The Role of NASA Safety Thresholds and Goals in Achieving Adequate Safety*

Homayoon Dezfuli[a†], Chris Everett[b], Allan Benjamin[c], Bob Youngblood[d], and Martin Feather[e]

[a] NASA, Washington, DC, USA
[b] Information Systems Laboratories, Rockville, MD, USA
[c] Independent Consultant, Albuquerque, NM, USA
[d] Idaho National Laboratory, Idaho Falls, ID, USA
[e] Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA

Abstract: NASA has recently instituted requirements for establishing Agency-level safety thresholds and goals that define long-term targeted and maximum tolerable levels of risk to the crew as guidance to developers in evaluating "how safe is safe enough" for a given type of mission. This paper discusses some key concepts regarding the role of the Agency's safety thresholds and goals in achieving adequate safety, where *adequate safety* entails not only meeting a minimum tolerable level of safety (e.g., as determined from safety thresholds and goals), but being as safe as reasonably practicable (ASARP), regardless of how safe the system is in absolute terms.

Safety thresholds and goals are discussed in the context 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. In this context, meeting of safety thresholds and goals is one of a number of distinct safety objectives, and the system safety analysis provides evidence to substantiate claims about the system with respect to satisfaction of the thresholds and goals.

Keywords: Safety Thresholds, Safety Goals, Safety Performance Margin, As Safe As Reasonably Practicable (ASARP), Risk-Informed Safety Case (RISC).

1. INTRODUCTION

NASA has recently instituted requirements for establishing Agency-level safety thresholds and goals that define "long-term targeted and maximum tolerable levels of risk to the crew as guidance to developers in evaluating "how safe is safe enough" for a given type of mission" [1]. Safety thresholds specify the minimum tolerable/allowable level of crew safety (maximum tolerable level of risk) for the design in the context of its design reference mission, and are to be used by the Agency as criteria for program acquisition decisions. Safety goals (and the accompanying requirements to implement safety upgrade and improvement programs) are motivated by the fact that the level of risk associated with initially flown designs is typically unacceptable in the long term and the fact that human spaceflight programs, informed by flight experience and analysis, can achieve significant reductions of risk over the life of a program.

This paper discusses some key concepts regarding the role of the Agency's safety thresholds and goals in achieving adequate safety. As discussed in the recently-released NASA/SP-2010-580, NASA System Safety Handbook [2], adequate safety involves not only meeting the minimum tolerable level of safety (e.g., as determined from safety thresholds and goals), but it also involves being as safe as reasonably practicable (ASARP), regardless of how safe the system is in absolute terms.

* This paper is based on the work performed by the NASA Office of Safety and Mission Assurance (OSMA) in support of the development activities for Volume 2 of NASA System Safety Handbook (NASA/SP-2014-612),

† hdezfuli@nasa.gov

This paper also provides a framework for implementing the safety thresholds and goals in a way that reflects expectations that the safety of a new system will improve over time, and is consistent with the technical challenges inherent in assessing the safety of such systems. In particular, synthetic[‡] methods of risk analysis are vulnerable to risk model incompleteness (i.e., only a subset of the system failure causes are identified and analyzed), especially in the early phases of the system life cycle. This paper outlines a method of accounting for these un-modeled sources of risk, providing a rational basis for determining whether or not a system meets the minimum tolerable level of safety.

Finally, this paper discusses the safety thresholds and goals in the context of the Risk-Informed Safety Case (RISC), defined in the NASA System Safety Handbook 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." In this context, meeting of safety thresholds and goals is one of a number of distinct safety objectives, and the system risk analysis provides evidence to substantiate claims about the system with respect to satisfaction of the thresholds and goals. The nature of a specific RISC depends upon the decision context in which it is developed. Different decision contexts produce different safety objectives, and use different sets of engineering observables as evidence to support the claim that the objectives have been met.

2. FUNDAMENTAL PRINCIPLES OF SAFETY

The NASA System Safety Handbook articulates two fundamental principles of safety that together constitute "adequate safety":

- An adequately safe system is assessed as meeting a minimum threshold level of safety, as determined by analysis, operating experience, or a combination of both. Below this level the system is considered unsafe. This minimum level of safety is not necessarily fixed over the life of a system. As a system is operated and information is gained as to its strengths and weaknesses, design (hardware and software), and operational modifications are typically made which, over the long run, improve its safety performance. 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.

- An adequately safe system is ASARP. 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 [3, 4]. A determination that a system is ASARP entails weighing its safety performance against the sacrifice needed to further improve it. The system is ASARP if an incremental improvement in safety would require a disproportionate deterioration of system performance in other areas including technical, cost, and schedule.

Figure 1, reproduced from the System Safety Handbook, illustrates application of these two principles of safety throughout the entire system lifecycle.

2.1. Meeting the Minimum Tolerable Level of Safety

In the context of the fundamental principles of adequate safety, Agency-level safety thresholds and goals express stakeholders' expectations about the minimum tolerable level of crew safety, where a safety threshold expresses an initial minimum tolerable level, and the goal expresses expectations about the safety growth of the system in the long term. As such, the safety thresholds and goals work together to establish a minimum tolerable level of safety that increases from the threshold to the goal

[‡] By "synthetic methods," we mean methods that produce risk estimates by explicitly constructing a scenario set and summing risk contributions to obtain an estimate of aggregate risk, as is typically done in probabilistic risk assessment (PRA).

over the life of the program. As discussed in NPR 8705.2B, safety thresholds are used by the Agency as criteria for program acquisition decisions, whereas safety goals specify the level of safety that is considered acceptable for repeated missions and serve as the long-term target for proactive safety upgrade and improvement programs that must be maintained for the duration of the program or until the safety goals have been met. The NPR does not specify the rate at which the minimum tolerable level of safety moves from threshold to goal, only that safety growth must be proactively pursued so long as the goal is unmet.

