requirement that an abort system be effective from ground to LEO for at least 80% of the accidents that would otherwise result in LOC is a probabilistic technical performance requirement.

Allocated Requirements

Volume 1 provided guidance on allocating failure probabilities from higher to lower levels using risk assessment models together with cost and schedule models (see Section 4.4.1 in Volume 1). The process is iterative and involves the following steps:

- Initial failure probabilities for subsystems and major components are assigned using historical experience.
- The cost and schedule impact for achieving each allocated failure probability are determined.
- The lower level failure probabilities, costs, and schedule impacts are entered into the risk and cost/schedule models to perform a bottom-up analysis for the safety, cost, and schedule of the overall system.

The allocations are iterated upon to achieve a system that satisfies the probabilistic requirement for safety and is ASARP.

A similar process for reliability allocation has been adopted by the Federal Aviation Authority (FAA) [80]. Their process is depicted in Figure 5-8.

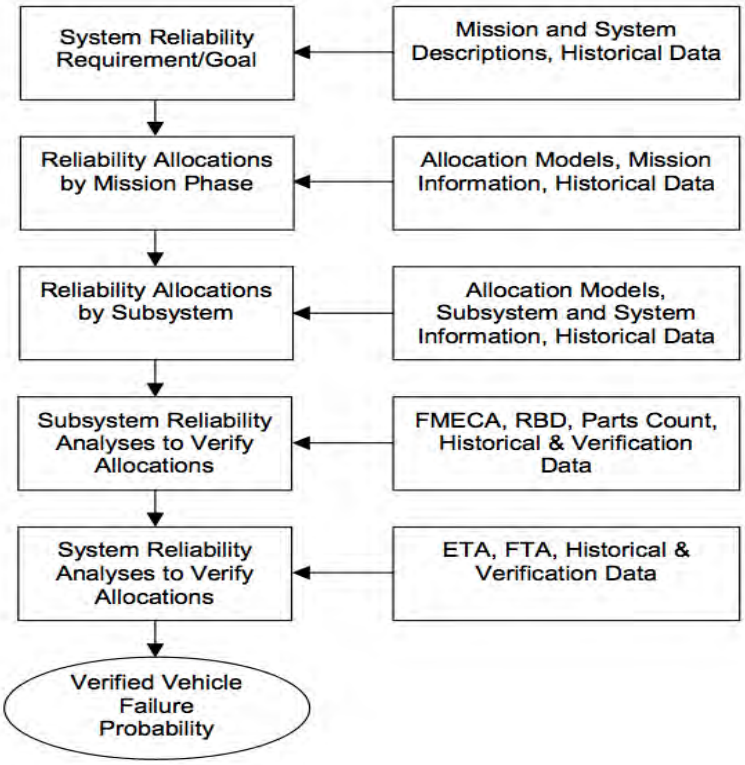


Figure 5-8. FAA Reliability Approach (copied from [80], Fig. 3.2-1)

The steps in the FAA reliability allocation approach, as described in [80], are as follows:

1. Set a vehicle reliability estimate based on a comparison to historical data from previous launches of vehicles developed and launched in similar circumstances.
2. Account for the differences in the mission and system parameters used for this vehicle compared to the previous launches, where possible. (For example, comparing the aerodynamic stresses on the vehicle or whether components with little test experience are being used.)
3. Allocate the vehicle reliability estimate among mission phases based on allocation models using historical data from launches of similar vehicles.
4. Allocate the vehicle reliability estimate among subsystems for each mission phase based on allocation models and historical data.
5. Employ bottom-up subsystem assessments to assess individual contributors to failure, to calculate probabilities when allocations are no longer reasonable (because the subsystems become interdependent or data are simply not available) or when it becomes impossible to propagate failures to a lower level, or to verify that the allocated reliability requirement has been met. Failure Modes, Effects, and Criticality Analyses (FMECA), Reliability Block Diagrams (RBD), and Parts Count analyses are often used. Verification data can also be used to provide additional data support or to disprove the analysis or the allocation.
6. Use system reliability analyses, such as Event Tree Analysis (ETA) and Fault Tree Analysis (FTA), to assist in verifying overall reliability estimates. The system reliability analyses are important. Reliability allocation approaches and bottom-up reliability analysis methods typically assume no interaction between components. In addition, they tend to ignore other systems factors, such as the environment, human interactions, and software, resulting in optimistic estimates of reliability. Therefore, system reliability analyses help ensure that the system safety goals have been achieved. Verification data can also be used to provide additional data supporting or disproving the analysis.

Derived Requirements

The process for deriving requirements is similar to the process for allocating requirements. However, derived requirements may include placing limits on important environmental parameters such as accelerations, temperatures, pressures, and fragment characteristics. They may also include placing limits on important functional parameters such as rates for mass flows and paths for information flows. Because of the fact that constraints on environmental and functional parameters may be included in specification of derived requirements, the logic models may need to be broadened to include phenomenological relationships. In particular:

- The risk models used for purposes of developing derived requirements need to capture the pertinent phenomenological relationships between component failure rates and/or load limits and the environmental and functional parameters mentioned above.
- These phenomenological relationships should include any significant cross-system effects such as the effects of accelerations in one part of the system on other parts of the system.
- The phenomenological relationships should also include any significant hardware-software-human interfaces and interactions.
- The cost and schedule models that are used to ensure that the derived requirements are cost- and schedule-effective need to account for how the placement of constraints on the environmental and functional parameters mentioned above affect cost and schedule.

- Technical performance models should also be used to ensure that technical objectives are met when constraints are placed on the environmental and functional parameters.

<u>The Need for Holistic Risk Modeling in Allocating and Deriving Lower Level Requirements</u>

Derived requirements may include probabilistic requirements that are placed on cross-subsystem functions. For example, in order to achieve an acceptably low probability of LOC, it may be necessary to specify a lower level requirement as follows: "The probability of SRB burn-through impinging on the mainstage shall be less than X." To determine an appropriate value for X, the risk model from which the lower level requirement is derived must reflect the system as a whole. It must in particular account for the likelihood of mainstage impingement given burn-through of the SRB[39].

In addition, when the development of derived requirements introduces constraints on environmental and functional parameters, the effects of such constraints tend to have system-wide effects rather than just local effects. This reinforces the need for a holistic risk modeling approach that includes all important interactions between different subsystems and between the hardware-software-human elements. Figure 5-9, which is similar to Figures 5-5 and 5-6, illustrates how such interactions may unfold and should be accounted for in the use of system-wide risk modeling to derive lower level requirements.

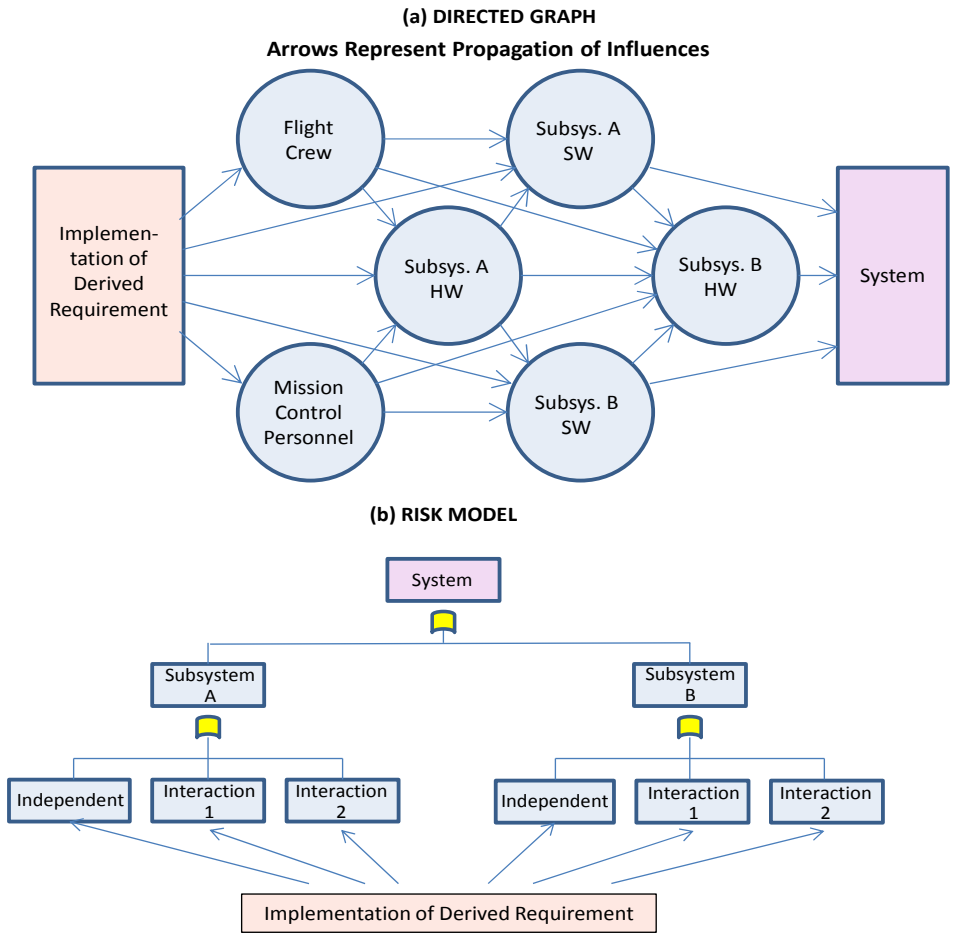


Figure 5-9. Schematic of Subsystem-Hardware-Software-Human Interactions Leading from Implementation of a Derived Requirement to Effects on the System

[39] The development of holistic risk models that include cross-system interactions was discussed in Sections 5.2.5 and 5.2.6.

Accounting for UU Risks

Based on the factors that tend to increase UU risks, as indicated in Section 3.1.1 and Table 4-2, it may be found that certain allocated or derived requirements may affect the amount of margin that needs to be applied to the loss probability from known risks in order to ensure that the total risk, including UU risks, is within the desired threshold or goal. The allocated or derived requirements may result in either an increase or a decrease in the needed margin. Therefore, the models used to allocate or derive lower level requirements need to account for the UU risks as well as known risks[40].

5.3.4 Conducting Ongoing Negotiations with the Acquirer

Negotiations between the Provider and Acquirer were brought up in Sections 3.1.1 and 4.4.3 in the context of rebaselining safety requirements and resetting verification procedures. This section concerns negotiations about allocated and derived requirements, both of which are set by the Provider at lower levels of the system as a means to satisfy the requirements levied by the Acquirer at higher levels.

Lower level allocated and derived requirements have to be reviewed by the Acquirer as part of the requirement development process because there is a chance that one or more of them might conflict with the higher level requirements. For example, a lower level requirement that a particular component be certified to 10 g's might conflict with a higher level requirement that the system as a whole be two-failure tolerant up to accelerations of 15 g's, even though the Provider might not consider the failure of the component in question to violate the failure tolerance requirement. If the Provider's argument for the lower level requirement cannot convince the Acquirer that failure tolerance is maintained up to 15 g's, then the lower level requirements may need to be changed. It is far better for the change to occur during the requirement development phase than later on after design decisions have already been implemented.

Even if the Acquirer accepts the Provider's logic for the allocated or derived requirements, the Provider of course still bears the responsibility of demonstrating through the RISC and the evidence developed for it that the system meets the Acquirer's safety requirements and is acceptably safe.

5.3.5 Considering Safety Performance Requirements and Levied Engineering Requirements as an Integrated Package

As mentioned earlier[41], one of the permissible reasons for waiving a levied requirement or adjusting it to make it less stringent, is that it does not provide an evident or discernible net increase in safety for the present mission. As an example, a requirement that all active parts of a system be two-failure tolerant might be waived or adjusted if it can be shown that it is not providing a discernible reduction in the probability of loss of crew, vehicle, or mission. To ascertain whether this is the case, it is advisable to utilize an integrated risk model in a trade-study fashion wherein, for example, the overall known risk for a comprehensively two-failure tolerant system (Option 1) is compared with the overall known risk for a system in which two-failure tolerance is applied only in areas where it is believed to be effective (Option 2). A result showing that Option 2 provides no discernible safety increase over Option 1 (or possibly even a net safety decrease) would provide a strong argument in favor of waiving two-failure tolerance in the areas that do not contribute to safety. Of course, as part of the argument, it would also have to be demonstrated that the possible benefit of retaining Option 1 as a protection against UU risks would not be significant. This is an example of the need to consider safety performance requirements and levied engineering requirements as an integrated package rather than as individual, unrelated entities.

---

[40] Section 4.1.3 has discussed approaches for estimating and implementing safety performance risk margins to ensure that thresholds and goals are met.

[41] See Section 4.3.3.

5.4 Supporting the System Design

5.4.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in supporting the system design can be summarized as follows:

- Determine how to implement the following design strategies for safety when making design decisions:
  - Elimination of hazards
  - Hazard control (via design for minimum risk, for example[42])
  - Failure tolerance (via design redundancy, for example)
  - System safing (via design provisions to accommodate degraded operation or changed operational objectives)
  - Emergency operations (via design for abort and/or destruct, for example)
- Determine how to prioritize safety performance during system design decision making in accordance with the ASARP principle, using findings from the ISA.
- Designate a set of SCIs, as described in Section 5.2.1.
- Plan for the specified levels of SCI performance (capability, reliability, and availability) to be verified through testing, analysis, and application of standard practices such as those reflected in consensus engineering standards.
- Plan for the design of the system to provide for SCI performance monitoring throughout system operation and sustainment.
- Risk-inform test plans and protocols for the purpose of focusing testing on priority system or subsystem designs, functions, and concerns most affecting safety. (Value of information (VOI) analysis is one method that provides a quantitative basis for decision making regarding information-gathering activities such as testing.)

The following topics pertaining to these responsibilities and areas of consideration are discussed in Sections 5.4.2 through 5.4.5:

- Using RIDM together with historical data to support system design concept decisions
- Applying ASARP principles in combination with the minimum tolerable level of safety performance to support system design improvement decisions

---

[42] The term "design for minimum risk (DFMR)" has a variety of specific meanings within the aerospace industry. The FAA System Safety Handbook [81] defines it as follows: "Design to eliminate risks. If the identified risk cannot be eliminated, reduce it to an acceptable level through design selection." This system-level definition is similar to the ASARP principle of this Handbook. The definition in the Air Force System Safety Handbook [82] is also similar. In contrast, at NASA, DMFR has traditionally been applied at the component level as an alternative to failure tolerance when non-safety considerations such as mass, render failure tolerance impracticable. In this context DFMR refers to (among other things) the inclusion of specific design features that minimize the probability of occurrence of failure modes, such as application of stringent factors of safety or other design margins. This Handbook adheres to the NASA conception of DFMR with respect to design features, but does not restrict its application only to situations where failure tolerance is impracticable. The appropriateness of DFMR in a specific application is a matter to be decided between the Provider and Acquirer.

- Designating and analyzing safety critical items and safety risk drivers as part of system design support
- Minimizing system design complexity

5.4.2 Using RIDM and Historical Data

This section pertains to cases when the Provider is asked to specify the design concept to be pursued, rather than simply implement the design concept that is specified by the Acquirer. The Provider in that case should exercise a RIDM approach that is based on the requirements that relate to RIDM in NPR-8000.4A [4] and the corresponding guidelines and recommendations for implementing them in the RM Handbook [28]. The main steps of the process may be stated as follows:

- Identification of Alternatives – Identify a set of alternative design concepts and the associated concepts of operation that have the potential to satisfy all the commitments requested by the acquirer with regard to each of the mission execution domains (safety, technical performance, cost, and schedule) in a manner that is as safe as reasonably practicable (ASARP).
- Risk Analysis of Alternatives – Develop a set of performance measures covering all the commitments requested by the acquirer and formulate/execute integrated analysis models to determine for each design concept the risk of not satisfying the acquirer's requirements.
- Risk-Informed Alternative Selection – Negotiate with the acquirer to establish risk tolerances for each commitment and provide a case for selecting the alternative that best provides the desired solution within these risk tolerances.

With regard to safety, the main commitment for the Provider is to ensure that the design satisfies the requirement for the probability of loss from known risks provided by the Acquirer as described in Section 4.2 and negotiated between the Acquirer and the Provider as described in Section 4.4.3. It is also important to the Acquirer that the UU risks be contained enough to provide confidence that the threshold for the probability of loss from all risks is also being satisfied at the time of the initial launch, or in other words that the Acquirer's conception of the margin for UU risks is corroborated by the Provider's attention to the factors that control UU risks[43]. Thus, there are two probabilistic safety commitments to be considered:

- The known risk (the part of the total risk that is determined from models and analyses) should be within the decision maker's risk tolerance for satisfying the *requirement* for the probability of loss from known sources as specified for each key mission objective.
- The total risk (from both known and UU sources) at the time of the first flight should be within the decision maker's risk tolerance for satisfying the *threshold* for the probability of loss from all risks, and after a sufficient number of flights have been completed, should be within the decision maker's tolerance for satisfying the *goal* for the probability of loss from all risks.

The two commitments together satisfy the need for the selection of the design concept and management plan to be informed by analysis of the known risks as well as by historical experience pertaining to UU risks.

The second of the two probabilistic commitments can be examined by multiplying the known risk obtained from models and analysis by the appropriate safety performance factor, from Table 4-2 or elsewhere, to obtain an estimate for the total risk that can be compared with the threshold.

[43] These factors were described in Section 3.1.1 and elsewhere.

The following illustration shows how the estimate can rather simply be obtained by multiplying point estimates:

- Suppose that the loss probability from known sources, P_{KNOWN}(Loss) is calculated to be 0.01, that the system is to be developed under significant time pressure, with reliability and safety having equal priority with cost and schedule, and that the design involves new integration concepts and significantly new technology.
- An estimate of the total risk using a safety performance factor of 5 from Table 4-2 would be P_{TOTAL}(Loss) = 0.01 x 5 = 0.05.
- If the safety requirement for the initial launch is that the loss probability from known sources be no greater than P_{REQT}(Loss) = 0.03 and the safety threshold commitment is that the loss probability from all sources to be no greater than P_{THRESH}(Loss) = 0.06, then both are satisfied by this design.
- If, however, the minimum requirement for the loss probability from known sources is P_{REQT}(Loss) = 0.01 and the threshold commitment is P_{THRESH}(Loss) = 0.02, then the requirement for the known risk is satisfied but the commitment for the total risk is not satisfied. In that case it would be incumbent on the Provider to modify the design and/or management plan, or if necessary seek a waiver so that both the requirement and the threshold commitment could be satisfied.

5.4.3 Applying ASARP in Combination with the Minimum Tolerable Level

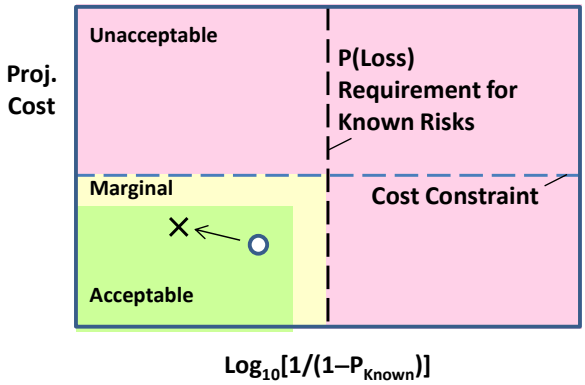
General ASARP Considerations

According to Volume 1 (Appendix B) of this handbook, being "as safe as reasonably practicable," or ASARP, is "a philosophy that safety should be increased as opportunities arise if the impact on cost, schedule, technical performance, or any other domain of interest to NASA is reasonable and acceptable." ASARP therefore entails a judgment from the decision maker about how much sacrifice in areas other than safety can be tolerated for a given amount of improvement in safety. In making this judgment, the decision maker needs to be presented with quantitative metrics for the increment in safety that results from a proposed change to the system, along with the decrement that this entails in other domains (cost, schedule, and technical). He/she also needs to know where the measure of each metric lies with respect to its associated constraint or requirement, since the decision maker's tolerability for cost, schedule, or technical decrements for the sake of improving safety would be affected by that consideration. Such information might be presented as shown conceptually in Figure 5-10, which depicts the loss probability versus the total project cost. A similar figure could be prepared for loss probability versus launch date, loss probability versus vehicle mass, and loss probability versus any other affected performance metric.

The safety metric in Figure 5-10 is labeled as the probability of loss from known risks. Its value is deemed to be acceptable if it lies within the safety performance requirement for the probability of loss from known risks, derived in the manner of Section 4.2. As mentioned several times earlier, it is important to look not only at P(Loss) from known risks but also at P(Loss) from all risks, known plus UU.

Figure 5-11 (left-hand chart) shows similar results where the abscissa depicts the probability of loss from all risks at the time of the first flight. Its value meets the minimum tolerable level of safety performance if it lies within the safety performance threshold derived as in Section 4.1.4. The total risk in this figure is calculated as the known risk times the safety performance factor, which is derived in the manner of Section 4.5.3. Finally in this sequence, Figure 5-11 (right-hand chart) shows results depicting the probability of loss from all risks when the system has matured after many flights. Its value meets the minimum tolerable level if it lies within the safety performance goal as derived in Section 4.1.4. The total

risk is equal to the known risk, because UU risks are assumed to have been wrung out for a mature system.



Notes:

1. **The use of a logarithmic scale for the abscissa in this figure is a surrogate scale that reflects the user's choice to present the information in a form where both scales vary from zero to infinity. Other scales could also be used.**
2. **Analogous charts would apply for loss probability versus other domain performance measures, such as launch date, vehicle mass, etc.**
3. **The region labeled "marginal" reflects the decision maker's view.**

Figure 5-10. Schematic of the Effect of a Design Change on the Probability of Loss from Known Risks at the Time of the First Flight and on the Project Cost

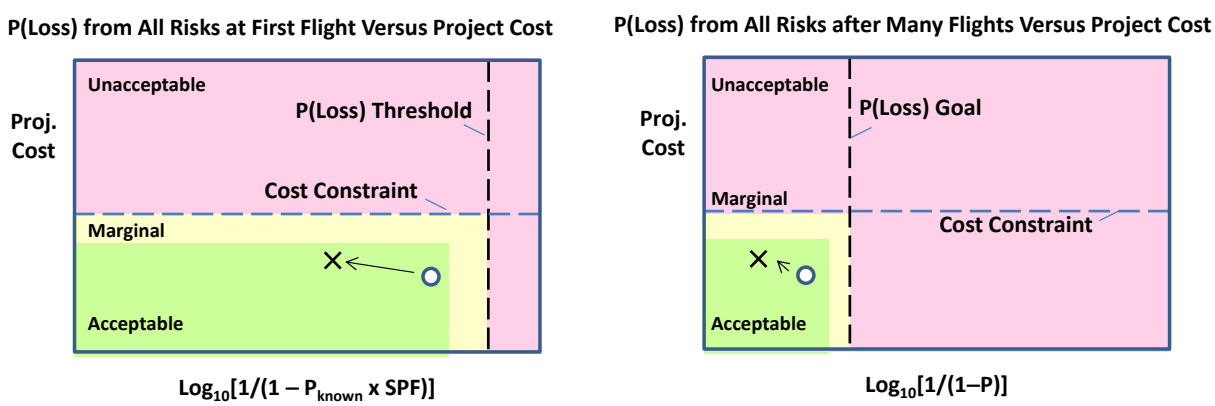


Figure 5-11. Schematic of the Effect of a Design Change on the Probability of Loss from All Risks and on the Project Cost

For all cases shown so far, the principle of ASARP according to the definition in the first paragraph of this section is preserved in going from the present system, denoted by a circle, to a changed system, denoted by an X. In all cases, safety is improved without causing cost to become unacceptable. If this were true as well for the other competing metrics (launch date, vehicle mass, etc.), then the proposed change to the system could be said to be within the ASARP principle.

ASARP in Terms of Joint Confidence Levels and Risk Tolerances

The NASA Cost Estimating Handbook [39] stipulates that the risks of exceeding cost constraints and schedule constraints should be evaluated at a 70% joint confidence level (JCL). This means that the cost and schedule performance is considered to be acceptable only if the joint likelihood of the cost exceeding the cost constraint and the schedule exceeding the schedule constraint is less than 30%. To state it another way, the cost and schedule performance is considered **un**acceptable if the likelihood of exceeding **either** the cost constraint **or** the schedule constraint is greater than 30%. The 30% value is in effect the decision maker's risk tolerance for cost overruns and schedule slippages.

The term "risk tolerance" is defined in the RM Handbook [28], and is basically the same conceptually for cost and schedule as it is for safety risk. Just as cost and schedule performance are considered to be acceptable if the joint likelihood of exceeding the cost and schedule constraints is less than the risk tolerance (30%), safety performance is considered acceptable only if the following arguments can be justified:

- The likelihood that the loss probability from known risks, P_{KNOWN}(Loss), exceeds the loss probability requirement, P_{REQT}(Loss), is less than the decision maker's risk tolerance for exceeding the requirement at all times during the operational timeline.
- The safety performance factor attributable to the factors discussed in Sections 3.1.1 and 4.5.3 is not greater than the value assumed by the Acquirer in deriving the margin between the safety performance requirement and the safety performance threshold in Section 4.1.3.

This interpretation of ASARP is illustrated schematically in Figure 5-12. The RM Handbook provides techniques for propagating uncertainty distributions for input variables to determine the uncertainty distribution for P_{KNOWN}(Loss), the probability of loss from known risks (referred to in the RM Handbook as the aggregate risk). The probability that P_{KNOWN}(Loss) exceeds the requirement is calculated from that distribution.

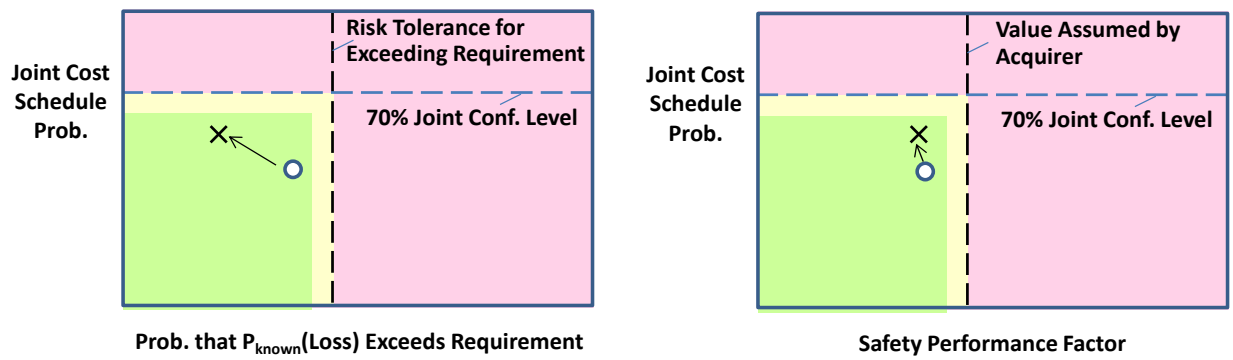


Figure 5-12. Schematic of the Effect of a Design Change on the Probabilities of Exceeding Joint Confidence Levels and Risk Tolerances

ASARP in Terms of Utility Functions

As mentioned above, the literal interpretation of ASARP given in the first paragraph of this section would imply that a proposed change is desirable if safety is improved without causing any other constraint (cost, schedule, mass, etc.) to be violated. From practical considerations, however, a change would not be implemented if the amount of safety improvement was disproportionately small compared to the amount of sacrifice in cost or other metrics of concern, regardless of whether or not the cost and other measures remained acceptable. For this reason, there is another criterion given in Volume 1 (Section 2.1) which provides a slightly different basis for making an ASARP related judgment. It states: "The system is ASARP if an incremental improvement in safety would require a disproportionate deterioration of system performance in other areas." The notion that a "disproportionate" sacrifice should not be made to achieve a small increase in safety introduces the idea of value versus impact into the ASARP implementation.

One way of introducing a value-impact context for ASARP is by using utility functions within a decision analysis framework. A utility function is a statement of the decision maker's preference for different values of a parameter or metric, in relation to other parameters or metrics. In its most general form, a utility function is a multi-dimensional monotonic function of a number of variables, $U(X_1, X_2, \ldots, X_N)$. That is, it either increases or decreases monotonically as each X_i is increased with all the other variables held constant. An example utility function of two variables is shown in Figure 5-13. In this example, X_1 is

a safety performance metric (the total probability of loss including both known and UU risks), and X_2 is a cost metric (the cost expended for reducing the probability of loss). In a higher dimension example, other variables in $U(X_1, X_2, \ldots, X_N)$ might include other cost metrics such as the total cost of the project, other safety performance metrics such as the probability of loss from only known risks, schedule metrics such the time to complete the project, and technical performance metrics such as the amount of data collected. In general, the variables considered in decision making are diverse and not necessarily independent.

To illustrate further, suppose a decision maker is queried and the following equation is derived to represent his/her utility for the two variables cited above (denoted henceforth as P and C instead of X_1 and X_2):

$$U(P,C) = -[3.0 + 2.0 \log P]\left[1 - \frac{C}{\$40\text{M}}\right] \text{ when } C \leq \$20\text{M}$$

$$U(P,C) = -[3.0 + 2.0 \log P]\left[3 - \frac{C}{\$8\text{M}}\right] \text{ when } C \geq \$20\text{M}$$

Figure 5-14 displays this equation, first in the form of U versus P for various values of C, and then in the form of U versus C for various values of P.

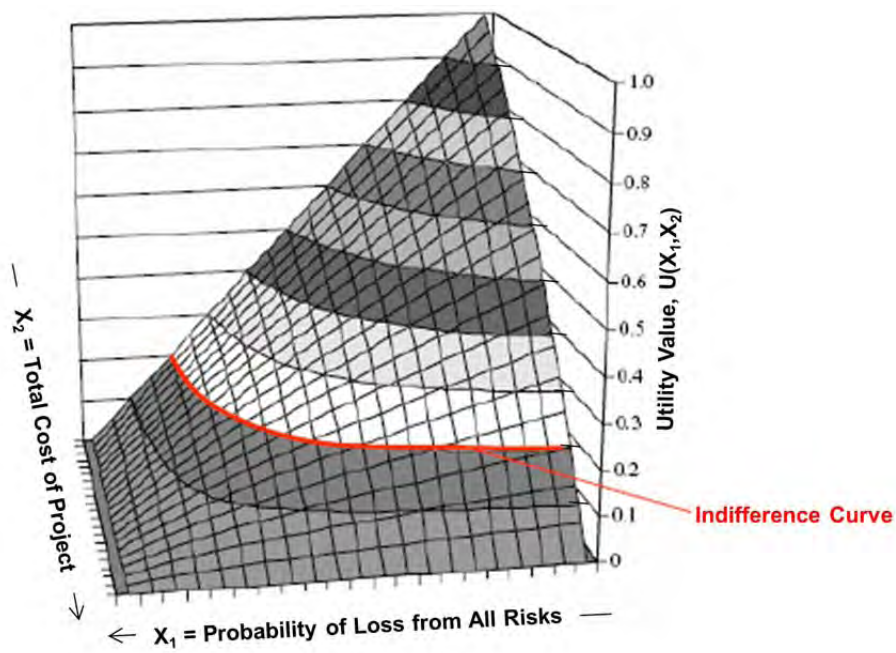


Figure 5-13. An Example Two-Dimensional Utility Function (Adapted from [83])

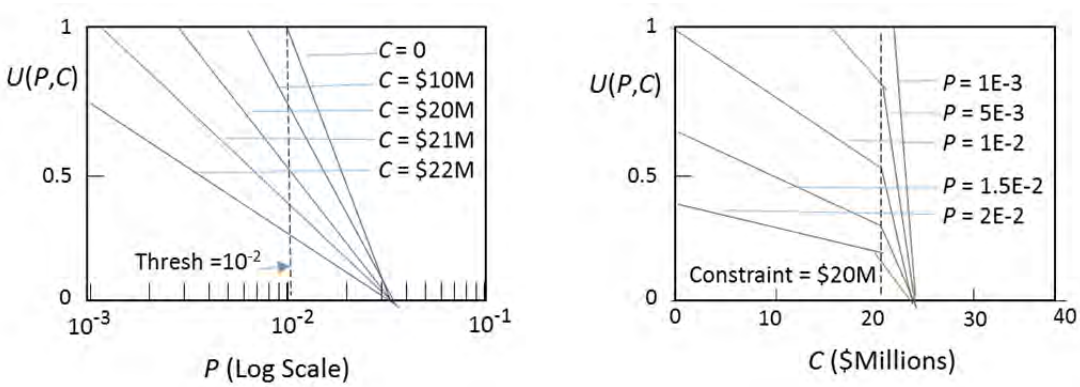


Figure 5-14. Example Contours of a Two-Dimensional Utility Function Used for Illustration

The slopes in a utility function for a performance metric depict the decision maker's comfort factor with respect to changes in the value of the performance metric. In the case of cost in Figure 5-14, for example, the decision maker is only mildly averse to cost increases when the cost is less than the constraint placed on the project for improvements in safety ($20 million), but very averse to cost increases when the cost exceeds the constraint. In the case of total loss probability, the decision maker is more sanguine about exceeding the threshold (1×10^{-2}) as long as it is not exceeded by more than a factor of about 3.

To pursue this illustration further, suppose that according to a current assessment, the expected total probability of loss for an existing system is 2×10^{-2}, which is larger than the threshold value of 1×10^{-2}. Substituting $P = 2 \times 10^{-2}$ and $C = 0$ into the first of the two equations above, the utility to the decision maker for the existing system is $-[3 + 2 \log 2\text{x}10^{-2}]\ [1 - 0] = 0.40$. If an opportunity arises to improve the expected total loss probability from 2×10^{-2} to 1×10^{-2} at a cost of $10 million, the utility to the decision maker increases to $-[3 + 2 \log 1\text{x}10^{-2}]\ [1 - \$10\text{M}/\$40\text{M}] = 0.75$. Since U has increased, a decision to exploit the opportunity at the cost of $10 million would be a positive step toward making the system ASARP.

Suppose, in addition, a different opportunity arises to improve the expected total loss probability from 2×10^{-2} to 1×10^{-3} at a cost of $22 million. Using the second of the two equations, the utility to the decision maker increases to $-[3 + 2 \log 1\text{x}10^{-3}]\ [3 - \$22\text{M}/\$8\text{M}] = 0.75$. Thus, either option represents an improvement over the *status quo*, and the choice between them is a wash.

In general, the variables, X_i, in the utility function are random variables with uncertainty distributions which, as shown in Section 4.5, may be rather large. Thus, in formal decision analysis, an "expected utility" is calculated by integrating the utility function over the probability distributions of the variables. Procedures for doing this may be found in [84]. The value of selecting one alternative action or response over another alternative action or response or over no response is based on the change in the expected utility calculated for each alternative. Integrating over the input probability distributions is a more rigorous process than calculating a utility based on the mean values of the input variables.

5.4.4 Designating and Analyzing Safety Critical Items and Safety Risk Drivers

Most people think of a critical items list (CIL) as being derived from the results of a Failure Modes and Effects Analysis (FMEA). An FMEA is a bottom-up, inductive analytical method that may be performed at either the functional or piece-part level. FMECA extends FMEA by including a criticality analysis, which is used to chart the probability of failure modes against the severity of their consequences. An FMEA matrix may be sorted by severity level to identify critical items and critical failure modes for which controls are desired.

