UNIVERSITY OF VIRGINIA

SCHOOL OF ENGINEERING AND APPLIED SCIENCE

DEPARTMENT OF SYSTEMS AND INFORMATION ENGINEERING

NASA AVIATION SAFETY PROGRAM SYSTEMS

ANALYSIS/PROGRAM ASSESSMENT METRICS REVIEW

For

NASA-LANGLEY RESEARCH CENTER /

AVIATION SAFETY PROGRAM

SHARON M. JONES – PROGRAM OFFICER

By

GARRICK E. LOUIS, PRINCIPAL INVESTIGATOR

KATHERINE ANDERSON

TISAN AHMAD

ALI BOUABID

MAYA SIRIWARDANA

PATRICK GUILBAUD

FINAL

TABLE OF CONTENTS i

LIST OF FIGURES

ABBREVIATIONS AND ACRONYMS

AHP	Analytic Hierarchy Process
AM	Accident Mitigation
ATM	Air Traffic Management
ARIBA	ATM system safety criticality Raises Issues in Balancing Actors responsibility
ASAFE	Aviation Safety Analysis and Functionality Evaluation
ASIST	Aviation Safety Investment Strategy Team
ASMM	Aviation System Modeling and Monitoring System
AvSP	Aviation Safety Program
BBN	Bayesian Belief Network
CAST	Commercial Aviation Safety Team
CFIT	Center Flight Into Terrain
CI	Consistency Index
CR	Consistency Ratio
DoD	Department of Defense
FAA	Federal Aviation Administration
FOOQA	Flight Operational Quality Assurance
GA	General Aviation
GAIN	Global Aviation Information Network
IRL	Implementation Risk Level
JSAT	Joint Safety Analysis Team
JSIT	Joint Safety Implementation Teams
LaRC	Langley Research Center
LMI	Logistics Management Institute
MMH/FH	Maintenance Man Hours/Flight Hours
M&T	Method and Tools
NAS	National Aerospace System
NASA	National Aeronautics and Space Administration
OAST	Office of Aero-Space Technology
OAT	Office of Aerospace Technology

PAT	Program Assessment Team
PCA	Program Commitment Agreement
PRA	Probabilistic Risk Assessment
ROI	Return-on-Investment
R&D	Research and Development
R&T	Research and Technology
SAAP	Single Aircraft Accident Prevention
SV	Synthetic Vision
SWAP	System Wide Accident Prevention
WG B	Working Group B
WxAP	Weather Accident Prevention

EXECUTIVE SUMMARY

The goal of this project is to evaluate the metrics and processes used by NASA's Aviation Safety Program in assessing technologies that contribute to NASA's aviation safety goals.

There were three objectives for reaching this goal. First, NASA's main objectives for aviation safety were documented and their consistency was checked against the main objectives of the Aviation Safety Program. Next, the metrics used for technology investment by the Program Assessment function of AvSP were evaluated. Finally, other metrics that could be used by the Program Assessment Team (PAT) were identified and evaluated.

This investigation revealed that the objectives are in fact consistent across organizational levels at NASA and with the FAA.

Some of the major issues discussed in this study which should be further investigated, are the removal of the Cost and Return-on-Investment metrics, the lack of the metrics to measure the balance of investment and technology, the interdependencies between some of the metric risk driver categories, and the conflict between "fatal accident rate" and "accident rate" in the language of the Aviation Safety goal as stated in different sources.

BACKGROUND

The goal of the NASA Aviation Safety Program (AvSP) is to develop and demonstrate technologies that contribute to a reduction in the aviation fatal accident rate by a factor of 5 by year 2007 and by a factor of 10 by year 2022.[1] The program is a partnership that includes NASA, the Federal Aviation Administration (FAA), the aviation industry and the Department of Defense. NASA's role is to develop technology and research needed to help the FAA and industry partners to achieve the President's challenge to improve aviation safety. The DoD's main role is to provide access to useful data and certain technologies. The NASA Aviation Safety Program has defined products that will possibly modify airline and/or air traffic control (ATC) operations, enhance aircraft systems, and improve the identification of potential hazardous situations within the National Aerospace System (NAS).

GOAL

The goal of this project is to evaluate the metrics and processes used by NASA's Aviation Safety Program in assessing technologies that contribute to NASA's aviation safety goals.

OBJECTIVES

There are three primary objectives in fulfilling this goal:

1. To document NASA's three main objectives for aviation safety and check their consistency with the three main objectives of the Aviation Safety Program.
2. To evaluate the metrics used for technology investment by the Program Assessment function of AvSP.
3. To identify and evaluate other metrics that could be used by the Program Assessment Team (PAT).

PROJECT SCHEDULE AND DELIVERABLES

Below is a table outlining our project schedule.

Project Schedule		
Date	Deliverable	Status
7/5	1st Draft of Written Report Due	Completed
7/12	LaRC comments	Completed
7/19	Oral Presentation	Completed
7/26	LaRC comments	Completed
8/2	Final Written Report Due	Completed

1.0 OBJECTIVES OF NASA AND AVSP

1.1 NASA'S THREE MAIN OBJECTIVES IN IMPROVING AVIATION SAFETY

NASA's three main objectives in improving aviation safety are to:

1. Increase accident survivability *(Figure 1: 3.1)*
2. Eliminate targeted accident categories *(Figure 1: 3.2)*
3. Strengthen safety technology foundation *(Figure 1: 3.3)* [2]

The following NASA and AvSP investment areas are supported by the above objectives:

1. Accident Mitigation *(Figure 1: 3.1.1)*
2. Accident Prevention *(Figure 1: 3.2.1)*
3. System monitoring and modeling *(Figure 1: 3.3.1)* [3]

The Aviation Safety Program's two main objectives in helping to achieve the accident reduction goal are:

1. Develop technologies that reduce aviation injuries and fatalities when accidents do occur *(Figure 1: 3.1.1.1)*
2. Develop and demonstrate technologies that reduce aircraft accident rates *(Figure 1: 3.2.1.1)*[4]

1.2 SOURCES

The sources of information used to present these objectives are:

- NASA Aviation Safety Program
  Program Plan, 08/01/1999.

- NASA Aerospace Technology Enterprise
  Website: www.aero-space.nasa.gov/goals/index.htm
  Website: www.aero-space.nasa.gov/goals/safety.htm

- Toward a safer 21st Century: aviation safety research baseline and future challenges
  Website: http://www.aero-space.nasa.gov/library/safer21C.htm
  This report explains the baseline upon which the current NASA/FAA Partnership for Aviation Safety Research was developed.

- NASA Program Commitment Agreement (PCA), Program Plan, Technical Integration Plan versions 1.0 and 2.0, Projects and Element Plans
  - Signed PCA 7/6/00
  - AvSP Program Plan 8/1/99 (access 06/20/2002)

4

- Technical Integration Plan
- project plans (AM, SWAP, SAAP and WxAP, ASMM, SVS, Aircraft Icing) see also http://icebox.grc.nasa.gov/

Website:
https://postdoc.arc.nasa.gov/postdoc/t/folder/main.ehtml?url_id=6460
These documents were present the objectives and the metrics used in the AvSP

- NASA ASIST
  Website: http://avsp.larc.nasa.gov/pdfs/ASIST.pdf
  This presentation presents the criteria for NASA Investment
  Website: http://avsp.larc.nasa.gov/about.html

- The Three Pillars for Success
  Website: http://oea.larc.nasa.gov/news_rels/1997/May97/97_35.html
  http://stdweekly.msfc.nasa.gov/techpapers.html
  http://www.aero-space.nasa.gov/goals/index.htm

- Federal Aviation Administration (FAA)

  Website: http://www.faa.gov/AviationSafety/index.htm

  This document presents the FAA's strategic goal and objectives for improving aviation safety.

- Department of Defense (DoD)
  Website: http://www.aero-space.nasa.gov/library/dod.htm
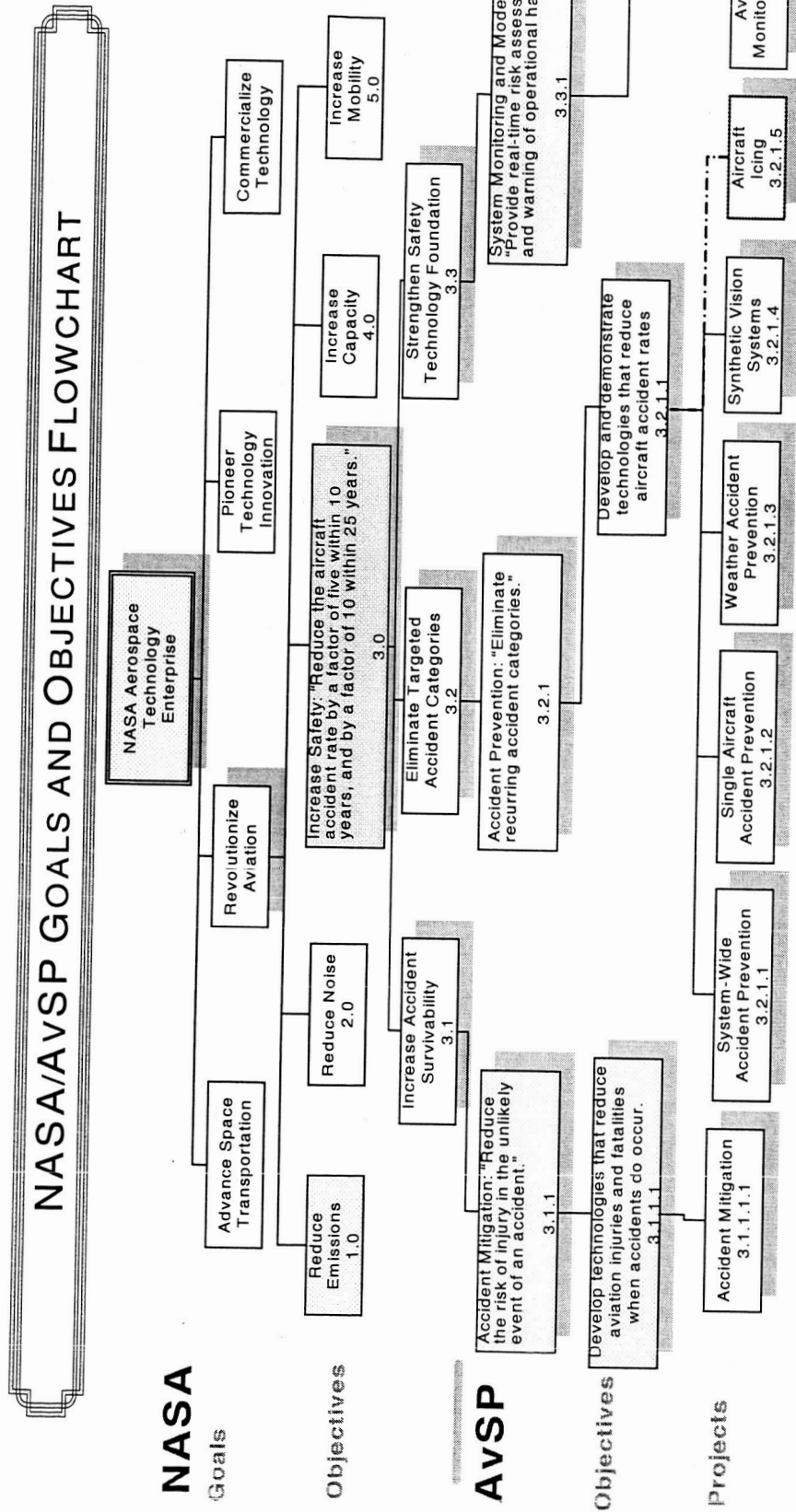
1.3 NASA/AvSP Goals and Objectives Flowchart