Figure 1: Fundamental Principles of Adequate Safety

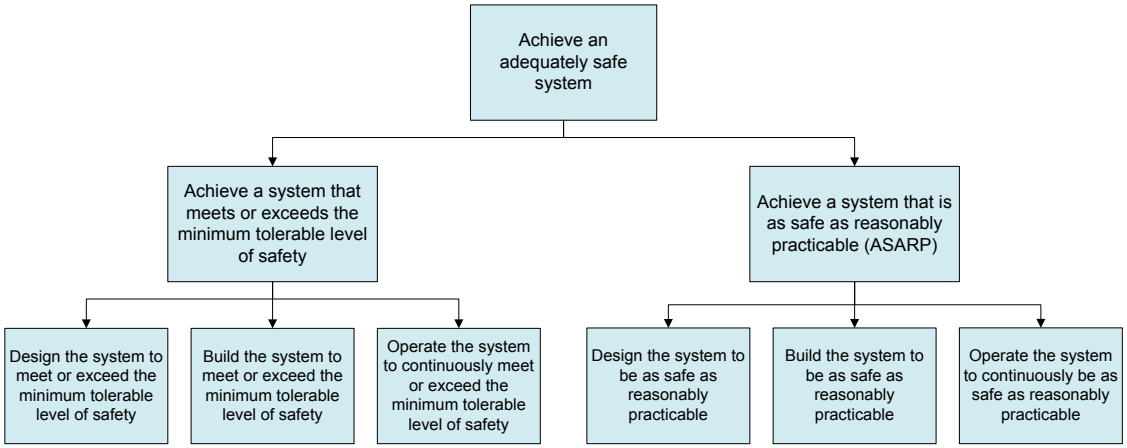
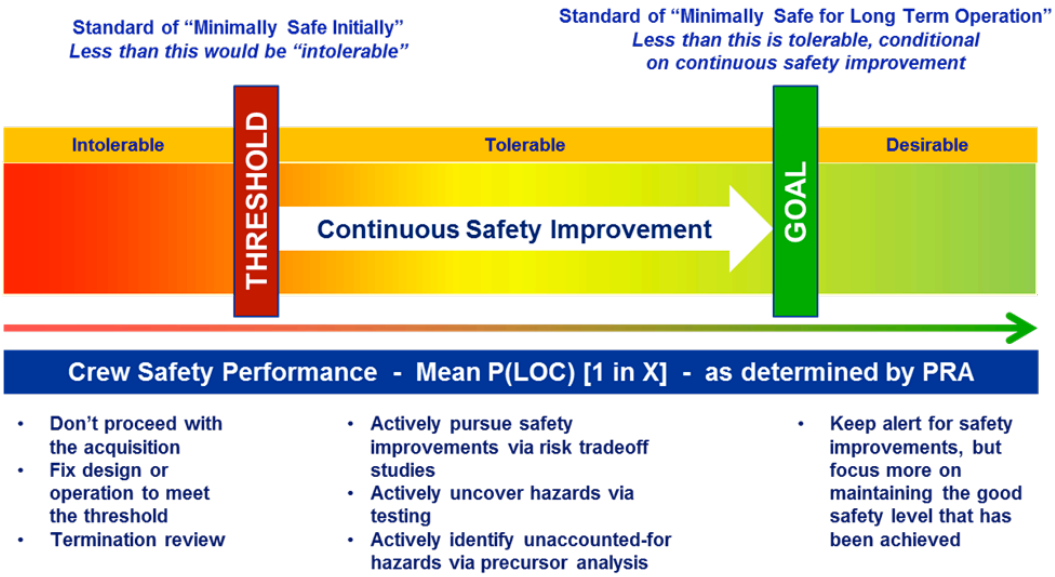


Figure 2, reproduced from the NASA System Safety Handbook, illustrates the NASA safety thresholds and goals.

Figure 2: NASA Safety Thresholds and Goals



2.2. Being As Safe As Reasonably Practicable

It is ethically problematic to accept a given level of safety if improvements to safety can be easily made. An adequately safe system is therefore one that is ASARP. Being ASARP is a separate and distinct consideration from meeting a minimum tolerable level of safety. It reflects a mindset of continuous safety improvement regardless of the current level of safety. It is an integral aspect of good systems engineering process that guides risk-informed decision making throughout the system lifecycle, beginning in formulation.

The ASARP concept is illustrated graphically in Figure 3, which is adapted from the System Safety Handbook. The curve represents the efficient frontier of the trade space of identified alternatives, and shows the tradeoff between safety performance and performance in other mission execution domains (cost, schedule, technical).[§] The ASARP region contains those alternatives whose safety performance is as high as can be achieved without resulting in intolerable performance in one or more of the other domains.