In the context of system safety, critical items have a broader meaning. As discussed in Section 5.2.5, safety is an emergent system-level property, and scenarios that challenge safety frequently involve cross-system interactions between subsystems and between the hardware, software, and human elements. Therefore, critical items for system safety have to evolve from a top-down approach that includes the kind of integrated system modeling that was discussed in Section 5.2.5 and illustrated in Figures 5–4 and 5–5. Safety-critical items (SCIs), then, can include any element or attribute of the system that is critical to safety, including hardware, software, interfaces between hardware and software, the human interface, operating procedures, and management practices.

In addition, the SCIs need to satisfy the risk target at their credited levels of performance, and the one-at-a-time FMEA- or FMECA-based approach does not necessarily achieve that property. The major vehicle for SCI identification is therefore the ISA, which is used to identify the hardware, software, human, operational, and managerial system features upon which safe system operation depends. As mentioned above, such items can be explicit in the ISA (e.g., redundancies, backup systems) or they can be implicit (e.g., assumptions regarding component structural integrity). In either case, designating these items as

safety-critical protects their safety functions by imposing rigorous and highly visible safety management provisions on them.

Expressed in slightly different terms, if a failure tolerance requirement has been levied, it follows that the complement of features needed to satisfy that requirement needs to be included in the set of SCIs. This includes items that were explicitly modeled, and items that are functionally in series with those items, whether or not they were modeled explicitly. Moreover, the set of SCIs includes embedded assumptions about the performance capability of those items. For example, if the failure tolerance requirement is satisfied assuming that one subsystem has the needed capability, then that capability, with suitable margin, becomes an SCI as well. Finally, the levels of reliability and availability credited to these items to address safety requirements become part of the performance baseline for purposes of risk management.

Note that in some domains, safety classification categories (which subsequently lead to groupings of safety critical items) are predefined and are associated with programmatic and quality assurance requirements that are meant to provide assurance of notionally high levels of performance and reliability. Designation of a particular item in that category then implies that it is subject to all such requirements.

In general, the designated SCIs should have the following characteristics:

- High safety function availability (as demonstrated through analysis, testing, or operational experience)
- Safety function verifiability
- High inspectability
- High maintainability

and, in the interest of promoting safety performance, should favor the following preferences:

- Broad coverage of accident scenarios over narrow coverage
- Accident prevention over mitigation
- Passive features over active features
- Engineered features over human factors

In practice, conformance to these characteristics and preferences usually has to be traded off against the need to satisfy constraints on weight and internal space. For example, a passive feature such as a massive heat sink might not be practical from the standpoint of weight and space requirements when compared with an active feature like an electrically operated heat exchanger.

Safety critical items are not the same as safety risk drivers. A safety risk driver is operationally defined in the RM Handbook (Section 4) as being a performance parameter, event, or set of performance parameters and/or events that, when varied over their range of uncertainty, cause the safety performance risk to change from tolerable to intolerable. The term “critical items” is used to denote elements of the system that are critical to the success of the system. These differ from “risk drivers” in the sense that risk drivers are defined by the combination of likelihood and consequence, whereas critical items are identified only by the consequence that could result if the item did not function properly, independent of the probability of that happening.

As discussed in Section 3.2.4, the “S” in SCI pertains to safety in the context of freedom from harm to humans or to the environment, but it may not always be necessary to apply SCIs in the context of safety that pertains to freedom from loss of equipment, property, or mission objectives. The decision maker may determine that safety in the latter context may be assured by attending to risk drivers rather than SCIs. Risk drivers are a subset of SCIs, since in addition to being critical to safety, they have to have a high enough failure probability or probability of occurrence to be significant contributors to the safety

performance risk. SCIs that are not risk drivers have to be continually monitored to ensure that the basis for their not being risk drivers remains intact.

Examples of risk drivers include:

- The reliability of a critical hardware or software component
- Human reliability for a task that is conducted under stressful conditions
- The environment within which the system operates
- The environment produced on a part of the system by another part of the system
- The accuracy or robustness of a model used to determine the risk associated with a type of hazard
- Experimental error for a particular test
- The ability of a test to simulate as-flown conditions

While both safety risk drivers and safety critical items are derived from integrated safety analysis, the former are deduced from risk importance calculations whereas the latter are deduced from sensitivity analyses.

Designation of SCIs and risk drivers represents a solution to an optimization problem: how best to achieve required system-level safety performance, given the constraints and priorities that operate within the given program. Designation of SCIs and risk drivers is not just labeling a collection of components or performance parameters for tracking purposes; it is a considered decision to rely upon certain items for certain levels of performance, and to invest in those items to make sure that the needed performance is attained. It is a system-safety ensurance activity, carried out as part of design within a systems engineering approach to development and deployment of a system, which in turn is executed within a risk management framework. Designation of SCIs and risk drivers is an instance of RIDM, and the subsequent assurance of their performance is an instance of CRM.

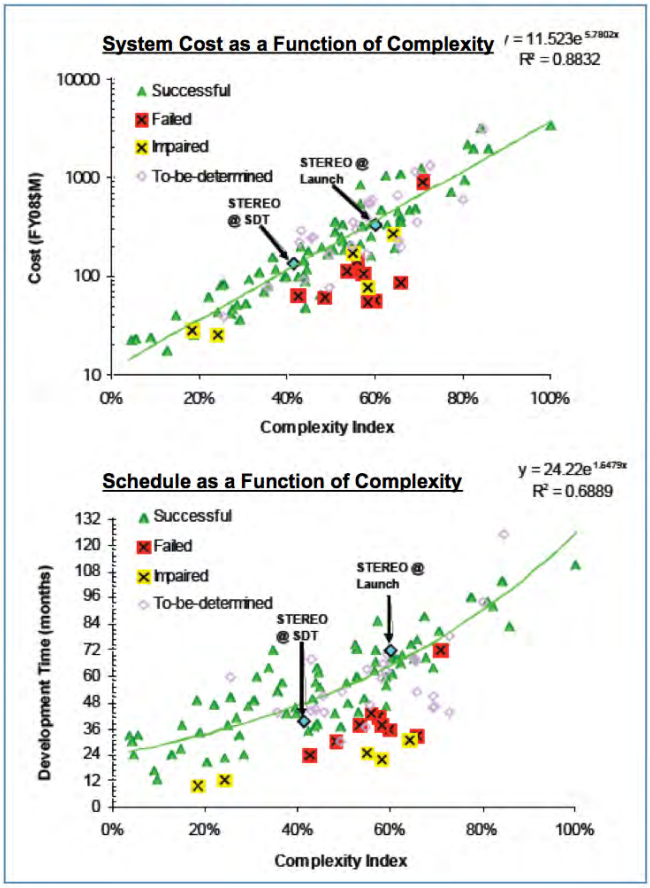
5.4.5 Minimizing Design Complexity

The concept of complexity is a term used by Perrow [23] to mean “baffling, hidden interactions” not anticipated in the original design, that have the potential to “jump” from one subsystem to another. High-risk technologies are complex in that a single component often serves more than one function. Perrow suggests that when a subsystem shares pipes, valves, and feed-lines, and when feedback mechanisms automatically control key processes, accidents are to be expected, even inevitable – and hence ‘normal.’ Moreover, components in different subsystems are often in close operational proximity. If a component fails in one subsystem, the disruption might ‘jump over’ into another subsystem, causing unplanned disruptive consequences. For Perrow, technical systems more prone to failure are complex, tightly coupled systems that make the chain of events leading to a disaster incomprehensible to the operators.

Leveson [64], speaking of software and its interactions with hardware, points out that software “allows us to build systems with a level of complexity and coupling that is beyond our ability to control, where the interactions among the components (often controlled by software) cannot all be planned, understood, anticipated, or guarded against.” She notes that the use of redundancy only makes the problem worse—the added complexity introduced by redundancy has resulted in accidents that otherwise might not have occurred.

Freaner, et al [17], notes a strong correlation between cost and schedule overruns and design complexity. These overruns affect safety by adding pressure to stay within budget and schedule. An example of their results is shown in Figure 5-15.

Actual Cost and Schedule Increase Exponentially with System Complexity



Aerospace Corporation Study

- **The Complexity Index refers to design complexity and is calculated by the Complexity Based Risk Assessment (CoBRA) tool**
- **For example, factors in the Complexity Index for satellite missions include: satellite launch mass, design life, maximum orbit distance from earth, beginning and end of life power, solar array area, cell type, and array/antenna configuration, battery type and capacity, structural material, Attitude Determination & Control System type, number of payload instruments, pointing accuracy and knowledge, number of thrusters, propulsion type, downlink communications band, and maximum data rate, solid state recorder memory, and thermal type**
- **Note that the correlations are both exponential functions, although one chart shows the result on a log-linear scale and the other on a linear-linear scale**
- **Important point: These data do not include very large projects (i.e., those costing over $3 billion)**

Sources: C. Freaner, et al., "An Assessment of the Inherent Optimism in Early Conceptual Designs and its Effect on Cost and Schedule Growth", European Aerospace Cost Engineering Working Group, May 2008. Also, D. Bearden, "Perspectives on NASA Mission Cost and Schedule Performance Trends", NASA Goddard Space Flight Center System Engineering Symposium, June 2008.

Figure 5-15. Effect of System Complexity on Program/Project Cost and Schedule

Skakoon [16] cites a number of guidelines for reducing complexity in mechanical design. Examples include:

- Give priority to purchasing rather than making components.
- Specify components by standards.
- Keep the functions of a design independent from one another by not using the same parts for different functions.
- However, while maintaining functional independence, combine multiple functions into single units when possible to reduce unit count.
- Avoid including functions that are not needed to accomplish the goals of the mission.
- Achieving the simplest system, of course, is not in itself NASA's ultimate goal. It may conflict with higher level Agency goals, such as extensibility (ability to apply the system to other as yet undefined missions). Such higher level goals should be specified in the program/project requirements to ensure that a proper balance between high level goals and simplicity of design is achieved. If not specified as a requirement, however, they should not take precedence over design simplicity.
- The Provider should show how their design achieves all the requirements while avoiding unnecessary functionality and using accepted guidelines for minimizing complexity.

5.5 Maintaining Adequate Safety Performance throughout the Life Cycle

5.5.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in maintaining adequate safety performance throughout the life cycle can be summarized as follows:

- Use the ISA to risk-inform the development of program controls and commitments, which include cost, schedule, and human resources.
- Implement and maintain the safety function(s) of all SCIs at levels of capability, reliability, and availability consistent with the ISA.
- Implement a risk management process in a manner that assures the satisfaction of safety performance requirements defined for future milestones in the system life cycle. This responsibility includes the following objectives:
    - Assure that the system risk management process is resourced and implemented in a manner sufficient to address any safety growth requirements levied by the Acquirer.
    - Assure that the risk of shortfalls relative to safety performance requirements is within the Acquirer's risk tolerance.
- Prioritize safety during instrumentation and performance monitoring decisions making.
- Monitor SCI performance and risk drivers as part of risk management.
- Implement a closed-loop process for identifying test and operational anomalies and assessing their safety risk significance.

An anomaly may be a near-miss or an accident precursor, i.e., an indication of a problem that could recur with more severe consequences. Closed-loop anomaly identification and resolution, e.g., via accident precursor analysis [29], provides a mechanism for managing discrepancies between the Provider's understanding of the behavior of the system, and the actual behavior of the system.

Maintenance of the system's safety performance is a strong function of the performance of the system's SCIs and control of the risk drivers. Maintaining SCI performance at levels consistent with the ISA is essential to the continuing safety of the system, as well as to the continuing validity of the RISC. SCI performance at a level below that which is asserted in the ISA may imply that acquirer safety requirements on the system are no longer being met. Therefore, if monitoring of SCI performance reveals a performance concern with a particular item, it may be necessary to rebaseline allocated performance in order to compensate for that concern. This is part of the "consequence" portion of the risk accepted by the Acquirer in approving the Provider's deliverable.

The following topics pertaining to these responsibilities and areas of consideration are discussed in Sections 5.5.2 through 5.5.8:

- Using CRM to manage emerging risks throughout the life cycle
- Monitoring and managing safety critical items and risk drivers throughout the life cycle
- Monitoring and correcting for anomalies and precursors throughout the life cycle
- Identifying and justifying departures from the plan throughout the life cycle
- Keeping the design within the validated domain throughout the life cycle
- Maintaining realistic budgets and schedules to adequately support safety throughout the life cycle
- Risk-informing support activities that are important to safety

5.5.2 Using CRM to Manage Emerging Risks

The Provider should have a plan for exercising a Continuous Risk Management (CRM) approach that is based on the requirements that relate to CRM in NPR 8000.4A [4] and the corresponding guidelines and recommendations for implementing them in the RM Handbook [28]. Because of the ASARP objective, wherein safety improvements are judged against the impacts to technical performance, cost, and schedule, CRM is executed across all four mission execution domains and all the results are used within the System Safety framework.

The main steps of the CRM process, encapsulated by the "CRM wheel," include the following segments that are repeatedly exercised during the various phases of the program/project as new information comes to light:

- *Identify* new risk scenarios and changes to existing risk scenarios as they emerge within each of the mission execution domains: safety, technical performance, cost, and schedule.
- *Analyze* the new or changed risk scenarios using aggregate risk models to determine how they affect the ability to satisfy each of the performance requirements or commitments in each of the mission execution domains.
- *Plan* responses to these new or changed risk scenarios, selected from among the following disposition options: accepting, mitigating, watching, researching, elevating, and closing.
- *Track* the status of each risk scenario and the resultant aggregate risks as the planned responses are implemented.
- *Control* each risk scenario and the resultant aggregate risks by applying contingencies where needed and by reentering the *Analyze* and *Plan* steps when necessary.
- *Document* and *Communicate* the status of each risk scenario and the resultant aggregate risks.

In the same sense as described for RIDM in Section 5.4.2, the CRM process considers both the P_{KNOWN}(Loss) requirement for known risks and the P_{TOTAL}(Loss) threshold commitment for total risks, and tries to assure that both are satisfied within the decision maker's risk tolerance. The intent of CRM is for UU risks to be converted to known risks as expeditiously as possible as time progresses.

5.5.3 Monitoring and Managing Safety Critical Items and Safety Risk Drivers

Because of their importance to maintaining safety performance, safety-critical items warrant heightened systems engineering, risk management, and safety management attention, in the form of derived requirements on their design, manufacture, maintenance, etc., that maximize their functional reliability and availability, consistent with the ASARP principle. These derived requirements frequently involve more rigorous quality control, maintenance, testing, and inspection procedures[44].

Because of their importance relative to the ability to achieve the probabilistic safety requirements, risk drivers warrant heightened tracking attention within the CRM process. In the event that the probabilistic risk requirements become jeopardized as a result of new information uncovered during tracking, mitigation responses should be implemented. Such responses generally involve the implementation of contingencies developed during the Plan step of CRM[45].

[44] Refer to Section 5.5.8 of this handbook and Section 4.2.4.2 of Volume 1 for more information pertaining to risk-informed prioritization of quality control, maintenance, testing, and inspection activities.
[45] Refer to Sections 4.5 and 4.6 of the RM Handbook for more information on tracking and controlling risk drivers.

5.5.4 Monitoring and Correcting for Anomalies and Precursors

The Provider should have a plan for ensuring that precursors to the breaching of any safety requirement, as well as other anomalies that could degrade the overall safety framework, are:

- Identified
- Evaluated with respect to their potential criticality
- Incorporated into the risk list if deemed important
- Incorporated when appropriate into the ISA models and analyses to update both P_{KNOWN} and P_{TOTAL}
- Tracked and controlled

There should be provisions for updating, or rebaselining, the risk-informed decisions if the precursor or anomaly results in a change to the identification and/or prioritization of risk drivers, and/or a change in the assessment of how well UU factors are being controlled[46]. Decisions to be considered for updating or rebaselining should include:

- Waivers and adjustments to requirements
- Derivation of allocated and derived requirements
- Prioritization of support activities such as QA, testing, training, maintenance, inspections, and supply of parts

5.5.5 Identifying and Justifying Departures from the Plan

The System Safety Management Plan should document the approach for handling expected and unexpected departures from the plan. Such departures could result from external factors, such as:

- Changes in program/project funding requiring a re-scoping of the technical objectives or leading to increased budgetary and schedule pressures

or from internal factors such as:

- New risk scenarios requiring implementation of mitigation
- Changes in existing risk scenarios requiring implementation of a contingency

The approach for formulating and instituting corrections should take account of all the factors that can lead to increased likelihood of UU risks and should try to minimize these factors[47].

5.5.6 Keeping the Design within the Validated Domain

It has been observed that the risk of catastrophic accidents can increase unexpectedly when designs are scaled beyond the knowledge or experience of the designer, even if the amount of departure from existing knowledge or experience seems small [24]. The added risk can result either from scaling up existing satisfactory designs to achieve operational parameters beyond the original design (known as incremental design), or from scaling down existing satisfactory designs usually to save on cost (known as streamlining or fine-tuning). As an example of the latter case, Starbuck and Milliken [85] argue that leading to the Space Shuttle Challenger accident, twenty-four previous successful flights had created such confidence at NASA that they began systematically "fine-tuning" the technology and design of Challenger and its rockets until it "broke." As another example, the switch to a freon-free foam application process for the

[46] Refer to Section 5.3 and its subsections 5.3.2 and 5.3.3 for guidance on risk-informing these decisions based on minimizing both known and UU risks

[47] See Sections 3.1.1 and various sections in Chapter 5 for a discussion of these factors.

Shuttle external tank resulted in an unexpected increase in foam spalling during the next flight, which could have caused an accident like Columbia to have occurred sooner than it did.

The following guidelines should apply when making incremental changes in design, fabrication, or operation:

- Make no assumptions about the robustness of the system when extrapolating beyond current knowledge and experience, even if the extrapolation appears to be small.
- Test the actual affected component(s) with changes incorporated in the as-flown environment and over all functional modes before putting it into operation.
- Make sure that all models used for safety analysis have been revalidated for the actual system in the as-flown environment.

The Provider should give evidence that these guidelines are being followed and that sufficient budget and time exists to support the recommended testing and analysis

5.5.7 Maintaining Realistic Budgets and Schedules

It has been observed many times that catastrophic accidents are much more probable when programs/ projects are beset by excessive budget and schedule pressures. For example, pressures to meet schedules and budget constraints were cited in various reports on the Challenger and Columbia accidents as a principal causative factor. In addition, as mentioned in [27], early launch vehicles based on ballistic missile technology suffered a large number of launch failures that can be attributed to the fact that launch costs and schedules had a higher priority for these early vehicles than launch quality and reliability.

As further evidence of the risk produced by inadequate budgets and schedules, time pressures are recognized as a fundamental reason for high human error rates in every model that is currently used for human reliability analysis (HRA). For example, among the performance shaping factors used in the Cognitive Reliability Error Analysis Method (CREAM) [86], available time is the most critical one. A continuously inadequate availability of time is assessed to result in a factor-of-5 increase in the human error probability for all four types of cognitive activities considered by CREAM: observation, interpretation, planning, and execution. This is especially relevant to present concerns because many of the UU risks that have come to fruition in the space program have involved human errors of one kind or another.

To ensure that budgets and schedules are adequate, sufficient reserves must be employed for both. These reserves should be derived in a manner that is consistent with knowledge gained from past programs/ projects that have experienced budget overruns and schedule slippages. There are a number of NASA references dealing with how to develop realistic margins for budget and schedule, among which are the following:

- C. Freaner, et al., “An Assessment of the Inherent Optimism in Early Conceptual Designs and Its Effect on Cost and Schedule Growth,” European Aerospace Cost Engineering Working Group, May 2008 [17]. See also Chapter 1 of: “Controlling Cost Growth of NASA Earth and Space Science Missions,” National Research Council, the National Academies Press, 2010 [18].
- D. Bearden, “Perspectives on NASA Mission Cost and Schedule Performance Trends”, NASA Goddard Space Flight Center System Engineering Symposium, June 2008 [19]. See also: “A Complexity-Based Risk Assessment of Low-Cost Planetary Missions: When is Mission Too Fast and Too Cheap?” 4th AIAA International Conference on Low-Cost Planetary Missions, JHU/APL, May 2000 [20].
- A. Chmielewski and C. Garner, “How to Calculate Budget Reserve for Your Project,” Presentation at the 6th NASA Project Management Challenge, February 24, 2009 [21].

- "NASA'S Challenges to Meeting Cost, Schedule, and Performance Goals," Rept. IG-12-021, Sept. 27, 2012 [14].
- NASA Advisory Council Meeting: Report of Audit and Finance Committee, Kennedy Space Center, February 5, 2009 [15].

In the event that budgets and/or timeframes are decreased by a large amount outside of NASA's jurisdiction, for example as a result of political decisions, then it may well be necessary to de-scope the programs/projects that are funded from these budgets so that the budgets and schedules remain commensurate with the tasks to be performed. The act of reducing budgets and/or timeframes without providing for a commensurate change in scope will almost always result in an increase in the UU risks. Unless the initial margins in all four mission execution domains are higher than needed, it is usually not possible to stay within the thresholds and constraints for safety, technical performance, cost, and schedule by simply making incremental changes to the system or adding controls.

The Provider should show how the scope of work to be performed is realistic based on the budget and schedule available for each task, and that there are realistic reserves in the budgets and schedules to accommodate unanticipated conditions or events.

5.5.8 Risk-Informing Support Activities

This section of the handbook addresses the use of results from risk analyses to prioritize support activities that affect the overall safety of the system.

There should be a process for applying a risk-informed approach together with use of best practices to prioritize the tasks pursued in performing support activities such as the following:

- Quality assurance and management
- Qualification testing
- Training and certification of crew, mission control personnel, and launch control personnel
- Daily operational support activities during an extended mission (e.g., ISS)
- Maintenance activities
- Inspections and audits
- Sparing provisions (i.e., ordering and stockpiling of spare parts)

Activities in these areas are prioritized so that actions that are important for reducing safety risk are conducted first and most thoroughly. Both known risks and UU risks should be considered in the process of prioritizing activities. The risk-informed process for known risks consists of placing highest priorities on the activities that are associated with risk drivers, as defined in Section 5.4.4 and determined from integrated safety analysis. Ideally, the support activities that affect safety should be prioritized in a manner that collectively best reduces the product of likelihood and consequence for each risk driver.

Minimization of UU risks is also addressed in the overall conduct of the support activities that are important to safety. Part of the purpose of these activities is to ensure that the generic factors that contribute to UU risks are being controlled. The following rules of conduct, if successfully implemented, are known to have a positive effect on reducing UU risks[48]:

- Budgets and schedule are realistic and do not lead to unreasonable pressures.
- Unneeded complexity is being avoided in the design, realization, and operation of the system.

[48] These factors were also discussed in Section 5.1.6.

- New technology and new applications of existing technology are being adequately tested within the larger system before becoming operational.
- Management is promoting a safety culture in which information about safety risks is discussed openly and inclusively between levels of the organization.
- Management oversight is being maintained over distributed sources and suppliers.
- Critical parts and services are readily available when needed.
- There is a process for communicating and correcting the deficiencies uncovered by the inspections and audits.

5.6 Taking Advantage of Emerging Opportunities to Improve Safety Performance

5.6.1 The Provider's Responsibilities and Areas to Address

It was mentioned in Section 2.1.5 that ASARP reflects a mindset that values safety improvement regardless of the current level of safety. This entails making a concerted attempt to seize opportunities as they emerge for improving safety performance above and beyond the minimum requirements. The Provider's responsibility, therefore, is to seek these emerging opportunities throughout the lifecycle, including during system operation and sustainment decision making, and to implement them when practicable in order to improve safety in accordance with the ASARP principle.

The following two topics pertaining to this responsibility are discussed in Section 5.6.2 and 5.6.3:

- Identifying and assessing safety improvement opportunities
- Testing safety improvement interactions with the whole system

5.6.2 Identifying and Assessing Safety Improvement Opportunities

New opportunities for improving safety may arise from various sources, including the following:

- Design improvements enabled by new technology. For example, availability of new lightweight materials with equal or better strength than existing materials may make it possible to add safety features that would not have been possible because of vehicle weight and space limitations.
- Diagnostic improvements enabled by new technology or by new applications of an existing technology. For example, new instrumentation with capability of detecting previously undetectable symptoms may make it possible to initiate fault correction and/or abort procedures more quickly.
- Testing improvements enabled by availability of new test equipment. For example, a new facility capable of more closely simulating extreme flight environments at larger scales than previously may make it possible to conduct testing of integrated systems at as-flown conditions.

Framework for Managing Safety Improvement Opportunities

The management of new safety opportunities should be integrated with the management of risks. Therefore, taking advantage of new opportunities to improve safety is best handled by expanding the framework of the Continuous Risk Management (CRM) plan to become a Continuous Risk and Opportunity Management (CROM) plan. The framework for risk and opportunity management should be integrated because new opportunities frequently evolve from new risks, and new risks are an expected byproduct of new opportunities.

The process for managing opportunities is analogous to the process for managing risks in the following respects:

- The activities in the plan for continuous opportunity management are basically the same as for continuous risk management: Identify, Analyze, Plan, Track, Control, Document and Communicate.
- The Risk Management team becomes a Risk and Opportunity Management team.
- Risk Statements become Risk and Opportunity Statements.
- The planning options are expanded to include: researching the opportunity, implementing the opportunity, elevating the opportunity, and closing consideration of the opportunity.
- All continuous opportunity management activities and reporting procedures are conducted together with the continuous risk management activities and reporting procedures.

Guidance for Managing Safety Improvement Opportunities

The guidance in the RM Handbook for managing risks is in practically all respects relevant to the management of risk and opportunities within this integrated framework. The same guidelines apply to calculating how proposed options for implementing opportunities affect the performance measures:

- The aggregate risks of not meeting performance requirements or commitments are evaluated over all the mission execution domains.
- An integrated safety analysis approach is used.
- The rigor of the analysis is graded according to the criticality of the mission.
- The amount of detail for various parts of the analysis is graded according to the importance of that part of the analysis for making the decision.

The Provider should document their Continuous Risk and Opportunity Management plan and should provide evidence that the guidelines from the RM Handbook are being followed.

5.6.3 Testing Interactions with the Whole System

While the use of heritage technology where possible is a valid means for minimizing the likelihood of UU risks (and has proved particularly effective for the development of launch systems, such as Delta which was derived from Thor), it should not be assumed that if there is a record of success for a heritage technology in one application, that record will carry over for the heritage technology in a new application [87].

The following guidelines should apply for a new technology after it has reached a high technology readiness level (TRL), as well when using a new technology in a new application:

- Conduct integrated system tests with the full-up configuration in the as-flown environment to look for new system interactions.
- If new system interactions are uncovered, perform integrated risk analyses to assess the risk impact of these interactions and develop controls to minimize this impact if needed.
- Make sure that all models used for safety analysis have been revalidated for the full-up system in the as-flown environment.

The Provider should give evidence that these guidelines are being followed and that sufficient budget and time exists to support the recommended testing and analysis.

5.7 Example for Chapter 5 – ASARP Principles Applied to a Proposed Space Shuttle Escape Pod

Statement of the Problem

In this example, the year is 1978 and it is proposed that a 7-person escape pod be added to the Space Shuttle design to protect the astronauts against unknown and underappreciated (UU) risks (see Figure 5-16). The escape pod is to be designed to safely return the crew at any time during launch, assent, orbit, or reentry, The detailed design of the Shuttle has already been developed, and the proposed escape pod would necessitate a modification of the design. However, no prototype hardware has been developed, and so there would not be a need for retrofitting the proposed escape pod into an existing system

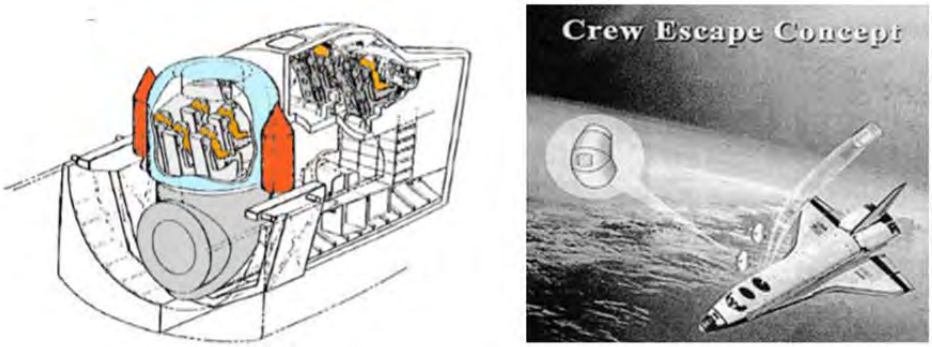


Figure 5-16. Schematic of Proposed Space Shuttle Escape Pod

For this example, we assume a detailed PRA has been performed on the existing design (obviously this is a departure from the actual situation in 1978). We further assume that the detailed PRA is on a level that is consistent with the Integrated Shuttle PRA that was completed after the Columbia accident in 2003. Therefore, we assume that the possibility of accidents similar to the Challenger and Columbia accidents was recognized in the 1978 PRA but the magnitudes of these risks were underappreciated.

Furthermore, we assume an analysis of prior launch vehicle failure frequencies was available, and that it was therefore known that UU risks can be large contributors to the total risk. From that information and the published literature at that time, the influence of organizational and programmatic factors on the potential magnitude of the UU risks was also known, at least qualitatively.

For this example, the safety benefit of including the escape pod is to be weighed against the additional cost of implementing it and the reduction in payload weight that would have to be accepted in order to accommodate the weight of the escape pod.

Background Information Used in the Execution of this Example

NASA had been evaluating Shuttle escape methods even before the fatal breakup of Columbia. After the Columbia accident, it was estimated that the retrofit of a 7-person escape pod capable of surviving both the Challenger and the Columbia accidents would cost $1 billion to $5 billion in 2003 dollars ($1.3 to $8 billion in 2011 dollars) and would take 6.0 to 8.5 years to complete (Houston Chronicle 2004). In comparison, the total cost of the Space Shuttle program in 2011 dollars from conception to retirement has been $196 billion. It was originally estimated that the total cost in 2011 dollars including 135 flights would be about $50 billion (Wikipedia). The escape pod retrofit was considered plausible when NASA envisioned extending the life of the Shuttle fleet until 2020, but the idea was dropped when it was decided to retire the Shuttle fleet by 2010.

Based on estimates for an escape capsule for the Apollo Command Module, the weight of an escape pod is estimated to be about 1650 lb per crew member, or about 11,600 lb for 7 Shuttle crew members (Encyclopedia Astronautica). In comparison, the maximum total payload weight in the final Shuttle design was about 55,000 lb (Wikipedia).

The integrated Shuttle PRA was not completed until prior to the 120th chronological flight. Results from the integrated Shuttle PRA applied retroactively to earlier flights indicated that the total probability of LOC prior to the fifth flight (the first flight without seat ejection capability) was about 0.1, and the probability of LOC prior to the fifth flight that would have been predicted by the integrated Shuttle PRA based on information that was known at the time was about 0.02 (see Figure 4-3 presented earlier). The difference is attributable to underappreciated risks that later manifested themselves in the Challenger and Columbia accidents.

The observation that underappreciated risks were about 5 times as large as appreciated risks is consistent with the programmatic, organizational, and design aspects of the mission: namely, moderately high schedule and cost pressures, top-down decision making process, and complicated design interfaces (see Section 4.5.3 and Table 4-2).

Results of a Hypothetical Safety Performance Evaluation

Suppose that the stated threshold for the first flight is for the total probability of LOC to be less than 0.01 (consistent with current safety threshold guidelines for LEO missions to the ISS). Assume that 80% of the requirement is to be reserved as a safety performance risk margin to accommodate possible unknown and underappreciated risks (consistent with a safety performance factor of 5). This fraction is consistent with the programmatic factors (schedule and cost pressures) and the design interface complexities being similar to those for the Shuttle program, but the top-down decision making process having been replaced by a more participatory one. The associated requirement for the maximum value of P(LOC) from known risks, therefore, is (0.2)(0.01) = 0.002.

Suppose that the PRA without the escape pod included indicates that the predicted probability of LOC accounting for known risks is 0.016 (slightly better than the result for the actual shuttle which was 0.024, see Table 4-1) The predicted value of P(LOC) without the escape pod is thus 8 times the required maximum value of 0.002. Suppose we assume that the addition of an escape pod will reduce the predicted risk by 90%,. to a value of 0.1 x 0.016 = 0.0016 (This 90% reduction is consistent with assumptions that have been made in other analyses for the effectiveness of abort systems, but should be calculated more rigorously for the present application using integrated safety analysis models.) The predicted value of P(LOC) with the escape pod is thus 0.8 times the required maximum value, and therefore the addition of the escape pod satisfies the requirement for P(LOC).

In addition, the Provider claims that there have been improvements in the safety management culture since the time of the Columbia accident, owing to the adoption of a more participatory safety process and better oversight of subcontractor activities. For this reason, the Provider claims that the ratio of total loss probability to known loss probability can be reduced from a factor of 5 to a factor of 3, consistent with Table 4-2. Furthermore, the Provider claims that this improvement in the safety performance factor (denoted in subsequent figures as SPF) applies both to the present design and to the design with the escape pod added, because the modification is being incorporated very early during the design process and there is plenty of time available for testing and validating the capability of the escape pod. Thus, the total loss probability including UU risks is claimed to be around 0.016 x 3 = 0.048 without the escape pod, and 0.0016 x 3 = 0.0048 with the escape pod. The value of P(LOC) with the escape pod including both known and UU risks is 0.48 times the threshold value of 0.01, and therefore the addition of the escape pod satisfies the threshold for P(LOC).

Results of a Hypothetical Cost Evaluation

Suppose that the acceptable total program cost is $200 billion (consistent with the total actualized cost of the Shuttle Program). Suppose also that 67% of the total program cost is to be reserved as a cost margin to accommodate possible unknown and underappreciated risks. (This is consistent with the fact that for the Shuttle program, about 75% of the total cost was underappreciated at the beginning of the program, but some of that would be reduced by virtue of the reduced risk of catastrophic accidents afforded by a more participatory decision making process.) The associated requirement for the calculated, or predicted, total cost is that it be less than (0.33)($200 billion) = $66 billion.

Suppose that a detailed cost analysis has been performed and that the predicted total cost accounting for known cost risks is $50 billion (consistent with the initial total cost estimate for the Shuttle), The predicted total cost without the escape pod is therefore $50B / $66B or about 0.76 times the margin-adjusted requirement.

Suppose that the addition of an escape pod will add a cost of $3 billion (less than the estimate for retrofitting an escape pod system because the escape pod is included in the original design). The predicted total cost with the escape pod is therefore $53 billion, which is 0.80 times the margin-adjusted requirement. Therefore, the addition of the escape pod does not cause the total cost requirement to be exceeded.

Figure 5-17 shows the loss probability from known risks versus the total program cost, and the loss probability from all risks versus total program cost, both without and with the escape pod.