Figure 1: Goals and Objectives Flowchart

1.4 SUMMARY DISCUSSION OF THE AvSP OBJECTIVES AND THEIR RELATIONSHIP TO NASA'S AVIATION SAFETY GOAL

The NASA goal of improving aviation safety is to reduce the aircraft accident rate by a factor of five within 10 years, and by a factor of 10 within 25 years. NASA has identified three objectives to reach this goal:

1. Eliminate targeted accident categories
2. Increase accident survivability
3. Strengthen safety technology foundation[5]

The first objective will be accomplished through key technical developments such as precision approach and landing technologies, affordable technologies and systems for data-linked communication and on-board graphical display of critical aviation weather information, turbulence modeling and detection technologies, and synthetic vision technologies.[6]

The second objective involves the development of advanced structural and material designs and fire hazard mitigation products. This objective does not appear to directly contribute to the reduction of the accident rate, which is the NASA goal that led to the creation of AvSP. However, it does increase safety by mitigating the consequence of an accident, which is a general NASA goal.

The third objective is achieved through aviation system modeling, human-error assessment methodologies, and integrated aviation system monitoring tools.

The Aviation Safety Program objectives are derived from NASA's three main safety improvement strategies: Accident Mitigation, Accident Prevention, and System Monitoring and Modeling, which in turn are derived from the objectives above.

The Aviation Safety Program's two main objectives are:

1. Develop technologies that reduce aviation injuries and fatalities when accidents do occur.
2. Develop and demonstrate technologies that reduce aircraft accident rates.

The program is structured around six projects. The first project, Accident Mitigation, is focused on increasing accident survivability and on reducing fatalities when accidents do occur. The decrease in the number of fatalities and injuries that will result from this reduction in risk will lead to NASA's objective of increasing accident survivability. The AvSP's objective of developing technologies that reduce aviation injuries and fatalities when accidents do occur is an attempt to satisfy this top-level NASA objective.

The next four projects, System-Wide Accident Prevention, Single Aircraft Accident Prevention, Weather Accident Prevention and Synthetic Vision Systems, support the accident prevention. They are focused on eliminating target accident categories.

Accident Prevention is defined as identifying interventions and developing technologies to eliminate the types of accidents that can be categorized as "recurring."[7] The AvSP's second objective is a response to this NASA strategy, although there is no emphasis on "recurring" accidents in the wording of AvSP's objective. According to the AvSP Program Commitment Agreement, the second objective encompasses not only the development of accident reduction technologies; it also includes the development of information technologies needed to build a safer aviation system. This particular aspect can be connected to the NASA objective of strengthening the safety technology foundation. As stated in the PCA, these four projects are intended to satisfy this objective. However, the connection of these projects to Accident Prevention is more obvious than the connection to strengthening the safety technology foundation.

The last project, Aviation System Monitoring and Modeling, is focused on strengthening the overall aviation system foundation. System Monitoring and Modeling seeks to provide real-time risk assessment and warning of operational hazards. This is the main part of NASA's effort to strengthen the technology foundation. There is no apparent link between this project and the objectives of the AvSP. However, this project supports one of the three objectives of NASA in improving aviation safety.

The above discussion illustrates the manner in which each of the AvSP projects contributes to the overall NASA goal in improving aviation safety by reducing the aircraft accident rate by a factor of five within 10 years and a factor of ten within 25

years. The Accident Mitigation project is not directly linked to this goal as stated, but instead helps to achieve NASA's objective of increasing accident survivability.

2.0 CURRENT METRICS USED WITHIN AvSP

2.1 PROGRAM ASSESSMENT OBJECTIVES AND METRICS

The figure below represents the Technical Integration Work Breakdown Structure used in the AvSP. There are four functional elements:

- Systems engineering
- Program assessment
- Product assurance
- Flight integration

In this project, we have focused our research on the functional element: Program assessment. This element has two primary objectives. The first one is the impact on safety, or the assessment of the impact of each product on accident reduction and/or future impact on aviation safety. The second objective is the balance between investment and technology focus. For that, the Program Assessment team periodically reviews the AvSP portfolio to ensure that a proper balance of investment and technology focus remains.

The Intermediate Program Assessment defines three main metrics used by the Program Assessment Team to determine the projected impact of safety technologies upon increasing aviation safety and to ensure that the AvSP research portfolio remains properly balanced between focused and broad-based solutions. These three metrics are (1) Implementation Analysis, (2) Technical Development Risk, and (3) Safety Benefit. Previously there were five metrics, but Technology Lifecycle Cost and Return-on-Investment (ROI) have been eliminated. The following two sections will offer a discussion of the three current metrics, and then the two former metrics.

Figure 2: PAT Objectives and Metrics

2.2 LIST OF THE CURRENT METRICS

The Program Assessment Team has used the three following metrics in the Intermediate Program Assessment:

1. Implementation Risk
2. Technical Development Risk
3. Safety Benefit

The assessment process is common to these three metrics and the two others, Cost and ROI, used in the preliminary program assessment. For each product of the AvSP, there are two main assessments:

1. Capability
2. System impact

By using a risk categorization, these assessments allow an evaluation of the impact of the product on reduction of the accident rate. The Risk level categorization is based on three levels of risk:

10

Low (Green), Medium (Yellow) and High (Red).

The criteria for level of risk are specific to each metric.
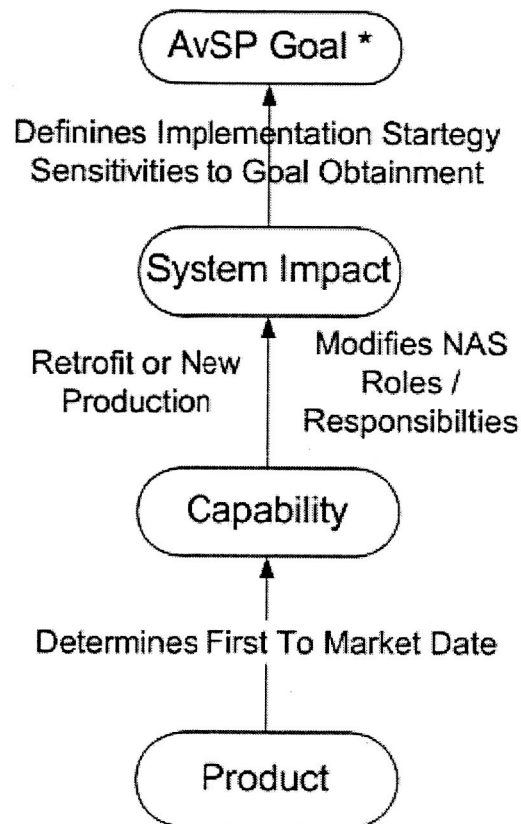
IMPLEMENTATION RISK METRIC

The implementation risk integrates the effects of deployment strategies on the safety benefits derived from each capability.[8] This metric measures the constraints to implementation and the sensitivity of those constraints to various factors. The Technical Integration Plan defines three variables that drive the implementation strategies model: (1) First unit to market date, (2) penetration rate, and (3) maximum penetration level. The following figures show the assessment process of this metric, its risk categorization and how the implementation analysis ratings are derived.

AvSP Goal *

Definines Implementation Startegy
Sensitivities to Goal Obtainment

System Impact

Retrofit or New
Production

Modifies NAS
Roles /
Responsibilties

Capability

Determines First To Market Date

Product

* Accident Rate reduced by xx %

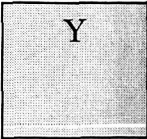**Figure 3: Implementation Risk Assessment Process**

Current certification process easily adaptable
Use acceptance high (customer pull/shared cost)
"Business as usual" level of stakeholder investment requirements
Airline operations impact minimal (retrofit during scheduled downtime)
FAA mandate with retrofit training subsidization program
 No FAA mandate; advisory only
Includes transfer of improved processes to establish programs such as in-house
safety, training, and maintenance functions.

Certification process historically difficult and/or rigorous
Airline operations impacts (unscheduled fleet downtime)
Additional automation/IT infrastructure required for transfer of NASA R&T
FAA mandate without subsidization
Additional training requirements (i.e. ATC, flight crew,...)
Initiation of applicable programs required for transfer of NASA R&T

Certification may be controversial, precedent-setting, or untried
Requires FAA regulation modification
Infrastructure builds dependent upon or diverse from FAA NAS Modernization Plan
Large stakeholder investment requirements
International rule-making required

Figure 4: Implementation Risk Categorization

Implementation Risk Assessment Criteria:

RISK-DRIVER CATEGORY (R_n)	RISK LEVEL		
	Low	**Medium**	**High**
IRL Impacts *	Current certification process easily adaptable	• Certification process historically difficult and/or rigorous • No FAA mandate; advisory only	• Certification may be controversial, precedent-setting, or untried • International rule-making required
Dependencies	No new training or infrastructure requirements	• Dependent on immature technologies • Additional automation/IT infrastructure required for transfer of NASA R&T	• Requires FAA regulation modification • Infrastructure builds dependent upon or diverse from FAA NAS Architecture Study
Market Penetration	• Business as usual level of stakeholder investment requirements • FAA mandate with retorfit or training subsidization program	• Provides product line growth in established market • FAA mandate without subsidization	Large stakeholder investment requirments
Market Impacts	• User acceptance high (customer pull/shared costs) • Airline ops impacts minimal • Includes transfer of improved processes to established programs • Decreased DOC	• Airline ops impacts • Additional training requirements • Initiation of applicable programs required for transfer of NASA R&T • Increased DOC	• User acceptance low • Entrepreneur market

* IRL Impacts has been changed to Certification Impacts.

Table 1: Implementation Risk Assessment Criteria

The risk levels determined for each risk-driver category are combined to form the Overall Risk Rating. The table below illustrates the relationship between the qualitative risk level and the quantitative risk rating.