Figure 3: As Safe As Reasonably Practicable (ASARP)

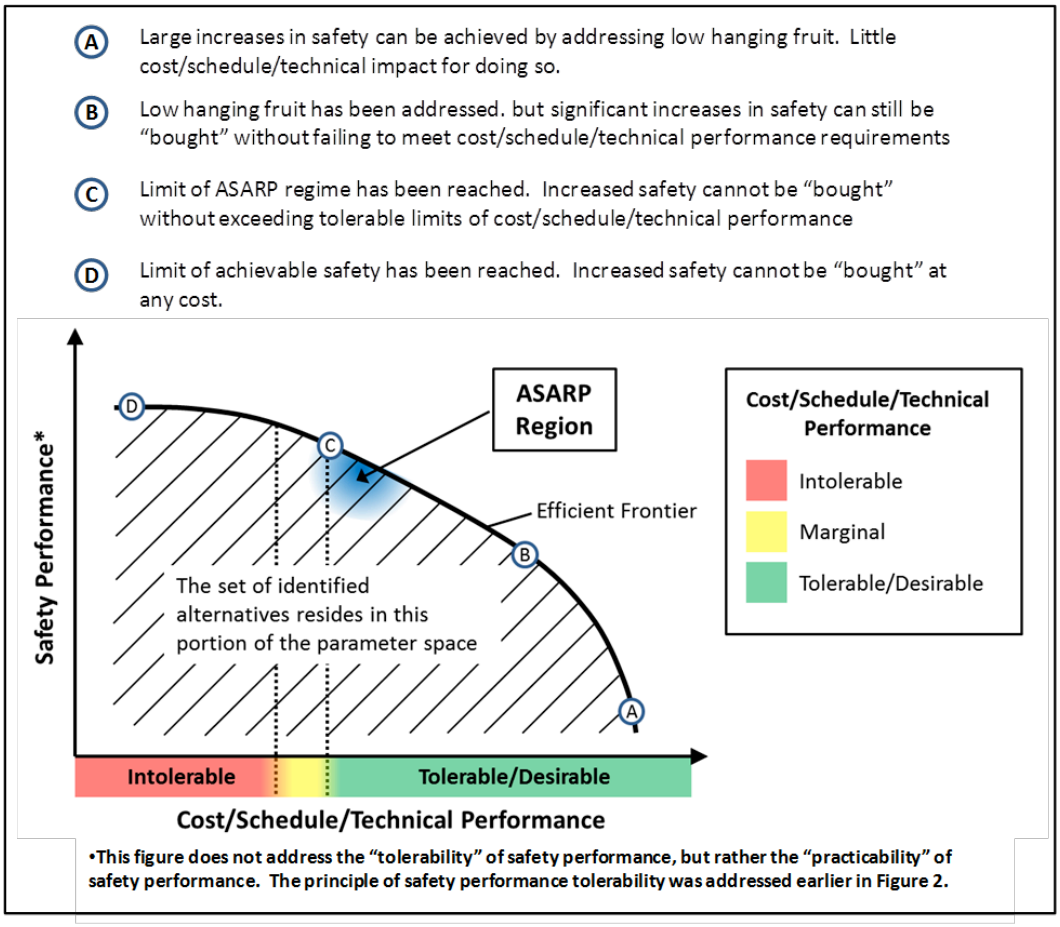


Figure 3 illustrates that:

- Improvements to cost, schedule, or technical performance beyond minimum tolerable levels are not ASARP if they come at significant expense in safety performance.

- The ASARP concept 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 that of other identified alternatives.

- ASARP is a region of the trade space and can contain more than one specific alternative. Moreover, the boundaries of that region are not sharply defined. Determining that a system is ASARP entails the prudent application of engineering and management judgment.

[§] Figure 3 is a two-dimensional representation of a space that involves four dimensions: safety, cost, schedule, and technical performance. For conceptual purposes, the cost, schedule, and technical performance dimensions are combined onto a single axis, since in this context they are all regarded as impacts incurred in order to increase safety. Each point on the efficient frontier may be interpreted as the maximum level of safety performance that may be achieved for a given level of cost, schedule, and technical performance.

The means of achieving a system that is ASARP include those for meeting the safety goal (i.e., implementation and maintenance of proactive safety upgrade and improvement programs), as well as meeting other applicable safety criteria; applying best practices in system design (e.g., appropriate margins and failure tolerance); high quality manufacture; robust operational procedures (including contingencies); and maintaining a safety analysis process to identify, evaluate, and, as appropriate, mitigate risks or deficiencies for the life of the program regardless of the satisfaction of the safety goal. In general, ASARP results from systems engineering activities that prioritize safety throughout the system lifecycle.

The concept of ASARP is necessarily in play not only during the development of a selected design alternative, which is usually the context within which ASARP is discussed, but also during the time in which the selection among competing design alternatives is made. In the latter case, the ASARP concept is imbued within the Risk-Informed Decision Making (RIDM) process [5], which has as its objective the selection of a design alternative based on risk-informed trade-offs involving safety, cost, schedule, and technical performance. The principle of RIDM calls for a decision maker who is risk-informed to apply his/her value system to the weighting of the trade-offs. In so doing, however, the decision maker is guided by the ASARP principle, which states that safety is to have a higher weight than any of the other three dimensions of the trade-off (cost, schedule, and performance).