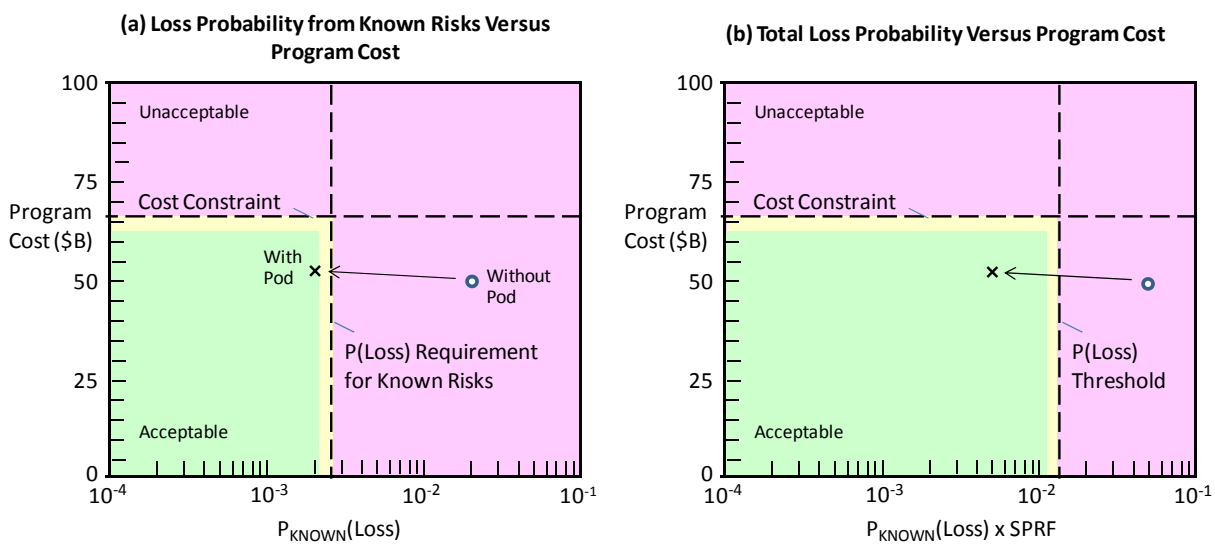


Figure 5-17. Effect of Adding an Escape Pod on Loss Probability and Program Cost

Results of a Hypothetical Payload Weight Evaluation

Suppose it is required by specification that the transport system carry a payload weighing at least 50,000 lb (this number is hypothetical). Suppose that 10% of the payload weight requirement is reserved as a margin to accommodate possible mass growth and contingencies. The associated minimum requirement for the calculated, or predicted, payload weight is (1.1)(50,000 lb) = 55,000 lb. Suppose that the predicted maximum payload weight is 60,000 lb (consistent with the final Shuttle design). The predicted total maximum payload weight without the escape pod is thus 1.091 times the margin-adjusted requirement.

Suppose that the addition of an escape pod will add 16,000 lb (slightly less than the estimate based on the Apollo CM ejection capsule), all of which must be subtracted from the available payload weight. The

predicted maximum payload weight with the escape pod is 60,000 – 16,000 = 44,000 lb, which is 88% of the minimum requirement. Therefore, the addition of the escape pod causes the requirement for payload weight to be violated.

Figure 5-18 shows the loss probability from known risks versus the total payload weight, and the loss probability from all risks versus total payload weight, both without and with the escape pod

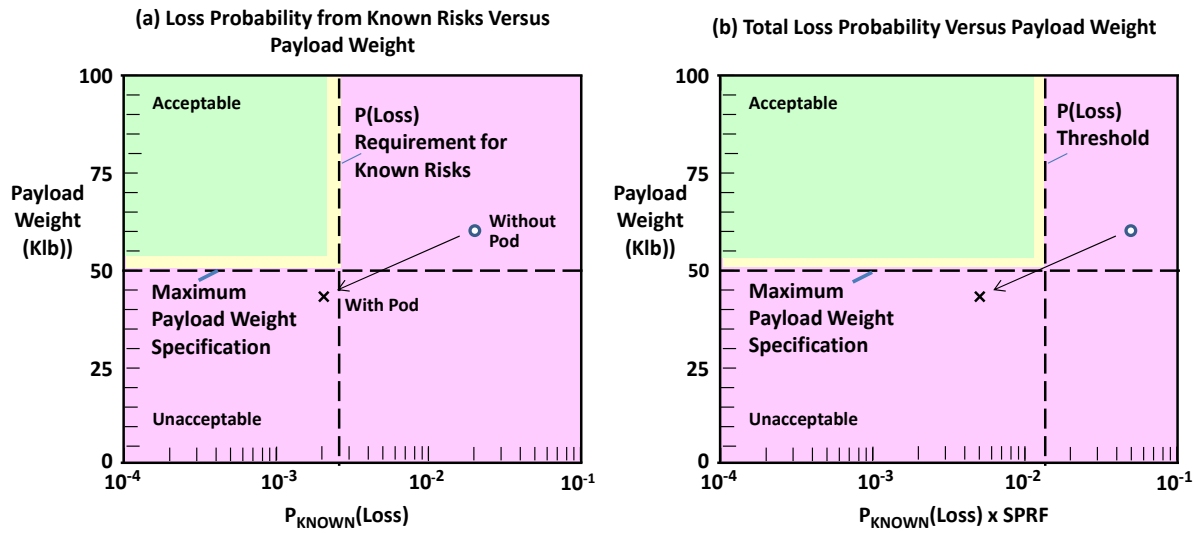


Figure 5-18. Effect of Adding an Escape Pod on Loss Probability and Payload Weight

Analysis

Although the addition of the escape pod reduces the loss probability from an unacceptable to an acceptable value, it also reduces the maximum payload weight from an acceptable to an unacceptable value. The decision maker in the Acquirer's organization has to determine whether a forfeiture in maximum payload weight to a value that is 12% lower than the 50,000 lb specification is "practicable," and if so, whether the forfeiture is more than made up for by the factor of 10 reduction in the loss probability afforded by an escape pod. If not, then the Acquirer will have to find other ways to reduce the loss probability.

6. Developing the Risk-Informed Safety Case (RISC): The Provider's Role

A safety case is commonly defined as "a structured argument, supported by a body of evidence that provides a compelling, comprehensible and valid case that a system is, or will be, adequately safe for a given application in a given environment" [88]. A *Risk-Informed Safety Case* (RISC) is a specialization of the standard safety case, in that it is intended specifically to be part of a deliberative decision making process that is informed by an assessment of the multi-attribute risk to a system (safety, technical, cost, and schedule) and tries to achieve a balance between performance in each of these other areas. Furthermore, a RISC attempts to demonstrate not only that the system is adequately safe, but also that plausible, knowable scenarios that could lead to a risk of system failure have been identified, rigorously analyzed, and conscientiously responded to.

In the context of NASA systems engineering, a RISC refers to the totality of safety-related documentation submitted to a given technical review in the project life cycle. As such, the documentation requirements of the RISC are consistent with the entrance criteria for the relevant review, as itemized in NPR 7123.1B, NASA Systems Engineering Processes and Requirements [3]. Similarly, the criteria for evaluating the adequacy of the RISC are consistent with the corresponding technical review success criteria in the NPR. The RISC addresses each of the operational safety objectives of the system, and includes a roadmap for achieving safety objectives that are applicable to later phases of the system life cycle. The primary focus of the RISC is on hazards that have a significant impact on safety, and while being as comprehensive as possible, should avoid focusing on hazards that have little impact on the safety of the system.

This chapter provides guidance and examples that expand upon the overview for developing and documenting the RISC presented in Section 3.3. The following subjects from that overview are discussed sequentially in Sections 6.1 through 6.5:

- Developing and documenting the RISC
- Assigning responsibilities and integrating the parts
- Exercising a graded approach
- Maintaining and updating the RISC and addressing future life-cycle phases
- Addressing weaknesses, limitations, and significant unresolved safety related issues

6.1 Overall Approach to RISC Development and Documentation

6.1.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in developing and documenting the RISC can be summarized as follows:

- Develop a Risk-Informed Safety Case to clearly argue for the safety of a system to be acquired by NASA. The RISC should explain why there is high confidence that the system is adequately safe and the risks are adequately controlled; provide substantial evidence to support this confidence; and explain actions taken and commitments needed to ensure the further continuance of this high confidence over the life cycle of the system.

- Develop a Risk-Informed Safety Case report. In general, the report should contain the following information: 1) an executive summary; 2) a description of the program, project, and system; 3) documentation of system safety objectives, requirements, policies, regulations, and standards; 4) the safety argument for the system with supporting evidence; 5) a roadmap for ongoing system safety activities; 6) a description of the change and configuration management procedures for the RISC; 7) a description of the audit and review process employed; and 8) a discussion of plans to resolve significant unresolved issues.

- In the case that a system is already in operation, is built but not yet in operation, or is in some stage of development/production without an earlier RISC having been submitted, prepare a RISC report representing the current state of knowledge of the system under development or operation, and update the RISC report for relevant subsequent system reviews.

The RISC report is the means through which the Risk-Informed Safety Case is communicated across Program/Project and between Provider and Acquirer, and provides the material by which the Acquirer may evaluate the RISC. For systems in operation, or at some stage in the development or production process, the RISC should reflect the current state of the system. However, regardless of the state of development/operation of the existing system, a complete and sufficiently comprehensive RISC should be developed particularly for systems of high value to NASA. An understanding of why earlier decisions were made and how effective they were is important for assessing lessons learned.

The following topics pertaining to development and documentation of a RISC are discussed in Section 6.1.2 through 6.1.6:

- Principles for the overall approach to RISC development
- Process for deriving safety claims for the RISC starting from objectives and requirements
- Process for developing evidence for the RISC
- Claims tree structures for the RISC (including optional use of goal structuring notation)
- Documentation of the RISC

6.1.2 Principles for the Overall Approach

According to Volume 1, the approach to demonstrating that a system is safe starts with developing a safety claims tree. The top claim in the tree is that the system is adequately safe, and the process for demonstrating confidence in that claim is based on devising a hierarchy of lower level sub-claims that support the top claim. The tree is developed down to a level where there is sufficient evidence to support the sub-claim at that level. Inference arguments are then provided to support the assertion that satisfaction of the sub-claims at the lowest level leads to satisfaction of the claim at the top level (i.e., that the system is adequately safe).

Processes for rigorously developing claims trees and constructing arguments based on evidence and inference are described by Hawkins [66] and by Denney [89, 90], among others. This volume of the handbook borrows from that work but presents it in a simplified form. We start with the observation that there are two types of claims in the claims tree, which we call “intermediate claims” and “base claims” (consistent with the terms “intermediate events” and “base events” in fault tree notation):

DEFINITION OF INTERMEDIATE AND BASE CLAIMS

- An intermediate claim is a claim that is further decomposed into lower level sub-claims that feed into it. It is demonstrated to be true as asserted to a high degree of confidence by demonstrating that all of the sub-claims feeding into it are true as asserted to a high degree of confidence.
- A base claim is a claim that is not decomposed to lower levels. It is demonstrated to be true as asserted to a high degree of confidence by providing evidence and by showing that all deficits in the evidence that erode confidence in the base claim are sufficiently minimal.

Similarly, there are two types of evidence for each base claim, which we call "direct evidence" and "supporting evidence":

DEFINITION OF DIRECT AND SUPPORTING EVIDENCE

- Direct evidence consists of information that is mostly quantitative and that supports the base claim by showing that the risk of not meeting it is acceptably low. Examples of direct evidence include failure rates from test data or operational experience, analyses of system response to various environments, results of probabilistic risk assessments, analysis of precursors and anomalies, and adherence to best practices.
- Supporting evidence consists of information that is mostly qualitative, provides confidence in the direct evidence, or demonstrates a general responsiveness to safety concerns. Examples of supporting evidence include personnel qualifications, verification and validation of analysis tools, applicability of experiments, quality of documentation, quality of external reviews, effectiveness of communication protocols, and safety culture of the organization.

The strength of the evidence is judged in terms of "assurance deficits" [66], which were defined in Section 5.1.4. In the approach advocated here, assurance deficits are scored by rating the degree to which the assurance deficit sources affect the confidence of achieving the base claim that the evidence pertains to. An example of a possible ranking scale might be as follows:

- Deficit Rank = 1 implies very low assurance deficit, corresponding to confidence of around 95% to 100% that the base claim is justified by the evidence.
- Deficit Rank = 2 implies low assurance deficit, corresponding to confidence of around 85% to 95% that the base claim is justified by the evidence.
- Deficit Rank = 3 implies moderate assurance deficit, corresponding to confidence of around 65% to 85% that the base claim is justified by the evidence.
- Deficit Rank = 4 implies high assurance deficit, corresponding to confidence of around 35% to 65% that the base claim is justified by the evidence.
- Deficit Rank = 5 implies very high assurance deficit, corresponding to confidence of around 0% to 35% that the base claim is justified by the evidence.

The percentile values suggested here are notional. For any particular program or project, the range of confidence percentiles corresponding to each rank would be selected by the Acquirer's decision maker on the basis of the criticality of the mission.[49]

The term "confidence" in the above list is interpreted to be equivalent to "degree of belief." Because the ranking of degrees of belief requires broad knowledge of the system and of the evidence presented to support each base claim, it should be a task that is assigned to highly qualified subject matter experts.

The ideas expressed in this section are illustrated in Figure 6-1, which depicts a conceptual claims tree consisting of intermediate claims and base claims, direct and supporting evidence that feeds into the base claims, and assurance deficits in the demonstration of the claim that emanate from deficits in the evidence[50].

[49] Mission criticalities are discussed in Section 5.2.4.

[50] We have not yet discussed how the ranking of assurance deficits at the base claim level can be propagated up through the tree to infer the resultant degree of confidence that has been demonstrated for the top claim. Processes for accomplishing this will be taken up in Section 7.2.5.

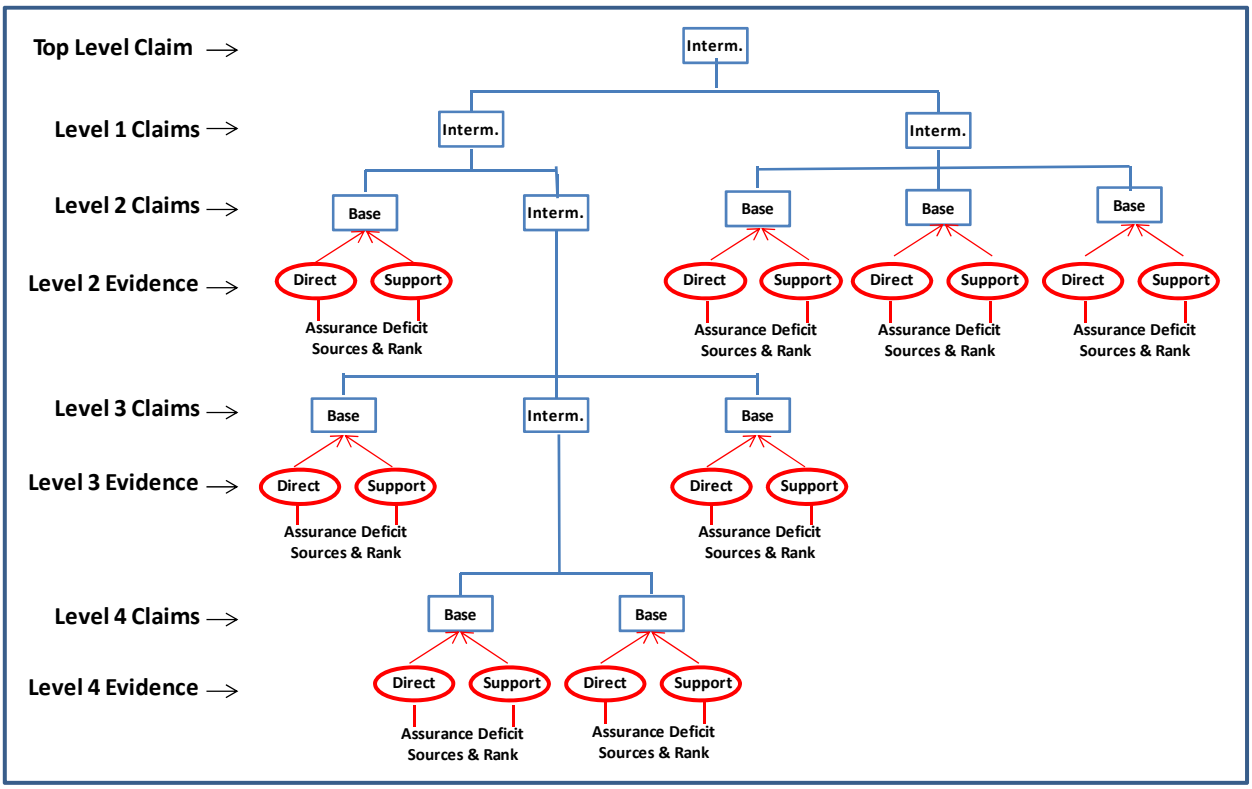


Figure 6-1. Claims Tree Conceptual Diagram

6.1.3 Deriving Safety Claims from Objectives and Requirements

Section 2.1.3 discussed how the top safety objective is to achieve an adequately safe system, and how there are two fundamental principles beneath the top objective:

- Achieve a system that meets or exceeds the minimum tolerable level of safety.
- Achieve a system that is as safe as reasonably practicable (ASARP).

These two principles are further decomposed into specific safety objectives, carried down to a level where they can be clearly addressed by systems engineering processes. This decomposition results in a set of operational safety objectives presented in the form of a generic operational objectives tree in Figure 3-2.

As discussed in the preceding section, the first tasks for the Provider include developing an operational objectives tree that is specific to the mission under consideration and a corresponding mission-specific claims tree that can be supported by evidence. As discussed in Volume 1 of this handbook, the safety claims are developed from the hierarchy of safety objectives and are therefore hierarchical themselves. Assurance that all the claims are true within acceptable risk tolerance limits implies that all of the safety objectives have been satisfied, and therefore that the system is adequately safe.

Because the claims tree is devolved down to the level of base claims that can be validated by evidence, the claims tree will generally be much lengthier than the objectives tree. Each base claim, however, should be traceable to one of the objectives in the operational objectives tree. Because the objectives tree, like the claims tree, is constructed in a top-down manner, a base claim that can be traced to an objective at the lowest level of the objectives tree is traceable to higher level objectives through the branching structure.

Two illustrations of how base level claims may be traceable to the operational objectives are provided in Tables 6-1 and 6-2. The Provider should produce similar mappings that depict the mission-specific objectives and the mission-specific base claims.

Table 6-1. Illustration of the Relationship between Base Claims and Operational Objectives for Claims Pertaining to the Acquirer's Roles for Setting Probabilistic Requirements

Operational Objective	Base Claim(s)
Establish safety performance margins.	A set of key mission objectives (KMOs) has been established that aptly represents all the key mission phases and evolutions where separate safety cases are needed.
	Results from relevant previous missions have been analyzed and the differences between the predicted (known) and actual (total) loss probabilities are understood.
	For each KMO, a margin has been established for the calculated probability of loss for a new system consistent with the understanding of UU risks from relevant previous missions.
	For each KMO, a safety threshold has been developed for the total loss probability that represents an achievable expectation for a new system.
	For each KMO, a safety goal has been developed for the total loss probability that represents an achievable expectation for a mature system.
Establish minimum tolerable levels of safety for known risks.	The requirement for the known loss probability is consistent with the threshold, margin, and goal, and decreases with multiple flights according to a realistic burn-down rate.
	The known loss probability requirement is accompanied by verification procedures according to which the Provider may argue compliance, and according to which the Acquirer may deem the requirement to have been satisfied.
	Provisions have been specified by which the known loss probability requirement and its accompanying verification procedures may be rebaselined in the event that they become unachievable due to the evolution of conditions and risks.

Table 6-2. Illustration of the Relationship between Base Claims and Operational Objectives for Claims Pertaining to the Provider's Roles for Managing Unknown and Underappreciated (UU) Hazards

Operational Objective	Base Claim(s)
Incorporate safety-related best practices into system design.	All relevant best practices for combating UU risks learned from previous missions have been identified and documented.
	The relevant best practices and lessons learned either have been incorporated into the present system or are demonstrated to be not practicable.
	Budgets and schedules are adequate for the tasks to be performed and include realistic reserves.
	Roles and responsibilities for resolving safety concerns are clearly defined in accordance with the program/project and RM plans, provide effective interaction, and resolve problems as they arise.
	Design complexity is minimized without violating safety.
	Any changes from previous validated designs, fabrication methods, and operational procedures remain within the domain for which the performance of the system is known.
	Provisions are made (and documented within the Plan) for testing any new technology or new application of an existing technology within the context of its interactions with the whole system.
Minimize the Introduction of hazards during system realization & operation.	Means are in place for identifying and evaluating any departures from the program/project and RM plans that could increase UU risks and for instituting corrections when needed.
Be responsive to new safety-relevant information.	Management is focused toward continuous safety improvement and inclusiveness in the resolution of risks.

6.1.4 Developing Evidence

As mentioned earlier (Section 6.1.2), direct evidence consists of information that is mostly quantitative and that supports the base claim by showing that the risk of not meeting it is acceptably low. Supporting evidence consists of information that is mostly qualitative, provides confidence in the direct evidence, or demonstrates a general responsiveness to safety concerns. Table 6-3 provides an illustration of the types of direct and supporting evidence that generally are expected to accompany various types of base claims. The Provider should show a similar mapping of the direct and supporting evidence used for each base claim in the RISC.

Table 6-3. Illustration of the Types of Evidence that Pertain to Various Potential Base Claims

Item	Example Base Claim	Example Direct Evidence	Example Supporting Evidence
1	Results from relevant previous missions have been analyzed and the differences between the predicted (known) and actual (total) loss probabilities are understood.	Observed mission failure rates and/or anomaly rates for comparable systems (Anomaly rates, while not directly leading to a catastrophic loss, provide important insights into the relative fraction of occurrences that are attributable to known and fully appreciated risks as opposed to unknown or underappreciated risks. It has been shown that in the past, this fraction tends to be similar for anomalies as for actual catastrophic failures.)	The quality of the records kept for past failures and anomalies, and the experience of the analysts relative to understanding the systems involved in previous missions and the causes of the anomalies and failures
2	For each KMO, a margin has been established for the calculated probability of loss for a new system consistent with the understanding of UU risks from relevant previous missions.	Correlations of historical occurrences of failures and/or anomalies with qualitative factors that tend to produce UU risks. Identification, analysis, and ranking of such factors	The rigor exercised in identifying the governing factors, the quality of the data used to develop the correlations, the standard errors of the correlations, the degree to which the correlations have been verified and validated for relevant applications, and the experience and commitment of the analysts
3	For each key mission objective, a safety threshold has been developed for the total loss probability that represents an achievable expectation for a new system.	1. Documentation of the selected safety threshold for the initial loss probability and the rationale for selecting that value 2. The models and results from a variety of PRAs to show that the threshold is achievable. Could include PRAs for previous systems and missions with additions and modifications to account for differences between the previous and present systems and missions 3. Mission failure and/or anomaly rates for comparable new systems if there is a sufficient amount of applicable data 4. The results of analyses of effectiveness of the launch abort system (in the case of crewed flights): Includes analyses of accident initiators in the launch vehicle including those that might be caused by the presence of the LAS and its interactions with the launch vehicle; Includes modeling and simulation of the environments produced by each accident scenario and their effects on the LAS and its passengers.	2. The quality of the documentation of the previous PRAs, their applicability to the present mission;.the analysts' understanding of the previous systems and missions and how they differ from the present system/mission 3. The amount and quality of failure and anomaly data from relevant previous missions that employed new systems;.the analysts' understanding of the previous systems and missions. 4. The quality of the models used to support the abort analysis (as measured by credibility assessment factors) and the qualifications of the analysts

Table 6-3. Illustration of the Types of Evidence that Pertain to Various Potential Base Claims (Cont.)

Item	Example Base Claim	Example Direct Evidence	Example Supporting Evidence
4	Processes to ensure that requests for waivers or modifications of safety requirements are risk-informed.	Waiver/modification approach documented in SSMP; documented rationale based on correlation of safety requirements with risk drivers	Demonstrated understanding of safety requirements and their relation to risk drivers; appreciation for potential UU risks; knowledge of historical precedence and best practices
5	Models used to calculate the probability of loss from known risks are appropriately graded according to the importance of the mission, the criticality of each risk scenario, and the maturity of the design.	The mission criticality ranking and associated rationale for that ranking; the list of safety risk scenarios and evidence that it is a comprehensive list; the criticality ranking for each risk scenario and associated rationale for each ranking; the assumptions and approximations contained in the bounding models and evidence that these assumptions and approximations are realistic as well as bounding; the basis for establishing a margin for the loss probability when using bounding estimates to accommodate a short time window; and the assumptions and degree of resolution utilized in the more rigorous models together with evidence that the assumptions and resolution are appropriate for the criticality of the mission, the criticality of the risk scenario, and the maturity of the design	The qualifications and experience of the analysts regarding the bounding, deterministic, and probabilistic models and the use of graded analysis approaches; the use of experts in specifying the assumptions, approximations, and degrees of resolution to be used in the models; the processes used to verify that the models are sufficiently accurate considering the criticality level for which they are being applied; and the processes used to validate the models against real data and against other models that are accepted in the community and can be used as benchmarking models.
6	Models used for high criticality risk scenarios within high criticality missions satisfy the Modeling & Simulation Credibility Assessment Scale (CAS) criteria and are backed by ISA-informed tests.	Depends on the factor being ranked, but in general includes comparisons of M&S results to an acceptable referent, comparison of input data with measured data, quantitative uncertainty estimates, repeatability of the M&S results, and sensitivity of the M&S results to input and model parameters for the real-world system.	Use of reliable error estimation methods, community acceptance of the model as a *de facto* standard, and availability of personnel with advanced engineering or science degrees or extensive work experience in M&S in general, and with extensive experience in the development and use of the M&S being reviewed in particular
7	Reasonable estimates of the uncertainty distribution and their correlation coefficients have been obtained.	Statistical analysis, applicability/ completeness of testing and modeling	Robustness of expert elicitation. Qualifications of participants; documentation; use of qualified independent reviewers

Table 6-3. Illustration of the Types of Evidence that Pertain to Various Potential Base Claims (Cont.)

Item	Example Base Claim	Example Direct Evidence	Example Supporting Evidence
8	Plans are in place for identifying/evaluating the potential risk significance of precursors and anomalies, and for instituting appropriate contingencies and controls when needed.	Plans documented in SSMP for precursor analysis, anomaly and problem reporting, and implementation of corrective actions	Quality and completeness of record keeping of precursors and anomalies; staffing plan; experience of analysts
9	Budgets and schedules are adequate for the tasks to be performed and include realistic reserves.	Monte Carlo analysis of cost and schedule; budgets and deliverable dates consistent with analysis results; inclusion of realistic reserves based on historical experience	Quality of analytical models; qualifications of analysts
10	Management is focused toward continuous safety improvement and inclusiveness in the resolution of risks.	Program/project management plan	Past performance of management in present and previous programs
11	Any changes from previous validated designs are small enough to remain within the domain for which the performance of the system is known.	Integrated safety analysis and testing of the final system over all mission parameter values	Quality of analytical models and tests; qualifications of analysts
12	Provisions are made for testing any new technology or application of an existing technology within the context of its interactions with the whole system.	Plans for integrated testing of the final system over all mission parameter values	Evidence of sufficient budget and time to support such testing

6.1.5 Optional Use of Goal Structuring Notation in Developing Claims Trees

Goal Structuring Notation (GSN) has become an accepted format, and in some circles a standard format, for presenting safety cases in a rigorous and organized manner [66, 91, 92]. It was introduced to make safety arguments easier to develop and easier to evaluate through providing a clear graphical structure (see Figure 6-2). Toward this end, the structure includes the following elements:

- Goals at various levels. Note that the term "goal" in traditional GSN notation is taken to be equivalent to the term "claim" in this handbook, as it is in many applications that use GSN. We prefer "claim" to "goal" because the latter has a particular connotation within NASA referring to a maximum allowable probability of loss.
- Strategy, which refers to the process of decomposing a higher level claim to a set of lower level claims in such a way that the probable truth of the lower level claims is sufficient to establish the probable truth of the upper level claim. Verification that the decomposition is appropriate and sufficient is accomplished by inference arguments.
- Context, which provides a list of the present conditions and sources of information that pertain to a particular goal or strategy
- Assumptions, which are the hypotheses that must hold for the safety case to be valid
- Evidence, which has the same meaning as used earlier in this handbook

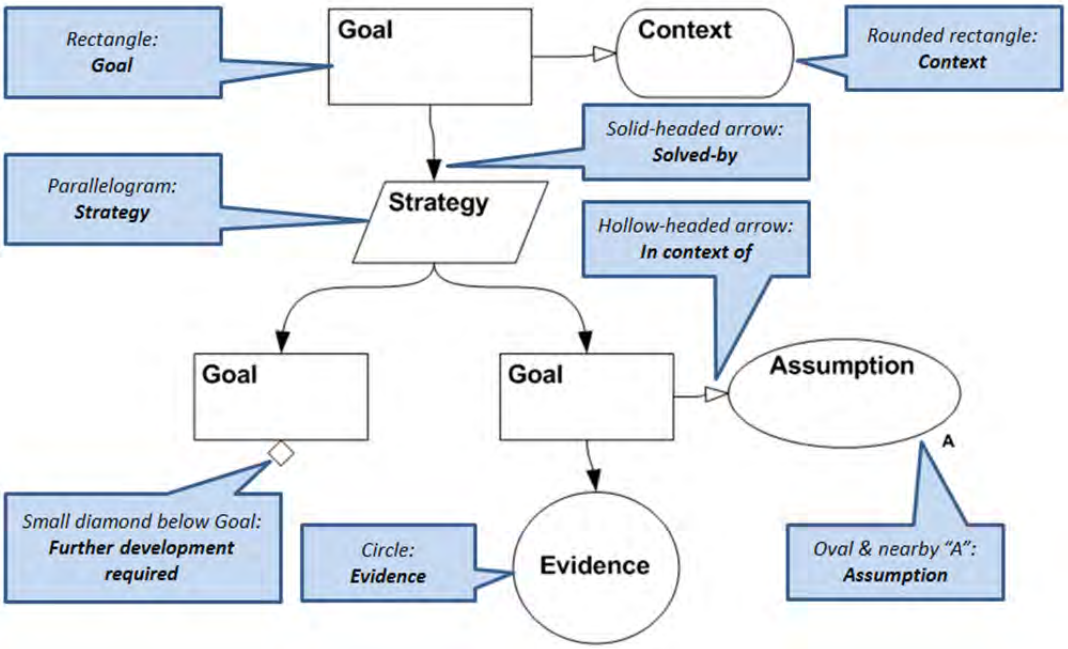


Figure 6-2. Goal Structuring Notation (GSN)

The use of GSN helps prevent the user from getting into common traps that pertain to the preparation of a safety case, such as [93]:

- Circular reasoning, which occurs when an argument is structured so that it reasserts its claim as a premise or defines a key term in a way that makes its claim trivially true
- Diversionary arguments, which contain excessive amounts of irrelevant material that could distract a reader from a weakly supported claim
- Fallacious appeals, which invoke irrelevant authorities, concepts, or comparisons as evidence
- Mathematical fallacies, which describe common pitfalls in probabilistic and statistical inferences
- Unsupported assertions, which are claims stated without evidence
- Anecdotal arguments, which show that their claims hold in some circumstances but not necessarily in general
- Omission of key evidence, which occurs when an otherwise complete argument omits evidence that is necessary to establish its validity
- Linguistic fallacies, which concern the use of misleading language that might lead the reader to an unwarranted conclusion. These fallacies may appear in any informal argument.

In addition, GSN helps to better document safety arguments by introducing common forms, and may assist in better management and maintenance of the safety case.

On the other hand, GSN has certain potential drawbacks. Among them is the fact that construction of large safety arguments using GSN could be complex, difficult to manage, and require specialized software (e.g., AdvoCATE [90] or commercial software). It is for this reason that the present handbook does not presume the use of GSN in developing a safety case. In developing their own safety cases, users should decide whether GSN adds enough benefit for their application to justify the added complexity or whether a more simplified approach is sufficient.

6.1.6 Documenting the RISC

This subsection provides guidelines for each of the eight sections of the RISC report enumerated in Section 6.1.1.

RISC Executive Summary

The executive summary should specify the context of the RISC, the review for which the RISC was prepared, key assumptions about the system, and key results and conclusions, including mention of significant unresolved issues.

Program, Project, and System Descriptions

This section should give sufficient detail about the program and project to unambiguously set the context in which the RISC has been developed. Generally, this section should give summary information pointing readers to specifics of the program/project/system via references.

Documentation of System Safety Objectives, Requirements, Policies, Regulations, and Standards

A subsection on safety objectives should discuss the safety performance objectives, safety thresholds, margins and goals, and assumptions about the projected rate of safety risk reductions (burn-down) over the life cycle of the project/program.

A subsection on system requirements should detail the system requirements appropriate to the review phase for which the RISC has been developed. Typically, this will be summary information with references to system and subsystem requirements documentation.

A subsection on policy, regulations, and standards should detail lower level requirements, policies, regulations, standards, NPRs, etc., which the program/project and system must satisfy. Examples include non-NASA policies, regulations and requirements to which the system must adhere (OSHA, DOD, DOE, etc.), and relevant NASA standards, NPRs, etc.

Where appropriate, accordance with and deviations from the objectives and guidelines in Chapter 4 and Section 5.3 of this handbook should be noted and explained.

The Safety Argument for the System with Supporting Evidence

A variety of approaches may be taken in presenting the safety argument for the system. Typically, summaries of the safety argument claims, and supporting evidence are provided in this section of the RISC report, with the detailed safety argument structure and evidence provided in appendices and references. Where appropriate, accordance with and deviations from the Provider's responsibilities and objectives in Chapter 6 of this handbook should be noted and explained.

Roadmap for Ongoing System Safety Activities

This section should present a plan for the various safety-related activities that will be taken in the future (after the current life-cycle review). The intention of the roadmap is to instill confidence in the Acquirer that an acceptable plan for ensuring the continued safety of the system is in place. As such, the plan should include activities that are currently in process and will continue past the current life-cycle review, as well as activities planned to start in the future. The plan should include a schedule showing when activities will begin and end, and dependencies between activities. Where appropriate, accordance with and deviations from the Provider's responsibilities and objectives in Chapter 5 of this handbook should be noted and explained.

Description of the Change and Configuration Management Procedures for the RISC

This section should describe in detail the change and configuration management processes, procedures, and system(s) employed to ensure that changes to the RISC are adequately tracked and controlled. The description should include an explanation for why the configuration and change control process is considered to be adequate, including a description of the process employed to incorporate changes and updates to the RISC.

Description of the Audit and Review Process Employed

This section of the RISC Report should describe the internal RISC audit and review process. In particular, the Provider should document the following when conducting an audit:

- When audits were conducted
- The results of the audits
- Audit team members and their qualifications, and
- Any significant concerns raised in the audit

Similar detail should be provided about the internal RISC review process employed by the Provider.

Discussion of Plans to Resolve Significant Unresolved Issues

Especially in the early phases of spaceflight system development, it is normal to have unresolved issues at review points. This section of the RISC is intended to give the Provider the opportunity to report on significant unresolved safety-related issues and discuss plans to resolve them. The plan for resolution should be as detailed as possible to allow the reviewers of the RISC to understand the issues and gain confidence that they have a high likelihood of being resolved.

6.2 Assigning Responsibilities for RISC Development and Integrating the Parts

6.2.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in assigning responsibilities for RISC development and integrating the parts can be summarized as follows:

- Identify and assign appropriately qualified personnel to be responsible for the development, integration, management, and maintenance of the RISC. Personnel should be chosen with the necessary experience, education, or training to effectively develop and update the RISC; to manage the integration of lower-level RISCs (e.g., those from subcontractors or other organizations) and supporting documentation from other subsystems (if required); and to manage the change and configuration control of the RISC Report.