Overall Risk Rating $= (R_1 + R_2 + R_3 + R_4)$

Overall Risk Score	Risk Rating
0.7-1.0	High
0.4-0.6	Medium
0-0.3	Low

Table 2: Translating Risk Rating to Risk Score

TECHNICAL DEVELOPMENT RISK

The Technical Development Risk assessment estimates the probability of successfully meeting a technology goal.[9] This metric takes into account two individual risk areas: the probability of failure and the severity of the impact. These two measures are then averaged to come up with an overall risk rating.

The first component of the technical development risk metric is the probability of failure(P_f). This involves five risk driver categories: (1) required technology advancement, (2) current technology status, (3) technology complexity, (4) technology dependencies, and (5) testability/verifiability. Each project is given a probability rating of high, medium, or low based on each risk driver category following Table 3. A numeric probability score is then assigned with 0.8 for high, 0.5 for medium, and 0.2 for low. Each risk driver is next given a risk weight between 0 and 1 and finally a weighted probability is computed by multiplying the probability score and the risk weight. Finally, these weighted probability scores are summed to arrive at the probability of failure.

$$P_f = \sum_i w_i p_i$$

$$0 < w_i < 1$$
$$P_i = 0.2 \quad low$$
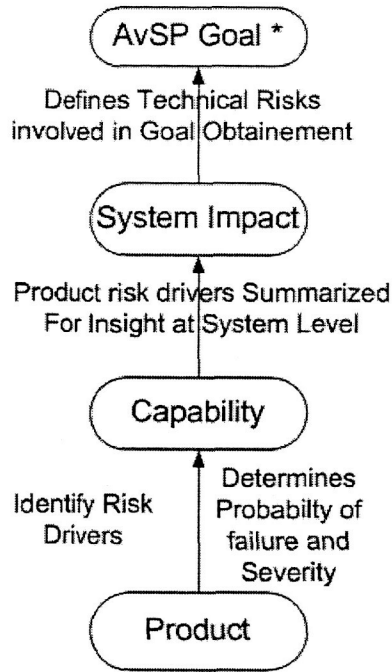$$P_i = 0.5 \quad medium$$
$$P_i = 0.8 \quad high$$

Note: The following condition must also be added to make the equation work:

$$\sum_i w_i = 1$$

The following figures show the assessment process for this metric, the risk categorization and the technical development risk criteria.

Figure 5: Technical Development Assessment Process

Figure 6: Technical Development Risk Categorization[10]

Technical Development Risk Assessment Criteria:

RISK-DRIVER CATEGRORY(R_n)	RISK LEVEL		
	Low	Medium	High
Required Technical Advancement	Minor modifications	Major modifications	State of the art or beyond
Technology Status	In use or prototype exists	Under development	Concept stage
Complexity	Simple	Moderately complex	Highly complex and uncertain
Dependencies	Independent of other technologies	Dependent on proven technologies	Dependent on unproven technologies
Testability/Verifiability	Can be fully tested using existing info	Requires development of new data/information	Can not be tested/verified under all adverse conditions

Table 3: Estimating the Probability of Failure   (Pf = R1 + R2 + R3 + R4 + R5)

Note: The following condition must also be added to make the equation work:

$$\sum_i w_i = 1 \qquad (1)$$

The second factor in computing the technical development risk is the severity of the impact of the technology goal (C_f). Each project is assigned a risk level according to Table 4. This level is then converted into a numeric number with 0.8 for high, 0.5 for medium, and 0.2 for low. This number is the severity of impact for a given technology.

	RISK LEVEL		
	Low	Medium	High
Impact On Technology Goal	Nonessential or minimum impact on technology performance	Partial technology performance can be obtained or alternatives available	"Show Stopper" – Technology cannot be developed and is infeasible

Table 4: Estimating the Severity of Impact (Cf)

Finally the Overall risk rating is computed by adding together the probability of failure with the severity of impact and dividing by two. This numeric overall rating is then converted back into a risk rating as per Table 5.

Overall Risk Score	Risk Rating
0.7-1.0	High
0.4-0.6	Medium
0-0.3	Low

Table 5: Overall Risk Rating   =   (Pf + Cf)/2

$$C_f = C_i = \begin{cases} 0.2, \, low \\ 0.5, \, medium \\ 0.8, \, high \end{cases} \quad (2)$$

$$R_f = P_f * C_f \quad\quad\quad (3)$$

Risk is the product of the probability and consequence of an undesirable event such as failure. In this case the calculation of R_f might follow equations 1, 2, and 3

SAFETY BENEFIT

The safety benefit analysis evaluates the effectiveness of a given technology. It determines how well a technology eliminates a hazardous condition and then its impact on the overall fatal accident rate. This metric also analyzes the relationship between a technology and various precursors.[11]

The tool that this metric employs is entitled the Aviation Safety Analysis and Functionality Evaluation (ASAFE). This tool inputs a technology's domain and evaluates the potential impact areas. It then reviews accident reports to evaluate a technology's effectiveness in relation to the change in the system.

This metric focuses on the change in the system when a control or intervention is put into place to mitigate a hazard. It uses four constraint areas, (1) environment, (2) system design, (3) systems operation, and (4) human involvement. The controls are placed within these four categories to increase understanding of the risk within the system. The following figures show the assessment process and the risk categorization for the safety benefit metric.
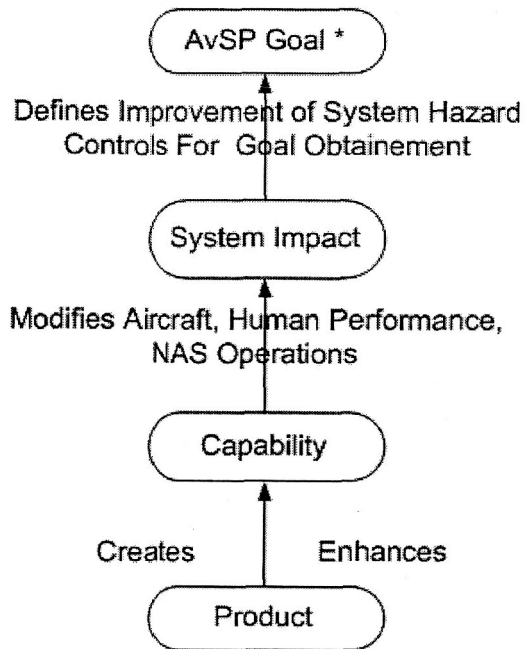
* Accident Rate reduced by xx %

Figure 7: Safety Benefit Assessment Process
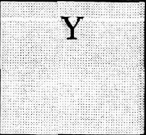
Hazard coverage
- Intervention/Prevention/Mitigation addresses cause, factors, findings across multiple accident
- Intervention/Prevention/Mitigation exclusively addresses a hazard category

System Impact
- Intervention/Prevention/Mitigation addresses areas creating redundant coverage
- Intervention/Prevention/Mitigation addresses areas currently not included in other safety activities
- Intervention/Prevention/Mitigation addresses hazard coverage beneficial to national and international space.

Hazard severity
- Intervention/Prevention/Mitigation addresses cause, factors that are considered the pivotal link in the accident chain
- Intervention/Prevention/Mitigation addresses accident categories that result in largest percentage of deaths and injuries.

Figure 8: Safety Benefit Risk Categorization

RETURN-ON-INVESTMENT

The ROI metric uses the same process for assessment as the three current metrics used by the PAT.[12] It uses also the same program assessment categorization. The Return-On-Investment analysis uses models that describe the operations, financial and investment requirements.
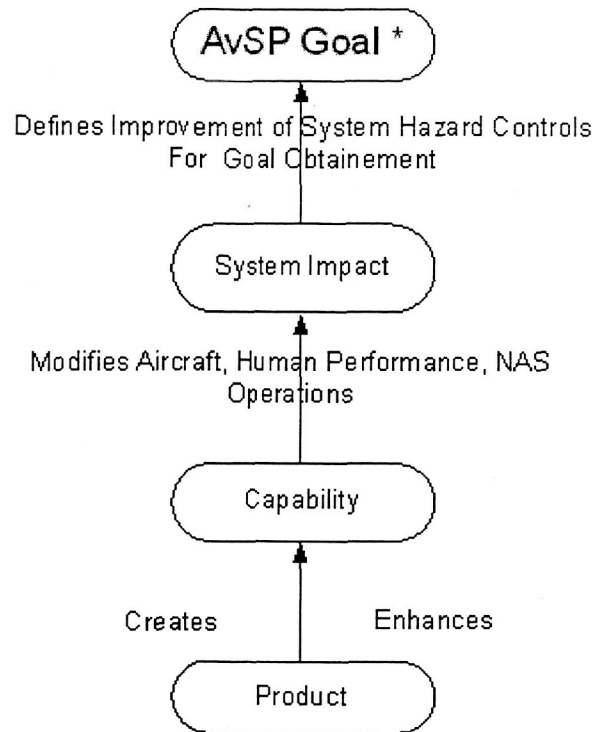
Figure 9: Return on Investment Assessment Process
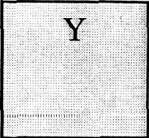
	"Business as usual" level of stakeholder investment requirements Investments tend towards near term and lower risk. May have narrow markets segments Current certification easily adaptable Regulatory requirements may already be in place Includes hardware add-on without system integration Includes transfer of improved processes to establish programs such as in-house safety, training, and maintenance functions. Includes transfer of NASA R&T such as weather phenomena of human factors model development.
Y	Stakeholder investment requirements and operational impacts can be compared with established historical development scenarios (e.g. propulsion and airframe retrofits,…) Business case scenarios needed to address affordability, market breath, fleet impact, and ROI questions Includes transfer of NASA R&T from higher risk subcategories of system-wide services and infrastructures (e.g. communications/ data link, network/database). Higher risk because impact requirements or architecture and future costs still unknown.
	Large stakeholder investment requirements Investments may be long term or high risk Major operational impacts and infrastructure investments Certification may be controversial, precedent-setting, or untried International rule-making required

Figure 10: Return On Investment Categorization

This metric also uses the ROI stakeholder matrix to calculate the economic impact given the operational impacts and investment required for utilization of a new capability.