3. ASSESSING THE SYSTEM RELATIVE TO THE MINIMUM TOLERABLE LEVEL OF SAFETY

NPR 8705.2B states that thresholds and goals are to be expressed in terms of an aggregate measure of risk such as the probability of a loss of crew (P(LOC))[**]. In order to compare a system's safety performance to the thresholds and goals, a risk analysis is performed that quantifies the relevant measure. The Administrator's letter on the Agency's safety thresholds and goals for crew transportation missions to the International Space Station (ISS) specifies using NASA-accepted probabilistic risk analysis (PRA) methods similar to those applied by the Space Shuttle, ISS, and Constellation programs, and using the mean P(LOC)[††] as the risk measure [6].

3.1. The Issue of Risk Analysis Incompleteness

However, although PRA methods [7] have a history of use at NASA for providing insight into the relative risk significance of potential accident scenarios that might occur in a system, and into the relative safety performance of different systems,[‡‡] it has long been recognized that there are challenges inherent in using synthetic methods such as those used in PRA to quantify a system's actual risk, due to the inherent incompleteness of the scenario sets identified by these methods [8-13]. The unaccounted-for (or insufficiently-accounted-for) scenarios typically involve organizational issues and/or 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 accidents. In the Challenger disaster, O-ring blow-by impinged on the external tank, leading to tank rupture and subsequent loss of crew. In the

[**] Other risk measures, such as the probability of loss of mission (P(LOM)), might be relevant in addition to or instead of P(LOC), for example in the case of robotic missions. However, for convenience, P(LOC) is used in this paper to refer to aggregate risk measures generally.

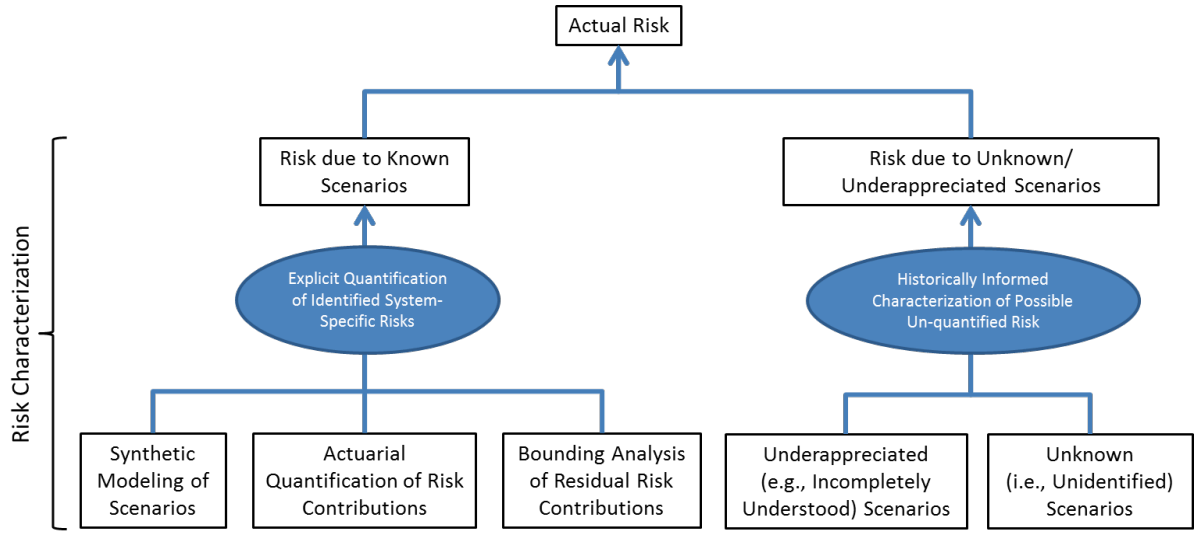
[††] PRA methods treat uncertainty quantitatively, e.g., by quantifying it at the basic event level and propagating it through the logic model to produce a probability distribution for the risk measure (e.g., P(LOC)). This enables statistics such as the mean value to be estimated. Additionally, the risk-driving uncertainties can be identified for prioritization of uncertainty reduction efforts.

[‡‡] PRA methods are generally effective at identifying system failures that result from combinations of failures events that propagate through the system due to the functional dependencies of the system that are represented in the risk model. Their use, especially during design, can contribute significantly to safety, since they enable risk reduction efforts to be focused on those known issues of greatest safety significance.

Columbia accident, insulating foam from the external tank impacted the wing leading edge reinforced carbon-carbon (RCC), puncturing it and allowing an entryway for hot plasma upon reentry into the Earth's atmosphere.[§§]

The situation is illustrated in Figure 4. On the left side of the figure is the risk due to known scenarios that are included in the risk analysis. Depending on the nature and magnitude of a given scenario, quantification may be based on synthetic, actuarial, or bounding methods. Synthetic methods are used when system-level risk data are scarce and must instead be constructed via probabilistic modeling of accident scenario initiation and propagation. Actuarial methods can be used when the volume of data supports quantification of demonstrated risk with reasonable degree of certainty. Bounding methods are typically used for residual risk contributors, provided that their aggregate contribution to explicitly quantified risk is small. In contrast, the right side of Figure 4 shows the risk due to the unknown and underappreciated scenarios (referred to throughout the remainder of this paper as "UU scenarios").[***] Their presence in the system is inferred by the historical observation that aerospace systems have consistently experienced scenarios that were either not identified during design, or whose probability and/or magnitude was underrepresented in analysis.