- Take responsibility for the accuracy of all RISCs developed by subcontractors or sub-organizations that are integrated into the RISC developed by the Provider. As the organization providing the system to NASA, it is the responsibility of the Provider to take full ownership of all evidence submitted by subcontractors and sub-organizations that argue for the safety of a part of the integrated system.

The use of qualified personnel is necessary to gain acceptance for the RISC and assure knowledgeable people that the system is adequately safe. This is particularly true because formulation of the safety case requires more care and rigor from the Provider than does fulfillment of a set of prescriptive process requirements.

This section of the handbook addresses the following topics pertaining to these responsibilities:

- Assigning qualified personnel for RISC development, integration, management and maintenance
- Integrating subsystem RISCs into a system level RISC

6.2.2 Assigning Appropriately Qualified Personnel

It would be expected that persons developing a risk-informed safety case for NASA programs/projects would have a degree in an engineering or science discipline along with significant experience in one or more of the following areas: aerospace engineering, system safety engineering, aerospace system safety certification, aerospace system operation, probability and statistics, risk analysis, failure analysis, and mishap investigation.

The individual should understand:

- The structure of modern standards and aerospace safety cases including deterministic and probabilistic safety assessment and engineering substantiation
- ASARP and its application throughout the safety case life cycle
- The engineering design and operation of the system being assessed and the mission for which it is used
- How the safety case can be implemented and how it integrates with the design and operation of the system being assessed
- Safety case standards and methodologies
- The need for and use of specialized analyses where needed

In addition, the Provider should use the following guidelines in the selection and training of personnel:

- Maintain an established and effective procedure for defining individual employee qualifications and for ensuring that selections are made against these criteria.
- Use a structured approach to identify all employee training needs.
- Ensure that training needs are satisfied before employees start executing the task.
- Maintain a system for following-up on newly placed personnel to verify their effectiveness and include a review of the need for changes to the employee qualification criteria and/or personnel training.
- Maintain a system for identifying personal training needs following changes to the system design and/or operational mode, safety standards, and procedures; ensure that training is implemented in a timely manner.
- Maintain a system regularly validating the competence of the training instructors.
- Apply the training standards to both company employees and contractors.
- Make use in the training programs of lessons learned and experience developed from previous projects and incidents.

6.2.3 Integrating Subsystem RISCs

Just as there is a need for subsystem analyses that feed into an integrated system analysis[51], there is a need for subsystem level RISCs that feed into a system level RISC. The system level RISC is prepared by the organization that is referred to in this report as the Provider. The subsystem level RISCs are prepared by the organization responsible for each subsystem. If subsystem development is allocated to subcontractors, each subcontractor should prepare a RISC that pertains to their assigned subsystem.

There are two general considerations for subcontractors preparing RISCs:

1. The subsystem RISC should address all subsystem level concerns that are identified in the system level RISC.
2. The RISCs at subsystem level should all be consistent with one another and with the system level RISC in terms of the claims that are made and the evidence that is used to substantiate the claims.

Regarding the first item, it has been pointed out earlier that many concerns identified in the system level RISC will cross over subsystem boundaries. For example, a claim in the system level RISC that the probability of loss of the system from known risks is within the probabilistic requirement, would likely require a risk assessment that would have to account not only for the failure of individual subsystems but also for cross-system scenarios. The approach advocated in Section 5.2.5 for identifying subsystem analyses that feed into an ISA would therefore apply to the formulation and substantiation of claims at subsystem level that feed into a system level RISC.

Regarding the second item, the claims at system level contain many elements that would also have to be addressed at subsystem level. For example, claims like the following would have to be included and substantiated at both levels:

- The agreed-to verification procedures have been executed, or if not, the reasons for any departures from the procedures have been adequately explained, justified, and agreed to by the Acquirer.
- Models used to calculate the probability of loss from known risks are appropriately graded according to the importance of the mission, the criticality of each risk scenario, and the maturity of the design.
- Models used for high criticality risk scenarios within high criticality missions satisfy the Modeling & Simulation Credibility Assessment Scale (CAS) criteria and are backed by ISA-informed tests.
- Fault management models and software models are adequately integrated with other ISA models.
- Configuration control and data management reflect NASA standards and guidelines and are maintained in all relevant areas, particularly where there are multiple suppliers and distributed resources.
- Relevant best practices and lessons learned either have been incorporated or are demonstrated to be not practicable.
- Budgets and schedules are adequate for the tasks to be performed and include realistic reserves.
- Management is focused toward continuous safety improvement and inclusiveness in the resolution of risks.
- Roles and responsibilities for resolving safety concerns are clearly defined in accordance with the

[51] The relationship between subsystem analyses and integrated system analysis was discussed in Section 5.2.5.

program/project and RM plans, provide effective interaction, and resolve problems as they arise.

- Design complexity is minimized without violating safety.
- Any changes from previous validated designs, fabrication methods, and operational procedures remain within the domain for which the performance of the system is known.
- A set of safety critical items (e.g., hardware features, software features, human actions, management practices, and administrative controls) has been identified whose performance at levels documented in the Integrated Safety Analysis assures the satisfaction of the system-level known loss probability requirement.

For these claims and others like them, it is incumbent on the Provider to ensure that consistency and relevance are maintained across the system RISC and all subsystem RISCs.

6.3 Exercising a Graded Approach in RISC Development

6.3.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibility in exercising a graded approach in RISC development can be summarized as follows:

- Develop the RISC to an appropriate level of detail to ensure that the safety of the system is fully and coherently documented, and in such a way as to allow for an independent review and evaluation of the RISC.

Small programs/projects do not require the same amount of analysis and documentation as ones that are larger or more complex. Correspondingly, the RISC should be developed to a level of detail that is sufficient to communicate the safety argument effectively. In all cases, for reasons of scrutability and to maximize the utility of finite resources, the RISC should focus on safety risk drivers, but should also provide rationale for why the other sources of safety risk not focused on are not considered to be significant.

6.3.2 Exercising a Graded Approach

The graded approach to be applied in a RISC pertains to two aspects of the RISC: the completeness of the claims tree and the completeness of the evidence. The criteria for determining how much rigor to apply in each of these two areas are similar to those discussed in Section 5.2.4, which addressed graded analysis for an ISA. Choices about the completeness of the claims tree depend on the criticality of the mission, analogous to choices about the scope of an ISA. Choices about the amount of effort to be applied in developing evidence for a claim in the claims tree depend on the mission criticality and on the criticality of the claim, just as the level of effort to be applied in analyzing a risk scenario in an ISA depends upon the mission criticality and the importance of the scenario.

In Section 5.2.4, mission criticality was taken to be equivalent to the project priority rankings as defined in the latest version of NPR 8715.3C [6]. The scheme for rating mission criticality can be used in this context as well. Namely:

- A mission has high criticality if it requires White House approval per PD/NSC-25, or is subject to planetary protection requirements, or has high strategic importance to the Agency, or has life-cycle cost exceeding $1 billion.
- A mission has medium criticality if it does not have high criticality and has a life-cycle cost between $250 million and $1 billion.

- A mission has low criticality if it does not have high or medium criticality and has a life-cycle cost less than $250 million.

The desired completeness of the claims tree parallels the criteria in the NPR and in Section 5.2.4 for the scope of the ISA:

- A mission with high criticality should entail full scope development of the claims tree according to the techniques described in Section 6.1 and subsections.
- A mission with medium criticality should entail claim development only in specific areas that are deemed to warrant it, together with a set of arguments to justify that the system as a whole is safe.
- A mission with low criticality should be accompanied by a set of arguments to justify the assertion that the system is safe but does not require formal development of a claims tree.

Risk scenario criticality, in Section 5.2.4, was measured in terms of rough estimates of the likelihood of the scenario becoming a reality, the corresponding severity of the outcome, and a qualitative estimate of the uncertainty. In the present context, the notion of risk scenario criticality is replaced by evidence criticality, which has the following rating scheme:

- The evidence criticality is "high" if the claim to which it applies is deemed to be important relative to the goal of achieving a safe system and the amount of assurance deficit for that claim using existing evidence is considered high[52].
- The evidence criticality is "moderate" if the claim to which it applies is important relative to the goal of achieving a safe system and the amount of assurance deficit for that claim using existing evidence is considered moderate but not critical.
- The evidence criticality is "low" if the amount of assurance deficit for that claim using existing evidence is considered low, regardless of the relative importance of the claim.

Table 6-4 provides a summary of the criteria for evidence completeness as a function of these two criticality factors.

Table 6-4. Amount of Additional Evidence Needed to Satisfy the Criteria for Evidence Completeness for Different Levels of Mission Criticality and Evidence Criticality

		Mission Criticality		
		Low	**Medium**	**High**
Evidence Criticality	**High**	No additional evidence required	As much as needed to reduce the assurance deficit to moderate (Rank 3 or less)	As much as needed to reduce the assurance deficit to low (Rank 1)
	Moderate	No additional evidence required	No additional evidence required	As much as needed to reduce the assurance deficit to low (Rank 1)
	Low	No additional evidence required	No additional evidence required	No additional evidence needed

[52] Sections 6.1.2, 7.2.3, and 7.2.4 provide more information on the identification and ranking of assurance deficits.

6.4 Maintaining and Updating the RISC and Addressing Future Life-Cycle Phases

6.4.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in maintaining and updating the RISC and addressing future life-cycle phases can be summarized as follows:

- For new systems, develop a RISC for the System Definition Review (SDR) and update it, as necessary, to reflect changes in the design or operation of the system[53]. At a minimum, the RISC should address all relevant milestone review entrance criteria specified in NPR 7123.1B [3]. Updates would typically be expected at the designated reviews in the NASA Systems Engineering Life Cycle (see Figure 6-3).

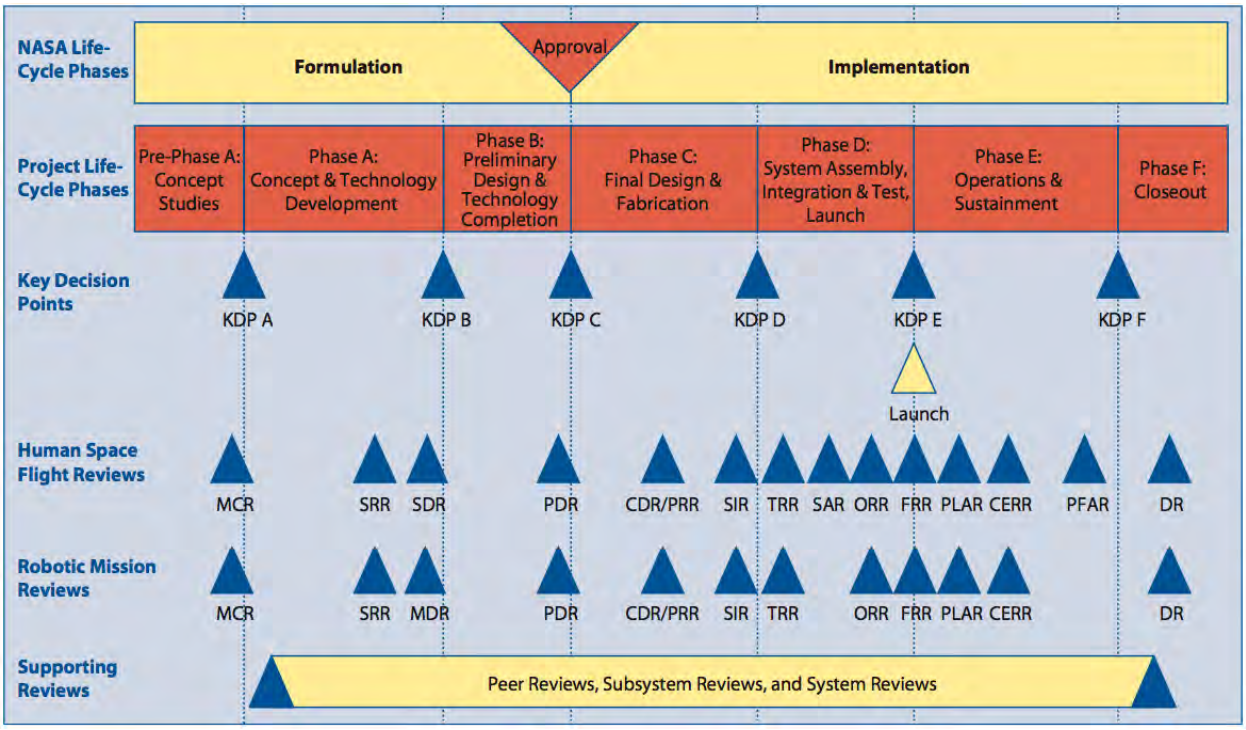


Figure 6-3. The NASA Systems Engineering Life-Cycle Phases

- For already existing systems to be acquired by NASA (for example, commercially supplied launch services), update the RISC as required by the program/project. (An initial RISC should already have been developed.)
- At scheduled life-cycle RISC reviews, delineate the burndown-to-date of safety risks associated with fulfillment of the levels of safety performance. As defined in Section 2.1.1, risk burndown refers to the expectation that as a program/project evolves over time, and as risk concerns are retired and the state of knowledge about the performance measures improves, uncertainty should decrease, with an attendant lowering of risk. The burndown of safety risks provides important evidence of the effectiveness of the Provider's RM process, as well as the support of RM by the Provider's management organization.

[53] See Section 6.1.1 for a brief discussion of RISC requirements for existing systems.

- In each RISC submittal, address safety-related objectives, plans, concerns, and scenarios for all future program/project life-cycle phases at an appropriate level of detail. It is understood that the level of detail presented in treating the safety-related objectives, concerns, and scenarios for future phases of the system will necessarily be of lesser detail than the current RISC for a given review phase, and that the level of detail may, in fact, be minimal for late life-cycle phases of a RISC submitted early in a Project/Program. For example, a RISC submitted at SDR would normally have only high-level discussion about system safety for the closeout phase of the program/project. This information would become more detailed in RISCs submitted later in the life cycle.
- Update the RISC when certain events occur. Circumstances that would warrant updating of the RISC include the following:
    - Significant unplanned updates to the system (e.g., unplanned block upgrade of avionics, propulsion, thermal protection system (TPS), upload of new or revised software, etc.)
    - Significant changes to the operation of the system not captured in the existing operative RISC
    - Failure of a critical subsystem that does not lead to catastrophic failure of the vehicle (e.g., failure of one component in a subsystem with redundancy)
    - A change of mission requirements
    - A change in the environment, the understanding of the environment, or the assumptions made about the environment that affects assertions made in the RISC
    - Design or operational changes resulting from return-to-flight decisions during hiatus following a catastrophic failure of the system

This section of the handbook addresses the following topics pertaining to the maintenance and updating of the RISC and the addressing of future life-cycle phases:

- Maintaining and updating the RISC
- Addressing future life-cycle phases

6.4.2 Maintaining and Updating the RISC

The normal processes of analysis, testing, and operation can lead to new information that affects the evaluation of the loss probability or the degree to which deterministic requirements are being met. For example, a new risk scenario may be uncovered, or the understanding of an existing risk scenario may need to be reassessed. Whenever this occurs, it is necessary to reevaluate whether the system still satisfies the probabilistic and deterministic requirements.

If a requirement is not satisfied for any key mission objective, then the first step is to try to regain compliance by applying controls that do not violate the criteria for minimizing UU risks, i.e., do not significantly increase the design complexity, reduce the effectiveness of defense-in-depth, or increase the pressures on time or budget. If this cannot be accomplished by introducing practicable controls, the alternative is to seek an adjustment (or rebaselining) of the requirement. In either case, the RISC should be updated accordingly. For launch vehicles and human-rated systems this may require re-certification of the system. The existing RISC will then need to be updated to reflect changes in the system design or operation.

6.4.3 Addressing Future Life-cycle Phases

The concept of adequate safety requires that safety be addressed throughout all phases of the system life cycle. Correspondingly, the RISC must also address the full system life cycle, regardless of the particular point in the life cycle at which the RISC is developed. This manifests in the RISC as two distinct types of safety claims:

- Claims related to the safety objectives of the current or previous phases argue that the objectives have been met.
- Claims related to the safety objectives of future phases argue that a 'roadmap' has been established for the satisfaction of objectives yet to be met, i.e., that necessary plans, preparations, and commitments are in place to meet safety objectives at the appropriate time.

The form of the RISC arguments for accomplished objectives vs. upcoming objectives is shown in Figure 6-4 for the point in time at which design has completed and realization is about to commence. As the system proceeds in the life cycle and RISCs are developed for successive milestones, arguments demonstrating an ability and commitment to meeting objectives are replaced by arguments demonstrating accomplishment of objectives.

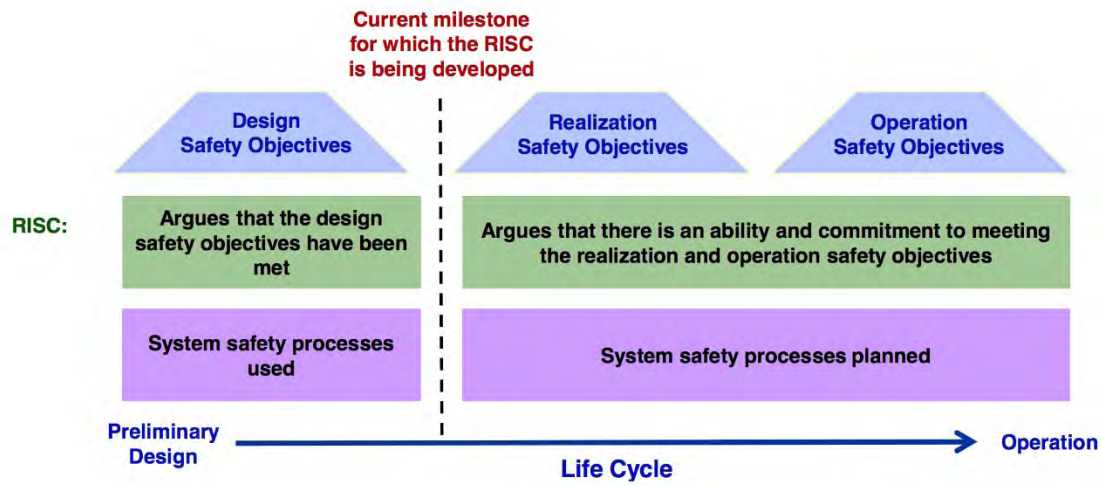


Figure 6-4. Coverage of the System Life Cycle in the RISC

The operational safety objectives tree presented in Section 3.1 (Figure 3-2) specifically highlights the responsibility of the Provider to address future life-cycle phases by separately referring to objectives that need to be addressed during design, others that need to be addressed during product realization, and still others to be addressed during operation and sustainment. To be consistent with this life-cycle perspective, the claims in the claims tree must cover these three phases at a minimum (as well as any other phases that require consideration, such as decommissioning). Frequently, the claims are worded without reference to life-cycle phase, but the understanding is that the evidence to be provided must substantiate each claim for each phase of the life cycle to which it is applicable. For example, a claim such as Item 4 in Table 6-3, "Processes to ensure that requests for waivers or modifications of safety requirements are risk-informed," must be presumed to apply anytime when waivers or modifications may be requested, whether during design, realization, or operation. The Provider should show that the premise of the claim is being addressed in the present phase by citing real evidence to that effect, and should also show how the claim will be addressed during future phases by referring to an appropriately documented section of the System Safety Management Plan (SSMP).

6.5 Addressing Weaknesses and Unresolved Safety Issues

6.5.1 The Provider's Responsibilities and Areas to Address

The Provider's responsibilities in addressing weaknesses and unresolved safety issues can be summarized as follows:

- Clearly address weaknesses and limitations in the RISC. This should include an assessment of the degree to which such weaknesses and limitations affect the overall assurance that the safety claims have been satisfied and the system is adequately safe.
- Discuss in the RISC the plan to address significant unresolved safety related issues and provide a schedule for achieving resolution of all such issues.

6.5.2 Addressing Weaknesses and Issues

For the Acquirer to have a full understanding of, and confidence in the RISC, it is critical that weaknesses and limitations in the RISC, such as shortcomings in the evidence, be identified and addressed. This process imparts confidence to the Acquirer that the RISC has been carefully prepared, that it is complete, and that the Provider and Acquirer understand its weaknesses and limitations.

In the development of new space flight systems that push the technological state-of-the-art, it is common to encounter problems in the design of the system during the development phase that may have impacts on the safety of the system. Identifying significant unresolved issues, and presenting a plan to address such issues, helps to focus attention on areas of the project that may require additional resources or that may cause schedule slips. Addressing unresolved safety-related issues helps with transparency in the Program/Project, increases Acquirer confidence in the system and the Provider, and increases the likelihood that the system will meet its safety thresholds and be ASARP.

The responsibility for identifying weaknesses and unresolved safety issues, assessing their effects on the claim that the system is adequately safe, and devising how to resolve those issues, lies principally with the Acquirer during the RISC evaluation step. However, the Provider is also expected to make their own assessment of these issues and correct them to their best ability prior to submitting the RISC to the Acquirer. The approach recommended for the Provider is basically the same as the approach recommended for the Acquirer. Refer to Sections 7.2.3 and 7.2.4 for guidance on assurance deficits, which also applies to this section of the handbook.

6.6 Example for Chapter 6 – RISC Fragment for New Technology on a Robotic System (Electric Ion Thruster)

As discussed by Brophy, et al., [94, 95], electric ion thrusters for use on deep-space science missions (see Figure 6-5) have challenging life qualification issues because they are expected to have operational lifetimes of tens of thousands of hours, operate over a broad range of input powers, and are subject to complex wearout failure modes. The customary approach of performing a single life test for the required number of hours plus margin provides insufficient information to characterize the failure risk.

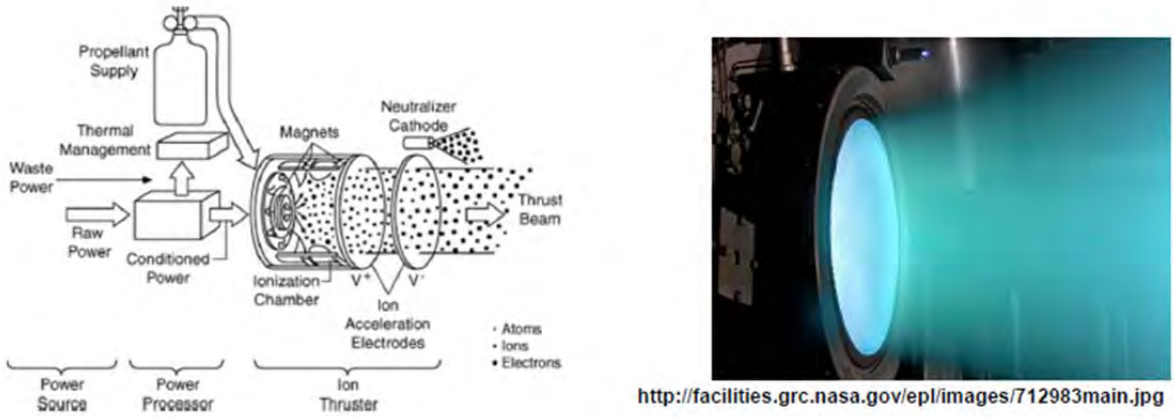


Figure 6-5. Schematic of an Electric Ion Thruster Subsystem and Photo of a Test Model

There are several questions to be addressed when gathering evidence to substantiate a claim that the thrusters will perform their mission without premature failure:

- Has the testing time and/or operational time been sufficiently long to establish that there are not unexpected wearout failure modes that could occur toward the end of life?
- Have there been a sufficient number of life tests to establish that there is not a wearout failure mode with random failure time that has not yet been observed?
- Have the environments for life testing and/or operation been identical to the environments that will be seen during actual operation (e.g., zero-gravity, vacuum, radiation, orientation)?
- Has the item been life-tested at conditions beyond the design basis environments to establish its robustness in the event of abnormal/unexpected environments?
- Have all functions that the item will have to perform during the mission been tested?
- Have all the above factors been tested in the full-up system configuration to account for interactive modes of wearout failure?
- Are there analytical models that successfully fill the gaps left by shortcomings in the testing and operational experience?

6.6.1 Safety Claims Tree

Figure 6-6, on the following two pages, shows an example claims tree for a proposed new electric ion thruster subsystem, in which the top claim to be demonstrated is that the as-designed spacecraft propulsion system is adequately protected against wearout failure.

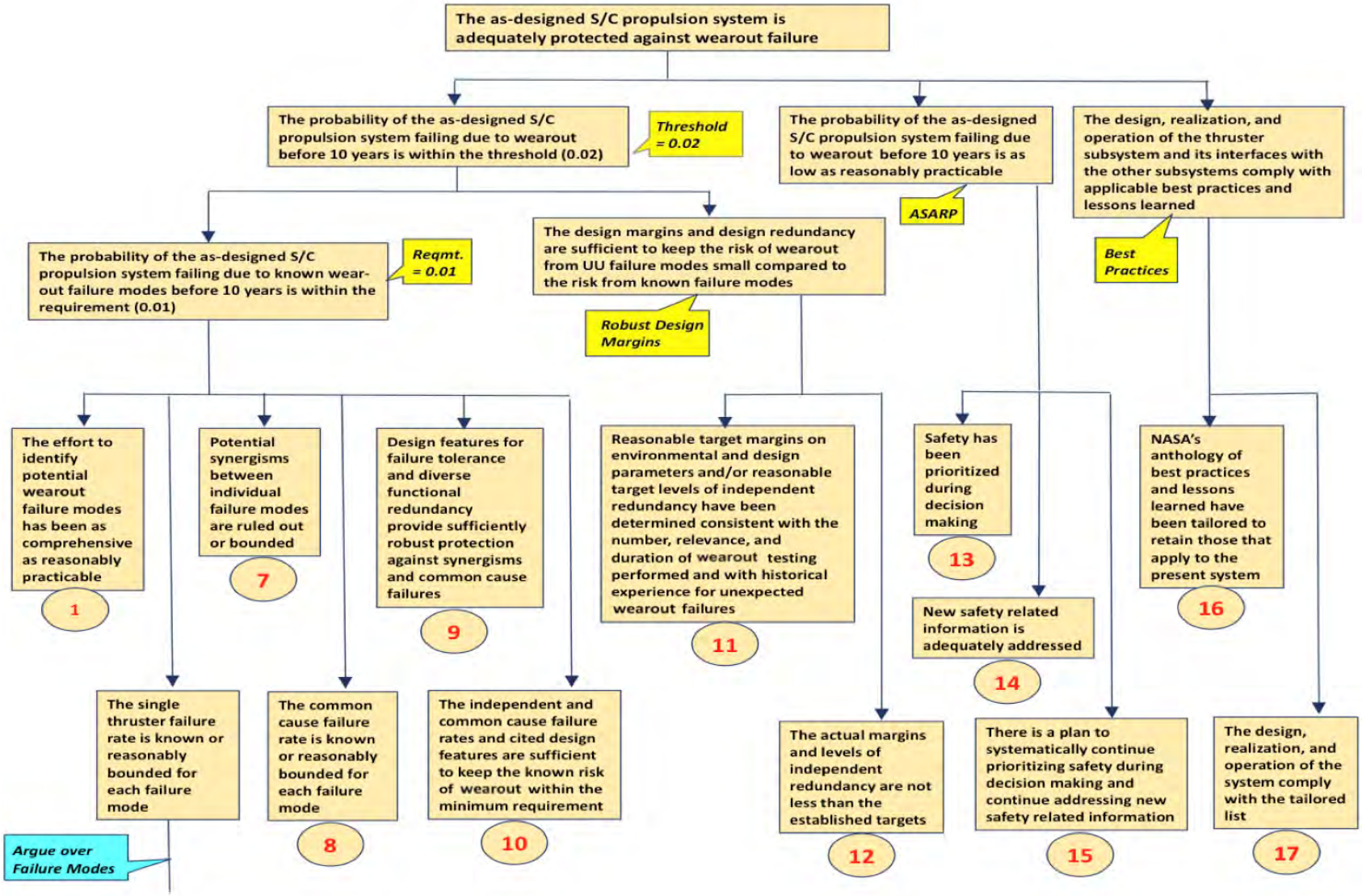


Figure 6-6. Claims Tree Pertaining to Wearout Failures for a New Technology Ion Electric Thruster (1 of 2)

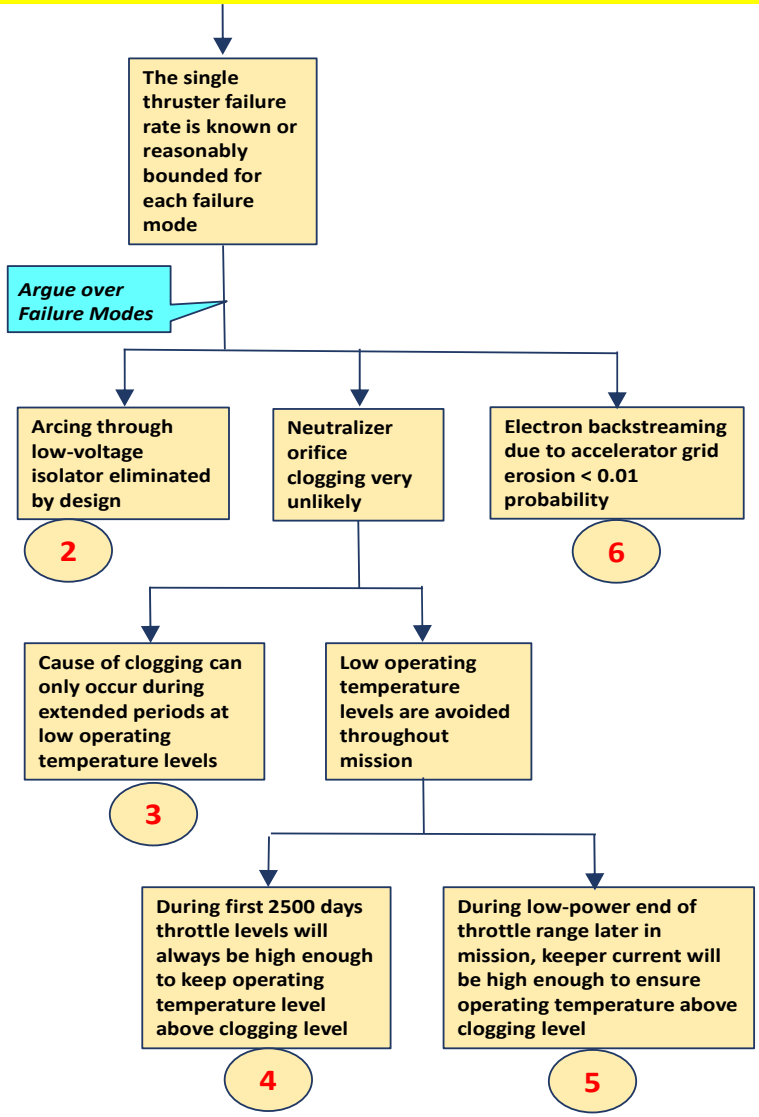


Figure 6-6 (Cont.). Claims Tree Pertaining to Wearout Failures for a New Technology Ion Electric Thruster (2 of 2)

6.6.2 Sources of Evidence

The safety claims tree has a total of 17 base claims. Table 6-5, below, presents the sources of evidence that are cited by the Provider to support these base claims. In addition to testing and operational results, the direct evidence includes the results of failure analyses, design inspections, historical reviews, and a literature review. Supporting evidence includes the quality of the tools used to perform the failure analysis and the qualifications of the analysts and expert consultants used

6.6.3 Hypothetical Suppositions, Assumptions, and Contexts

For purposes of this example, it is assumed that the following set of suppositions, assumptions, and contexts applies to the mission to be executed and the evidence that was obtained. These suppositions, etc., are stated purely for purposes of developing the example, are hypothetical, and are not intended to represent the actual reality for an existing mission or system.

- The mission lifetime is 10 years.
- There have been 65 full-up tests on the ion thrusters to be used, with test times between 2 years and 5 years, and there were no failures during these tests and no evidence of material degradation.
- Gridded ion thrusters of various kinds have been tested in microgravity, in vacuum, and in high radiation environments, but not all three at the same time; therefore, there is some question about their life expectancy in the combination of environments.
- Prior spaceflight history implies that the probability of experiencing an environment during a mission that would subject the ion thrusters to stressors beyond their design margins is on the order of 10%.
- There has been testing for every functional mode of the thrusters to be experienced during the mission.
- The only identified significant system interaction is a potential for impingement of eroded material from the ion thruster subsystem on other subsystems.
- The only identified significant hardware-software interaction is the possibility unusual cycling of the ion thrusters due to malfunction of the control logic.
- The statistical analysis performed in association with Base Claim 10 in Table 6-5 was a maximum likelihood estimation (MLE) analysis which was conducted using the data from the 65 real-system tests as input, to examine the implications of the limited test time. A Weibull failure distribution was assumed with a high shape parameter so that the resulting estimate could be considered bounding. Binomial statistics were used to examine the possible influence of randomness in the test results. A confidence level of 95% was assumed for this analysis to be conservative.

  Results from the expert judgment elicitation conducted in association with Base Claim 11 in Table 6-5 were used to estimate the probability that there may be unknown critical wearout failure modes from each of the following causes: (1) environments that have not been simulated during testing (e.g., the combination of microgravity, vacuum, and high radiation); (2) unanticipated abnormal environments (e.g., solar flares); (3) unintended functional modes for the ion thrusters, and (4) unknown or underappreciated system interactions (e.g., interactions between the electrical and chemical thrusters). Inputs to the elicitation included the results of testing on similar systems in a variety of environments, the historical record of occurrences of unanticipated environments and heretofore unknown interactions, and the amount of margin in the thruster design.

Table 6-5. Evidence for Selected Base Claims Pertaining to Wearout Failures for the Ion Thruster Subsystem

Item	Base Claim	Evidence
1	The effort to identify potential wearout failure modes has been as comprehensive as reasonably practicable.	1. **DIRECT:** Inspection following Extended Life Testing (ELT): 30,000 hour ELT of NSTAR thruster 2. **DIRECT:** Body of Evidence for similar thrusters: Ion thruster technology's long history of life testing and operational experience 3. **DIRECT:** Literature Review: Extensive published technical material for ion thrusters existing or currently under development 4. **SUPPORTING:** Competency of personnel: Identification activities performed by personnel with extensive background in theory and practice of Ion engines
2	Arcing through low-voltage isolator eliminated by design	1. **DIRECT:** Design inspection confirms no use of low-voltage isolators.
3	Cause of clogging can only occur during extended periods at low operating temperature levels	1. **DIRECT:** Literature search
4	During first 2500 days throttle levels will always be high enough to keep operating temperature level above clogging level.	1. **DIRECT:** Analysis of thrusting levels during mission profile 2. **SUPPORTING:** Favorable NASA-STD-7009 review of analysis
5	During low-power end of throttle range, keeper current will be high enough to ensure operating temperature above clogging level.	1. **DIRECT:** Design inspection confirms keeper current settings will be high.
6	Electron back-streaming due to accelerator grid erosion < 0.01 probability	1. **DIRECT:** Analysis of test results: No failures observed 2. **DIRECT:** Deterministic physics of failure modeling and Monte-Carlo simulation for back-streaming due to accelerator grid erosion: Failure within mission lifetime < 0.01 probability 3. **SUPPORTING:** Deterministic models verified to be correct and validated for ground test environments based on M&S Credibility Assessment Scale criteria in NASA-STD-7009 4. **SUPPORTING:** Probabilistic models consistent with best PRA practices in NPR 8705.5
7	Potential synergisms between individual failure modes are ruled out or bounded.	1. **DIRECT:** Literature review shows no interactions between different failure modes. 2. **DIRECT:** Expert judgment elicitation indicates very low likelihood of significant potential synergisms. 3. **SUPPORTING:** The experts are highly qualified.
8	The common cause failure rate is known or reasonably bounded for each failure mode.	1. **DIRECT:** Robust quality assurance and control program to protect against
9	Design features for failure tolerance and diverse functional redundancy provide sufficiently robust protection against synergisms and common cause failures.	1. **DIRECT:** The design includes two extra ion thrusters beyond the minimum needed.

