Current Issues of the Safety Goal and Safety Threshold Policy

Frank Groen - OSMA

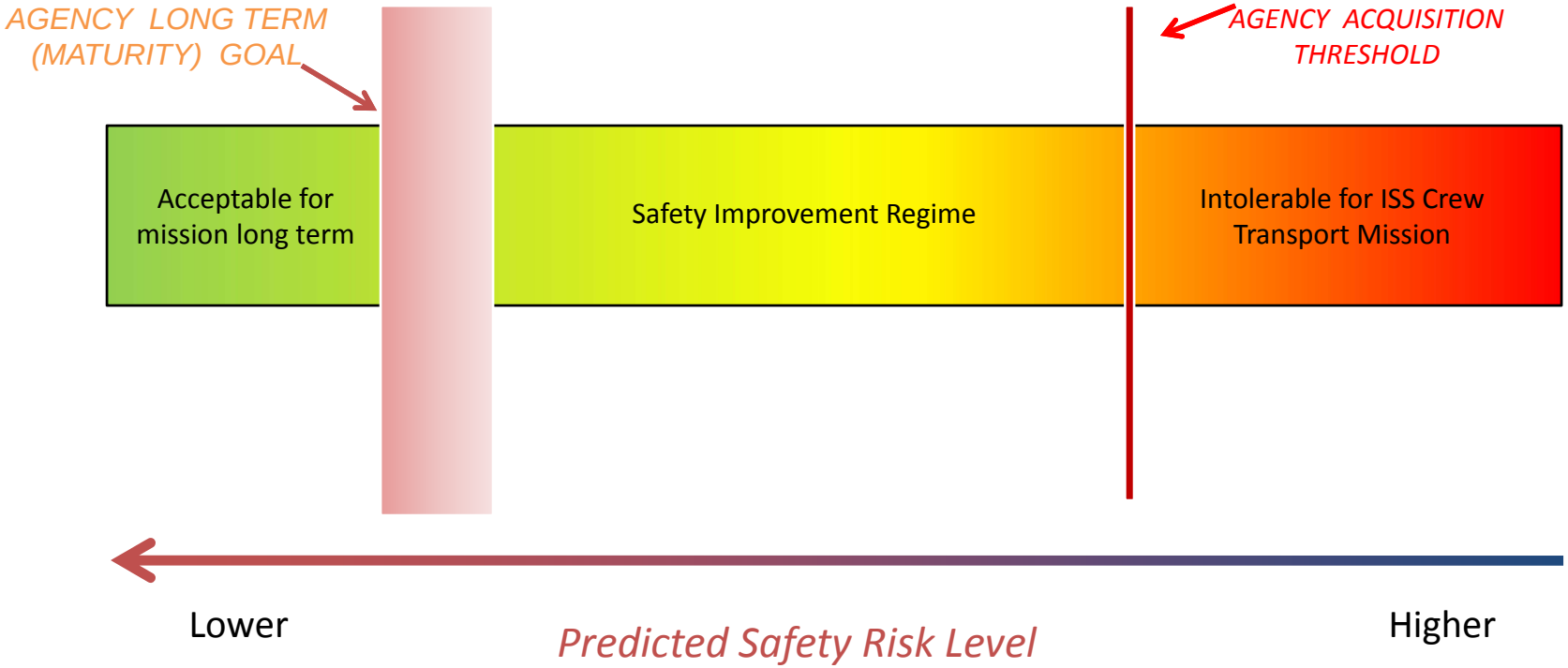
December 2011

Origin of the Policy

- NASA's Human-Rating NPR 8705.2B introduces a new failure tolerance requirement:
  - Level/implementation to be decided based on integration of safety and engineering analyses.

- NASA's oversight panel (Aerospace Safety Advisory Panel):
  - Approves of the shift from a prescriptive to a risk-informed approach to safety
  - Expresses concern that "approach is viable only if a common understanding of 'sufficiently safe' exists"
  - Recommends that "NASA stipulate directly the HRR acceptable risk levels including confidence intervals for the various categories of activities […] to guide managers and engineers in evaluating 'how safe is safe enough'"

Agency Risk Tolerance Policy for Mission Design



- Quantitative criteria to be applied as part of larger set of design and assurance methods, including evaluation and control of individual hazards

Example: ISS Crew Transport Threshold

Design Threshold: At a minimum, the spaceflight system designed for transport of the crew to the ISS shall be at least as safe for the combined ascent and entry phases as the Space Shuttle was at the end of its operational life, and in the aggregate, for a 210-day mission to ISS, the system shall be at least as safe as the Space Shuttle was at the end of its operational life on a 12-day mission to the ISS.