Figure 4: Characterizing the Contributors to Actual Risk



The Aerospace Safety Advisory Panel (ASAP) [8] and others have raised the need to consider the gap between actual risk and explicitly quantified risk when applying safety thresholds and goals. This concern reflects the expectation that during the early stages of operation there is likely to be significant risk due to UU scenarios. NASA's agency-level safety thresholds and goals do not explicitly address the question of how to account for UU scenarios. Nevertheless, the expectation is that the safety thresholds and goals refer to the system's actual risk, which includes both known, adequately modeled scenarios as well as UU scenarios.[†††]

[§§] Because of the often holistic and environment-dependent nature of such interactions, they tend not to be revealed by subsystem testing. 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 sufficient volume of testing. Consequently, they tend to remain unknown (or known but underappreciated) until they manifest as an accident.

[***] Note that the term "underappreciated" refers here to scenarios whose likelihoods of occurrence or severity of impact are underestimated because of a deficiency in knowledge about the etymology of the risk. This differs from risks that are underestimated because they have not been analyzed in detail. The Challenger and Columbia accidents would fall in the category of underappreciated risks as we define the term here.

[†††] The position that UU scenarios should be acknowledged and accounted for when evaluating whether safety thresholds and goals are satisfied is also reflected in the ASAP Annual Report for 2011 [8].

3.2. Addressing Risk Analysis Incompleteness via Safety Performance Margin

One possible approach for characterizing the contribution of UU scenarios to the actual risk is imbedded in the concept of *safety performance margin* (referred to in the NASA System Safety Handbook as *safety risk reserve*). In the safety performance margin approach, the actual risk of a system is understood to be the sum of the risk from known scenarios, as explicitly quantified using traditional risk analysis methods, plus the risk from UU scenarios, as characterized by the safety performance margin. Just as the minimum tolerable level of safety starts at an initial value (the safety threshold) and diminishes to a lower value as flight experience is gained (the safety goal), the safety performance margin starts an initial value that is consistent with historical information about the magnitude of risks from UU scenarios and subsequently diminishes with time as safety performance information is gained through system operation (including tests). The concept is illustrated in Figure 5.

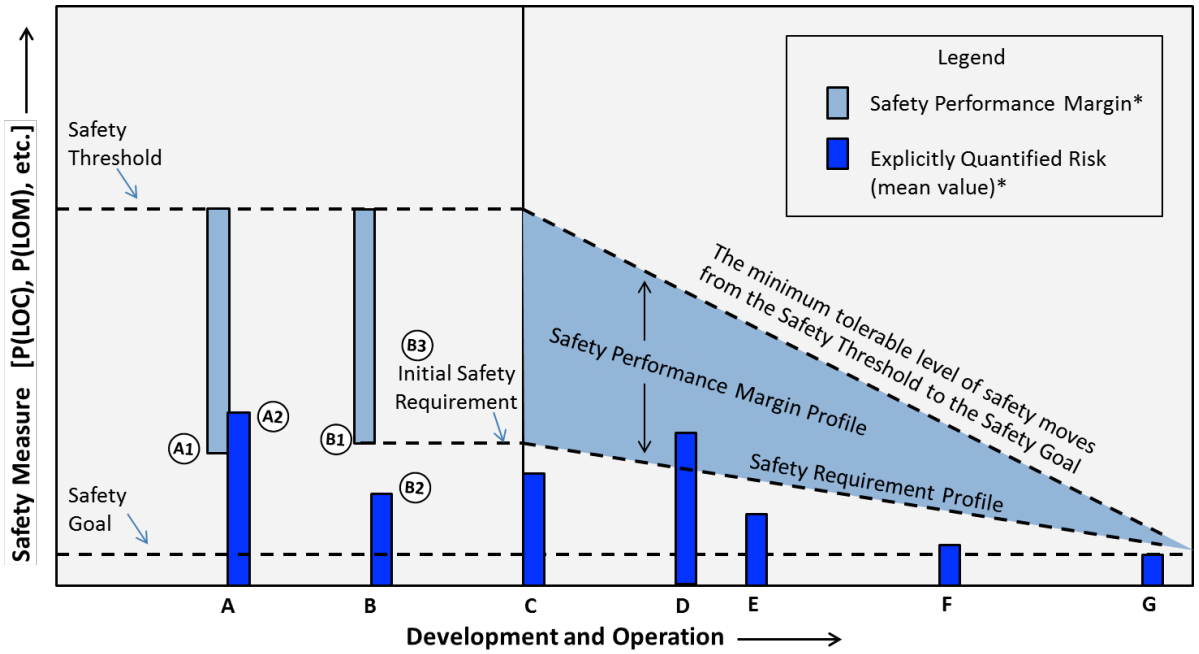
A limit on the allowable explicitly quantified risk can be derived by subtracting (in risk terms) the safety performance margin from the minimum tolerable level of safety. If the explicitly quantified risk is within this limit, then by implication there is reasonable assurance that the actual risk is within the minimum tolerable level of safety. Methods for establishing an initial safety performance margin and a margin draw-down profile based on historical data for similar systems are currently being investigated by the NASA Headquarters Office of Safety and Mission Assurance (OSMA).

As an example, Figure 5 illustrates system development and operation 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 length 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).[‡‡‡] At this point an initial safety requirement for explicitly quantified risks is specified (B3) as being equal to the value of the safety measure at point B1. Additionally, a decreasing safety performance margin profile can be derived, 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.[§§§] During system operation, proactive safety upgrade and improvement programs reduce the risk in increments (F) until the safety goal is met (G).

[‡‡‡] 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.