Table 6-5. Evidence for Selected Base Claims Pertaining to Wearout Failures for the Ion Thruster Subsystem (Cont.)

Item	Base Claim	Evidence
10	The independent and common cause failure rates and cited design features are sufficient to keep the known risk of wearout within the minimum requirement.	1. **DIRECT:** Statistical analyses of censored failure data from ion thruster testing indicates 80% confidence that wearout failure will not occur before mission is completed. 2. **SUPPORTING:** Statistical models are verified and validated for many hardware applications (but not specifically for ion thrusters). 3. **DIRECT:** Expert judgment elicitation indicates low likelihood (<10%) of exceeding minimum requirement. 4. **SUPPORTING:** The experts are highly qualified.
11	Reasonable target margins on environmental and design parameters and/or reasonable target levels of independent redundancy have been determined consistent with the number, relevance, and duration of wearout testing performed and with historical experience for unexpected wearout failures.	1. **DIRECT:** Margin added to the keeper current level to account for potentially underappreciating the failure rate due to clogging of the neutralizer orifice 2. **DIRECT:** Margins to account for UU failure modes are consistent with UU failure rates that have occurred for other systems. 3. **DIRECT:** Level of redundancy for ion thrusters is consistent with level of redundancy used for other new technology subsystems in spaceflight applications. 4. **DIRECT:** Expert judgment elicitation on the likelihood of unintended environments, functional modes, and system interactions indicates low probabilities (< 10%) 5. **SUPPORTING:** The experts are highly qualified.
12	The actual margins and levels of independent redundancy are not less than the established targets.	1. **DIRECT:** Inspection of the design and operational plan indicates targets will be met.
13	Safety has been prioritized during decision making.	1. **DIRECT:** Management structure, roles and responsibilities, safety policies, inclusiveness in decision making all indicate that safety has a high priority.
14	New safety related information is adequately addressed.	1. **DIRECT:** Risk management processes and database contents, status of unresolved safety issues, precursor analysis program, problem reporting and corrective action, and information sharing processes all indicate that new safety related information is being adequately addressed.
15	There is a plan to systematically continue prioritizing safety during decision making and continue addressing new safety related information.	1. **DIRECT:** The System Safety Management Plan and Risk Management Plan indicate that safety will continue to be prioritized and new safety information will be addressed.
16	NASA's anthology of best practices and lessons learned have been tailored to retain those that apply to the present system.	1. **DIRECT:** Corroborated by the status of the System Safety Requirement Analysis (SSRA) and adherence to best practices and lessons learned from the Goddard GOLD Rules, NASA Space Flight Program and Project Management Handbook, NASA Preferred Practices for Design and Test of Robust Systems (JPL), NASA Lessons Learned System, etc.
17	The design, realization, and operation of the system comply with the tailored list.	1. **DIRECT**: Inspections and audits concur that best practices and lessons learned are being adhered to.

Section 7.4 will continue with this example in providing results of the analysis and interpreting those results in terms of confidence in the safety case.

6.6.4 A Note on Margins

Safety margins in the context of the present example have a different connotation from safety performance margins as discussed in Section 4.1.3 and demonstrated in Section 4.5. For the present example, margins have been used in a more traditional sense. That is, the safety claim relating to margins in Figure 6-6 refers to factors of safety on environmental parameters (e.g., stress, temperature) and design parameters (e.g., material thickness, configuration), rather than on safety performance parameters such as the probability of loss of vehicle or mission. The higher level intermediate claim is that the design margins and design redundancy are sufficient to keep the risk of wearout from UU failure modes small compared to the risk from known failure modes.

The philosophy applied here is that when the source of risk is entirely phenomenological (i.e., associated with the understanding of the physics), it is not fruitful to attempt to develop margins directly for the safety performance risk based on experience. As opposed to program/project risks, where the principal sources of UU risk tend to be associated with organizational, management, and programmatic factors, the magnitude of phenomenological risks from unknown and underappreciated sources is basically unpredictable (e.g., [94]). For phenomenological risks, therefore, it is more fruitful to develop margins on environmental and/or design parameters and to make those margins large enough to gain confidence that the risk of wearout failure from UU risks is contained.

An example pertaining to the principle of margins for phenomenological risks is the risk of satellite control processors (SCPs) failing due to shorting as a result of the growth of tin whiskers. In the early 2000s, such failures were happening as early as four years after launch for satellites that were designed to last for 15 years [e.g., see NASA web site nepp.nasa.gov/WHISKER/failures/index.htm]. The principal cause was that the solder in the electronic circuitry was made of tin coated by a thin polymer film, and the relaxation of the tin solder from thermal stresses produced by cyclical exposure to the sun caused whiskers to grow. The whiskers eventually become long enough to short the energized bus terminal posts to the grounded relay case.

The tin whiskers phenomenon was known long ago, but what was not known was that the growth mechanism was aggravated by conditions characteristic of space that were not manifest in ground testing and operation on Earth (i.e., microgravity and vacuum). It is appropriate to ask how greater design margins might have helped for such a problem that was known but underappreciated. It can be postulated that greater margins would have been successful if applied to accomplish any of the following objectives:

- Lower thermal variations inside the spacecraft
- Greater thickness of conformal coating (polymer film)
- Greater independent redundancy (additional fully contained SCPs) available in a standby mode

For the ion thruster system, it may be postulated that greater margins should reduce the risk from UU scenarios if applied to accomplish the following:

- Operating temperature kept within a narrower range (low temperatures can increase orifice clogging whereas high temperatures can challenge material capabilities)
- Stronger grid design (addresses electron backscatter caused by grid erosion)
- Independent redundancy: the inclusion of standby ion thrusters that are physically separate and functionally independent from the active thrusters to allow for continued operation in case the active thrusters degrade or fail as a result of wearout

The first two of these are intended to protect against known risks that are underappreciated, whereas the last provides protection against both underappreciated and unknown risks.

7. Evaluating the RISC: The Acquirer's Role

Once the Provider has delivered the RISC to the Acquirer, it will typically be evaluated on behalf of the Acquirer by an Evaluator, which is an agent of the Acquirer who is tasked to evaluate the RISC and generate findings on its technical adequacy (i.e., validity and completeness). The Evaluator's findings, along with the RISC, are furnished to the Acquirer's decision-making entity, which will use them as a safety-specific technical basis to allow or disallow the system to move forward in its life cycle.

In general, an Acquirer has a safety risk acceptance decision to make at each KDP in a system development life cycle, and the RISC is a key input to that decision process. The Acquirer is responsible for the decision, and decides how much and what kind of review, or how much independent analysis, will be needed in order to create a technical basis for an acceptance decision. The Evaluator may need to review the RISC in very significant depth to warrant confidence in it.

The norms for safety risk acceptance itself are beyond the purview of this handbook. The balance of the discussion in this chapter, therefore, is focused on technical evaluation of the RISC, which entails reviewing the claims made in the RISC in order to:

- Understand the technical basis (i.e., arguments and evidence) behind the claim of adequate safety.
- Evaluate the technical basis of the claim to determine its validity.
- Judge the adequacy of the claim (validity and completeness).

In questioning the technical basis of the RISC and judging its adequacy, the Evaluator will analyze the RISC from a critical viewpoint. If all key arguments in the RISC are understood by the Evaluator and are logically compelling, such that all evaluation findings required of the Evaluator are manifestly supported, then there will be no need for clarification or supplementary argument during RISC evaluation. However, due to the complexity of space systems, gaps or differences in understanding are possible on either side. Since the Acquirer is ultimately accepting or not accepting the risk of not meeting safety objectives, the Acquirer determines the outcome of a gap or difference in understanding.

If gaps or differences in understanding warrant, or if the key arguments in the RISC are not logically compelling, the Acquirer may ask for information in addition to that submitted by the Provider, e.g., through a formal Request for Additional Information (RAI) process. RAIs are the mechanism through which the Acquirer is able to seek clarification of a particular concern during the RISC evaluation process. This additional information from the Provider will typically be integrated into the RISC by the Provider during the review process, and the final RISC and associated evaluation findings prepared by the Acquirer will reflect these Evaluator/ Provider interactions. The assumptions underlying the RISC and the effectiveness of the processes implemented in accordance with the RISC will be reviewed throughout every phase of the project.

This chapter addresses the processes the Evaluator may use when establishing and communicating expectations concerning a forthcoming RISC, and when evaluating a submitted RISC. It provides guidance and examples that expand upon the overview for developing RISC-specific evaluation processes and findings presented in Section 3.4. The following subjects pertaining to that overview are discussed sequentially in Sections 7.1 through 7.3:

- Interfacing with the Provider
- Reviewing, evaluating, and scoring the RISC
- Documenting the findings of the RISC evaluation

As discussed in earlier chapters, the evaluation processes described herein may be tailored by the Acquirer in accordance with the principle of implementing a graded approach to risk management.

7.1 Interfacing with the Provider

7.1.1 The Acquirer's Responsibilities and Areas to Address

The Acquirer's pursuits in interacting with the Provider to facilitate evaluation of the RISC can be summarized as follows:

- Before submittal of the RISC, request a System Safety Requirements Analysis (SSRA) and a finalized System Safety Management Plan (SSMP) from the Provider.
- Inform the Provider of the expected level of detail required in the development of the RISC and of any expectations with regard to structure and presentation of the RISC prior to its development, negotiating with the Provider as appropriate to most effectively satisfy the Acquirer's information needs within the context of the SSMP.
- Set periodic review meetings with the Provider during development of the RISC to gauge progress, receive/communicate updates, and determine if issues have arisen or been resolved.
- When necessary, initiate a Request for Additional Information (RAI), relating each RAI to a perceived weakness in the support for one or more safety claims in the RISC.

The SSRA serves to clarify what detailed requirements the Provider expects to address in the ensuing development of the system, and argues that the satisfaction of these requirements will provide evidence of satisfaction of the top-level requirements. Evaluation of the SSRA gives the Acquirer an early opportunity to ensure that the Provider is adequately addressing the safety performance requirements and is implementing a risk-informed process in development of the system (for example, through the use of tailored requirements).

When the scope of requirements to be addressed has been clarified through the SSRA process, the SSMP can be finalized, again through iteration between the Acquirer and the Provider. This plan ties specific processes to specific requirements. The iteration between Acquirer and Provider concerns the finalization of processes for implementing requirements. The commitments documented in the SSMP relative to the decision for which the RISC is developed provide context for evaluating the RISC.

During the evaluation of the RISC, the Acquirer may request additional information to clarify or further substantiate elements in the RISC, e.g., via RAIs. Note that if the Provider does not satisfy the Acquirer on a given issue, the Evaluator may not be able to find the technical basis of the RISC to be adequate. Consequently, the Acquirer's assessment of the safety risk of the system may exceed the Acquirer's risk tolerance.

7.1.2 Guidance on Interfacing with the Provider to Facilitate Evaluation of the RISC

Guidelines on the processes of interfacing with the Provider with respect to the RISC were presented in Section 3.6 and in Tables 3-1 and 3-2. More detailed guidance on this subject has not been developed, since the interactive process by its nature must be flexible enough to accommodate new or unexpected findings.

7.2 Reviewing, Evaluating, and Scoring the RISC

7.2.1 The Acquirer's Responsibilities and Areas to Address

The responsibilities of the Evaluator in reviewing, evaluating, and scoring the RISC can be summarized as follows:

- Review the RISC submittal by the Provider to determine that all necessary material has been provided.

- Conduct a Qualitative Scope/Methods Review of the RISC submittal.
- Conduct a quantitative evaluation of the RISC submittal. Within a graded approach, the quantitative evaluation focuses on areas of highest impact to the safety of the system and those with high uncertainties. Nevertheless, the evaluation is intended to support a safety risk acceptance decision, and the coverage of the evaluation should therefore be sufficiently comprehensive to support that decision.
- As part of the quantitative evaluation of the RISC, conduct sensitivity studies for key parts of the RISC submittal (e.g., sensitivity to key assumptions and models).
- Rate the RISC as *Acceptable or Unacceptable* based on the evaluation findings.

The process of evaluating the RISC falls into three phases: 1) acceptance review of the submitted RISC; 2) a qualitative scope/methods review (surveying); and 3) a quantitative evaluation, including judiciously selected peer review calculations and sensitivity studies to spot-check the credibility of the Provider's results. Once an evaluation has been completed, the results of the RISC evaluation are communicated by the Evaluator to the Decision Maker in the form of a RISC Evaluation Report.

Acceptance review is a step undertaken to make sure that the information needed for substantive review is present in the submittal, and is intended to save resources (including time). During the acceptance review the submittal will be compared with expectations created earlier in the development process.

A qualitative scope/methods review (surveying) refers to a process of comparing RISC elements to standard methods and to other analyses. Examples include comparing methods and results of hazard analyses, methods and qualitative results of probabilistic risk analyses, and comparing assessed risk contributions between launch vehicles. In this step, the evaluators derive their own perspective on the results. Evaluators should also assess the Providers' qualifications including experience with the operation of the system.

The quantitative evaluation process consists of independent checking of selected technical results by the Acquirer, in order to confirm the Provider's results and to establish operationally that the Acquirer understands, in detail, what methods, data, assumptions, and models were used to obtain the results. Evaluation may range from "peer review" (high-level checking of selected results) up through replication of key results in the RISC. For example, the Evaluator may develop simplified analytic models to confirm the Provider's results.

Figure 7–1 shows the flow of the evaluation process. The following topics pertaining to the review, evaluation, and scoring of the RISC are discussed in Section 7.2.2 through 7.2.6:

- Composition of the evaluation team
- Sources of assurance deficit caused by incomplete or inaccurate evidence
- Ranking of the severity of the assurance deficits and their importance relative to the safety of the system based on the judgments of subject matter experts
- Experts' qualitative ranking of the overall confidence that the system is adequately safe
- Using Value-of-Information (VOI) methods to analyze options for reducing uncertainty

7.2.2 Composition and Independence of the Evaluation Team

The Acquirer's Evaluation Team should consist of a corps of experts that includes not only subject matter experts but also experts with broad experience and knowledge of the system as a whole and the risks that challenge the system. They should be independent from the Provider and from the Acquirer's decision making authority. Independence from the Provider is needed to avoid the possibility of conflicts of

interest. Independence from the Acquirer's decision making authority is desirable to ensure that the process is devoid as much as possible from political influences.

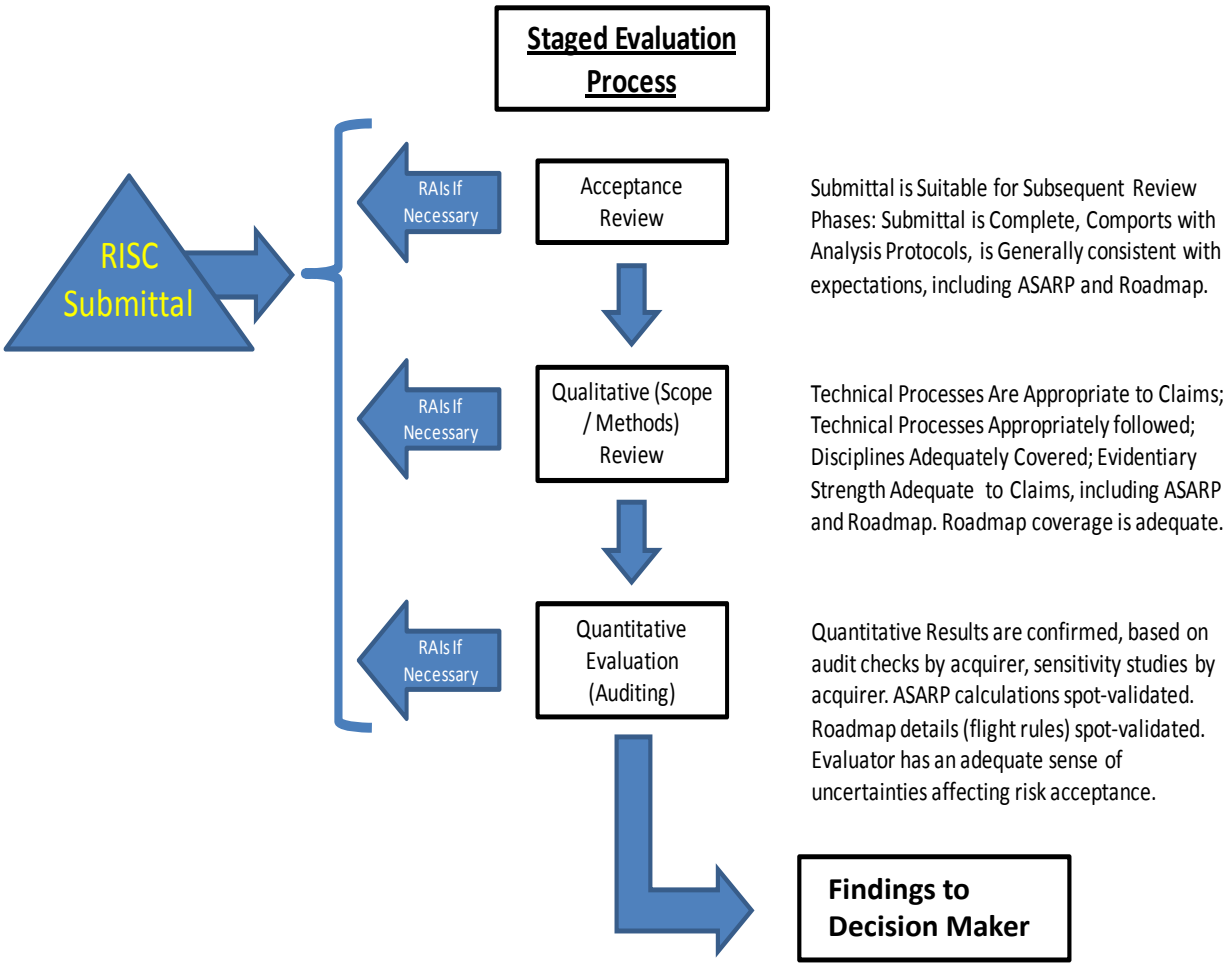


Figure 7-1. Phased Evaluation of the RISC

7.2.3 Sources of Assurance Deficit

As mentioned earlier, the strength of the evidence that supports a base claim is judged in terms of "assurance deficits," which were defined in Section 5.1.4 as "any knowledge gap that prohibits perfect (total) confidence" [66]. Assurance deficits relate to the sum of the evidence that is used to support a particular base claim. As suggested in Section 6.1.2, they can be ranked on a scale of 1 to 5 in terms of the confidence that the base claim is justified by the evidence.

Table 7-1, an extended version of Table 6-3, identifies some of the potential sources of assurance deficit that could apply to some of the base claims and associated evidence that were presented in the earlier table. This list is intended to be representative and is not intended to be an exhaustive list.

Table 7-1. Illustration of Assurance Deficit Sources that Pertain to Evidence for Various Potential Base Claims

Row No.	Example Base Claim	Example Direct and Supporting Evidence	Example Sources of Assurance Deficit
2	For each KMO, a margin has been established for the calculated probability of loss for a new system consistent with the understanding of UU risks from relevant previous missions.	**DIRECT:** Correlations of historical occurrences of failures and/or anomalies with qualitative factors that tend to produce UU risks. Identification, analysis, and ranking of such factors **SUPPORTING:** The rigor exercised in identifying the governing factors, quality of the data used to develop the correlations, standard errors of the correlations, degree to which the correlations have been verified and validated, experience and commitment of the analysts	Completeness of identification of the qualitative factors that tend to produce UU scenarios. Robustness of the correlation of historical failures/anomalies with the above qualitative factors; robustness of the analysis and of the ranking of the qualitative factors for the present program/project
4	Processes to ensure that requests for waivers or modifications of safety requirements are risk-informed.	**DIRECT:** Waiver/modification approach documented in SSMP; documented rationale based on correlation of safety requirements with risk drivers **SUPPORTING:** Demonstrated understanding of safety requirements and their relation to risk drivers; appreciation for potential UU risks; knowledge of historical precedence and best practices	Incomplete or inaccurate understanding of the risk drivers; inadequate appreciation for UU risks; lack of knowledge of historical precedence and best practices
6	Models used for high criticality risk scenarios within high criticality missions satisfy the Modeling & Simulation Credibility Assessment Scale (CAS) criteria and are backed by ISA-informed tests.	**DIRECT:** Comparisons of M&S results to an acceptable referent; comparison of input data with measured data; quantitative uncertainty estimates; repeatability of the M&S results; sensitivity of the M&S results to input and model parameters for the real-world system **SUPPORTING:** Use of reliable error estimation methods; community acceptance of the model as a *de facto* standard; availability of personnel with advanced engineering or science degrees or extensive work experience	Unfavorable comparisons of M&S results to an acceptable referent; unavailability of a suitable referent; unfavorable comparison of input data with measured data; lack of quantitative uncertainty estimates;. non-repeatability of the M&S results; lack of knowledge of the sensitivity of the M&S results to input and model parameters; lack of community acceptance of the model as a *de facto* standard; lack of a test plan that is realistically informed by the ISA models and results; unsuitable test facility or equipment; lack of personnel with needed educational and experience background

Table 7-1. Illustration of Assurance Deficit Sources that Pertain to Evidence for Various Potential Base Claims (Cont.)

Row No.	Example Base Claim (from Table 12-3)	Example Direct and Supporting Evidence (Synopsized from Table 12-3)	Example Sources of Assurance Deficit
7	Reasonable estimates of the uncertainty distribution and their correlation coefficients have been obtained.	**DIRECT:** Statistical analysis; applicability/completeness of testing and modeling **SUPPORTING:** Robustness of expert elicitation; qualifications of participants; documentation; use of qualified independent reviewers	Uncertainty associated with the analysts' and experts' understanding of the risk scenarios; uncertainty associated with differences between the experts on the uncertainty distributions for the significant known uncertainty sources and the correlations between them
8	Plans are in place for identifying/ evaluating the potential risk significance of precursors and anomalies, and for instituting appropriate contingencies and controls when needed.	**DIRECT:** Plans documented in SSMP for precursor analysis, anomaly and problem reporting, and implementation of corrective actions **SUPPORTING:** Quality and completeness of record-keeping of precursors and anomalies; staffing plan; experience of analysts	Uncertainty about the qualifications of the participants, the availability of resources, and the commitment of management
9	Budgets and schedules are adequate for the tasks to be performed and include realistic reserves.	**DIRECT:** Monte Carlo analysis of cost and schedule; budgets and deliverable dates consistent with analysis results; inclusion of realistic reserves based on historical experience **SUPPORTING:** Quality of analytical models; qualifications of analysts	Incomplete, biased, or inaccurate evaluation of costs and schedule; inadequate reserves; likelihood and severity of future budget cuts
11	Any changes from previous validated designs are small enough to remain within the domain for which the performance of the system is known.	**DIRECT:** Integrated safety analysis and testing of the final system over all mission parameter values **SUPPORTING:** Quality of analytical models and tests; qualifications of analysts	Insufficient testing or flight experience to demonstrate that the changes do not introduce new UU risks
12	Provisions are made for testing any new technology or application of an existing technology within the context of its interactions with the whole system.	**DIRECT:** Plans for integrated testing of the final system over all mission parameter values **SUPPORTING:** Evidence of sufficient budget and time to support such testing	Uncertainty about the likelihood and severity of budget cuts that could compromise the quality and scope of the testing

7.2.4 The Process of Ranking Assurance Deficits and Base Claim Importance

In Section 6.1.2, it was mentioned that assurance deficits can be scored by ranking the degree to which the assurance deficit sources affect the confidence of achieving the base claim that the evidence pertains to. In the illustration cited, Rank 1 was taken to imply a 95 to 100 percent level of confidence that the base claim is justified by the evidence, Rank 2 an 85 to 95 percent confidence, Rank 3 a 65 to 85 percent confidence, Rank 4 a 35 to 65 percent confidence, and Rank 5 a zero to 35 percent confidence. These high-end-weighted percentile values would likely apply for a mission with high criticality, whereas a more uniform distribution of percentile values (80 to 100 percent, 60 to 80 percent, etc.) would be more applicable for a lower criticality mission. The range of confidence percentiles corresponding to each rank is selected by the Acquirer.

Because of the qualitative nature of much of the evidence, assurance deficit rankings are unavoidably subjective. They should be based on the combined breadth and depth of knowledge and experience brought by the Acquirer's Evaluation Team. The Evaluation Team may access outside sources of information as appropriate to make these judgments.

The criticality of evidence is influenced by two factors: (1) the amount of assurance deficit in the evidence relative to acceptance of the base claim, and (2) the importance of the base claim relative to acceptance of the top claim (that the system is adequately safe).

The second of these two factors is also an item to be ranked by experts. A typical ranking scheme might be expressed as follow:

- Importance Rank = 1 implies that the base claim has a very small effect on the top claim, such that if the base claim were not satisfied, the effect on the confidence of the top claim would be less than 5%.
- Importance Rank = 2 implies that the base claim has a moderately small effect on the top claim, such that if the base claim were not satisfied, the effect on the confidence of the top claim would be around 5% to 15%.
- Importance Rank = 3 implies that the base claim has a mediocre effect on the top claim, such that if the base claim were not satisfied, the effect on the confidence of the top claim would be around 15% to 35%.
- Importance Rank = 4 implies that the base claim has a moderately large effect on the top claim, such that if the base claim were not satisfied, the effect on the confidence of the top claim would be around 35% to 65%.
- Importance Rank = 5 implies that the base claim has a very large effect on the top claim, such that if the base claim were not satisfied, the effect on the confidence of the top claim would be greater than 65%.

Again, the percentile range corresponding to each rank is selected by the Acquirer. As was the case for the deficit ranks, the experts' judgments on importance ranks are subjective but based on extensive knowledge and experience.

The experts are aided in their assessment of base claim importance by the claims tree, which shows the logic by which the base claims feed into higher level claims that ultimately feed into the top claim. Part of the assessment involves deliberations on whether or not the claims tree accurately and completely captures the top-down decomposition from the top claim down to the base claims. In addition to providing deficit and importance rankings, therefore, the experts' assessment should inform the Acquirer about the soundness of the logic in the claims tree and whether or not it is sufficiently inclusive to assess the confidence that the system is adequately safe.

7.2.5 Experts' Estimates of Overall Confidence

As mentioned above in Section 7.2.1, one of the responsibilities of the Acquirer (or an Evaluator appointed by the Acquirer) is to rate the RISC as *Acceptable* or *Unacceptable*. If a scoring scheme is used to accomplish this objective, the Acquirer should give the Provider direction regarding the approach prior to initiation of the RISC.

There are various means for scoring a safety case by propagating assurance deficits and importance measures through the claims tree, but all of them are open to the criticism that there are subtle interplays between the various claims that are mission-specific. Thus, there is no simple method of aggregation that can apply to all cases. However, it is instructive to discuss a variety of simple methods that can be used as rough indicators for various situations but with the caveat that their limitations be understood and appreciated. The next several paragraphs will look at the following four simple methods of approximation: (1) multiplying independent confidence levels together, (2) taking the weighted average of the confidence levels, (3) using the minimum of the confidence levels, and (4) multiplying weighted confidence levels together (weighted averaging of the logarithms). An overall recommendation will follow the exposition on these simplified methods.

Multiplying Independent Confidence Levels Together

Multiplying base-claim confidence levels together to aggregate them to a higher level is mathematically correct only when the base claims themselves are statistically independent. Given the large number of base claims that will typically exist in a RISC for a system of even moderate degree of complexity, it is virtually impossible to conjecture that all the base claims will be independent. For example, consider the following two base claims: (1) The system is appropriately modeled in the ISA, and (2) The ISA adequately addresses safety critical items. While both claims are appropriate in a RISC, they are not independent because the truth of Claim (1) is dependent on the truth of Claim (2). Put another way, both depend to some extent on the same evidence. The same is true for the following two base claims: (3) an effective quality control process is in place, and (4) an effective configuration control and change management process is in place. Claims (3) and (4) are not independent because both quality control and configuration control, while they may be under separate organizations at a lower level, are under the same management at the top level and both depend on the organizational management culture.

On the other hand, it *is* possible to estimate the confidence at the *base-claim level* by dividing the evidence that supports each base claim into independent attributes and multiplying the confidence in each attribute together. For example, the following three attributes for the evidence supporting a base claim can be considered to be independent:

1. The "credibility" of the evidence, denoted by symbol c. This is the level of assurance that the models and data used are accurate and comprehensive for the context in which they are being used.
2. The "applicability" of the evidence, denoted by symbol a. This is the degree to which the evidence is applicable to the system and mission being evaluated.
3. The "strength" of the result, denoted by symbol s. This is the level to which the probability of loss is claimed to be controlled on the basis of the evidence. (For example, if the evidence demonstrates that a failure rate is less than 10^{-3}, the strength of the evidence is $s = 10^{-3}$.)

Since the attributes are independent, it is possible to estimate the confidence in the base claim using the following equation:

$$C(B) = C(c)\ C(a)\ P(s)$$

where $C(B)$ is the confidence in the base claim, $C(c)$ is the confidence that the modeling and/or data are credible and comprehensive, $C(a)$ is the confidence that the modeling and/or data are applicable, and $P(s)$

is the probability that the probability of the base claim not being true is less than *s*, assuming the modeling and/or data are credible, comprehensive, and applicable.

It should be mentioned that multiplying confidence levels together assumes that a statement about the confidence level is equivalent to a statement of probability. The probabilistic interpretation of confidence (implied, for example, in Bayes' theorem) has for a long time been a subject of considerable debate.

Taking the Weighted Average of the Confidence Levels

When there are many base claims and it is clear that they are not mutually independent, it is not uncommon for people to assume that the confidence in the top claim can be estimated by taking a weighted average of the confidence levels for the base claims. The following equation expresses this process:

$$C(T) = \Sigma\, W_i\, C(B_i)$$

where $C(T)$ is the confidence in the top claim, and W_i is a weighting factor that in some sense represents the relative importance of Base Claim *i* in demonstrating the truth of the top claim. With insightful selection of the weighting factors, this process can be argued to be potentially applicable as an approximation only when no single base claim is essential to justify the top claim. Since the result (assuming that none of the weights is zero) is always higher than the minimum value of $C(B_i)$ and lower than the maximum value of $C(B_i)$, the weighted average does not allow for the possibility that any of the base claims could be a weak link which, if it were not true, would make the top claim not true.

Using the Minimum of the Confidence Levels

It is also not uncommon for people to estimate the confidence in the top claim by using the minimum value of the confidence levels for the base claims, as follow:

$$C(T) = \mathbf{min}\ [\ C(B_i)\ ]$$

This process can be applicable for trees where *all* of the base claims are potential weak links, such that a failure of any one of them to be true results in the top claim being untrue. It is commonly believed that taking the minimum base-claim confidence value provides a lower-bound estimate on the confidence of the top claim. However, this is not necessarily true. Taking the minimum implies that if the base claim with the lowest confidence level is untrue, it makes no difference whether any of the other base claims is also untrue. This may clearly be an optimistic assumption in many cases (e.g., cases for which two or more of the base claims are independent).

Multiplying Weighted Confidence Levels Together (Weighted Averaging of the Logarithms)

Still another approach is to estimate the confidence in the top claim by using a weighted average of the *logarithms* of the confidence levels for the base claims. The operative equation in this case is as follows:

$$\log\,[C(T)] = \Sigma\, W_i \log\,[C(B_i)]$$

Upon taking the anti-log of both sides, the result is as follows:

$$C(T) = \Pi\, [C(B_i)]^{Wi}$$

where the symbol Π denotes the product. This form has the advantage that when all of the weights are equal to 1.0, the result is mathematically correct form for the case where all the base claims are independent. Thus, the weights W_i are a measure of the degree of independence between base claims. The difficulty with this approach is in knowing *a priori* how to select values for W_i that accurately represent the degree of independence of the base claim relative to the other base claims.

Recommendation

In general, it is not fruitful to try to use a formulaic method for aggregating confidence levels from the base claims (where the evidence is) to the top claim (where the decision resides). Rather, the scoring of the RISC should be left to the experts and should be based on rational arguments that account for the uniqueness of and complex relationships within the mission. As mentioned in the preceding subsection, the corps of experts should include not only subject matter experts but also experts with broad experience and knowledge of the system as a whole and the risks that challenge the system.

Where justified, the experts should provide not only a two-tier ranking as suggested in Section 7.2.1 (acceptable and unacceptable), but also each expert should provide a numerical estimate of his/her confidence that the top claim has been demonstrated to be true (i.e., that the system is adequately safe). This estimate can be both subjective and rough. For example, one expert may conclude that he is 80% confident that the top claim has been proven true, while another may claim that she is 90% confident. Although the estimate is subjective, it should be backed by the expert's rankings of assurance deficit and base claim importance, and it should be accompanied by a set of arguments about how his/her confidence of the truth of the top claim is shaped by these rankings.

The action of providing a numerical estimate of confidence at the top claim level provides the following benefits:

- Top level confidence values calculated for one system can be compared to top level confidence values calculated for another system to compare which system is more robust from a safety viewpoint.
- The differences between the experts' confidence values provides the decision maker (in this case the Acquirer) with more information about the uncertainties associated with the safety case and how to interpret it.
- Inferences can be drawn from the experts' rationale about which base claims are driving the overall confidence, thereby providing a basis for selecting critical safety items.

A common challenge in expert judgment elicitation is what to do if the experts markedly differ from one another in their confidence assessment. A common approach when the elicitation results differ is to use the average of the results provided by all the experts participating in the elicitation (the so-called consensus approach or "delphi method"). However, the discipline of expert judgment elicitation provides for other possibilities, one of which is to drop the most pessimistic and optimistic responses (the outliers) before taking the average. There is a considerable litany of information on expert judgment elicitation that address questions like this [96].

If the experts' individual scores are widely divergent and a consensus cannot be reached, the team should attempt to agree on what remedies could be instituted so that the whole team would have high confidence that the system is safe. This information should be conveyed from the Acquirer to the Provider so that the Provider understands what needs to be done to obtain a favorable consensus from the members of the Evaluation Team. Such a consensus would help to convince the Acquirer's decision making authorities that the system successfully meets the safety requirements within the decision maker's risk tolerance.

7.2.6 Using Value-of-Information Methods to Analyze Options for Reducing Uncertainty

In essence, a decision maker's (DM) risk tolerance is an expression of the amount of uncertainty that the DM is willing to tolerate in regard to whether a safety performance requirement will be satisfied by the time the system becomes operational. For example, if a DM states at a particular point in time that he/she has a 10% risk tolerance for P(LOC) being greater than X, that means that he/she wants to be 90%

confident that P(LOC) will be less than or equal to X by the time of the first flight. In other words, he/she wants P(LOC) to be ≤ X with less than 10% uncertainty.