Stakeholders List	Impact on stakeholder day-to-day operations Comparative metric	Magnitude of Stakeholder investment required Comparative metric	Stakeholder point-of-view when FAA proposes change Eg risk, alternatives	Factors in stakeholder investment decision
1. Subsystem Manufacturing(Mnf)				
2. Engine Mnf				
3. Airframe Mnf				
4. Avionics Mnf				
5. Maintenance & Repair Provider				
6. Service Provider				
7. FAA (which functions)				
8. Airports (major or regional)				
9. Airlines -major,minor				
10. Fixed base operator or GA pilot				
11. Passenger				

Table 6: ROI Stakeholder Matrix

COST ANALYSIS

This metric principally defines the end user cost impacts of installing, maintaining and utilizing new technology solutions. This metric also uses the same process for assessment as the three current metrics used by the PAT.[13]

```
                    ┌─────────────┐
                    │  AvSP Goal *│
                    └─────────────┘
                           ▲
              Defines Direct Operating Costs
                 For Goal Obtainement
                           │
                    ┌─────────────┐
                    │ System Impact│
                    └─────────────┘
                           ▲
               Defines Implementation Costs
                           │
                    ┌─────────────┐
                    │  Capability │
                    └─────────────┘
                           ▲
                  Identifies Cost Drivers
                           │
                    ┌─────────────┐
                    │   Product   │
                    └─────────────┘
```
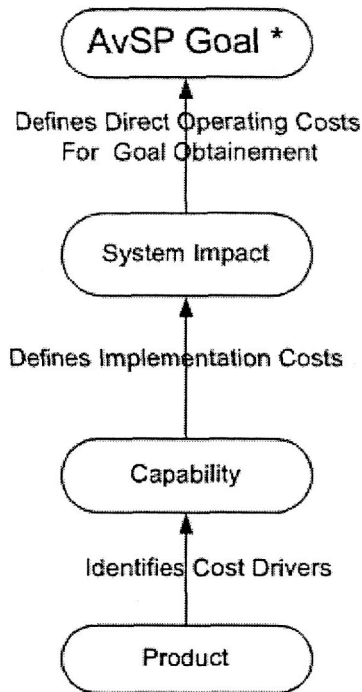
* Accident Rate reduced by xx %

Figure 11: Cost Analysis Assessment Process

▓▓▓	Increased utilization Reduction in insurance/ liability costs Reduction in number of flight tests needed for certification Reduction in design and development time Reduction in aircraft weight (e.g. less fuel, lighter material, etc.) Reduction in (MMH/FH) maintenance man hours/flight hours
Y	Increased training (pilot and maintenance) Increased material costs Retrofit costs Increased aircraft weight (additional material) Increased fuel costs
▓▓▓	Certification may be controversial Radical technology (aircraft, materials, fuel, etc.) changes Major infrastructure issues Significant change in manufacturing process

Figure 12: Cost Analysis Categorization

23

The cost analysis is performed using validated aircraft models (e.g., Tailored Cost Model, Aircraft Computerized Cost Evaluation Support System), using as a baseline targeted aircraft platforms. FAA economic analyst support is used for the assessment of capabilities produced by procedural changes that impact the NAS operational environment.

Analogous industry models may be used to assess the cost impact of capabilities produced by training programs, data sharing, and analysis tools/aids.[14]

The ROI and Cost Analysis (Lifecycle cost) were metrics used in the Preliminary Program Assessment but removed in the Intermediate Program Assessment. The reasons for taking out these metrics are not entirely clear, although one factor was the lack of resources.

3. 0 SURVEY OF OTHER METRICS

EARLY AvSP METRICS[15]

The metrics below are taken from a Systems Analysis Team workshop in 1998, and represent some early ideas for Program Assessment Metrics. Several of these criteria were eliminated from consideration when the first five official metrics were created. Some of those, such as political support, are used by other organizations in their safety investment decisions.

Available Resources • Personnel • Skills • Management • Money • Facilities Investment Balance • Long term v. Short Term • Research v. Technology • Technology • Partner / Customer • Customer • Facilities • Centers Visibility of Results • Customer • Stakeholder • Public	Agency Mission Appropriateness Previous Investment/Accomplishment Customer/Stakeholder Support • Political • Agency • Advisory Group • Partner Non-NASA R&D Investment • Government • Private • International Implementation • Contribution/Benefit • Cost • Produceability • Operability • Supportability • Profitability • Acceptability • Performance Risk

Figure 13: Early AvSP Metrics

ANALYTIC HIERARCHY PROCESS

Analytic Hierarchy Process (AHP) is a tool involving pairwise comparisons of criteria which can be used to help choose between investments. AHP is used to prioritize multiple objectives in choosing where to allocate resources. For example, NASA could use the process to help quantify the relative importance of its objectives; for example, pairwise comparisons of increasing capacity, increasing safety, and increasing mobility could be used to help weight these criteria in making investment decisions appropriately. The table below shows what such a comparison might look like if done by someone whose primary concern was safety.

	Safety	Capacity	Mobility
Safety	1	3	5
Capacity	1/3	1	7
Mobility	1/5	1/7	1

Table 7: Comparison Matrix

The next table illustrates the manner in which the numbers are assigned. The '3' in the safety row and the capacity column, for example, signifies that safety is slightly more important or preferred than capacity. The reciprocals on the left lower triangle of cells reflect the inverse relationship (i.e. "capacity to safety" as opposed to "safety to capacity").

Numerical Values	Definition
1	Equally as important or preferred
3	Slightly more important or preferred
5	Strongly more important or preferred
7	Extremely more important or preferred
9	Most important or preferred
2,4,6,8	Intermediate values to reflect compromise
Reciprocals	Used to reflect dominance of the second alternative as compared with the first

Table 8: Rank Assignments

The vector of priorities is determined through normalization of these rankings, and can then be used to prioritize the objectives for use in resource allocation decisions.

The prioritization can be checked for consistency through computation of the principal eigenvalue, λ_{max}. The closer this value is to the number of objectives being compared, the more consistent the result. The approximation of this value is found by multiplying the comparison matrix by the vector of priorities, then dividing each component of the new vector by the corresponding component of the priority vector. The average of the resulting components gives the approximation to λ_{max}.

In order for this method to be applicable the rank order of the matrices must be compatible and care must be exercised in placing the vector of priorities as a pre- or post-multiplier of the comparison matrix. In this example, the 3x3 comparison matrix may only be multiplied by a vector of 3 priorities.

A consistency ratio (CR) can be computed to measure the consistency of judgment. First, the consistency index (CI) is computed for the matrix:

$$CI = \frac{\lambda_{max} - n}{n - 1}$$

where n = number of objectives being compared

This value is then divided by a random index (RI) from a table of each possible number of objectives to compute the CR.[16]

JSIT AND JSAT METRICS

The Commercial Aviation Safety Team (CAST), an organization comprised of NASA, the FAA and aviation industry organizations, is aimed at developing and implementing a common safety agenda to help meet the 80% accident reduction rate challenge made in the report to the President[17]. CAST chartered a Joint Safety Analysis Team (JSAT) to develop a process for identifying interventions, or projects, with a high likelihood of improving aviation safety.

JSAT developed a process by which safety "interventions" were prioritized based on "Effectiveness" and on "Feasibility". Feasibility was defined as "the potential for widespread implementation of an intervention, including retrofit as necessary, within the ten-year time frame specified in Vice President Gore's committee report to the White House." Subsequent to the CFIT JSAT, the responsibility for assessing feasibility was transferred from JSAT to Joint Safety Implementation Teams (JSITs), so currently JSAT is only responsible for effectiveness.[18] The following section will discuss both aspects and the corresponding metrics.

JSIT's feasibility ratings are based on six elements:

1. *Technical feasibility* – The ability of the current project to take advantage of the current state of technology in pursuing further development.

2. *Financial feasibility* – Should consider the total cost of the implementation, including the planning process. Also involves the capability of the performing

organizations to make available the appropriate funds needed to implement the project.

3. *Operational feasibility* – Involves the "practicality" of the project within the context of the operating environment, including NAS, ground operations, maintenance, inspection, etc. Considers which operations within the aviation system are impacted.

4. *Schedule feasibility* – The ability of the project to contribute to achieving the goal in a selected time frame. Must consider implementation schedule by project.

5. *Regulatory feasibility* – Should be evaluated against current rules and certification process. Could be a deterrent due to a long approval process.

6. *Sociological feasibility* – Requires an evaluation of the compatibility of project goals with the prevailing goals of the political system. Worthy projects may face heavy opposition due to sociological factors alone, while a less meritorious project may receive support due strictly to the vision that is "politically correct." [19]

Part of the process developed by JSIT was the construction of logic trees to help determine the feasibility of an intervention. This piece of the process was not used in all cases, but is a useful tool. The logic tree originates with the language of the intervention itself, and then brainstorming helps identify follow-up actions or circumstances that could have some bearing on the outcome of the project. Feasibility ratings are determined for the various branches defined by these circumstances, and help in the formulation of the Feasibility value for the intervention as a whole.

Feasibility assessment in all cases was accomplished through the assignment of a numerical value for each of the six Feasibility elements. The JSIT assigns a value of 1, 2, or 3 under each Feasibility category. The following table shows the criteria for each.

Feasibility Type	3	2	1
Technical	Off-the-shelf technology, no development required	Some development required, not currently in public use	Major technology development effort required
Financial	Less than $100M to implement	Less than $250M, greater than $100M to implement	Greater than $250M to implement
Operational	Minimal change to entities within the operating environment	Modest change to operating environment	Major change to operating environment
Schedule	less than 2 years to full implementation	full implementation in 2-5 years	longer than 5 years to full implementation
Regulatory	no policy change	guidance change only (orders, handbooks, policy)	rule change
Sociological	positive push from political system	neutral	negative

Table 9: JSIT Feasibility Scoring

Effectiveness is defined as a measure of the potential impact of an intervention based on the breadth and depth of its relative potential for preventing accidents. Long-term value was also taken into consideration so as not too ignore projects with potentially high future safety benefits. The effectiveness ratings are determined according to the JSAT process, which takes into account three factors: Power, Confidence, and Future Global Applicability.

The first factor, Power, measures how well the intervention directly and definitively addresses the problems and contributing factors in the accident, and by doing so, would have reduced have reduced the likelihood of the type of accident in question, *if everyone or everything performed as the intervention intended.*[20] The power factor is divided into two sub-metrics:

- P_1: The importance of the problem or contributing factor at which the intervention is aimed in causing the accident in question. P_1 ratings are developed for each standard problem statement that is called out in each accident.

- P_2: The ability of the intervention to mitigate the problem or contributing factor. A general intervention should get a lower P_2 rating, and a more clearly focused intervention that directly addresses the problem and its characteristics should get a higher P_2 rating.

P_1 is rated on a scale of 0 to 6, with P_1 signifying that the problem had no influence in causing this accident and 6 signifying that the problem would have caused the accident all by itself, and without this problem or contributing factor this accident would not have happened. P_2 is rated on a scale of 0 to 6, with 0 meaning this intervention will have no effect on the problem or contributing factor in question and 6 meaning the intervention will completely eliminate the problem or contributing factor in all cases. To make the final Power assessment, the two types of Power are combined using the formula

$$\textbf{Power} = \textbf{(P}_1 \textbf{ x P}_2\textbf{)2 / (P}_1 + \textbf{P}_2\textbf{)}$$

Future Global Applicability is used to evaluate how frequently the recorded problems will continue to be present on a widespread basis in future operations. Applicability of a specific intervention is rated on a scale of 0 (no applicability; the problem will be virtually non-existent in future operations) to 6 (the problem will recur very frequently in future operations).