[§§§] 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 re-baselining the safety requirement, if organizational and/or programmatic changes provide an adequate basis for reducing the safety performance margin.

Figure 5: Conceptual View of Managing Safety Performance consistent with the Minimum Tolerable Level of Safety



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

4. MAKING THE CASE THAT THE SYSTEM MEETS THE MINIMUM TOLERABLE LEVEL OF SAFETY

The NASA System Safety Handbook introduces the construct of the risk-informed safety case (RISC), 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." The safety case addresses all aspects of safety, including the standing of the system relative to the minimum tolerable level of safety.[****] The elements of the RISC are [14]:

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

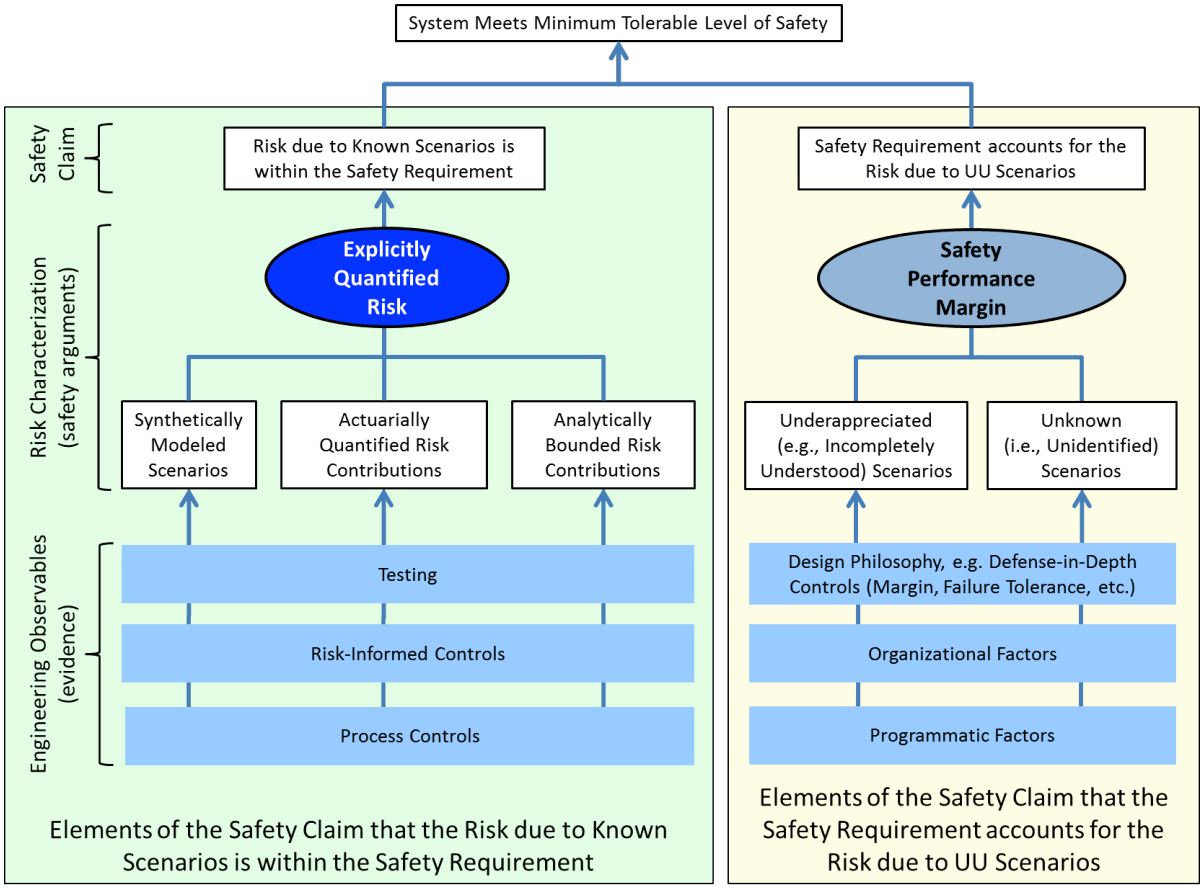
The claims made (and defended) by the RISC dovetail with the safety objectives negotiated at the outset of system formulation. In other words, the satisfaction of each distinct safety objective is stated as a corresponding distinct claim in the RISC. By substantiating each claim with appropriate arguments and supporting evidence, the RISC demonstrates that the corresponding objective has been met and, thus, that the system is adequately safe.

[****] For example, the claim that the system is ASARP is also within the scope of the RISC, as is the claim that proactive safety upgrade and improvement programs will enable the safety goal to be met at some (reasonable) point in the future.

Figure 6 builds on the decomposition of actual risk into known and UU components as shown in Figure 4, notionally illustrating the structure of the safety claim that the system meets the minimum tolerable level of safety. This claim is the conjunction of two sub-claims: the risk due to known scenarios is within the safety requirement; and the safety requirement accounts for the risk due to UU scenarios. If both of these sub-claims are valid, then by logical implication the top-level claim is valid. The claim that the known risk is within the safety requirement is addressed via quantitative risk analysis that analyzes observable system attributes, such as test results/plans and controls, to quantify the known risk. As indicated by the term "risk-informed" in the figure, known risk can be explicitly managed by applying controls that are specifically designed to effectively mitigate or eliminate identified scenarios. Indeed, quantitative risk analysis provides a basis for evaluating the potential benefits of different control strategies, particularly during system design.