For practical purposes, DMs have three choices when told that the risk of not meeting a safety performance requirement exceeds their risk tolerance:

1. They may accept the risk, which means that they increase their risk tolerance.
2. They may mitigate the risk, which means that they accept the costs of mitigation.
3. They may seek to reduce the uncertainty by seeking more information, which means that they accept the costs of obtaining the required information.

The decision between these three choices is governed by several factors: (1) the amount of conviction behind the DM's risk tolerance (or technically speaking, his/her "utility" for deviation from the stated risk tolerance); (2) the expected cost of mitigating the risk, (3) the likelihood that the mitigation will be successful; (4) the expected cost of obtaining the new information; and (5) the likelihood that the new information will be sufficient to reduce the uncertainty by the amount needed. Determining the trade-off between these factors is an optimization problem that requires tradeoffs between a combination of objective and subjective factors.

Methodologically, such tradeoffs are classified as Value-of-Information (VOI) analysis. Summaries of VOI methodology and implementation principles are available in [97-99].

7.3 Documenting the Findings of the RISC Evaluation

7.3.1 The Acquirer's Responsibilities and Areas to Address

The responsibilities of the Evaluator in documenting the RISC evaluation findings can be summarized as follows:

- Develop and report a set of summary evaluation findings to the Decision Maker. The summary evaluation findings act as the primary results from the RISC evaluation process.
- Develop and report a set of detailed evaluation findings to the Decision Maker. The detailed evaluation findings act as the foundational findings from the RISC evaluation.
- At the conclusion of the RISC evaluation process, present the results of the RISC evaluation to the Decision Maker, in the form of a RISC Evaluation Report.

The following topics pertaining to the RISC evaluation findings and their documentation are discussed in Section 7.3.2 and 7.3.3:

- Contents of the summary and detailed RISC evaluation findings
- Contents of the RISC Evaluation Report

7.3.2 Contents of the Summary and Detailed RISC Evaluation Findings

Table 7-2 and Table 7-3 provide guidance for summary evaluation findings and detailed evaluation findings, respectively. Because of the diverse range of programs/projects and possible RISC structures, it is not possible to give an exhaustive accounting of findings that will apply in all situations, but Tables 7-2 and 7-3 provide a recommended, generic set for Acquirers to use as a starting point.

The Summary Evaluation Findings are an overview of the more detailed evaluation process findings, present a simplified view of the evaluation of the RISC, and enable the Decision Maker to quickly comprehend the results of the RISC evaluation, especially areas of the RISC that have been found to be inadequate. As with other recommendations in this handbook, this set is subject to tailoring or revision in

a specific application; but the expectation is that the Provider should understand *in advance* the findings that will be used during the RISC evaluation process.

7.3.3 Contents of the RISC Evaluation Report

As with the RISC Report submitted by the Provider (see Section 6.1.6), the RISC Evaluation Report should contain a standard set of information to be submitted by the Evaluator to the Acquirer's decision makers. The RISC Evaluation Report should include at least the following sections:

- Executive Summary
- Composition of the Evaluation Team
- Description of the process employed to evaluate the Provider's RISC
- Description of the methodologies used to score and rate the RISC
- RISC Evaluation Findings
- Documentation of weaknesses, limitations, and any open issues relating to the Provider's RISC

The following paragraphs provide guidelines for the type of information to be included in these sections.

Executive Summary

The executive summary of the RISC Evaluation Report should specify the context in which the RISC evaluation is performed, the review for which the report was prepared, key assumptions about the evaluation, and key conclusions including any open or outstanding issues.

Composition of the Evaluation Team

This section should list the members of the Evaluation Team and any non-team personnel consulted, their roles in the evaluation process, and their qualifications.

Description of the Processes Employed to Evaluate the Provider's RISC

In this section the various processes employed to evaluate the Provider's RISC should be discussed, including the following:

- The acceptance review process for the Provider's RISC to determine that all necessary material was provided
- A description of the process used to qualitatively scope and review the methods use in the RISC
- A description of the process used to undertake quantitative evaluations of the RISC and sensitivity studies, i.e., which areas of the RISC were selected for quantitative review/audit and why
- The process for initiating RAIs
- Overall RISC evaluation scoring process, e.g., number of reviewers assessing each finding, etc.

Description of the Methodologies Used to Score and Rate the RISC

This section should give a description of the methodologies used by the Evaluation Team to score and rate the RISC, including:

- The methodology used to score the RISC Summary Evaluation Findings and RISC Detailed Evaluation Findings, as mentioned in the first paragraph of this section
- The methodology to determine the overall rating of the evaluated RISC (Acceptable or Unacceptable)

RISC Evaluation Findings

This section details the findings of the Evaluation Team. While this handbook mandates no specific format, it is recommended that the evaluation findings be presented in a tabular format; similar to the format presented in Tables 7-2 and 7-3. The evaluation findings should include identification of and rankings for the key sources of assurance deficit[54].

Documentation of Weaknesses, Limitations, and any Open Issues Relating to the RISC

Examples of weaknesses, limitations, and/or open issues to be discussed in this section include weak sub-claims, sub-arguments, or evidence; unresolved issues from a previous evaluations, etc. Where appropriate, weaknesses, limitations, and open issues should be consistent with the assurance deficits identified and discussed in the RISC evaluation findings.

[54] Key sources of assurance deficit were defined in Section 6.1.2 and discussed in Sections 7.2.3 and 7.2.4.

Table 7-2. Guidelines for Summary Evaluation Findings

	Index	Primary Findings	Secondary Findings	Acceptable	Unacceptable
	0	The Risk-Informed Safety Case is adequate.	N/A	The RISC provides a clear case that the system is safe and is supported by sufficient, verifiable evidence to support this claim.	The RISC does not provide a clear case that the system is safe and/or is not supported by sufficient, verifiable evidence to support this claim.
Provider has Adequate Organizational Capability	**1.0**	Provider is Capable of Designing/Building/ Operating System.	N/A	The Provider has clearly demonstrated that they have the expertise, infrastructure (capital, hardware, supply chain, etc.), and plans in place to successfully design/build/operate system.	The Provider has not demonstrated that they have the expertise, infrastructure (capital, hardware, supply chain, etc.), and plans in place to successfully design/build/operate system.
System Meets or Exceeds Minimum Tolerable Levels of Safety	**2.0**	System meets Safety Performance Requirements.	2.1 System as designed, built, or operated, meets levied safety performance requirements.	The ISA clearly demonstrates that the system meets levied safety performance requirements.	The ISA does not show that levied safety performance requirements are met.
			2.2 The Provider has an adequate plan to ensure the system will continue to meet levied safety performance requirements.	The Provider's System Safety Management Plan (SSMP) clearly demonstrates that the system is on track to meet levied safety performance requirements.	The SSMP does not demonstrate that the system will continue to meet levied safety performance requirements.
System is ASARP	**3.0**	Provider Risk-Informs Decisions.	3.1 The Provider prioritizes safety during design/realization/operation decision making.	The Provider shows that risk management decisions are subject to clear and consistent heuristics/criteria that prioritize safety within cost, schedule, and technical constraints.	The Provider lacks clear and consistent decision guidance that prioritizes safety.
			3.2. The Provider risk-informs safety requirement allocation decisions.	The Provider shows that safety performance requirements are allocated into subsystems in a balanced manner supported by appropriate risk analyses.	The Provider has not shown that safety performance requirements are allocated into subsystems in a balanced manner supported by appropriate risk analyses.
			3.3. The Provider is responsive to new safety-relevant information during system realization/operation/sustainment.	The Provider shows that mechanisms are in place to identify safety relevant information internal and external to the system into the RM process.	The Provider has not shown that mechanisms are in place to identify safety relevant information internal and external to the system into the RM process

	Table 7-2. Guidelines for Summary Evaluation Findings (Cont.)				
	4.0	The Provider has Adequately Addressed Unknown and Underappreciated Risks.	4.1 The Provider has incorporated safety-related best practices and lessons learned into system design/realization/operation.	The Provider's SSRA clearly shows the application of safety-related best practices and lessons learned that are relevant to the specific system.	The Provider has not shown the application of safety-related best practices and lessons learned that are relevant to the specific system.
			4.2. The Provider minimizes the introduction of hazards during system realization and operation.	The Provider clearly demonstrates the application of QA, configuration management, and other safety management best practices during system realization and operation.	The Provider has not shown the application of QA, configuration management, and other safety management best practices during system realization and operation.
System Complies with Safety-Related Engineering and Process Requirements	**5.0**	The Provider Complies with Safety-Related Engineering and Process Requirements.	5.1 System as designed, built, or operated, meets levied safety-related engineering and process requirements.	The requirements verification process clearly demonstrates that the system meets levied safety-related engineering and process requirements.	The requirements verification process does not show that the system meets levied safety-related engineering and process requirements.
			5.2 The Provider has an adequate plan to ensure the system will continue to meet levied safety-related engineering and process requirements.	The Provider's SSMP clearly demonstrates that the system will continue to meet levied safety-related engineering and process requirements.	The Provider's SSMP does not show that the system will continue to meet levied safety-related engineering and process requirements.

Table 7-3. Guidelines for Detailed Evaluation Findings

1.0 The Provider is institutionally capable of providing the desired system/service (cross-cutting claim)				
Index	**Findings**	**Related Evaluation Objectives**	**Typical Evidence from Provider**	**Common Assurance Deficits**
1.1	Provider has sufficiently qualified personnel conducting safety-related activities.	Verify that Provider-conducted system safety activities are performed by qualified personnel.	Personnel resumes, training programs, etc.	Inadequately qualified or trained personnel, lack of employer provided training
1.2	Provider is able to sustain design/realization/operational support for the required duration.	Verify that the Provider organization has the capability to provide appropriate support for the duration of the acquisition.	History of organization, past performance on similar scale projects, training programs, etc.	Lack of success on similar types of programs/projects, little operating history, no past experience with program/projects of similar scale
2.0 The System Meets Safety Performance Requirements				
2.1 The system as designed, built, or operated, meets all safety performance requirements				
Index	**Findings**	**Related Evaluation Objectives**	**Typical Evidence from Provider**	**Common Assurance Deficits**
2.1.1	The Provider demonstrates that the system as designed, built, or operated, meets safety performance requirements.	1. Verify that analysis methods are appropriate to: - The specifics of the system - The point in the life cycle at which the analysis is being conducted (i.e., maturity of the system design, realization, or operation) - The project category (refer to NPR 7120.5) - The risk classification	All analysis methods used in the ISA (e.g., hazard analysis, FMEA, HAZOP, PRA, statistical analyses of comparative systems or subsystems, physics-based modeling, analysis of test results, etc.)	Analysis method used is inappropriate for the specifics of the systems or the point in the life cycle at which the analysis is being made.
		2. Verify that the identified scenarios meet expectations of comprehensiveness, considering: - System design (e.g., internal hazards, external hazards, system complexity) - Mishap history of similar systems - Operational loads on the system - Possible subsystem interactions - Extent of application of deterministic standards (which potentially remove issues from analytic consideration by virtue of overwhelming margin)	Scenarios identified from and analyzed in the ISA	Scenarios are not sufficiently comprehensive.

		3. Verify that models are consistent with the requirements of NASA-STD-7009 Standard for Models and Simulation, and that their credibility is acceptable.	The analytic models, and/or documentation of the models, data and methodologies used	Models are not sufficiently documented or do not comply with NASA-STD-7009.
		4. Verify that numerical values and uncertainties are reasonable (i.e., check for indefensibly low numbers, missing common cause dependencies, unsubstantiated crediting of functionality, narrow uncertainty ranges, etc.).	ISA analysis results	Analyses over-credit new technology, lack bounding analysis to "sanity" check results, use data from significantly different systems, or use inappropriate mathematical, numerical, or statistical methods.
		5. Verifying that modeled safety performance measures are not unduly sensitive to analysis assumptions (i.e., by performing targeted sensitivity analyses).	ISA analysis results	Sensitivity of safety performance results to credible regions of the parameter space
		6. Verify that safety-critical items have been properly identified, i.e., at a minimum, that the explicit and implicit safety functions documented in the ISA are designated as safety-critical.	ISA analysis results, safety-critical items lists	Items relied on for safety but not explicitly credited in the ISA (e.g., due to assumed high reliability) are not identified as safety-critical.
		7. Verify compliance with all requirements derived from the safety performance requirements (e.g., safety-related engineering requirements imposed on safety-critical items).	Requirements database, requirements verification matrix	Unverified requirements, inadequate adherence to verification procedures
		8. Verify compliance with all safety performance requirements according to established verification protocols (analysis protocols, confidence levels, etc.).	ISA analysis results	Analysis result does not meet safety performance requirement, or does not adequately address the verification protocol associated with it.

2.2 The Provider has an adequate plan to ensure that the system will continue to meet levied safety performance requirements during realization and/or operation

Index	Findings	Related Evaluation Objectives	Typical Evidence from Provider	Common Assurance Deficits
2.2.1	System safety performance will be adequately maintained.	1. Verify that safety-critical item capability, reliability, and availability are maintained at levels consistent with their assessed performance in the ISA.	Appropriate safety-related engineering requirements Appropriate safety-related process requirements (surveillance, testing, maintenance, etc.)	Plans and processes do not adequately demonstrate that system safety performance will be adequately maintained; lack of adherence to standards; inadequate surveillance activities

2.2.2	The ISA will be adequately maintained.	1. Verify that a process is in place to update the ISA in light of operating and test experience (e.g., anomaly resolution/precursor analysis).	ISA Development Plan and Roadmap; appropriate safety-related process requirements (e.g., Accident Precursor Analysis process requirements)	Poorly documented process for updating ISA, or ISA is not updated
		2. Verify that the ISA is updated as appropriate to reflect risk scenarios identified by the CRM process.	Updated ISA; RM process requirements	" "
		3. Verify that the ISA is updated to reflect system design and operation changes.	Updated ISA; configuration management process requirements	" "
		4. Verify that the ISA is updated to address any identified potential inadequacies in the ISA.	Updated ISA	" "
2.2.3	System safety performance will be improved consistent with safety growth requirements (i.e., from threshold to goal).	1. Verify that an adequate program of continuous safety improvement is being implemented at a level of effort/commitment consistent with safety growth requirements, e.g., safety performance projections are consistent with safety growth requirements.	Safety improvement process requirements	Inadequate staffing/funding; inadequate linkage of safety improvement activities to safety goals
3.0 The Provider Risk-Informs Decisions				
3.1 The Provider prioritizes safety during design/realization/operation decision making				
Index	**Findings**	**Related Evaluation Objectives**	**Typical Evidence from Provider**	**Common Assurance Deficits**
3.1.1	An effective RIDM process has been (and continues to be) used to identify and analyze credible decision alternatives.	1. Verify that decision alternatives include a comprehensive set of safety-promoting options, e.g.: - Margin - Simplicity - Testing - Technology maturation - Use of proven technology - Automation over human-in-the-loop - Inhibits - Fault management (redundancy, defense in depth, contingency, etc.)	ISA, RISRs, RMP, SSMP; RM process requirements	Consideration/analysis of only a narrow set of alternatives
		2. Verify the adequacy of the ISA to characterize the safety performance of each alternative for decision making purposes.	" "	Inconsistency in assumptions and/or analysis methods across the alternatives space; graded approach to analysis not used to focus on concerns that are determinative

		3. Verify that the ISA is being used to characterize the safety performance of each decision alternative, across all safety performance measures.	" "	ISA not used for early direction-setting decisions; ISA is not integrated into day-to-day systems engineering.
		4. Verify that relevant safety performance information is communicated to decision makers in a timely and effective manner.	" "	Lack of standardized and understood communication aids; system safety information lags behind systems engineering decisions.
3.1.2	Decisions made are in the interests of safety performance, within technical, cost, and schedule constraints.	1. Verify that decision rationales are documented to a level appropriate to the significance of the decision.	ISA, RISRs, RMP, SSMP; RM process requirements	Inadequate documentation of rationale; decisions driven by management pressure
		2. Verify that safety adequately drives decision making, as documented in the decision rationales.	" "	" "

3.2 The Provider risk-informs safety requirement allocation decisions

Index	Findings	Related Evaluation Objectives	Typical Evidence from Provider	Common Assurance Deficits
3.2.1	The Provider risk-informs safety requirement allocation decisions.	1. Verify that a process is in place that assures that the ability of subordinate organizations to meet allocated safety requirements is consistent with the safety performance risk tolerances of the allocating organization.	ISA, RMP, SSMP; RM process requirements	ISA not used to allocate safety requirements; lack of margin over and above ISA results to account for UU hazards

3.3 The Provider is responsive to new safety-relevant information during system realization/operation/sustainment

Index	Findings	Related Evaluation Objectives	Typical Evidence from Provider	Common Assurance Deficits
3.3.1	The Provider actively seeks safety-relevant information from the system.	1. Verify that a test program actively seeks to improve safety (e.g., testing is risk-informed).	Test plans and process documentation	Poorly documented test plans, or test program does not adequately seek to improve safety.
		2. Verify that the system is instrumented/monitored/inspected adequately to reveal safety-relevant information.	Design specs, ISA	ISA not used to risk-inform instrumentation/monitoring decisions; opportunities to instrument/monitor against UU hazards are not taken.

		3. Verify that a process is in place that responds effectively to safety-relevant test and operation information, e.g., anomaly resolution/precursor analysis.	Safety analysis; appropriate safety-related process requirements (e.g., Accident Precursor Analysis process requirements	Process is beholden to schedule pressure; normalization of deviance
		4. Verify that an effective CRM process is in place (per NPR 8000.4A and the NASA RM Handbook).	RMP, RM process requirements.	RM disconnected from SE
3.3.2	Provider actively seeks safety-relevant information from sources external to the system.	1. Provider is responsive to safety-relevant information from industry sources, e.g., Government-Industry Data Exchange Program (GIDEP). 2. Provider is responsive to safety-relevant technology developments.	Documentation of active program to seek out and utilize safety-relevant information from external sources; appropriate safety-related process requirements	Lack of documented program, or poorly documented program, to seek out and utilize safety-relevant information from external sources
4.0 Provider has Adequately Addressed Unknown and Underappreciated Scenarios				
4.1 The Provider has incorporated safety-related best practices and lessons learned into system design/realization/operation				
Index	**Findings**	**Related Evaluation Objectives**	**Typical Evidence from Provider**	**Common Assurance Deficits**
4.1.1	Provider incorporates safety-related best practices and lessons-learned into system design/realization/operation.	1. Verify that decision alternatives include a comprehensive set of safety-related best practices, e.g.: - Margin - Simplicity - Testing - Technology maturation - Use of proven technology - Automation over human-in-the-loop - Inhibits - Fault management (redundancy, defense in depth, contingency, etc.)	Documentation showing a comprehensive review and application of relevant best practices throughout the system; associated safety-related engineering and process requirements	Poorly documented, or incomplete review of best practices; inadequate adoption of identified best practices
		2. Verify that the ability of each decision alternative to address unknown and underappreciated sources of safety risk is considered and communicated to decision makers in a timely and effective manner.	Documentation of process used to consider and communicate UU risks to decision makers (e.g., technical bases for deliberation)	Poorly documented, or inadequate processes for communicating the ability of the decision alternative to address UU safety risks
		3. Verify that there is a process to identify safety-critical items and preserve their safety function during operation/sustainment, e.g., via: - Appropriate deterministic standards, e.g., engineering standards - Appropriate surveillance, testing, maintenance, etc.	ISA, safety-critical items lists, appropriate safety-related engineering and process requirements	Safety-significant assumptions are made in the ISA that are not translated into SCIs (e.g., the functioning of a component whose high reliability depends on effective safety management).

4.2 The Provider has minimized the introduction of hazards during system realization and operation				
4.2.1	The Provider has minimized the introduction of hazards during system realization and operation.	1. Verify that effective quality assurance (QA) processes are in place, e.g.: - Qualification testing - Process inspections - Acceptance testing	QA plans and process documentation; associated safety-related process requirements	Inadequate QA plans and processes
		2. Verify that effective configuration control (CC) and data management are in place.	Documentation of CC and data management processes, protocols, and system; documentation of compliance with process standards such as ISO 9000, CMMI, etc.	CC processes are inadequate, poorly documented, or not followed; lack of compliance with industry standards such as ISO 9000 or CMMI
5.0 The Provider Complies with Safety-Related Engineering and Process Requirements				
Index	**Findings**	**Related Evaluation Objectives**	**Typical Evidence from Provider**	**Common Assurance Deficits**
5.0	The Provider Complies with Safety-Related Engineering Requirements.	1. Verification audit of selected safety-related engineering requirements	Verification matrices	Unverified requirements; inadequate adherence to verification procedures
5.1	The Provider Complies with Safety-Related Process Requirements.	1. Verification audit of selected safety-related process requirements	SSMP, RMP, RM Process requirements	Disconnect between processes as documented and processes as implemented

7.4 Example for Chapter 7 – RISC Evaluation for New Technology on a Robotic System (Electric Ion Thruster)

This section continues the example developed in Section 6.6. The RISC presented in that section pertained to the life expectancy of the ion electric thrusters that are to be used as part of a robotic interplanetary exploration system. Section 6.6 provided a statement of the problem, presented a safety claims tree, identified the sources of evidence, and discussed suppositions, assumptions, and contexts associated with the mission and the evidence needed to substantiate the RISC. This section identifies the sources of assurance deficit in that evidence, addresses the ranking of those assurance deficits, and discusses the experts' estimates of confidence that the risk of premature wearout is sufficiently low.

7.4.1 Identification and Analysis of Assurance Deficit Sources

Based on the information provided in Section 6.6, an analysis has been performed to identify the potential sources of assurance deficit for the example, and to rank them in terms of their potential effect on the confidence that the electric thrusters will last for their intended lifetime. Table 7-4 illustrates what the results of such an analysis might look like.

7.4.2 Experts' Assessment of Overall Confidence

Having developed the information in Table 7-4, each expert is now asked to provide his/her overall estimate of confidence that the electric ion thrusters will meet their lifetime requirements before wearing out. Example estimates are as follows:

- Experts 1 and 2 believe that the overall confidence ranking should be around 90%. They reason that the analyses employ a large amount of conservatism and the design margins for the thrusters are large. The fact that the models have not been validated in real world conditions for the actual hardware is compensated by the fact that the amount of testing on analogous hardware in a variety of environments is rather extensive.
- Expert 3 believes the overall confidence ranking should be around 50%. He believes that a higher confidence ranking would not be justified until there is operational data in environments that stress the system to its maximum. He believes that while there is some redundancy from the fact that there are more ion thrusters than needed to meet the minimum success criterion, the protection against wearout failures would be greater if some of the redundant thrusters were normally kept in a standby mode. The current design calls for all of them to be active. Furthermore, he believes that wearout data from operational military fighter aircraft should have been considered to provide perspective on the likelihood of UU wearout failure modes.

7.4.3 Request for Additional Information

Because of the concerns of Expert 3, an RAI is issued asking the Provider to explore the following possibilities and obtain the following additional evidence:

- Perform tests at beyond-design-basis environmental conditions to explore the possibility of unknown nascent wearout failure modes.
- Determine the feasibility of keeping redundant thrusters in standby mode during normal operation and isolating them from the active thrusters.
- Explore whether wearout experience from military fighter aircraft operation is relevant to the determination of margins to prevent UU wearout failure modes in ion thrusters.

The response to this RAI will enable the Evaluation Team to form a consensus on whether the system is adequately protected from wearout failure.

Table 7-4. Example Analysis of Assurance Deficits Affecting Confidence in Long-Term Thruster Operation.

Item	Source of Assurance Deficit	Hypothetical Ranking	Rationale
1	The testing time for *individual thrusters* may not have been sufficiently long to establish that there are no unexpected wearout failure modes that could occur toward the end of life.	2	Bounding analysis using Weibull maximum likelihood estimation indicates likelihood to be less than 0.20.
2	The number of life tests may not have been sufficiently large to rule out the possibility of a wearout failure mode not having been observed due to the luck of the draw.	2	Binomial statistical analysis indicates probability to be 10% at a 95% confidence level.
3	The environments for life testing may not have simulated the environments that will be seen during actual operation closely enough (e.g., the combination of zero-gravity, vacuum, radiation, and orientation of the thrusters to the direction of acceleration).	3	Based on expert judgment elicitation considering the results of testing on similar systems and components in a variety of environments
4	The item may not have been life-tested at conditions beyond the design basis environments to establish its robustness in the event of abnormal/unexpected environments.	2	Based on expert judgment elicitation considering the historical record concerning occurrences of unanticipated environments and the amount of margin in the thruster design
5	The testing in the *full-up system configuration* may not have covered the full range of operational time and environments.	3	Based on expert judgment elicitation considering the historical record concerning occurrences of unexpected interactions
6	Some relevant sources of wearout data many not have been examined and included in the analysis where relevant.	3	Wearout data for DoD military systems have not been examined.
7	Impingement of erosion products on other subsystems may cause adverse cross-system effects.	2	Significant grid erosion would lead to ion thruster failure before failure of other subsystems.
8	Control logic errors may cause deleterious cycling of the ion thrusters.	2	The control logic has been tested thoroughly.
9	Margins on environmental and design parameters are insufficient.	2	Experts agree that the margins are reasonable and sufficient.
10	Redundant ion thrusters intended for failure tolerance may be subject to wearout if used actively during the mission or collocated with active thrusters.	4	All of the thrusters are active during the mission and are collocated.

Table 7-4 (Cont.). Example Analysis of Assurance Deficits Affecting Confidence in Long-Term Thruster Operation.

Item	Source of Assurance Deficit	Hypothetical Ranking	Rationale
11	The failure models for electron back-streaming due to accelerator grid erosion and the failure rate statistical models in general may not successfully satisfy the M&S Standard verification criteria.	1	The tools have been verified to be free of error.
12	The failure models for electron back-streaming and the failure rate statistical models in general may not successfully satisfy the M&S Standard validation criteria.	3	Cannot validate completely because there is no real world or experimental data leading to failure for the present hardware.
13	The failure models for electron back-streaming and the failure rate statistical models in general may not successfully satisfy the M&S Standard criteria for people qualifications or for management oversight.	1	The analysts have adequate experience with failure analysis and the use of these tools, and the management oversight is sound.
14	The failure models for electron back-streaming may not successfully satisfy the M&S Standard criteria for input pedigree, results uncertainty, and results robustness.	2	The input data agree well with real world data, the uncertainty that these calculations are bounding is very small, and the sensitivity of the results to parameter variations has been well established.
15	The failure models for electron back-streaming may not successfully satisfy the M&S Standard criteria for use history.	3	The models have been used many times but not for the same or similar applications.

7.5 Generic RISC Evaluation Tree

This section presents an overview of a generic evaluation tree that the Acquirer's Evaluation Team can use to rate their confidence in the Provider's RISC. The evaluation tree in this section has the same form as a safety claims tree that a Provider might produce, but the claims on it are more generic and not specific to a particular program or project. The intent is to facilitate the Evaluator's task of determining where there may be gaps in the logic of the Provider's RISC or where there may be significant assurance deficits in the evidence presented by the Provider.

In addition to being an aid to the Evaluator, the generic RISC evaluation tree in this section could serve as a starting point for the Provider in developing a system safety claims tree for the system that the Provider is developing. If the Provider were to use this tree as a starting point, some tailoring of the tree would probably be necessary to make it specifically applicable to the Provider's system and suitable to the Provider's needs.

The Evaluation Tree uses a Goal Structuring Network (GSN) format, for which the relevant symbology is summarized in Figure 7-2. The Evaluation Tree itself is presented in Figures 7-3 through 7-15.

The following paragraphs provide a brief verbal summary of the tree.

The Overall Safety Argument for the System

The overall safety argument for the system (Figure 7-3) follows the top-level safety objectives from the operational objectives hierarchy presented earlier in Figure 3-2, specifically demonstrating that the system is safe through providing argument and evidence that:

1) "The system meets the minimum tolerable level of safety for known risks and will continue to do so" (G1), and
2) "The system is (and will continue to be) As Safe as Reasonably Practicable" (G2).

These top-level claims are then decomposed further to the point that evidence for their truth can be provided.

Claim G1: The System Meets the Minimum Tolerable Level of Safety

Figure 7-4 sets out the arguments and evidence that Claim G1 is met. The arguments for satisfaction of Claim G1 include:

1) The ISA demonstrates that the system meets all relevant safety performance requirements (G1.1).
2) The SSMP demonstrates that the system will continue to meet the minimum tolerable level of safety in the future (G1.2).

Sub-Claim G1.1: The ISA demonstrates that the system meets all relevant safety performance requirements

The arguments leading from sub-claim G1.1 down to the level where evidence is expected are presented in Figures 7-5 through 7-7. The base claims that eventually flow down from G1.1 in this set of figures are listed below:

1) ISA results show that system safety performance is within levied safety performance requirements (G1.1.1).
2) The system is correctly modeled in the ISA (G1.1.2.1).
3) The ISA is scenario-based and adequately accounts for all credible failure scenarios (G1.1.2.2).

4) The ISA is appropriately scoped (G1.1.2.3).

5) The tools and techniques used in the ISA are appropriate to the life-cycle phase of the system (G1.1.2.4).

6) Uncertainty and sensitivities have been adequately addressed in the ISA (G1.1.2.5).

7) The ISA is consistent with analysis procedures (G1.1.2.6).

8) The ISA adequately addresses Safety Critical Items (G1.1.2.7).

9) The ISA is consistent with applicable safety-related test results (G1.1.2.8).

10) All derived safety-related engineering requirements are met (G1.1.4).

NOTE: The next group of base claims (#11 through #22) descends from intermediate goal G1.1.2.8: ISA is consistent with applicable safety-related test results.

11) The system meets safety-related test requirements (G1.1.2.8.1.1).

12) Techniques used to analyze and report test results are correct and applied appropriately (G1.1.2.8.2.1).

13) Tests are appropriate to the life-cycle phase of the system (G1.1.2.8.2.2).

14) Tests have been designed to test, as closely as possible, the actual operating environment, including extremes of the performance envelope and off-nominal situations (G1.1.2.8.2.3).

15) Testing is appropriately scoped (G1.1.2.8.2.4).

16) Uncertainty and sensitivities have been adequately addressed in the test results (G1.1.2.8.2.5).

17) Test results are up-to-date, and for the current version of the system (G1.1.2.8.3.1).

18) Tests have been adequately documented (G1.1.2.8.3.2).

19) Models and simulations used to analyze test results adhere to the NASA Modeling and Simulation Standard (NASA-STD-7009) (G1.1.2.8.3.3)

20) An effective Configuration and Change Management process is in place for test results (G1.1.2.8.3.4).

21) Tests have been performed by appropriately qualified personnel (G1.1.2.8.3.5).

22) Bounding analyses using historical and similar systems are consistent with the safety performance claims derived via testing (G1.1.2.8.3.6).

NOTE: The next group of base claims (#23 through #28) descends from intermediate strategy S1.1.3: Argue that the ISA has been performed properly and is trustworthy.

23) The ISA is current and accurately represents the system (G1.1.3.1).

24) The ISA has been adequately documented (G1.1.3.2).

25) Models and simulations used in the ISA adhere to the NASA Modeling and Simulation Standard (NASA-STD-7009) (G1.1.3.3).

26) An effective Configuration and Change Management process is in place for the ISA (G1.1.3.4).

27) The ISA has been performed by appropriately qualified personnel (G1.1.3.5).

28) Bounding analyses using historical and similar systems are consistent with the safety performance claims (G1.1.3.6).

Sub-Claim G1.2: The SSMP demonstrates that the system will continue to meet all relevant safety performance requirements in the future

The arguments leading from sub-claim G1.2 down to the level where evidence is expected are presented in Figures 7-8 and 7-9. The evidence is expected to address the following base claims:

1) The SSMP provides the link between baselined safety requirements and system safety activities (G1.2.1.1).
2) The SSMP adequately details the specific actions and arrangements required to operate the System Safety Program and defines system safety milestones for the project (G1.2.1.2).
3) The SSMP has been adequately reviewed and updated regularly throughout the current life-cycle phase (G1.2.1.3).
4) The SSMP adequately delineates the framework for the Provider's organization to direct and control its safety management activities, including the organizational structure, processes, procedures, techniques, and methodologies (G1.2.1.4).
5) The SSMP describes, in appropriate detail, plans for development of the RISC for each major milestone review (G1.2.1.5).
6) The SSMP adequately addresses the activities necessary to ensure safety throughout the system life cycle (G1.2.1.6).
7) The SSMP is current and accurately represents the system (G1.2.2.1).
8) The SSMP has been adequately documented (G1.2.2.2).
9) An effective Configuration and Change Management process is in place for the SSMP (G1.2.2.3).
10) The SSMP has been performed by appropriately qualified personnel (G1.2.2.4).

Claim G2: The System is ASARP (As Safe as Reasonably Practicable)

Figure 7-10 sets out the arguments and evidence that Claim G2 is met. The arguments for satisfaction of Claim G2 include:

1) Best practices and lessons learned have been incorporated into the system (G2.1).
2) The introduction of hazards has been minimized (G2.2).
3) Safety has been prioritized during decision making (G2.3).
4) Safety requirements have been risk-informed (G2.4).
5) New safety-related information is adequately addressed (G2.5).
6) Organization functions effectively and prioritizes safety (G2.6).

The arguments leading from claim G2 down to the level where evidence is expected are presented in Figures 7-11 through 7-15. The evidence is expected to address the following base claims:

1) Decision alternatives include a comprehensive set of safety-related best practices (G2.1.1).
2) A process is in place to ensure that lessons learned are effectively incorporated into the system (G2.1.2).
3) A process is in place to actively seek and respond to safety-relevant information from the system (G2.5.1).

4) A process is in place to actively seek and respond to safety-relevant information from sources external to the system (G2.5.2).

5) Design complexity has been minimized (G2.2.1).

6) An effective quality control process is in place (G2.2.2).

7) An effective configuration control and change management process is in place (G2.2.3).

8) A process is in place to effectively monitor and manage SCIs (G2.2.4).

9) A system is in place to monitor, report, and correct anomalies and precursors (G2.2.5).

10) An effective CRM process is in place (G2.2.6).

11) Plans have been correctly implemented and adhered to (G2.2.7).

12) An adequate roadmap for achieving safety performance during the remainder of the life cycle is in place (G2.2.8).

13) The design is kept within the validated domain (G2.2.9).

14) Budgets and schedules are realistic (G2.2.10).

15) An effective RIDM process is used to identify and analyze credible decision alternatives (G2.3.1).

16) Decisions made are in the interests of safety performance, within technical, cost, and schedule constraints (G2.3.2).

NOTE: The next several claims (#17 through #22) pertain to the risk of not satisfying derived and allocated requirements. The evidence presented to support the claims associated with integrated analysis will include including engineering studies, detailed probabilistic safety analyses, test data, etc., that are part of the ISA. The evidence for demonstrating adequate margins for these types of claims associated with derived requirements will include similarity analyses, applicable and tailored best practices and lessons learned from relevant systems, and analyses of historical systems contained in the ISA. The evidence for demonstrating adequate assessment of the risk of qualitative or process related requirements will include the SSMP, the RM Plan, and RM process requirements.