The Confidence factor measures how strongly the scorer believes that everyone or everything will perform as expected. Confidence ratings should assume that the intervention has been implemented, so that feasibility issues to not get mixed in with this rating. Confidence is rated on a scale of 0 to 6, with 0 representing that the intervention will probably never work as intended and 6 representing that it will always perform as intended.

This is the reverse of the power scale for power, 0 is worst and 6 is best. However for future global applicability, 0 is best and 6 is worst. This could lead to inconsistency if these criteria are combined.

The process for combining the three factors is as follows. The highest Power rating an intervention was given is used. Then those problems with the highest Power

ratings are examined to find the ones with the highest Confidence ratings. Based on how an intervention works for the family of problems being addressed, the Confidence rating may be moved up or down. Then the scorer looks at the interventions with the highest Power ratings to find those with the "highest applicability ratings", checks the contributing factors or problems that it addresses, and considers raising the applicability accordingly. But this will cancel out the power and confidence as presently set up. The final numbers are then combined to form the Overall Effectiveness rating as shown:

$$OE = P \times C/6 \times A/6 = P \times C \times A/36$$

This method is inconsistent as presently formulated. In order to be an effective metric the applicability scale should be reversed so it is consistent with the power and confidence scales.

The results of this scoring are coupled with the feasibility scores and then used to generate color-coded spreadsheets, which help to visually code the numerical values. The prioritization of interventions is achieved through the creation of another spreadsheet based on the *product* of the effectiveness rating and the feasibility rating. Based on the sort of E x F, a cutoff value is determined to identify the highest leveraged products to reduce the accident rate. Research solutions are considered separately if they are of a long-term nature and are included in the final JSIT recommendations, due to the potential for high future safety benefit.
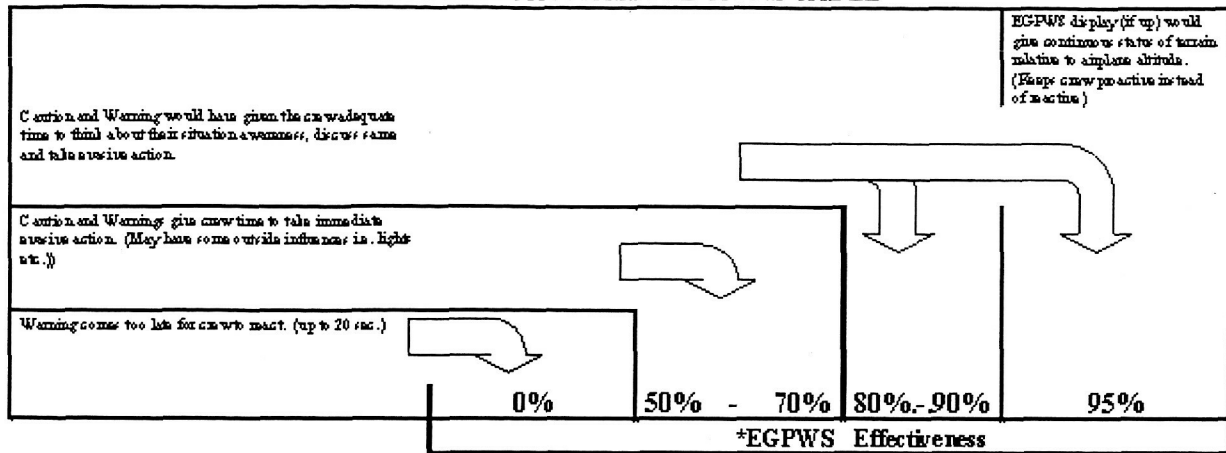
JIMDAT PROCESS

JIMDAT is a Prioritization Methodology team developed by the Commercial Aviation Safety Team (CAST) to evaluate, measure and track the accident reduction potential of safety enhancements.[21] The JIMDAT prioritization attempts to measure the effectiveness of an intervention against any selected historical dataset, allowing comparison between interventions. The process also provides identification of future areas for safety studies and, most importantly, the creation of a master strategic implementation plan based on safety effectiveness and resource considerations. The prioritization methodology is flexible enough to allow rapid evaluation of changes in the strategic plan and provides consistent estimates of the accident prevention potential of safety enhancements. Sufficient detail is included in the methodology to account for the benefit of a single intervention or a combined group of interventions and also to address any overlap with other interventions/technologies. Also, the JIMDAT process preserves analysis criteria and results, which allows for future adjustments and alterations when necessary.[22] The main assumption that the process relies on is that future incidents and accidents will occur at the same rates and with the same types of causal chains as historical accidents.[23]

The JIMDAT process calculates the potential safety benefit of an intervention using the formula[24]:

$$\text{Accident Rate Reduction} = f\left(\begin{array}{l}\text{Effectiveness that an intervention has for reducing the accident rate if incorporated}\end{array}, \begin{array}{l}\text{Portion of world fleet with intervention implemented}\end{array}\right)$$

The effectiveness factor is based on a set of historical accidents. Interventions are evaluated against each accident in that historical set to determine how effective the intervention would have been at preventing those accidents. The Portion of World Fleet factor is based on the portion of the "fleet" that either have the intervention currently incorporated or are expected to incorporate it by a future date. Table 10 displays the JIMDAT Effectiveness Rating System.

EGPWS EVALUATION/ASSUMPTIONS TABLE

			EGPWS display (if up) would give continuous status of terrain relative to airplane altitude. (Keeps crew proactive instead of reactive.)
Caution and Warning would have given the crew adequate time to think about their situation awareness, discuss same and take evasive action.			
Caution and Warning give crew time to take immediate evasive action. (May have some outside influence ie. light etc.)			
Warning comes too late for crew to react. (up to 20 sec.)			
0%	50% - 70%	80% - 90%	95%
		*EGPWS Effectiveness	

Sources:

Allied-Signal Flight Into Terrain and the Ground Proximity Warning System Report   Revised

Allied-Signal CFIT Engineering Report (8/21/97)

Allied-Signal Evaluation of Boeing data (817/98)

* Effectiveness values are based on initial good in-service experience reported by airlines equipped with EGPWS (Enhanced situational awareness of high terrain and low occurrence of false warnings).

Bender/Noges
March 1999

Table 10: JIMDAT Effectiveness Rating Table

The process also attempts to compute the effectiveness of combinations of interventions. A logic diagram model, similar to that used in fault tree calculations, is used to calculate this combined effectiveness level. The effectiveness evaluation in both the individual and combined forms follows the guideline of the chart above.

When new information becomes available, the Excel spreadsheets that the process uses can be easily modified to reflect the change.

Figure 11 illustrates the Accident Intervention Process developed by JIMDAT for the example of a training aid intervention. The same process can also be drawn as a fault tree.

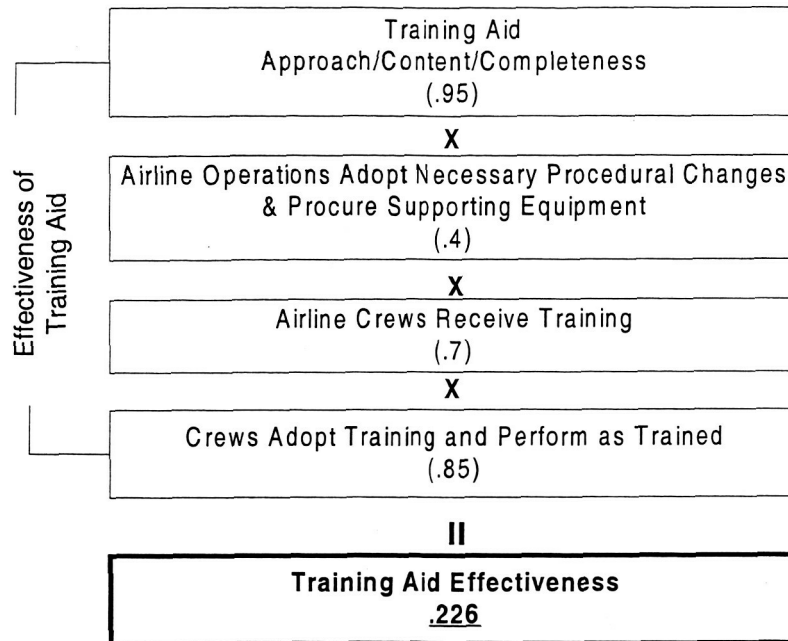## Training Aid Effectiveness/Accident Reduction

Figure 14: Training Aid Example

The product of all the individual components of effectiveness determines the final overall effectiveness level for the intervention.

$$TAE = 0.95 \times 0.4 \times 0.7 \times 0.85$$
$$TAE = 0.226$$
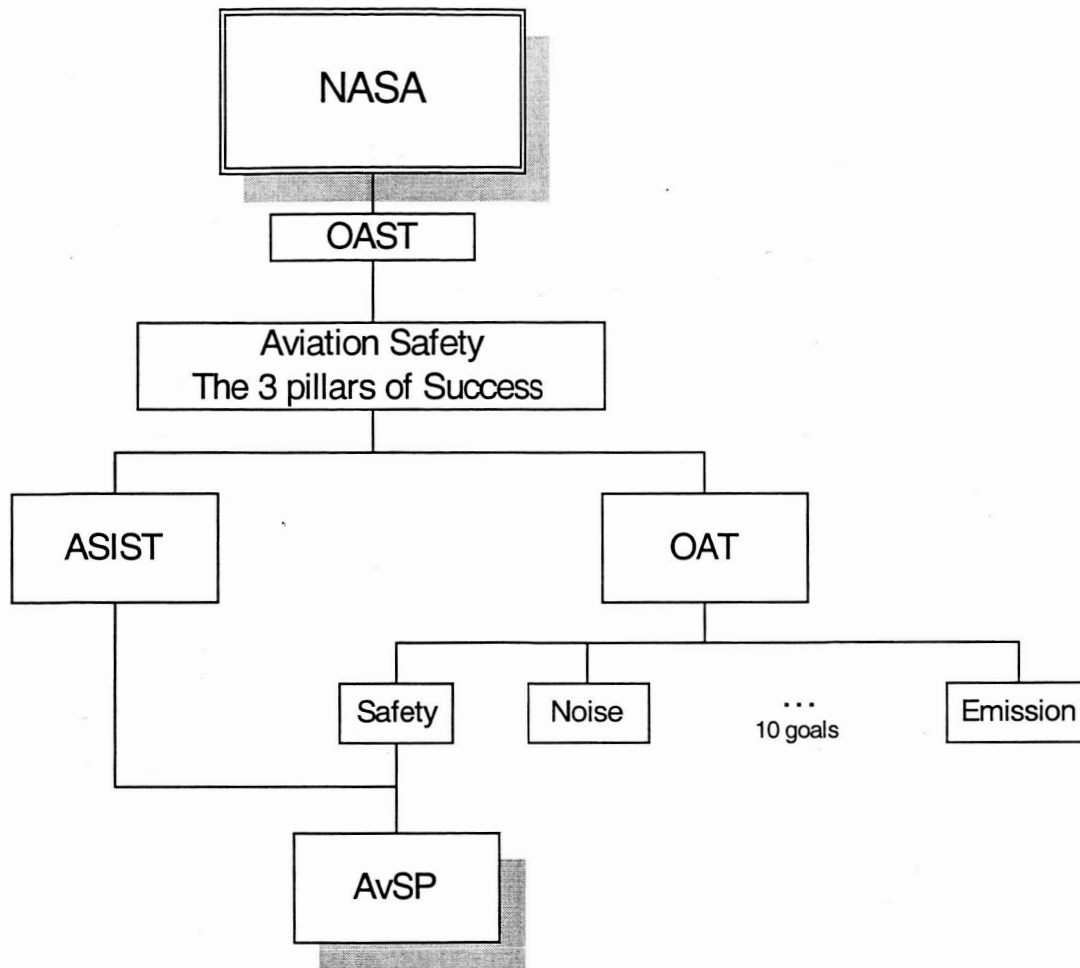
Figure 15: Relation between ASIST, OAT and the AvSP