In contrast, the minimization of UU risk chiefly relies on broad measures that reflect the degree to which safety considerations factor into the management of the program/project. Such considerations provide assurance that the system is robust against unexpected stresses, and can maintain operability (or at least safety) in the event of failures within the system. Of course, these factors also affect known risk, e.g., by assuring the functionality of risk-informed controls and/or by providing additional measures of defense in the event of control failure.

Figure 6: The RISC and the Safety Requirement Provide Assurance that the System Meets the Minimum Tolerable Level of Safety (notional)

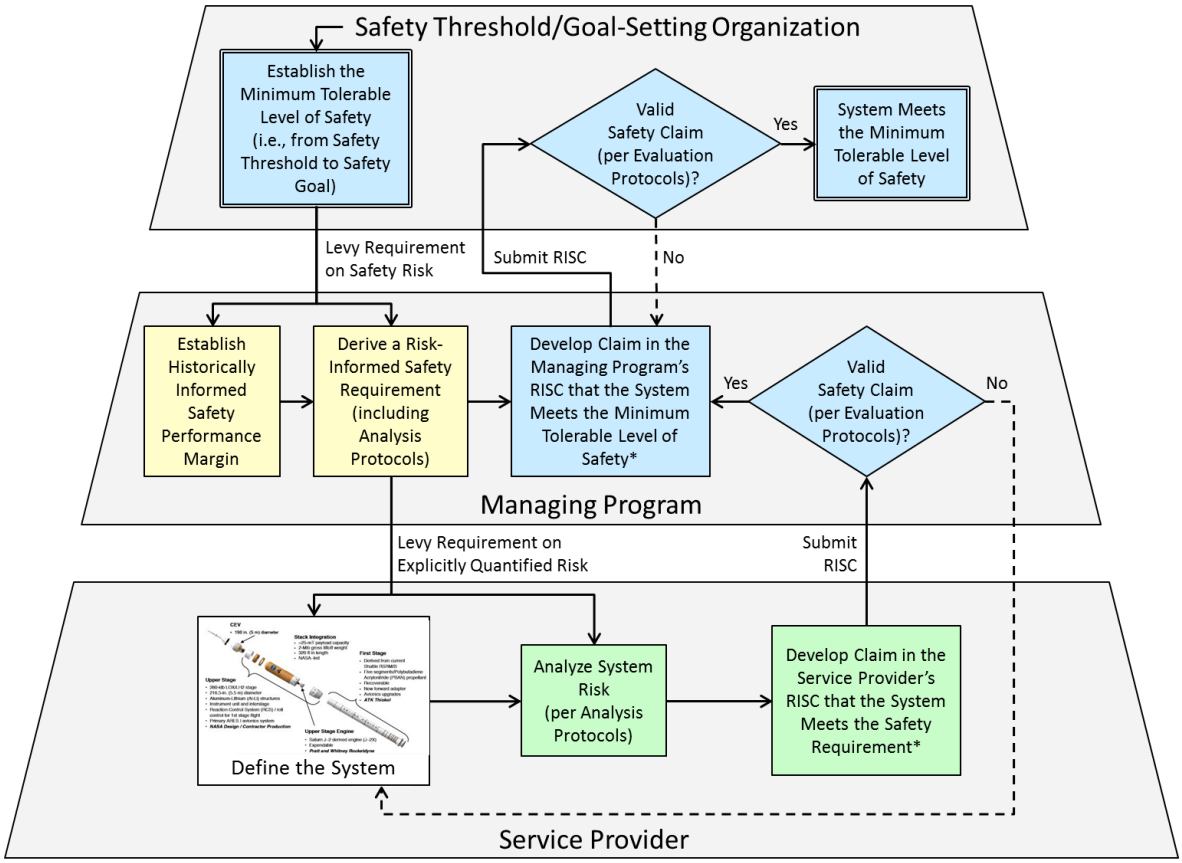


It is the responsibility of the organizational unit upon whom the safety requirement is levied to make the case that the system meets it. However, it is not the responsibility of that organization to make the

case that the requirement itself is appropriately derived.[††††] That responsibility rests with the organizational unit at the next higher level of the NASA hierarchy, who owns the requirement and is responsible for risk management oversight of the lower-level unit. Correspondingly, the case that the safety requirement accounts for the risk from UU scenarios is the responsibility of this higher-level unit.

The situation is illustrated in the notional example of Figure 7. In this figure, an upper-level organization establishes the minimum tolerable level of safety for a service of the type to be provided by a service provider. This minimum tolerable level of safety is levied on the program that is managing the service provider and focuses on actual safety risk. The managing program operationalizes this requirement by deriving a safety requirement that is levied on the service provider and focuses on explicitly quantified risks. This safety requirement, including quantitative risk analysis protocols for demonstrating that the requirement is met, is informed by a safety performance margin that is based on relevant historical data, supplemented with applicable system and organizational information from the service provider.

Figure 7: Activity Flowchart for Meeting the Minimum Tolerable Level of Safety



* The RISC is comprehensive with respect to safety, and addresses other aspects of safety, such as ASARP.