- Continuous safety upgrade and improvement program is also required

src: NASA/SP-2010-580
NASA System Safety Handbook

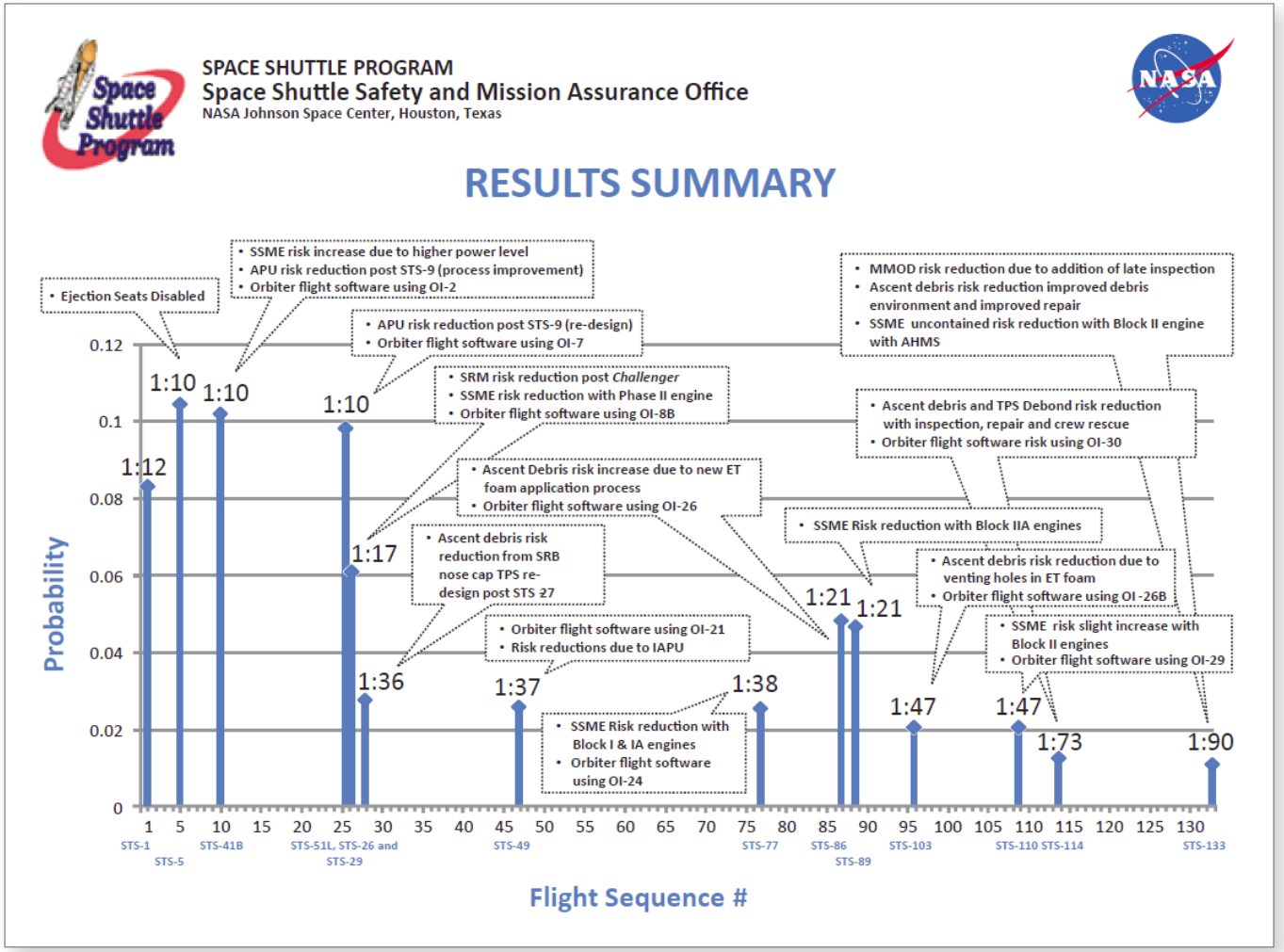
Impact of the Policy

- To greater or lesser extent, policy drives programs to:
  - Consider total amount of crew safety (and mission success) risk, not just individual risks
  - Consider impact of design decisions in early trades
  - Consider applicability of traditional design rules to new mission types
  - Select processes to identify and manage sources of risk based on safety performance criteria
- Created technical and organizational challenges