The Three Pillars for Success

Both Aviation Safety Investment Strategy Team (ASIST) and Office of Aerospace Technology(OAT) are part of the NASA's "The Three Pillars for Success" initiative defined by the Office of Aero-Space Technology (OAST) to establish three major goals in terms of technology. This plan is articulated around three technology "pillars:" Global Civil Aviation, Revolutionary Technology Leaps, and Access to Space. In each pillar, the following enabling technology goals are defined:

1. Global Civil Aviation: make dramatic improvements in safety, environmental compatibility, and affordability of air travel.
2. Revolutionary Technology Leaps: overcome barriers to high-speed travel, revitalize the U.S. general aviation industry, and develop next-generation design tools and experimental aircraft.
3. Access to Space: make access to space significantly more affordable and reliable.

AvSP is geared toward answering the Pillar One safety goal.

NASA, the Federal Aviation Administration (FAA), the Department of Defense (DOD), industry, and academia have to find the necessary technology solutions to turn these goals into reality.[25]

ASIST

Aviation Safety Investment Strategy Team (ASIST) is a tri-lateral group made of members from the NASA, FAA, DoD.

The process for answering the Aviation Safety Initiative is as follows:

- Analyze the industry input and identify the major accident causes and issues
- Identify underlying problems
- Identify some solutions
- Propose a set of integrated solution and investment options to the Office of Aeronautics and Space Transportation Technology (OASTT)[26]

In 1997, ASIST defined the main metric as fatal accidents. The goal was to link this metric to the precursors of incidents and accidents. ASIST defined the aviation safety research investment strategy. The three areas of investment are:

- Accident prevention
- Accident mitigation
- Aviation system monitoring and modeling.

The five focus areas (human error consequences; weather; flight critical systems and information integrity; human survivability; and aviation system-wide monitoring, modeling and simulation) are allocated as illustrated in the figure below.[27]
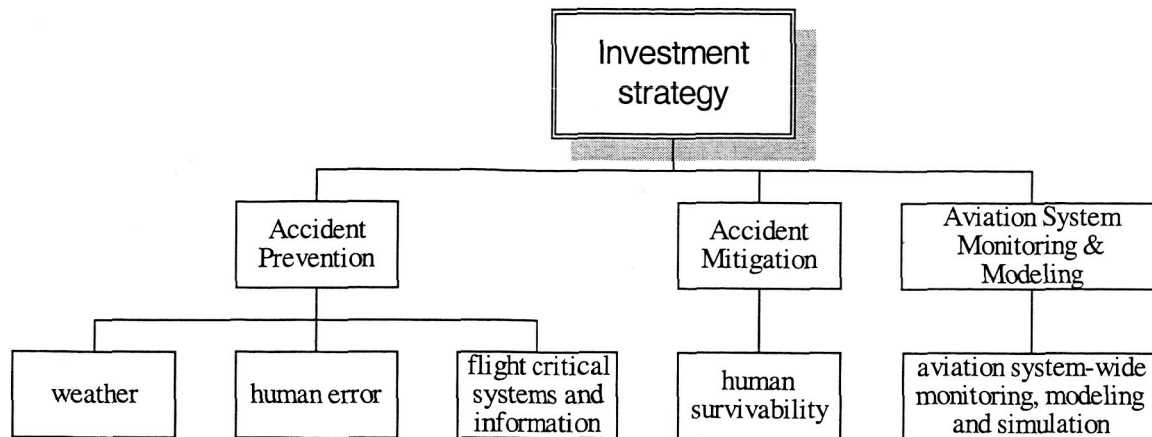
Figure 16: ASIST focus areas

The AvSP is part of the "Three Pillars for Success" initiative that spells out what NASA will do to achieve national priorities in aeronautics and space transportation technology that is defined by ASIST.

OAT

The Office of Aerospace Technology (OAT) manages the Aerospace Technology Enterprise. The Aviation Safety Program is a Level I program of NASA's Office of Aerospace Technology (OAT).

The OAT answers the Three Pillars Aerospace goals and the aerospace industry needs by addressing 10 goals: safety, noise, emission, cost of air travel, capacity, general aviation, supersonic travel, design and test, space access, and in-space transportation. OAT has a Program Assessment function based on 4 teams: Technical Evaluation & Integration Team, Vehicle/Fleet Team, Airport/Airspace Team, Spaceports/Operations Team. They assess the areas defined by the 3 Pillars for each program. One of these areas is Safety. The concept of a safety data analysis framework is to create metrics from projections for 2007-2022. These projections are based on intervention/technologies analyses and FAA and DoD forecasts. In November 1999, the four metrics were:

✓ accident rate

✓ fatal accident rate

✓ number of fatalities

✓ number of injuries.

There is a 100% overlap in accident coverage allowed due to multiple technologies impacting individual accidents which is consistent with AvSP philosophy of increased reliability through redundant technology impacts.

The technology impacts to different aircraft classes are analyzed separately. Transports, commuters, GA and rotorcrafts are the different aircraft classes.

The forecasts in 1999 are listed below for transport and commuter aircrafts.

Transport aircrafts:
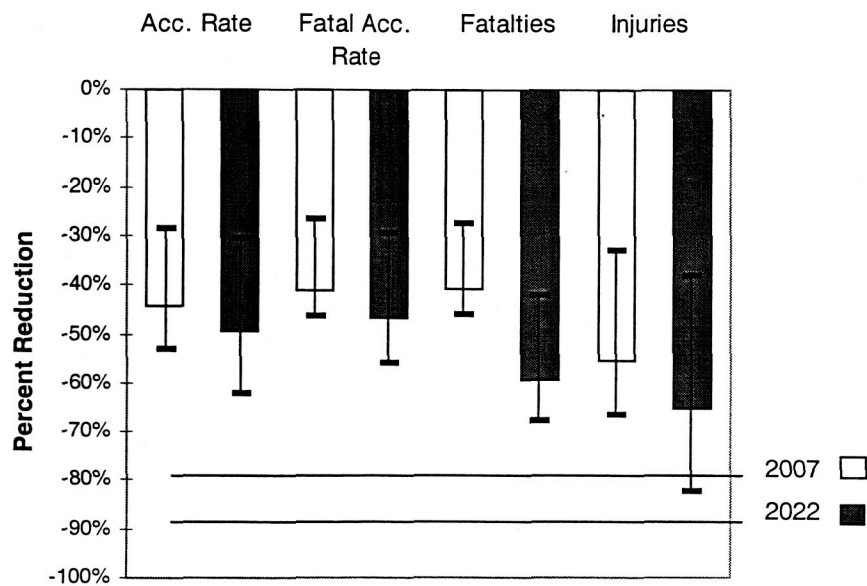
Figure 17: Metrics forecast for transport aircrafts

Commuters aircrafts:
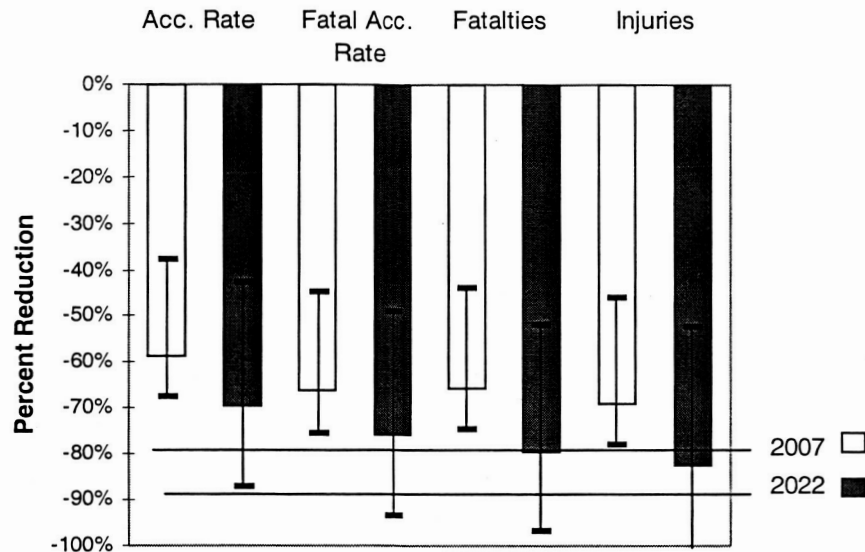
Figure 18: Metrics forecast for commuters aircrafts

NLR ARIBA PROCESS

NLR is an independent European non-profit research institute focused on five areas:

1. Civil Aviation (**Safety**, Noise and Emissions, Air Traffic Management)

2. Military Aviation

3. Aircraft Development

4. Space Technology

5. Non-aerospace Applications We found one interesting document, the ARIBA project.

In the context of Air Traffic Management (ATM) operation certification, an accident risk assessment is performed with comparison from other industries such as petrochemical, nuclear industries. This process is illustrated in Figure 19:

40
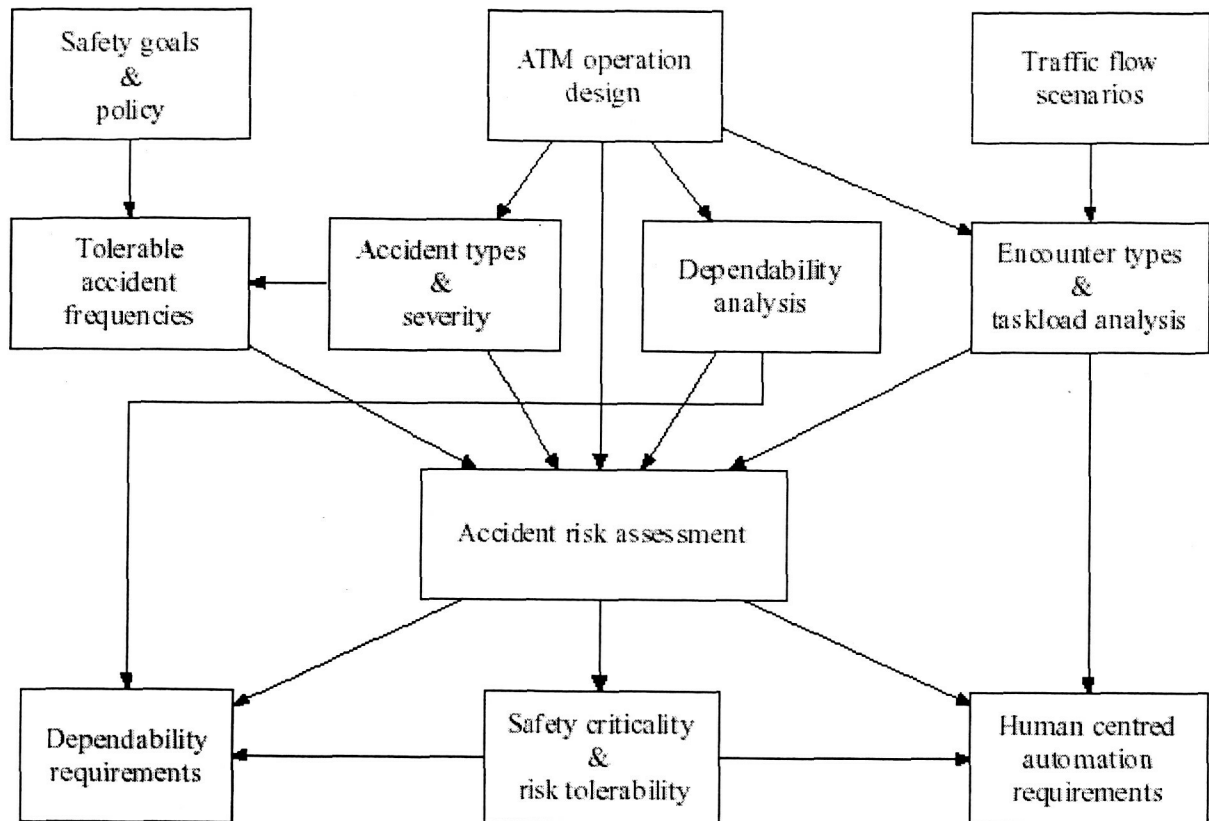
Figure 19: Accident risk Assessment Process (NLR-ARIBA)

The boxes at the top are the advanced operations to be certified. The second level represents the various safety related assessments. The third level is accident risk assessment and the fourth level (the boxes at the bottom) is the outputs of the risk assessment.

The three advanced operations to be certified are:
- ✓ Safety goals and policy
- ✓ ATM operation design
- ✓ Traffic flow scenarios

The four steps in the second level of safety risk assessments are discussed below:

1. Accident type and severity

   The first step in the process is to define the types of accidents during various flight phases (e.g. collision on ground or in flight, with an aircraft or with ground or other

ground based object, incident induced by expedite deceleration…) and to assess the severity of the consequences of each accident type in terms such as:

✓ the expected number of fatalities,

✓ the expected number of injuries, and

✓ the expected material damage.

2. Tolerable accident frequencies

Next in order to incorporate the concept of tolerating some risk, a frequency requirement by 3 regions is defined for each accident risk as shown in Figure 20:
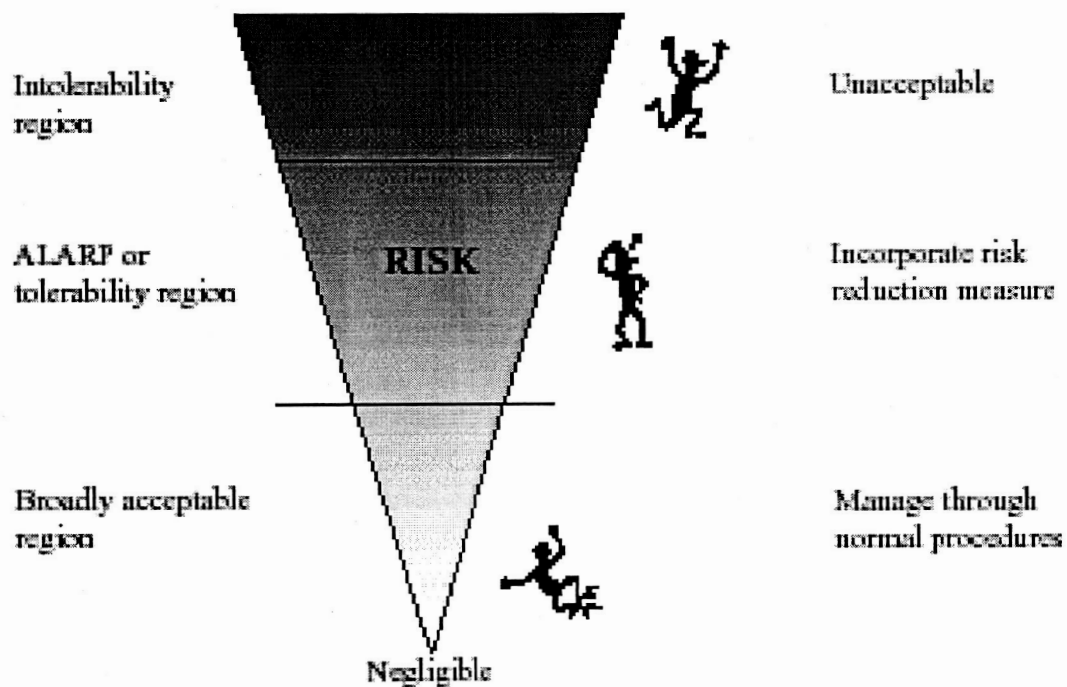

Figure 20: Risk Regions (NLR ARIBA)

The final step is to combine the accident severity classes and the accident frequency classes into an accident risk tolerability matrix (see Figure 21 for an illustration).

Severity of accident		Frequency of accident				
Expected material damage	Expected injury or fatalities	1x / yr in civil aviation	>1x / yr in civil aviation	1x / yr per large airline	>1x / yr per large airline	1x / yr per aircraft
No damage	No injury					
Minor damage	Minor injury					
Serious damage	Major injury	Manage through normal procedures		Incorporate risk reduction measures		
Major damage	Single fatality					
Hull loss	Many fatalities					Intolerable
Hulls loss	Hundred(s) of Fatalities					

Figure 21: Accident risk tolerability matrix (NLR ARIBA)

3. Encounter types and tasks load analysis

The aim of the encounter types and task load analysis is to characterize the encounter types and frequencies, and the related controller and pilot tasks and workloads for the advanced ATM operation considered.

4. Dependability

Dependability is studied as the ability of a technical system to perform one or several required functions under given conditions. Dependability assessment methodology incorporates severity-frequency criteria for the tolerability of failure conditions for safety-critical technical systems. These criteria can be expressed in the form of a tolerability matrix:

	Severity			
Probability Level	Catastrophic	Hazardous	Major	Minor
Probable	*Intolerable*	*Intolerable*	*Intolerable*	*Tolerable*
Remote	*Intolerable*	*Intolerable*	*Tolerable*	*Negligible*
Extremely remote	*Intolerable*	*Tolerable*	*Negligible*	*Negligible*
Extremely improbable	*Tolerable*	*Negligible*	*Negligible*	*Negligible*

Figure 22: Tolerability Matrix (NLR ARIBA)

Next, the accident risk assessment is performed to answer level 3 of the procedure. This involves three steps:

1. Identify and qualify hazards: hazard identification during hazard brainstorming sessions helps to identify all possible hazards, hazardous events and their causes and consequences from various viewpoints. The goal of these brainstorming is to generate various viewpoints: an operational experience viewpoint (what went wrong in the past), a functional viewpoint (failure conditions, human errors), a cognitive viewpoint (operator internal states/strategies, experience/training issues), an organizational viewpoint (general working conditions, CRM issues, culture), and a safety management viewpoint (both proactive and reactive).

2. Qualitative risk assessment: it consists of a preliminary analysis of the hazards identified.

3. Quantitative accident risk assessment: it consists of four complementary risk modeling approaches to provide a clear insight into the safety, as part of follow-up activities:
   - ✓ Dependability and human reliability
   - ✓ Human operator cognitive models
   - ✓ Aircraft evolution, incident and accident models
   - ✓ Co-ordination and controlAccident risk evaluationThe accident risk evaluation consists of two subsequent steps. First, develop an appropriate model using all

44

information collected. Second, use the model to evaluate the accident risks involved with the various encounter types of the advanced operation.

✓ Model development: after all relevant information is identified, a specific risk model is developed iteratively; each iteration consists of a model synthesis step and a model verification step.

✓ Model based evaluation: this process evaluates, in a quantitative way, the frequencies of various accidents happening during particular flight phases using: stochastic analysis to decompose the accident risk estimation; and Monte Carlo simulation to evaluate the probability distribution for the identified classes of event and to evaluate conditional accident probabilities.Potential proactive and reactive measuresPotential proactive and reactive measures are conceptualized based on the Bow-Tie approach illustrated in Figure 23.

However, if there are "adverse conditions" classified as being "Tolerable" or "Intolerable" then it is necessary to develop risk reducing measures for these "adverse conditions". The objective is to collect for each key "adverse condition" two types of measures:

- Proactive measures; to improve the chances to avoid entering the adverse condition at all, and

- Reactive measures; to improve the chances to escape from the adverse condition prior to its escalation.