The service provider demonstrates adherence to the safety requirement by conducting a risk analysis per the specified analysis protocols, and, in the service-provider-generated-RISC, makes the claim that the safety requirement is met. This RISC is submitted to the managing program, who evaluates the claim according to established protocols. If it is sound, the managing program uses it, in conjunction with the claim that the safety requirement has been appropriately derived, to make the claim in its own RISC that the system meets the minimum tolerable level of safety established by the safety

[††††] Although it is the responsibility of that organization to establish the feasibility of successfully meeting it, per NPR 8000.4A.

threshold/goal-setting organization. This RISC is submitted to the upper-level organization, who evaluates the managing program's claim according to appropriate protocols. If it is sound, then there is a sound basis for concluding that the system meets the minimum tolerable level of safety.

5. CONCLUSION

This paper discusses the role of the Agency's safety thresholds and goals in the context of "adequate safety" as discussed in the NASA System Safety Handbook. The following bullets summarize the discussion and reiterate some of the challenges and potential approaches to its implementation:

- "Adequate safety" for a given technology means both (1) meeting a minimum tolerable level of safety that is acceptable to the decision maker, and (2) being as safe as reasonably practicable (ASARP) such that safety performance is given priority relative to technical, cost, and schedule performance, given the subject technology and program constraints.

- NASA has instituted requirements for establishing safety thresholds to be used by the Agency as criteria for program acquisition decisions, and safety goals that reflect expectations about the long-term safety performance that is achievable from the system as its design and technology mature. The goal and threshold are currently specified in terms of a risk measure: the probability of loss of crew (P(LOC)). It is necessary to operationalize the thresholds and goals by providing analysis protocols explaining agency expectations regarding the analysis that supports the performance claims, as well as evaluation protocols explaining how the decision-making entity will review the analysis and apply its results. In principle, these protocols need to be based on applicable decision situations.

- Experience with various technologies has shown repeatedly that risk analysis incompleteness is a serious issue. Especially for new systems, the experience base is too meager to provide strong evidence of low risk, and UU (unknown/ underappreciated) scenarios may exist that are not adequately modeled in the analysis. In fact, a recent retrospective of Space Shuttle flight experience performed by NASA has shown that the risk from UU scenarios was initially at least three times as large as the risk from known and adequately appreciated scenarios [15]. The NASA System Safety Handbook introduces the concept of safety performance margin as one possible approach for characterizing the risk from UU scenarios. In this approach, the safety requirement is derived as the difference between the minimum tolerable level of safety (initially the safety threshold) and a safety performance margin derived from historical data for similar systems and adjusted based on system/organizational-specific factors. This requirement is used to assess and manage safety performance as explicitly quantified via traditional risk analysis methods.

- Demonstration of adequate safety to decision makers requires making a coherent case, supported by evidence, that all relevant safety objectives have been met, including, but not limited to, meeting minimum tolerable levels of safety. The NASA System Safety Handbook has introduced the construct of the risk-informed safety case (RISC) as the vehicle by which a claim of adequate safety is conveyed. The RISC serves as a comprehensive proxy for the safety of the system, stressing appropriate processes, clearly stating the assumptions that must be actualized if the safety claims are to remain valid, and committing to ongoing analysis of operating experience, so that safety performance improves continuously. The RISC is not a radically new idea. Rather, it is proposed as a formalization and integration of processes and ideas that are already in place or being incorporated to support decision contexts such as certification. The RISC is meant to subsume those processes, and furnish a coherent argument for how safe the system is or will be.

Acknowledgements

The research was carried out at NASA, and at Information Systems Laboratories, Idaho National Laboratory, and the Jet Propulsion Laboratory, California Institute of Technology, under contract with NASA.

References

[1] NASA. NPR 8705.2B, *Human-Rating Requirements for Space Systems*, Washington, DC. 2008.
[2] NASA. NASA/SP-2010-580, *NASA System Safety Handbook*, Washington, DC. 2010.
[3] U.S. Code of Federal Regulations, 10 CFR 20, *Standards for Protection Against Radiation*, Washington, DC. 1991.
[4] Parliament of the United Kingdom, *Health and Safety at Work etc. Act*, London, UK. 1974.
[5] NASA. NASA/SP-2010-576, *NASA Risk-informed Decision Making Handbook*, Washington, DC. 2010.
[6] 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.
[7] NASA. NASA/SP-2011-3421, *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners*, Second Edition, Washington, DC. 2011.
[8] ASAP. *Aerospace Safety Advisory Panel Annual Report for 2011*, Washington, DC. 2012.
[9] Review of U.S. Human Spaceflight Plans Committee, "Seeking a Human Spaceflight Program Worthy of a Great Nation," Washington, DC, October 2009.
[10] Kaplan, S., et al., "On the Quantitative Definition of Risk," Risk Analysis, Vol. 1, No. 1. 1980.
[11] Leveson, N., "The Use of Safety Cases in Certification and Regulation," Journal of System Safety, Vol. 47, No. 6. 2011
[12] Vesely, W., et al., "Demonstrating the Safety and Reliability of a New System or Spacecraft: Incorporating Analyses and Reviews of the Design and Processing in Determining the Number of Tests to be Conducted," NASA Office of Safety and Mission Assurance, Washington, DC. 2010.
[13] Morse, E., et al., "Modeling Launch Vehicle Reliability Growth as Defect Elimination," Valador Inc., Herndon, VA. 2011.
[14] Bishop, P., and Bloomfield, R., "A Methodology for Safety Case Development, Safety-Critical Systems Symposium," Birmingham, UK. 1998.
[15] Hamlin, T., et al., "Shuttle Risk Progression: Use of the Shuttle PRA to Show Reliability Growth," AIAA SPACE Conference & Exposition. 2011.