Basis for Estimation

- Different bases for estimation are considered:
  - First Flight: no flight demonstration
  - Operational Flight: mission capability validated
  - Design Capability: only uncontrollable risks remain
- Require consideration of different sources of risk, e.g.,
  - Potential for systemic issues, incorrect assumptions
  - Random conditions and events
- Would result in varying risk estimates
- Design capability assumption used in past programs
  - “Mature system”

2010 Shuttle Risk Evolution Study



- Shows progression due to design changes, not knowledge improvements (e.g., system validation)

src: ASAP Annual Report 2011

Consideration of Risk for New Designs

- For new designs, most risk due to lack of validation rather than random failures
  - Lack of knowledge rather than random processes
- Many accident scenarios cannot be modeled explicitly
  - "Completeness" issue
- Estimated risk reduces faster than statistical growth models would indicate, but initial risk is very high
- Both technical and (external) communication issue

Final Considerations

- Policy has emphasized safety as key design attribute
  - Next to technical performance, cost, and schedule
- Modeling and test methods still subject of debate
  - Allowances, discount factors, new model types
- System acceptance decisions will be strongly influenced by other assurance methods
  - Design standards, identification and control of hazards, quality programs, test/demonstration, …

BACKUP

ASAP Recommendation 2006-03-02

"During our meeting at the Kennedy Space Center, the ASAP noted, with regards to risk assessments that are being made to support launch decisions, it appears that a series of fragmented, non-standardized tools and methodologies are in use. The result is that risk recommendations to senior management concerning individual hazards effecting launch are sometimes made in isolation without consideration of overall launch risk. For example, the most recent Shuttle launch focused heavily on two of the 569 potentially catastrophic hazards currently known to exist, without any assessment of the overall likelihood of such a catastrophic failure. A lack of confidence in the technical basis for the assessments also appears to sometimes exist, and variations in risk matrix definitions among programs have been observed. Lastly, only limited guidance is available concerning Agency policies on what risks should be accepted under what conditions.

The ASAP recommends that a comprehensive risk assessment, communication and acceptance process be implemented to ensure that overall launch risk is considered in an integrated and consistent manner The process should be sound, mature, consistently implemented to yield high confidence and consistent results that are generally accepted by the majority of the community."

[emphasis added]

ASAP Recommendation 2009-01-02a

"The recently revised HRR standard represents a fundamental shift from telling developers how to create a safe design (by relying primarily on redundancy) to establishing a process for using a risk-informed design approach to produce a design that is optimally and sufficiently safe. The ASAP applauds switching to such a performance-based approach because it emphasizes early risk identification to guide design, thus enabling creative design approaches that might be more efficient, safer, or both. However, this approach is viable only if a common understanding of "sufficiently safe" exists, and the current HRR procedures leave that determination to individual programs, which could lead to inconsistent "safe-enough" thresholds among various developers if not carefully managed. This consequence could be especially problematic for development (and possible future use by NASA) of potential future human-rated vehicles produced by organizations external to NASA, such as Commercial Orbital Transportation System (COTS) firms or the programs of other nations.

The ASAP recommends that NASA stipulate directly the HRR acceptable risk levels including confidence intervals for the various categories of activities (e.g., cargo flights, human flights)-to guide managers and engineers in evaluating "how safe is safe enough." These risk values should then be shared with other organizations that might be considering the creation of human-rated transport systems so that they are aware of the criteria to be applied when transporting NASA personnel in space. Existing thresholds that the CxP has established for various types of missions might serve as a starting point for such criteria."

[emphasis added]

Technical Requirement in NPR 8705.2B w/change 2

3.2.2 The space system shall meet probabilistic safety criteria derived from the Agency-level safety goals and safety thresholds with a specified degree of certainty (Requirement).

Rationale: Probabilistic safety analysis methods provide one basis for the comparison of design options with regards to safety (see paragraph 2.3.7.1). Probabilistic safety requirements defined in accordance with paragraph 1.4.6 establish criteria for safety metrics such as loss of crew probabilities that are an outcome of such analyses. The analyses must consider the uncertainty associated with calculated values and the degree of certainty that the probabilistic criteria are met. The required degree of certainty is specified as part of the probabilistic safety requirements.

Even when these metrics are determined in accordance with accepted analysis protocols, it is recognized, however, that as an analytical tool, probabilistic safety analysis methods rely on assumptions and are subject to uncertainties. Calculated values of such safety metrics are therefore not in themselves sufficient to determine that a system is safe. Consequently, compliance with probabilistic requirements can only be an element of the case to be made that a system provides an acceptable level of safety.

A.46 Probabilistic Safety Requirement: The specification of a criterion for a probabilistic safety metric (e.g., the probability of a loss of crew) and the degree of certainty with which such criteria must be met.