17) Integrated analysis has been used to assess the known risks associated with derived requirements (G2.4.1.1).

18) Sufficient margin has been included in the derived requirements to account for UU risks (G2.4.1.2).

19) A process is in place to adequately assess the risk of qualitative or process related derived requirements (G2.4.1.3).

20) Integrated analysis has been used to assess the known risks associated with allocated requirements (G2.4.2.1).

21) Sufficient margin has been included in the allocated requirements to account for UU risks (G2.4.2.2).

22) A process is in place to adequately assess the risk of qualitative or process related allocated requirements (G2.4.2.3).

NOTE: The last several base claims (#23 through #27) descend from intermediate claim G2.6: The organization functions effectively and prioritizes safety.

23) Management clearly demonstrates a commitment to safety through policies that promote safety (G2.6.1).

24) Organization is motivated and communicates safety-related concerns effectively (G2.6.2).

25) Organization is involved in safety-related concerns and is committed to safety (G2.6.3).

26) Roles and responsibilities have been clearly defined and are adequate to ensure safety (G2.6.4).

27) Effective oversight of subcontractors is maintained (G2.6.5).

Symbol	Meaning
G1 Claim	Claim for the safety argument
S1 Strategy	Strategy or solution for the argument
Context	Context for a claim
E1.2 Evidence	Evidence to support the claim
	Argument continued from above
	Argument completed below
	Argument yet to be completed

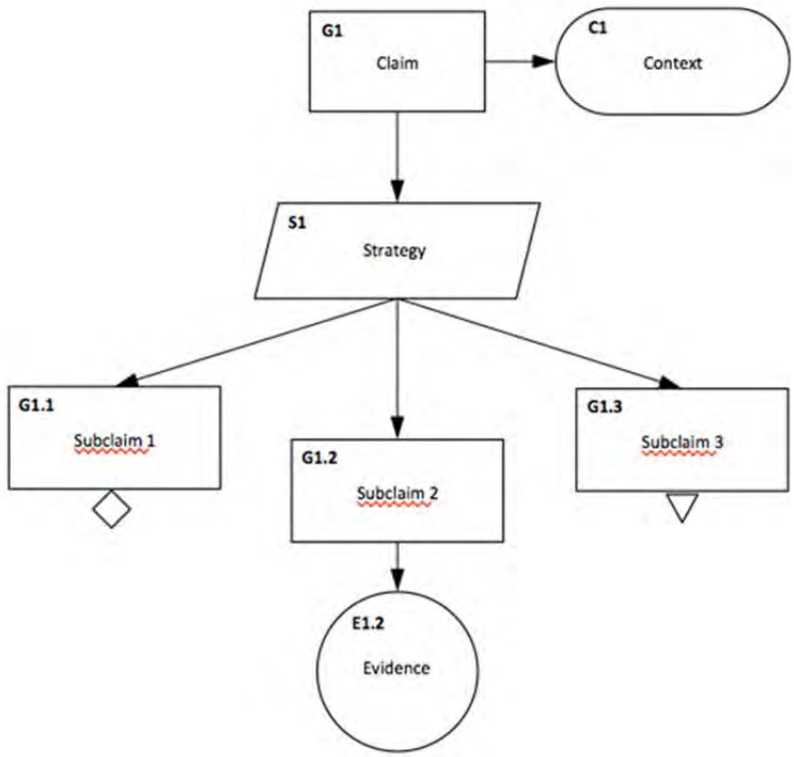


Figure 7-2. Illustration of Symbols Used in Goal Structuring Notation

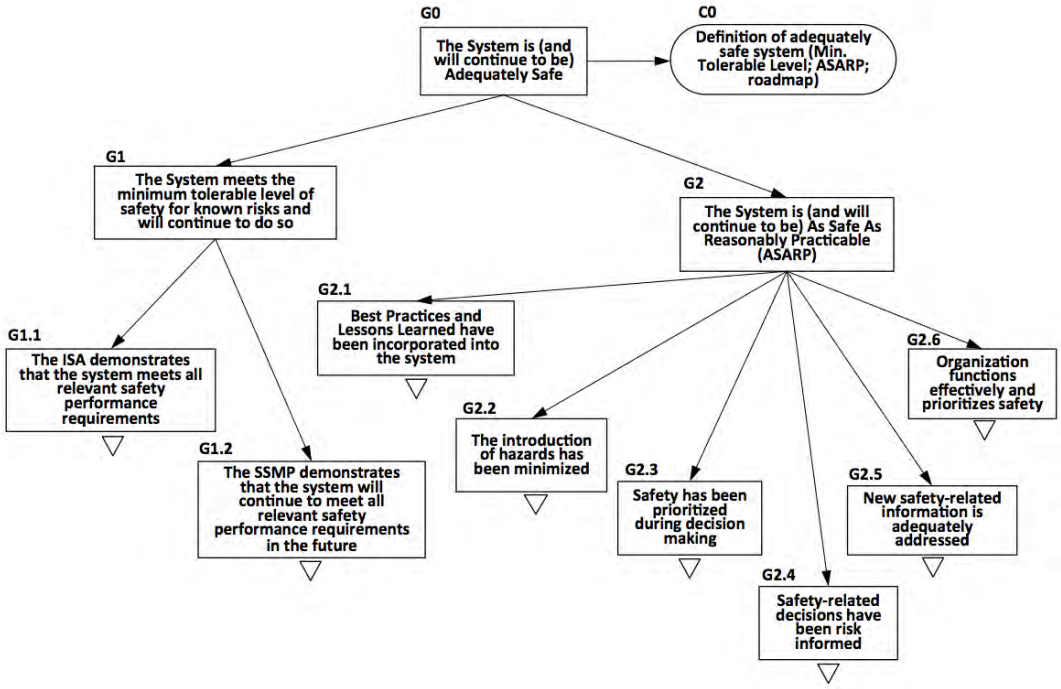


Figure 7-3. Overall Safety Argument

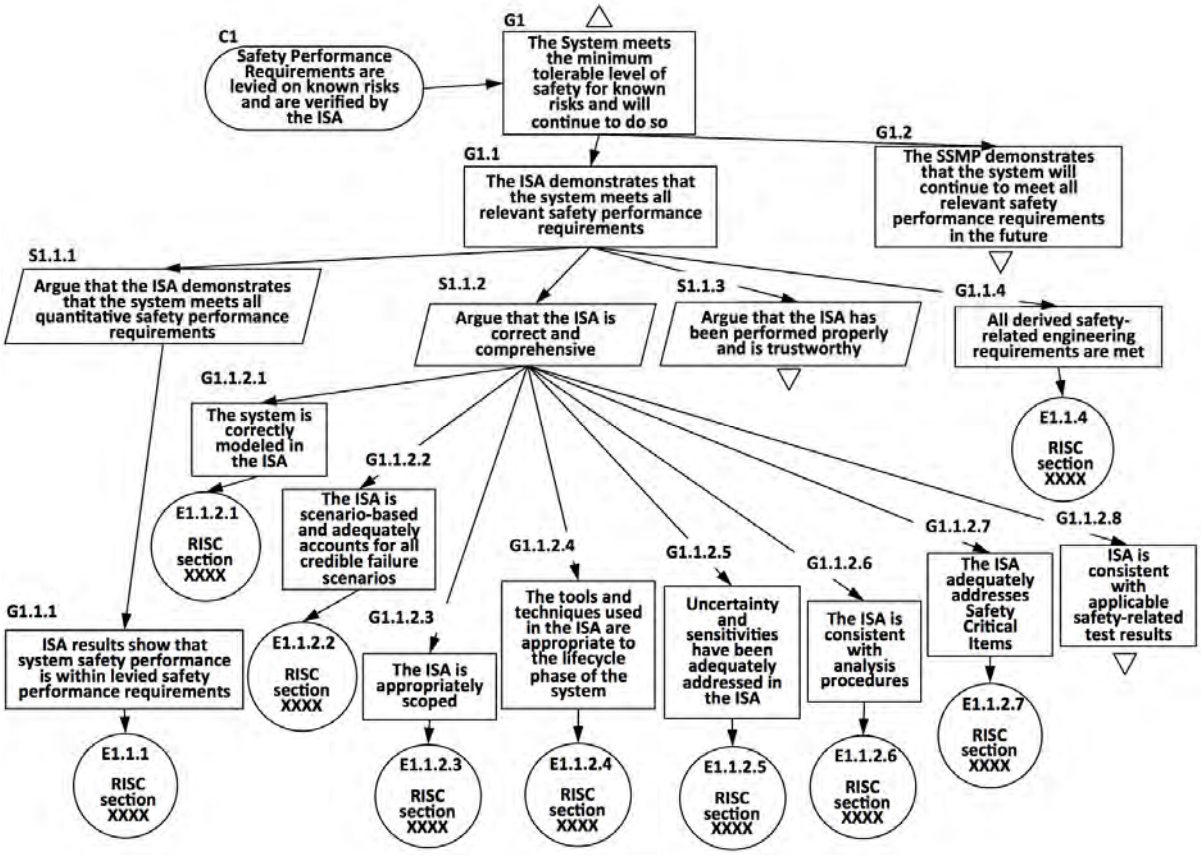


Figure 7-4. The System Meets the Minimum Tolerable Level of Safety for Known Risks

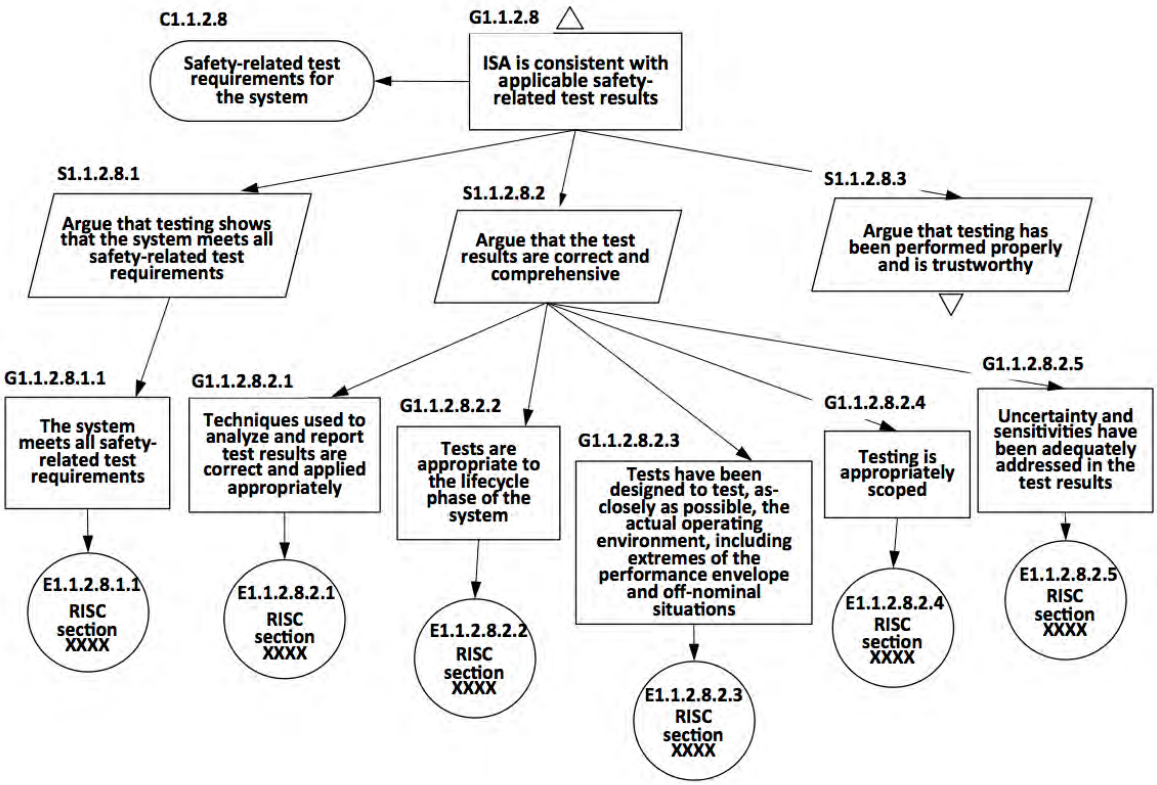


Figure 7-5. ISA is Consistent with Applicable Safety-related Test Results

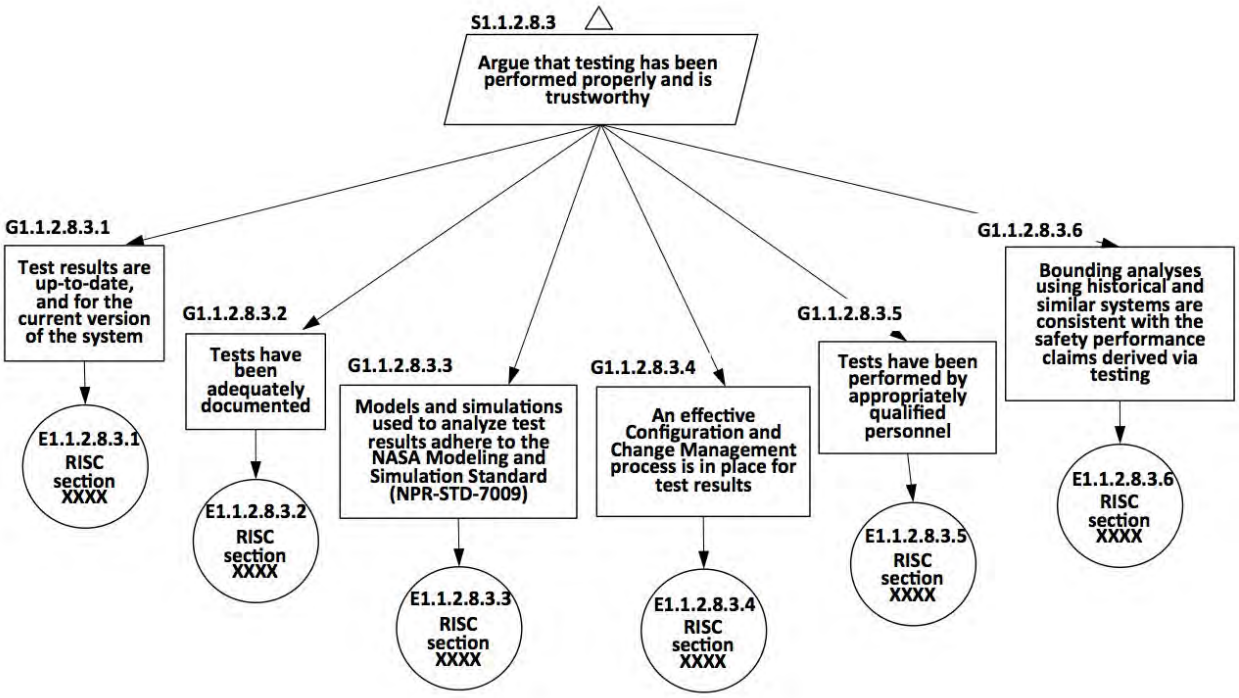


Figure 7-6. Argue that Testing Has Been Performed Properly and Is Trustworthy

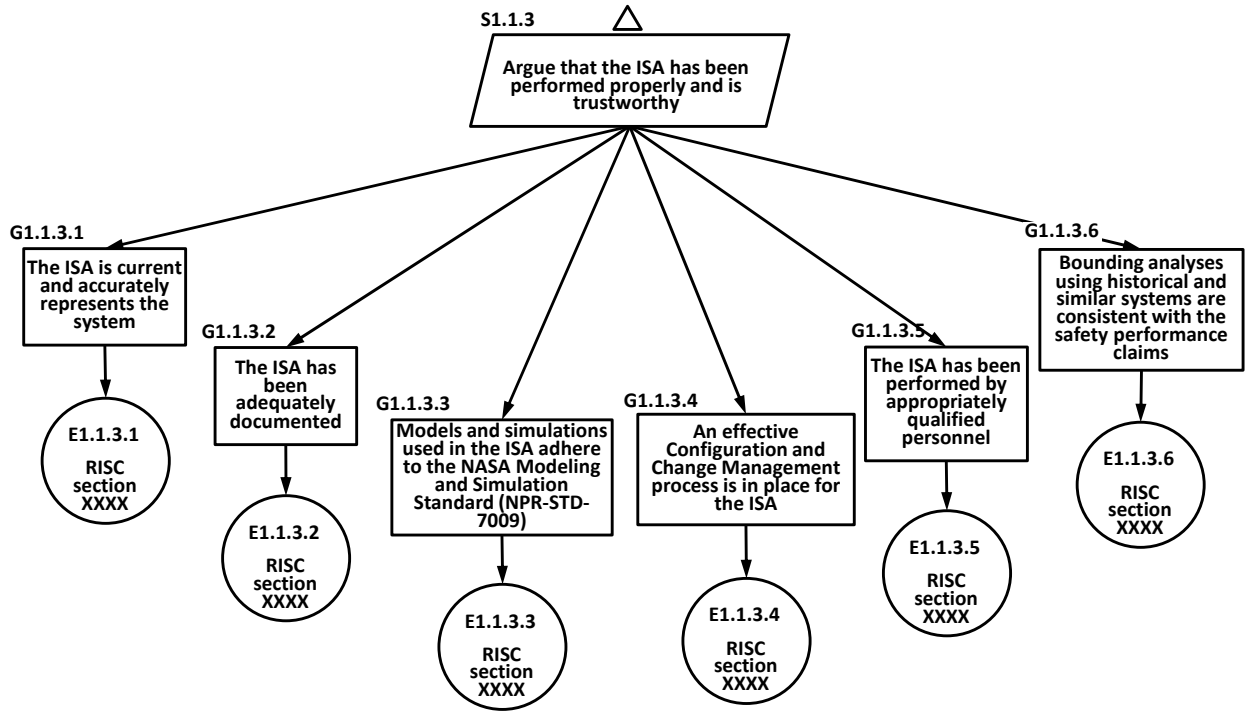


Figure 7-7. Argue that the ISA Has Been Performed Properly and Is Trustworthy

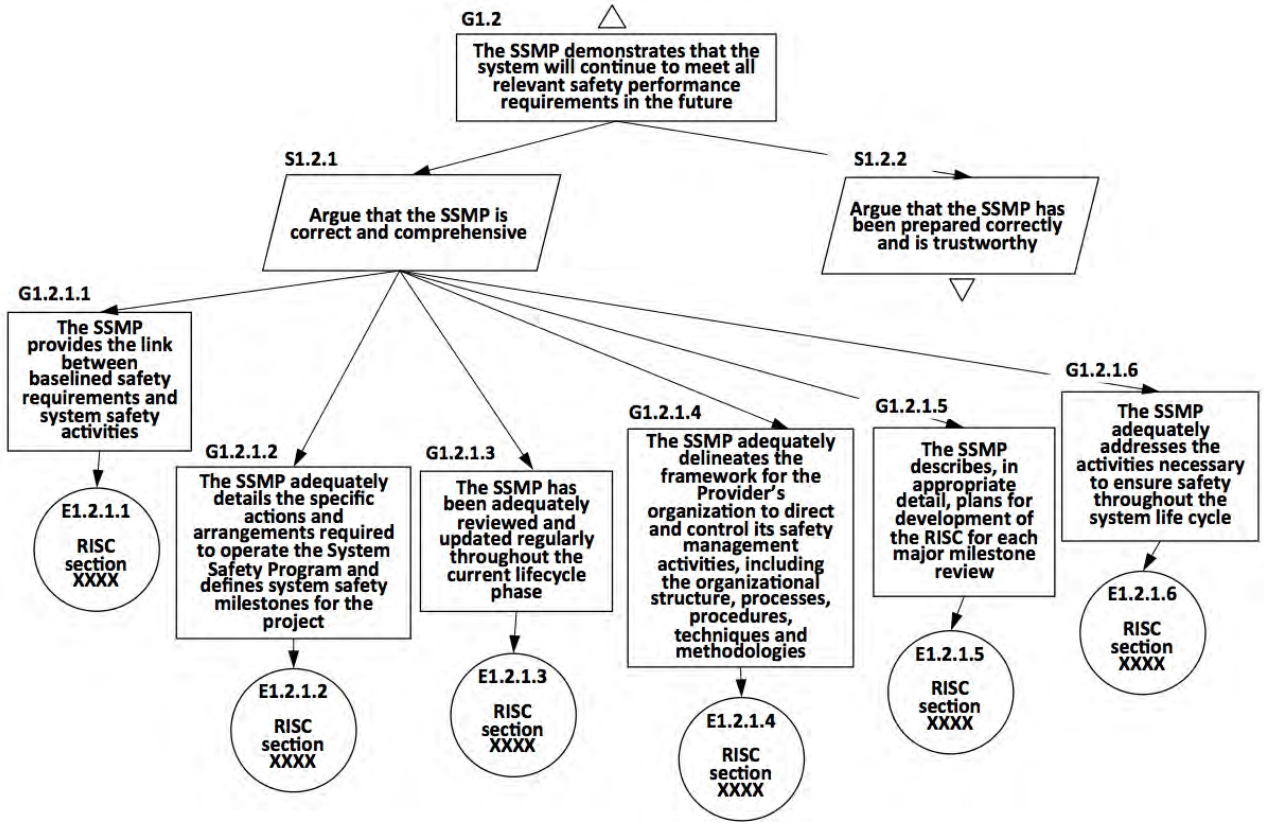


Figure 7-8. The SSMP Demonstrates that the System Will Continue to Meet all Relevant Safety Performance Requirements in the Future

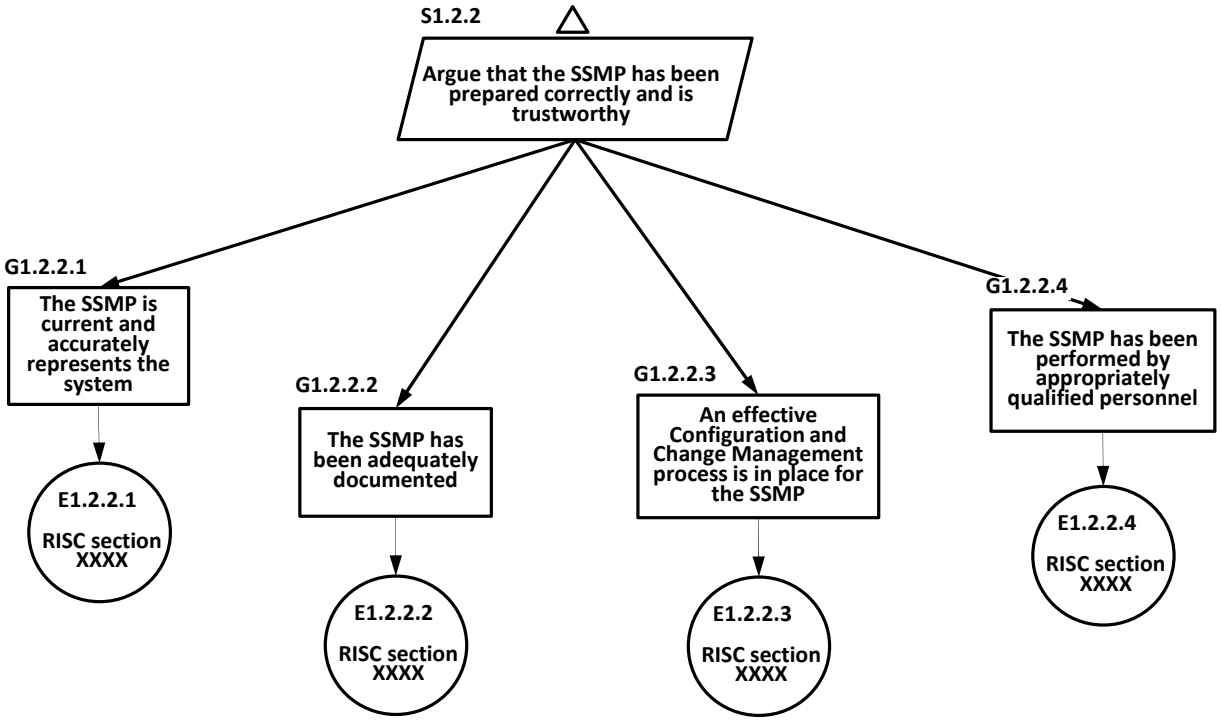


Figure 7-9. Argue that the SSMP Has Been Prepared Correctly and Is Trustworthy

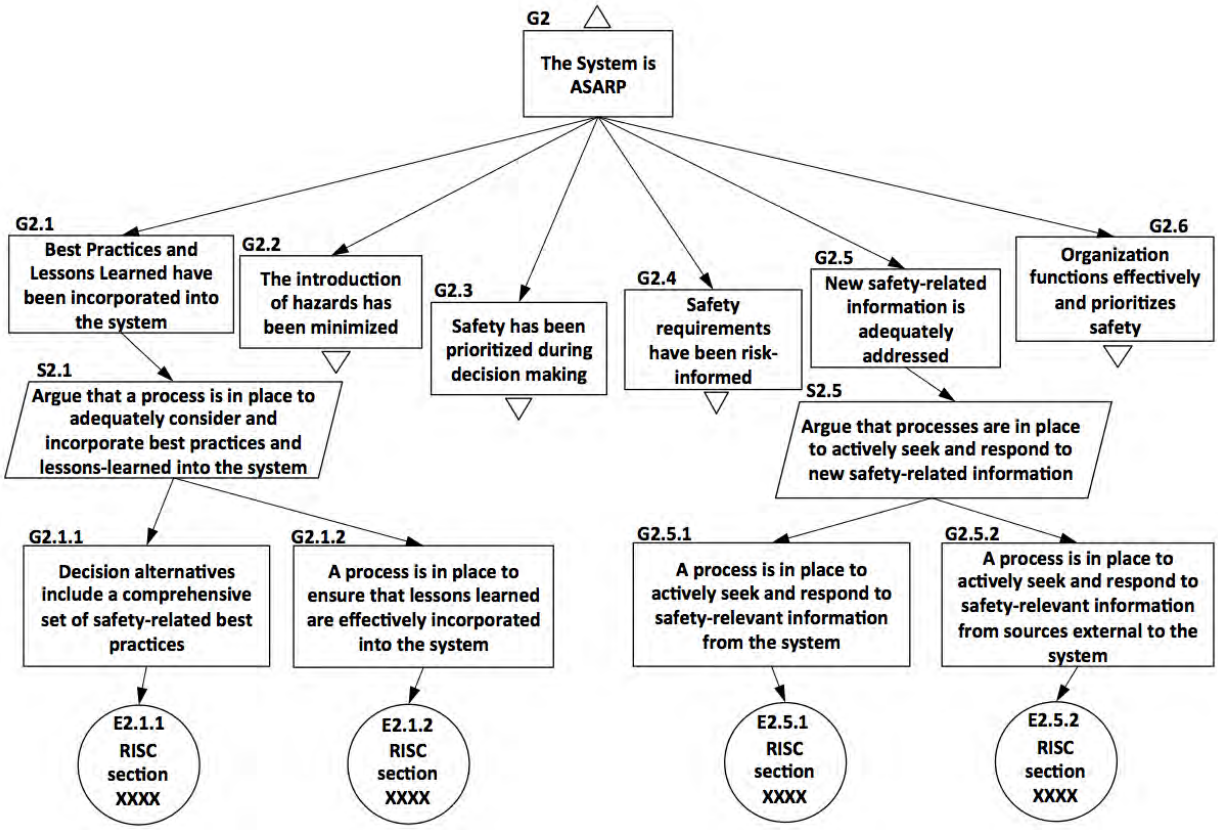


Figure 7-10. The System Is ASARP

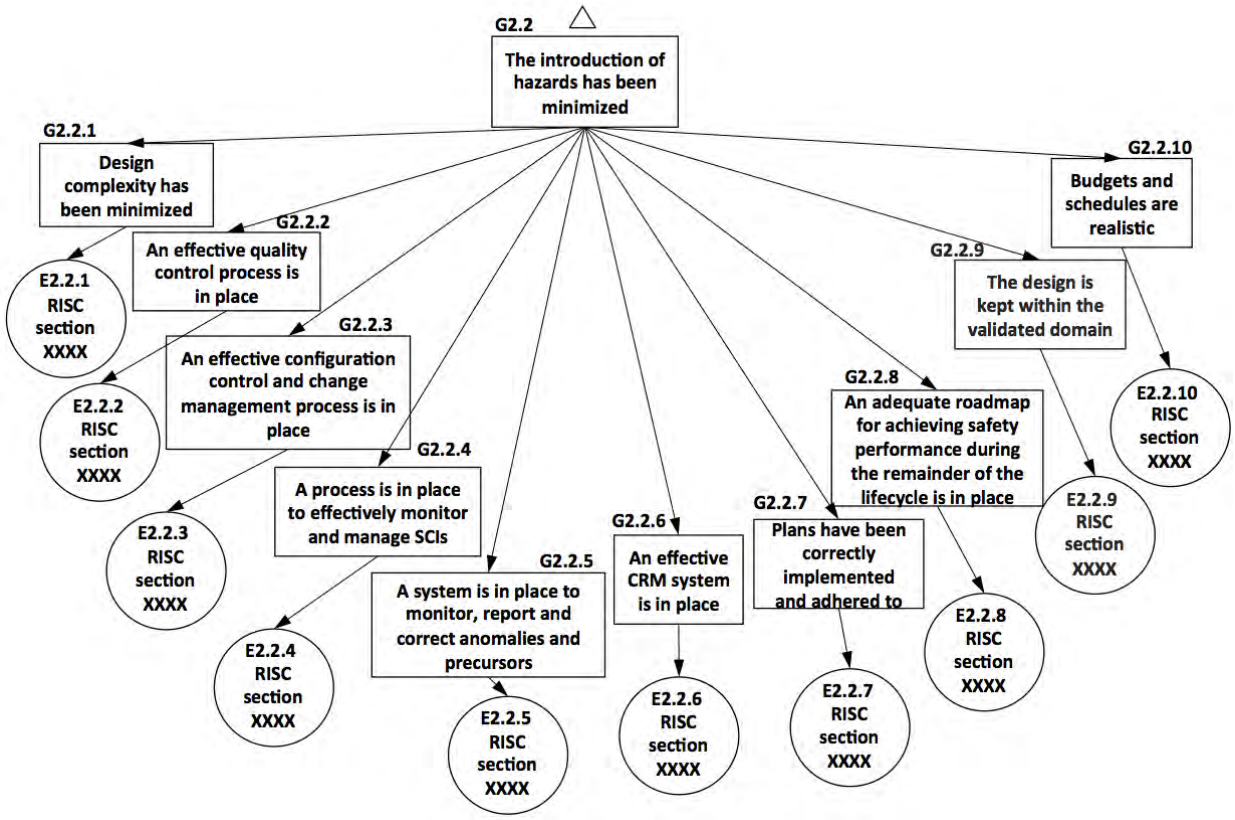


Figure 7-11. The Introduction of Hazards Has Been Minimized

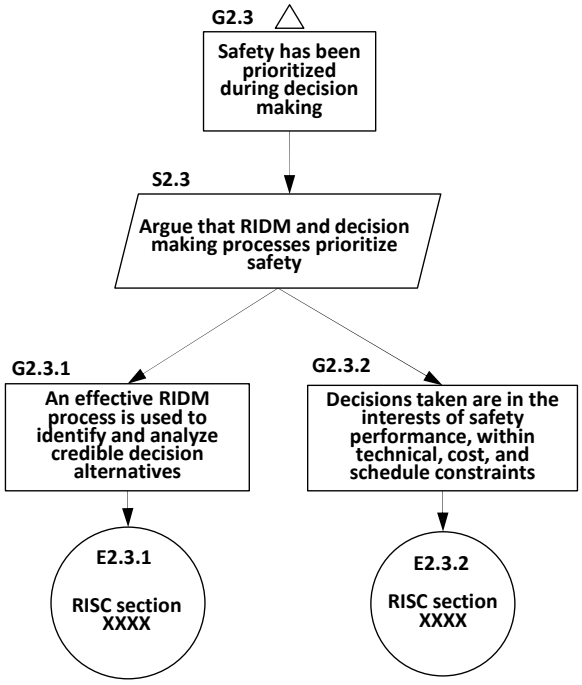


Figure 7-12. Safety Has Been Prioritized during Decision Making

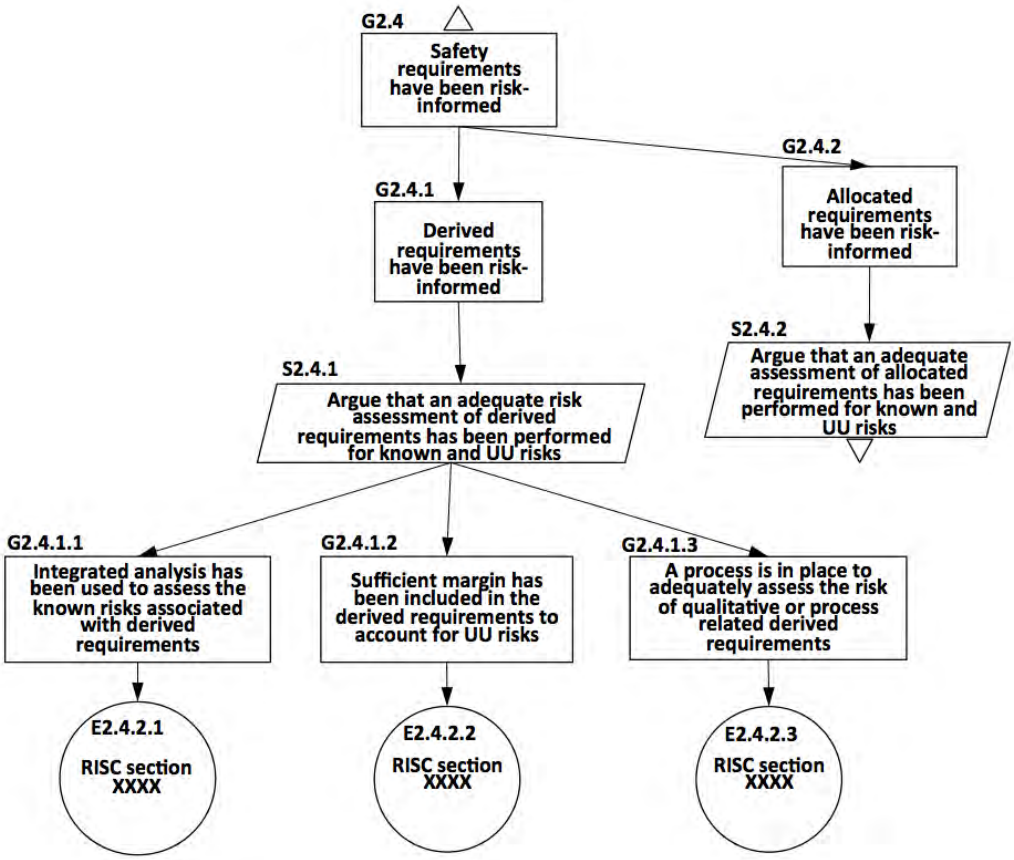


Figure 7-13. Safety Performance Requirements Have Been Risk-Informed

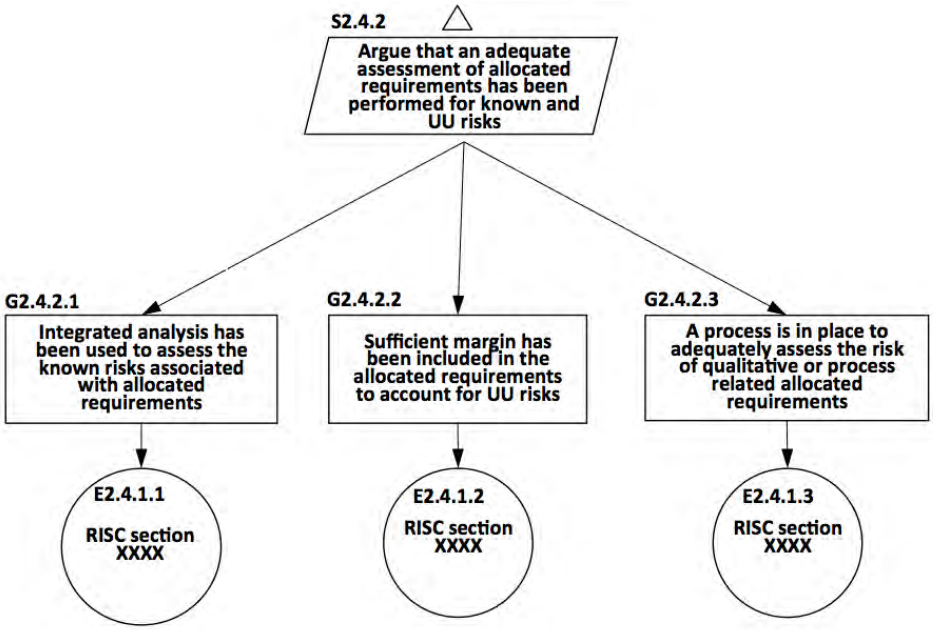


Figure 7-14. Argue that an Adequate Assessment of Allocated Requirements Has Been Performed for Known and UU Risks

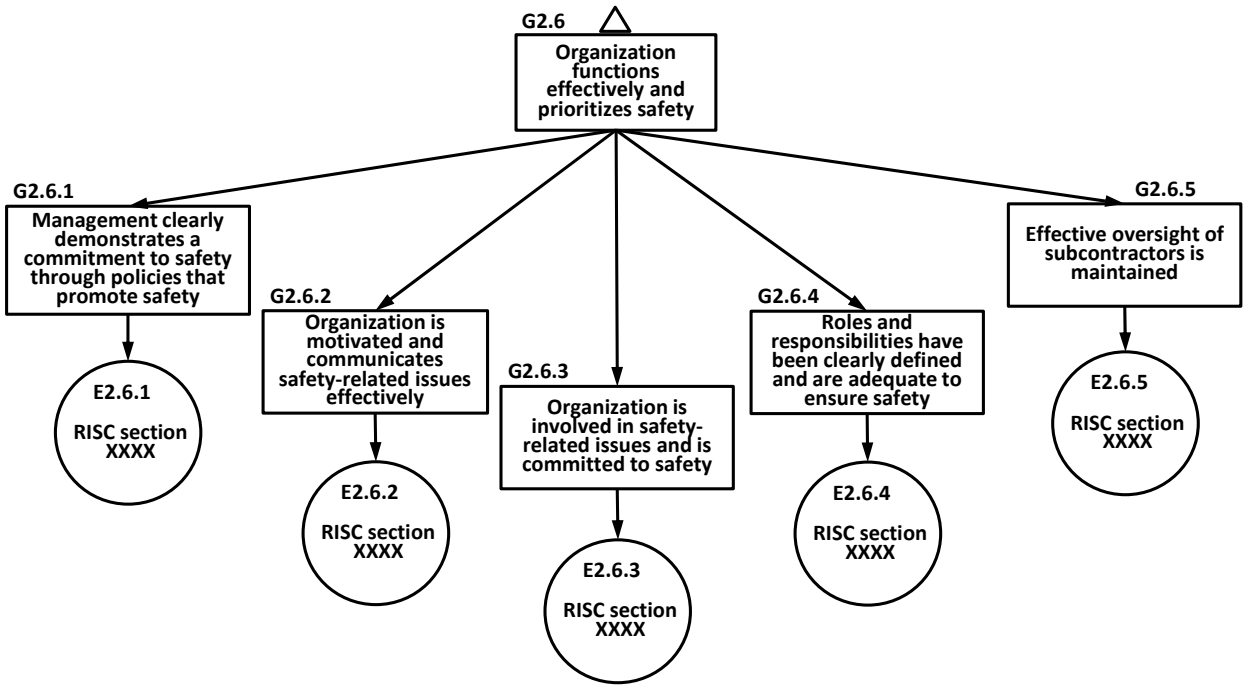


Figure 7-15. Organization Functions Effectively and Prioritizes Safety

References

[1] NASA/SP-2010-580, NASA System Safety Handbook Vol. 1, NASA, November 2011.

[2] MIL-STD-882E, Department of Defense Standard Practice – System Safety, May 2012.

[3] NASA NPR 7123.1B, NASA Systems Engineering Processes and Requirements, April 2013.

[4] NASA NPR 8000.4A, Agency Risk Management Procedural Requirements, December 2008.

[5] S. Insley, et al., "RISC Evaluation Management Tool," Information Systems Laboratories, Inc., to be published.

[6] NASA NPR 8715.3C, NASA General Safety Program Requirements, with Change 9, February 2013.

[7] NASA Decision Memorandum for the Administrator, "Agency's Safety Goals and Thresholds for Crew Transportation Missions to the International Space Station (ISS)," Washington, DC. 2011.

[8] NASA/SP-2011-3421, Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners, Second Edition, December 2011.

[9] U.S. Code of Federal Regulations, 10 CFR 20, Standards for Protection Against Radiation, Washington, DC. 1991.

[10] Parliament of the United Kingdom, Health and Safety at Work etc. Act, London, UK. 1974.

[11] NASA NPD 1000.0A, Governance and Strategic Management Handbook, August 2008.

[12] N. Leveson, "White Paper on the Use of Safety Cases in Certification and Regulation," MIT website, also see Journal of System Safety, Vol. 47, No. 6, Nov-Dec 2011.

[13] NASA/SP-2007-6105, NASA Systems Engineering Handbook, 2007.

[14] "NASA's Challenges to Meeting Cost, Schedule, and Performance Goals," Rept. IG-12-021, Sept. 27, 2012.

[15] NASA Advisory Council Meeting: Report of Audit and Finance Committee, Kennedy Space Center, February 5, 2009.

[16] J. Skakoon, "The Elements of Mechanical Design," ASME Press, 2008.

[17] C. Freaner, et al., "An Assessment of the Inherent Optimism in Early Conceptual Designs and Its Effect on Cost and Schedule Growth," European Aerospace Cost Engineering Working Group, May 2008.

[18] National Research Council, "Controlling Cost Growth of NASA Earth and Space Science Missions," National Academies Press, 2010.

[19] D. Bearden, "Perspectives on NASA Mission Cost and Schedule Performance Trends," in NASA Goddard Space Flight Center System Engineering Symposium, 3 June 2008.

[20] D. Bearden, "A Complexity-Based Risk Assessment of Low-Cost Planetary Missions: When is Mission Too Fast and Too Cheap?" 4th AIAA International Conference on Low-Cost Planetary Missions, JHU/APL, May 2000.

[21] A. Chmielewski and C. Garner, "How to Calculate Budget Reserve for Your Project," Jet Propulsion Laboratory, 6th NASA Project Management Challenge, 2009.

[22] N. Leveson, "A New Accident Model for Engineering Safer Systems," Massachusetts Institute of Technology, Safety Science, Vol. 42, No. 4, April 2004.

[23] C. Perrow, *Normal Accidents: Living with High-Risk Technologies*, Princeton University Press, 1984.

[24] B. Turner, *Man-Made Disasters*, Wykam Press, London, 1984.

[25] S. Sagan, *The Limits of Safety*, Princeton University Press, 1993.

[26] W. Evan, and M. Manion, *Minding the Machines: Preventing Technological Disasters*, Prentice Hall, 2002.

[27] I-Shih. Chang, "Space Launch Vehicle Reliability," Aerospace Corp. Magazine, 2001.

[28] NASA/SP-2011-3422, NASA Risk Management Handbook, November 2011.

[29] NASA/SP-2011-3423, NASA Accident Precursor Analysis Handbook, December 2011.

[30] Bishop, P. and Bloomfield, R., "A Methodology for Safety Case Development," Safety-Critical Systems Symposium, Birmingham, UK. 1998.

[31] Report of the Presidential Commission on the Space Shuttle Challenger Accident, Washington DC, June 1986.

[32] Columbia Accident Investigation Board Report, NASA, Washington, DC, August 2003.

[33] N. Leveson, "Engineering a Safer World: Systems Thinking Applied to Safety," MIT Press, 2011.

[34] NASA NPR 7120.5E, NASA Space Flight Program and Project Management Requirements w/Changes 1-10, August 2012.

[35] AIAA S-120-2006, Mass Properties control for Space Systems, Standard of the American Institute of Aeronautics and Astronautics, December 2006.

[36] SMC-T-002, Tailoring Instructions for AIAA-S-120-2006, AFSC Space and Missile Systems Center, August 8, 2008.

[37] NASA NPD 1000.5A, Policy for NASA Acquisition, January 15, 2009.

[38] NASA Presentation, "Joint Cost and Schedule Confidence Level (JCL), A Status Report," February 2010.

[39] NASA Cost Estimating Handbook, Washington, DC, 2008.

[40] NASA NPR 8705.2B, Human-Rating Requirements for Space Systems, 2008.

[41] D. Mathias, et al., "Simulation Assisted Risk Assessment Applied to Launch Vehicle Conceptual Design," NASA Ames Research Center, Also Reliability and Maintainability Symposium, RAMS, January 2008.

[42] A. Shih, et al., "Probabilistic Design Analysis (PDA) Approach to Determine the Probability of Cross-system Failures for a Space Launch Vehicle," NASA Langley Research Center; also Probabilistic Safety Assessment and Management Conference, PSAM10, June 2010.

[43] K. Gee and S. Lawrence, "Launch Vehicle Debris Models and Crew Vehicle Ascent Abort Risk," NASA Ames Research Center, 2013.

[44] S. Lawrence, et al., "Simulation Assisted Risk Assessment: Blast Overpressure Modeling," NASA Ames Research Center, May 2006.

[45] P. Prassinos, et al., Constellation Probabilistic risk Assessment (PRA): Design Considerations for CEV, NASA Office of Safety and Mission Assurance, OSMA-PRA-07-01, May 2006.

[46] J. Hanson, et al., "Launch Vehicle Failure Dynamics and Abort Triggering Analysis," NASA Marshall Spaceflight Center, August 2011.

[47] GSFC-STD-1000F, "Rules for the Design, Development, Verification, and Operation of Flight Systems" (The Goddard Open Learning Design (GOLD) Rules), Feb 2013.

[48] "NASA Preferred Practices for Design and Test of Robust Systems," Jet Propulsion Laboratory, http://oce.jpl.nasa.gov/preferred_practices.html#test.

[49] NASA Space Flight Program and Project Management Handbook, Washington, DC, May 2013.

[50] NASA Technical Memorandum 4322, "NASA Reliability Preferred Practices for Design and Test," http://www.hq.nasa.gov/office/codeq/rm/prefprac.htm.

[51] NASA Technical Memorandum 4628, "Recommended Techniques for Effective Maintainability," http://www.hq.nasa.gov/office/codeq/rm/prefprac.htm.

[52] NASA Lessons Learned System, NASA Engineering Network, http://llis.nasa.gov.

[53] CxP 70000, Constellation Architecture Requirements Document (CARD), December 2006.

[54] A. Benjamin, et al., "Developing Probabilistic Safety Performance Margins for Unknown and Underappreciated Risks," PSAM-12 International Conf. on Probabilistic Safety and Management, June 2014.

[55] T. Hamlin, et al., "Shuttle Risk Progression: Use of the Shuttle PRA to Show Reliability Growth," AIAA SPACE Conference & Exposition. 2011.

[56] E. Thigpen, "Shuttle PRA Iteration 3.3 Changes Notebook," NASA Internal Document, Johnson Space Center, Houston, TX, November 2010.

[57] J. Wiggins, "Space Shuttle Range Safety Analysis," Technical Report 81-1329 prepared for NASA Kennedy Space Center, July 1981.

[58] R. Weatherwax and E. Colglazier, "Review of Shuttle/Centaur Failure Probability Estimates for Space Nuclear Mission Applications," Sierra Energy and Risk Assessment, December 1983.

[59] S. Isakowitz, et al., International Reference Guide to Space Launch Systems, 3rd ed., American Institute of Aeronautics and Astronautics, 1999.

[60] I-S Chang, “Space Launch Vehicle Reliability,” Aerospace Corp. Crosslink, Oct. 2008, also website www.ewp.rpi.edu.

[61] S. Guikema. and M. Pate-Cornell, "Probability of Infancy Problems for Space Launch Vehicles," Reliability Engineering and System Safety, Vol. 87, 2005 (303–314).

[62] E. Morse, J. Fragola, and B. Putney, “Modeling Launch Vehicle Reliability Growth as Defect Elimination,” AIAA SPACE 2010 Conference & Exposition, AIAA 2010-8836, Aug-Sep 2010.

[63] S. Guarro, “Quantitative Launch and Space Transport Vehicle Reliability and Safety Requirements: Useful or Problematic?” PSAM-12 International Conf. on Probabilistic Safety and Management, June 2014.

[64] N. Leveson, “The Role of Software in Spacecraft Accidents,” AIAA Journal of Spacecraft and Rockets, Vol. 41, No. 4, July 2004.

[65] NASA JPR 1700.1, JSC Safety and Health Handbook, Appendix 2B, Johnson Space Center, April 2008.

[66] Richard Hawkins, et al., “A New Approach to Creating Clear Safety Arguments,” University of York and University of Virginia, 2010.

[67] A. Stephenson, et al., "Mars Climate Orbiter Mishap Investigation Board Phase I Report," NASA, November 1999.

[68] NASA-STD-0005, NASA Configuration Management (CM) Standard, September 2008.

[69] “Checklist for an Audit of Safety Management,” International Association of Oil and Gas Producers, Report No: 6.15/160, February 1990.

[70] NASA NPR 8705.5A, Technical Probabilistic Risk Assessment (PRA) Procedures for Safety and Mission Success for NASA Programs and Projects, June 2010.

[71] T. Kurtoglu, et al., “Integrating System Health Management into the Early Design of Aerospace Systems Using Functional Fault Analysis,” NASA Ames Research Center, 2008.

[72] I. Tumer, “Design Methods and Practices for Fault Prevention and Management in Spacecraft,” NASA Ames Research Center, 2005.

[73] S. Johnson, et al., System Health Management with Aerospace Applications, Chapter 8, “System Design and Analysis Methods,” by I. Turner, Wiley & Sons, 2011.

[74] A. Nikora and N. Green, “Software Anomaly Trends in JPL Missions,” Jet Propulsion Laboratory Presentation, 2007.

[75] N. Green and A. Hoffman, “Anomaly Trends for Missions to Mars: Mars Global Surveyor and Mars Odyssey,” Jet Propulsion Laboratory Paper, 2007, also published in Journal of Spacecraft and Rockets.

[76] “Context-based Software Risk Model (CSRM) Application Guide,” ASCA, Inc., Report AR 12-07, Draft May 15, 2012.

[77] D. Montgomery, Design and Analysis of Experiments, 8th Edition, Wiley and Sons, 2012.

[78] NASA-STD-7009, Standard for Models and Simulations, July 2008.

[79] NASA-HDBK-7009, NASA Handbook for Models and Simulations: an Implementation Guide for NASA-STD-7009, Draft (Not Yet Approved), December 2012.

[80] "Guide to Reusable Launch and Reentry Vehicle Reliability Analysis," Federal Aviation Authority (FAA), April 2005.

[81] FAA System Safety Handbook, 2000.

[82] Air Force System Safety Handbook, 2000.

[83] Yan Liu, "Multi-Attribute Utility Models with Interactions," Department of Biomedical, Industrial & Human Factors Engineering, Wright State University.

[84] G. Hazelrigg, "Systems Engineering: An Approach to Information-Based Design," Prentice-Hall, 1996.

[85] W. Starbuck and F. Milliken, "Challenger: Changing the Odds until Something Breaks," Journal of Management Studies, Vol. 25, 1988.2.59.

[86] E. Hollnagel, Cognitive Reliability and Error Analysis Method (CREAM), Elsevier, 1998.

[87] D. Lawson, "Engineering Disasters – Lessons to be Learned," ASME Press, 2005.2.60.

[88] U.K. Ministry of Defence, Defence Standard 00-56, "Safety Management Requirements for Defence Systems," London, UK. 2007.2.61.

[89] E. Denney, et al., "Towards Measurement of Confidence in Safety Cases," NASA Ames Research Center, 2011.

[90] E. Denney, et al., "AdvoCATE: An Assurance Case Automation Toolset," NASA Ames Research Center, 2011.

[91] T. Kelly and R. Weaver, "The Goal Structuring Notation – A Safety Argument Notation," Proc. of Dependable Systems and Networks Workshop on Assurance Cases, 2004.

[92] J. Spriggs, GSN - The Goal Structuring Notation: A Structured Approach to Presenting Arguments, Springer, January 2012.

[93] W. Greenwell, et al., "A Taxonomy of Fallacies in System Safety Arguments," University of Virginia, also Proceedings of the 2006 International System Safety Conference, 2006.

[94] J. Brophy, et al., "Lifetime Qualification of Electric Thrusters for Deep-Space Missions," 44th AIAA/ASME/SAE/ASEE Joint Propulsion Conference, AIAA-2008-5184, July 2008.

[95] J. Brophy, et al., "Implementation of the Dawn Ion Propulsion System," 41st AIAA/ASME/SAE/ASEE Joint Propulsion Conference & Exhibit, AIAA 2005-4071, July 2005.

[96] M. Meyer and J. Booker, Eliciting and Analyzing Expert Judgment: A Practical Guide, Society for Industrial and Applied Mathematics (SIAM) Publication, 2001.

[97] J. Keisler, et al., “Value of Information Analysis: The State of Application,” Environment Systems and Decisions, Springer, 2014.

[98] R. Howard, “Information Value Theory,” IEEE Transactions on Systems Science and Cybernetics (SSC-2), 1966.

[99] C. Kirkwood, “Decision Tree Primer,” Chapter 3, Arizona State University, 2002.

Appendix A – Abbreviations and Acronyms

ALARA	As Low As Reasonably Achievable
ALARP	As Low As Reasonably Practicable
APU	Auxiliary Power Unit
ASAP	Aerospace Safety Advisory Panel
ASARP	As Safe As Reasonably Practicable
CAS	Credibility Assessment Scale
CDR	Critical Design Review
CERR	Critical Events Readiness Review
CREAM	Cognitive Reliability Error Analysis Method
CRM	Continuous Risk Management
CSRM	Concept-based Software Risk Model
DFMR	Design for Minimum Risk
DM	Decision maker
DoD	Department of Defense
DR	Decommissioning Review
ETA	Event Tree Analysis
FAA	Federal Aviation Administration
FEA	Finite Element Analysis
FFA	Functional Fault Analysis
FMEA	Failure Modes and Effects Analysis
FMECA	Failure Modes, Effects, and Criticality Analysis
FTA	Fault Tree Analysis
GIDEP	Government-Industry Data Exchange Program
GSN	Goal Structuring Notation
HA	Hazard Analysis
HAZOP	Hazard and Operability Analysis
HRA	Human Reliability Analysis
IHA	Integrated Hazard Analysis
ISA	Integrated Safety Analysis
ISS	International Space Station
JCL	Joint Confidence Level
JSC	Johnson Space Center
KDP	Key Decision Point
KMO	Key Mission Objective
LAS	Launch Abort System
LEO	Low Earth Orbit
LOC	Loss of Crew
LOM	Loss of Mission
LOV	Loss of Vehicle
M&S	Modeling and Simulation
MBSE	Model-Based Systems Engineering
MGA	Mass Growth Allowance
MLE	Maximum Likelihood Estimation
MMOD	Micrometeoroids and Orbital Debris
MOP	Measure of Performance
NASA	National Aeronautics and Space Administration
NPR	NASA Procedural Requirements
OSMA	Office of Safety and Mission Assurance

PDR	Preliminary Design Review
PPE	Personal Protective Equipment
PRA	Probabilistic Risk Assessment
QA	Quality Assurance
RAI	Request for Additional Information
RBD	Reliability Block Diagram
RISC	Risk-Informed Safety Case
RIDM	Risk-Informed Decision Making
RM	Risk Management
R/M	Reliability/Maintainability
S3G	System Safety Steering Group
SCI	Safety-Critical Item
SCP	Satellite Control Processor
SDR	System Definition Review
SE	Systems Engineering
SIR	System Integration Review
SMA	Safety and Mission Assurance
SPF	Safety Performance Factor
SRB	Solid Rocket Booster
SS	System Safety
SSMP	System Safety Management Plan
SSRA	System Safety Requirements Analysis
STS	Space Transportation System
T&E	Test and Evaluation
TA	Technical Authority
TPM	Technical Performance Measure
TPS	Thermal Protection System
TRL	Technology Readiness Level
UFE	Unallocated Future Expenses
UU	Unknown and/or Underappreciated

Appendix B – Definitions

Acquirer – A NASA organization that tasks a subordinate organization to produce a product or deliver a service.

As Safe as Reasonably Practicable – A philosophy that states that safety should be increased as opportunities arise if the impact on cost, schedule, technical performance, or any other domain of interest to NASA is reasonable and acceptable.

Assurance Deficit – Any knowledge gap that prohibits perfect (total) confidence. Assurance deficits are caused by variability or lack of knowledge concerning the data or models being used to produce the evidence, the parameter inputs to the models, and the interpretation of model outputs.

Availability – The fraction of time that an item is in-service.

Base Claim – A safety claim that is not decomposed to lower levels. It is demonstrated to be true as asserted to a high degree of confidence by providing evidence and by showing that all deficits in the evidence that erode confidence in the base claim are sufficiently minimal.

Capability – The maximum operating load or stress for which an item is certified.

Condition – A current fact-based situation or environment that is causing concern, doubt, anxiety, or uneasiness.

Consequence – The foreseeable, credible negative impact(s) on the organizational unit's ability to meet its performance requirements.

Continuous Risk Management (CRM) – A specific process for the management of risks associated with implementation of designs, plans, and processes. The CRM functions of identify, analyze, plan, track, control, and communicate and document provide a disciplined environment for continuously assessing what could go wrong, determining which concerns are important to deal with, and implementing strategies for dealing with them.

Control – In the safety context, any provision taken to reduce the likelihood and/or severity of an accident. Controls can include design modifications to address specific risks, improvements in quality assurance, modification of procedures, improvements in personnel training, provisions to improve management oversight where needed, etc.

Departure Event – An undesired event that might occur at a future time representing a change from the current plan and leading potentially to a consequence. It is the uncertainty in the occurrence or non-occurrence of the departure event that is the initially identified source of risk.

Design for Minimum Risk – The inclusion of specific design features that minimize the probability of occurrence of failure modes, such as application of stringent factors of safety or other design margins.

Direct Evidence – Information that is mostly quantitative and that supports a base claim by showing that the risk of not meeting it is acceptably low. Examples of direct evidence include failure rates from test data or operational experience, analyses of system response to various environments, results of probabilistic risk assessments, analysis of anomalies, and adherence to best practices.

Evaluation Protocol – A set of techniques, standards, and practices to be applied in demonstrating the level of satisfaction of a performance requirement (e.g., a safety goal). An evaluation protocol may include mandated assumptions, may specify a particular process of analysis, and may limit the degree of credit that can be taken for a particular design feature.

Hazard – A state or a set of conditions, internal or external to a system, that has the potential to cause harm. Examples of hazards include materials, energy sources, or operational practices that in uncontrolled

situations can lead to scenarios that could produce death, injury, illness, equipment loss or damage, or damage to a protected environment.

Hazard Analysis – An application of systematic and replicable methods to identify and understand hazards, and to characterize the risk of mishaps that involve hazards. Risks originate from hazards – the absence of a hazard implies a freedom from the associated risk.

Initiating Event – A departure from a desired operational envelope to a system state where a control response is required either by human, software, or machine intervention.

Integrated Safety Analysis – The development and analysis of scenarios that may lead to undesirable consequences with respect to safety. ISA includes both hazard-centric and non-hazard-centric methods for identifying and characterizing potential accident scenarios. This includes accident causes, contributing factors, effectiveness of controls (both existing and proposed), analysis of physical responses of the system to the environments it encounters, and analysis of the probability that the undesirable consequences will be realized. The analysis of any particular scenario can be either quantitative or qualitative, as appropriate for the scenario being considered and the nature of the undesired consequence.

Intermediate Claim – A claim that is further decomposed into lower level sub-claims that feed into it. It is demonstrated to be true as asserted to a high degree of confidence by demonstrating that all of the sub-claims feeding into it are true as asserted to a high degree of confidence.

Key Decision Point – The event at which the Decision Authority determines the readiness of a program/project to progress to the next phase of the life cycle (or to the next KDP).

Key Mission Objective – Different multiple objectives within a single mission for which the decision maker needs to be presented with safety sub-cases in order to make decisions affecting the program/project as a whole. For a planetary exploration mission that has many rendezvous events, each major rendezvous event would be considered a key mission objective. For a mission consisting of many flights with each flight having one destination, any flight that marks a significant change from previous flights would constitute a separate key mission objective.

Known Risk – A scenario affecting safety performance that has been correctly identified and accurately assessed with respect to its likelihood of occurrence and potential severity of harm or loss.

Limiting Condition for Operation – A limiting condition for operation defines the limits that represent the lowest functional capability or performance level of an SCI required to perform safely.

Limiting Control Setting – A limiting control setting defines the setting on an SCI that controls process variables to prevent exceeding a safety limit.

Minimum Tolerable Level of Safety – The level of safety performance below which a system is considered unsafe.

Model – A description or representation of a system, entity, phenomenon, or process.

Objectives Hierarchy – An arrangement where objectives are decomposed into a set of quantifiable sub-objectives, each of which is implied by the top-level objective.

Operational Safety Objective – A safety objective that has been decomposed to a level where it can be clearly addressed by systems engineering processes.

Performance Measure – A quantifiable attribute of a decision alternative, used to support decision-making. Performance measures are typically defined for all mission execution domains and for institutional performance. For purposes of System Safety at NASA, performance measures include metrics related to human safety, asset protection, and environmental protection.

Performance Requirement – A value of a performance measure to be achieved by an organizational unit's work that has been agreed-upon to satisfy the needs of the next higher organizational level. [NPR 8000.4A]

Probabilistic Risk Assessment – A structured, probabilistic treatment of scenarios, likelihoods, consequences using a graded approach. Within this approach, the word "probabilistic" refers explicitly to a Bayesian treatment of uncertainty.

Provider – A Provider is a NASA or contractor organization that is tasked by a customer or supervising organization (i.e., the Acquirer) to produce a product or service.

Reliability – The probability that an item will perform its intended function for a specified interval under stated conditions.

Risk – The potential for shortfalls, which may be realized in the future, with respect to achieving explicitly stated performance requirements. Risk is characterized by a set of triplets: 1) the scenario(s) leading to degraded performance in one or more performance measures, 2) the likelihood(s) of those scenarios, and 3) the consequence(s) of the impact on performance that would result if those scenarios were to occur.

Risk Driver – A significant source of performance risk. Operationally, a risk driver can be a single performance parameter, a single event, a set of performance parameters collectively, or a set of events collectively that, when varied over their range of uncertainty, causes the performance risk to change from tolerable to intolerable (or marginal).

Risk-Informed Decision Making – A decision making approach that uses a diverse set of performance measures (some of which are model-based risk metrics) along with other considerations within a deliberative process to inform decision making.

Risk-Informed Safety Case – A risk-informed safety case (RISC) is a structured argument, supported by a body of evidence, that provides a compelling, comprehensible and valid case that a system is or will be adequately safe for a given application in a given environment.

Risk Management – A process that includes risk-informed decision making and continuous risk management in an integrated framework. This integration is done in order to foster proactive risk management, to better inform decision making through better use of risk information, and then to more effectively control implementation risks by focusing the continuous risk management process on the baseline performance requirements emerging from the RIDM process. [NPR 8000.4A]

Risk Statement – A statement of a concern about a scenario that could affect the ability to achieve one or more safety requirements. Each risk statement contains a *condition*, a *departure*, an *asset*, and a *consequence*.

Safety – Freedom from those hazards that can result in failure to meet one or more safety objectives by causing death, injury, or illness in humans, adversely affecting the environment, and/or causing damage to or loss of equipment or property.

Safety Assurance – The development of confidence that safety has been sufficiently ensured.

Safety Case – A documented body of evidence that provides a convincing and valid argument that a system is adequately safe for a given application in a given environment.

Safety Claim – A statement asserting the level of safety of a system or subsystem.

Safety-Critical Item – A system feature whose performance (e.g., capability, reliability, and availability) at levels documented in the ISA is necessary for the satisfaction of system-level safety performance requirements, or which is designated as such.

Safety Ensurance - The reduction and elimination of system hazards and the achievement of adequate safety performance.

Safety Factor – The structural capacity of an item beyond the expected loads or actual loads.

Safety Goal – A target level of safety performance that is expected from continuous safety upgrades and improvements to the system.

Safety Growth – The rate at which safety performance increases due to safety upgrades and improvements to the system.

Safety Limit – A limit on a process variable that, if exceeded, could directly cause the failure of a barrier.

Safety Margin – Extra performance allocated to systems, structures, and components to preserve safety over the range of credible variations in the loads (stresses, temperatures, etc.) to which they will be subjected.

Safety Performance Margin – An incremental margin subtracted from the evaluated safety performance to account for the estimated total effects of unknown and underappreciated hazards. It is estimated from analysis of historical experience with similar technologies taking into account the complexity of the system, the degree to which new technology is being used, and the degree to which new operating environments are being introduced.

Safety Performance Factor – The ratio of the loss probability from all risks to the loss probability from known risks.

Safety Threshold – The level of safety performance against which initial system performance is assessed.

Scenario – A sequence of credible events that specifies the evolution of a system or process from a given state to a future state. In the context of risk management, scenarios are used to identify the ways in which a system or process in its current state can evolve to an undesirable state.

Sensitivity – The variation in the output of a model as a function of variation in the model inputs and parameters.

Supporting Evidence – Information that is mostly qualitative, provides confidence in the direct evidence, or demonstrates a general responsiveness to safety concerns. Examples of supporting evidence include personnel qualifications, verification and validation of analysis tools, applicability of experiments, quality of documentation, quality of external reviews, effectiveness of communication protocols, and safety culture of the organization.

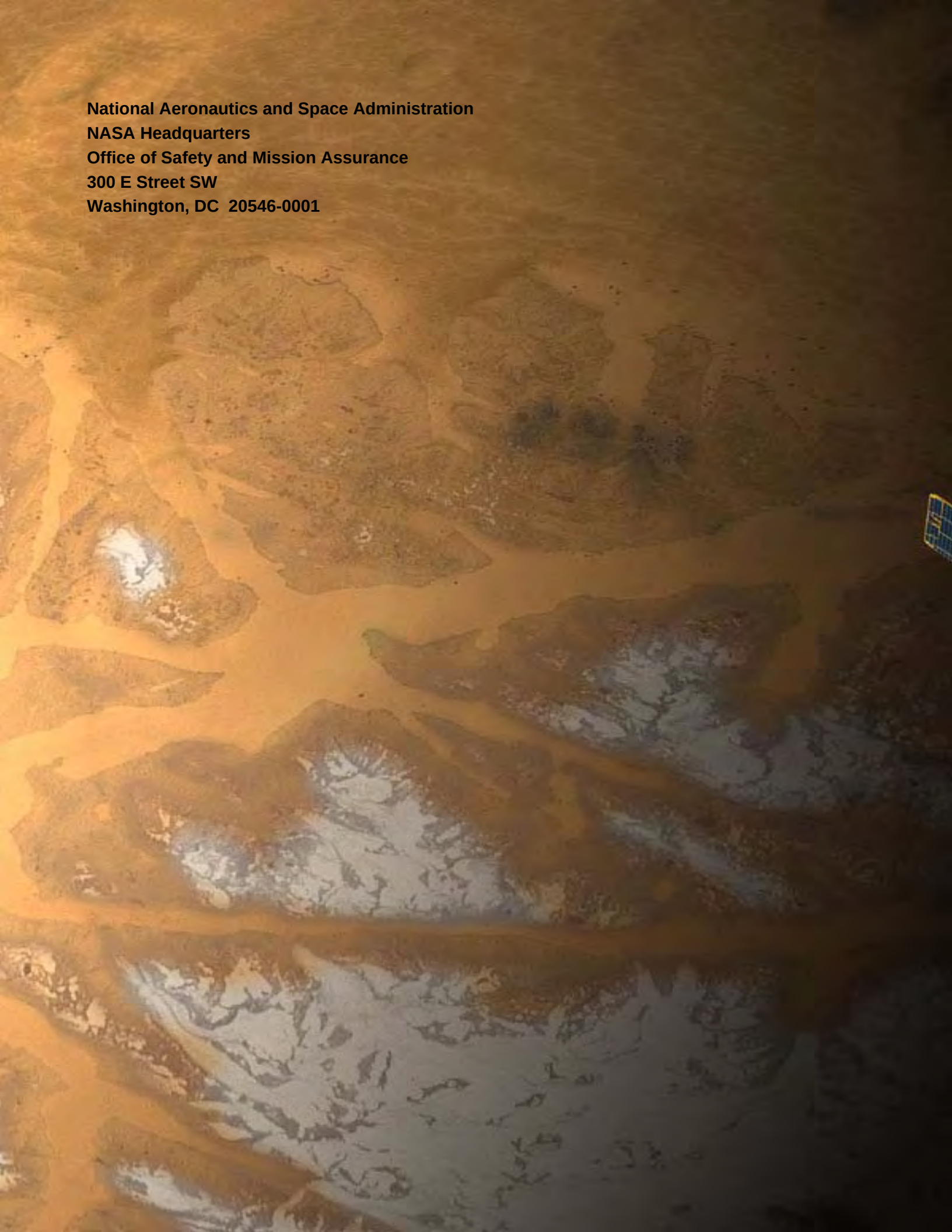
Synthetic Analysis Methods – Methods that produce system-level risk estimates by aggregating the effects of explicitly identified individual contributors to that risk.

System Safety – A disciplined, systematic process for the consideration of risks resulting from hazards that can affect humans, the environment, or mission assets. Per NPR 8715.3C, System Safety is the rational pursuit of safety within a systems perspective, where the degree of "safety" is to be understood in the context of a particular application. The system safety process does not expect to attain absolute safety, but strives to attain a degree of safety that fulfills obligations to the at-risk communities and addresses Agency priorities.

Uncertainty – An imperfect state of knowledge or a physical variability resulting from a variety of factors including, but not limited to, lack of knowledge, applicability of information, physical variation, randomness or stochastic behavior, indeterminacy, judgment, and approximation.

Underappreciated Risk – A scenario affecting safety performance that has been correctly identified but for which the likelihood of occurrence and/or potential severity of harm or loss are underestimated.

Unknown Risk – A scenario affecting safety performance that has not been identified and is therefore unknown at the time of analysis.



National Aeronautics and Space Administration
NASA Headquarters
Office of Safety and Mission Assurance
300 E Street SW
Washington, DC 20546-0001