To collect proactive and reactive measures, brainstorm sessions are organized and the outcomes are documented.
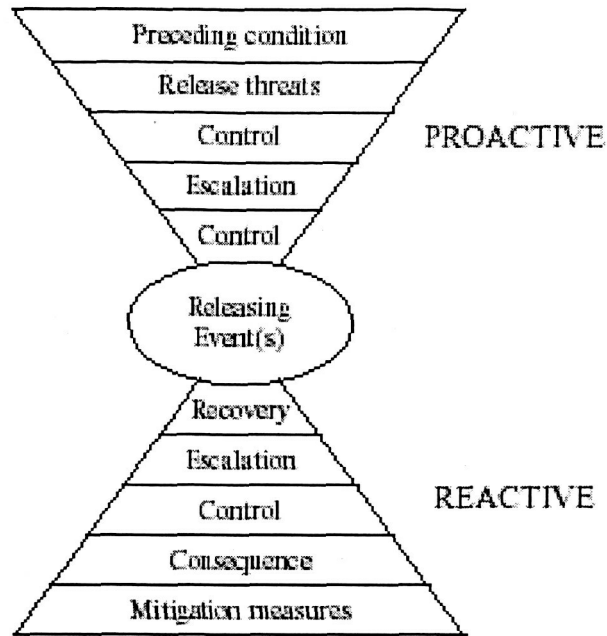
Figure 23: Adapted Version of the Bow-Tie (NLR ARIBA)

The last level is the feedbacks of this process. The results of these risk tolerability and safety criticality assessments provide feedback at the three levels of advanced operation design.

- ✓ Safety management of the advanced operation. The risk tolerability specifies how well the advanced ATM operation considered satisfies the Safety goals.

- ✓ Dependability requirements. The dependability assumptions form a useful basis for setting better requirements on the technical systems, and to feedback these findings to the designers/manufacturers of technical systems.

- ✓ Human centered automation requirements. If human cognitive workload decreases safety, it is important to feedback these findings to the designers of the advanced ATM operation.

Specific points of interest:

The part III of the ARIBA report tackles the subject of Safety Validation Criteria. Quoted from Section 2.3 of this document (Ref.: ARIBA/AAF/WP6/FR-III):

*Safety criteria must be translated into metrics before safety can be assessed. However it is impossible to **prove** rigorously that an ATM system will have some level of **safety** in the future. So direct safety metrics, such as the number of fatalities by passenger-kilometre, have to be ruled out, except possibly for already operational systems.*

The only way to assess safety is to find factors that have a (more or less direct) impact on safety, and then, when possible, to define metrics for each of these factors. Such safety factors are, for example, reliability or availability of the automated system.

But even for such metrics, it is often difficult to get figures. Therefore practical metrics often has to be still more indirect, and this analysis must be iterated until measurable indicators are found. For example, such measurable indicators may be test coverage, code complexity (measured through standard metrics), methods used for development and for ensuring safety and to what degree they were used, etc. Feedback from operational systems is required for this analysis

This document presents a safety case (that is "a consistent and coherent set of arguments and evidence that the system meets or exceeds the system safety standard or target, used to justify the safety of a system.") for an automated system by a manufacturer.

This gives an idea about how easy applicable is the methodological framework proposed[28].

GLOBAL AVIATION INFORMATION NETWORK(GAIN)

GAIN is an industry and government led initiative to promote and facilitate the collection of safety information in the international aviation community to improve safety[29]. GAIN is composed of three working group that are interdisciplinary teams that work toward the action plans of the GAIN steering committee. The three working groups are: Working Group A: Aviation Operator Safety Practices, Working Group B: Analytical Methods and Tolls, and Working Group C: Global Information Sharing Systems.

Working Group B(WG B) was made to provide members of the aviation community better information about the tools and analytical techniques that can help airlines turn their data into valuable information to improve safety. As a step towards reaching this goal, WG B created the *Guide to Methods & Tools for Airline Flight Safety Analysis*(M & T) and published it in December 2001. This guide summarizes 50 methods and tools that can be used to analyze flight safety data. It provides tools that could be useful primarily to airlines.

The M & T guide is organized into three areas: Flight safety event reporting and analysis systems, General methods and tools for event analysis, and Flight operational quality assurance (FOQA)/Digital Flight Data Analysis Tools. The second category is split up further into six categories: Descriptive Statistics & Trend Analysis, Cost benefit analysis, Risk Analysis, Text/Data Mining & Data Visualization Occurrence Investigation, and Human Factors Analysis. Of the tools outlined in this guide, three of them should be looked at further by the AvSP Program Assessment Team.

Two of the tools fall under the cost benefit analysis category. First, the Airbus Service Bulletin Cost Benefit Model. It is build to assist the decision to apply or not apply a Service Bulletin on a given fleet or aircraft. The result of this analysis is a Return on Investment figure.

Secondly, the Boeing Digital Technologies Cost Model is also an effective tool for performing a cost benefit analysis. This tool quantifies the financial impact of delays and cancellations due to accidents and incidents on airlines. It enables a manager to quickly and easily begin assigning dollar costs to accidents. This multi-purpose tool can

also determine costs to the out of service times of any aircraft type. It is used by many airlines and is given out freely by Boeing.

The third tool, Fault Tree (Fault Tree Module) is a risk analysis software package. The purpose of this software is to assess a system by identifying an undesirable end event and examining the range of potential events that could lead to that condition. Fault tree is also a graphical method used in reliability engineering. Probabilistic Risk Assessment (PRA) is a method of conducting risk analysis. This method quantifies the probabilities and consequences associated with accidents by applying probability and statistical techniques. Furthermore, it provides a systematic framework for estimating risks and evaluating them before making decisions.

Working Group B has created the M & T guide that gives summaries of 50 methods and tools that could be used to analyze flight safety data. There are three tools that the AvSP should look further into namely, the Airbus Service Bulletin Cost Benefit Model, Boeing Digital Technologies Cost Model, and Fault Tree (Fault Tree Module). For more information, please view GAIN's Guide to Methods & Tools pages 19, 37, and 39.

LOGISTICS MANAGEMENT INSTITUTE (LMI) SAFETY BENEFIT METHODS

The Logistics Management Institute (LMI) is currently performing a safety-benefit analysis of three of the AvSP's projects: synthetic vision systems, weather accident prevention, and system-wide accident prevention. LMI uses an integrated safety analysis method that comprises of two components, a reliability model and a simulation model. In the reliability model, technology is broken down into components, such as hardware, software, and human agents, then define how those components interact, and finally determine the failure rates of the components. In the simulation, an operational scenario is modeled using Monte Carlo methods to investigate the performance of the technology. The algorithm that LMI uses for estimating safety is:

$P(Accident) = P(Hazard) * P(Accident \mid Failure\ and\ Hazard)$

where

$P(Accident)$ is the probability of an accident,

P(Hazard) is the total probability of a hazardous condition,

P(Accident | Failure and Hazard) is the conditional probability of an accident given a failure when a hazard exists

In this preliminary report, the results show that the AvSP technologies provide significant safety benefits.[30]

BAYESIAN BELIEF NETWORK (BBN)

Dr. James T. Luxhoj of Rutgers University published a paper in 2001 that identifies organizational factors that may lead to aircraft failure. [31] Among these factors are communication, management structure, processes and culture. Using these possible accident causing factors and the Bayesian Belief Network, Luxhoj is developing a metric for the Aviation Safety Program. On June 28, 2002, Luxhoj gave a presentation to the LaRC in which he explained how the BBN could be applied to flight safety. The Risk Intensity Level Metric evaluates the impact of technology insertion on system risk. Please see Professor Luxhoj's June 28 presentation for more information.

4.0 Conclusions and Recommendations

This study sought to identify the main objectives of NASA of improving aviation safety in the National Airspace System (NAS) and of the AvSP. The FAA's aviation safety objectives were also identified. The next step involved checking these groups of objectives for consistency and compatibility. This investigation revealed that the objectives are in fact consistent across organizational levels at NASA and with the FAA.

One main issue uncovered through the study was a conflict in the statement of the aviation safety improvement goal. The original called for a reduction in the *"fatal accident rate"*, whereas now the goal is stated as fatal accident rate by AvSP but simply "accident rate" by NASA.

The review of current metrics and survey of potential new metrics has revealed a few areas that the Program Assessment Team should investigate further. The removal of the Cost and Return-on-Investment metrics for the Intermediate Program Assessment should be reconsidered. The Program Plan calls for "affordable technologies" to meet their safety goals, and yet the metrics specifically addressing this issue have been eliminated. Certain elements of the ROI and Cost categories have been accounted for through the Implementation Risk metric, but others are not included anywhere.

In addition, there is no apparent way to measure the portfolio's effectiveness in meeting the objective of the Program Assessment Team objective calling for a balance of investment and technology.

One other issue is the interdependencies between some of the metric risk driver categories. In order to use an additive model like the one currently in place, the metrics would need to be completely independent. AvSP should examine the way that other groups, such as JSAT, maintain such independence in their safety assessment processes.

The AvSP should also consider adapting some of the methods to assess aviation safety used by NLR, such as the Accident Risk Assessment. Another possible area for AvSP to incorporate into their metrics would be the question of what to do when there are conflicts between the AvSP goal of improving safety and other NASA goals, such as increasing capacity. The Analytic Hierarchy Process could be a useful tool to help prioritize those objectives.

REFERENCES

[1] http://www.aero-space.nasa.gov/goals/index.htm

[2] AvSP Program Plan, August 1, 1999

[3] http://www.aero-space.nasa.gov/goals/safety.htm

[4] Aviation Safety Program Program Commitment Agreement, Version 1.2, 07/06/00.

[5] http://www.aero-space.nasa.gov/goals/index.htm

[6] AvSP Program Plan, August 1, 1999

[7] http://www.aero-space.nasa.gov/goals/safety.htm

[8] Technical Integration Plan, Baselined Version 1.0. May 25, 2001

[9] Intermediate Program Assessment Package, Synthetic Vision Systems

[10] Intermediate Program Assessment Package, Synthetic Vision Systems

[11] Technical Integration Plan, Baselined Version 2.2. January 18, 2002

[12] AvSPEC Briefing Program Assessment Team

[13] AvSPEC Briefing Program Assessment Team

[14] Technical Integration Plan, Baselined Version 2.2. January 18, 2002

[15] Systems Analysis Team Workshop LeRC 10/29/98

[16] Haimes, Yacov Y. *Risk Modeling, Assessment, and Management*, 1998.

[17] Toward a safer 21st Century: aviation safety research baseline and future challenges, http://www.aero-space.nasa.gov/library/safer21C.htm

[18] JSAT Process Handbook, May 17th, 2001

[19] JSIT Process Handbook, February 28th, 2000

[20] JSAT Process Handbook, May 17, 2001

[21] Joint Implementation Measurement Team Boeing Presentation, 05/09/01

[22] www.casa.gov.au/hotopics/faa-jaa_conf/CAST-JSSI_Safety_Agendas/FAA_Pardee/FAA_Pardee_Narrative.doc

[23] Joint Implementation Measurement Team Boeing Presentation, 05/09/01

[24] Noges, Rob. Joint Implementation Measurement Team, Boeing Presentation: Accident Intervention Process. May 9 2001

[25] http://oea.larc.nasa.gov/news_rels/1997/May97/97_35.html, http://stdweekly.msfc.nasa.gov/techpapers.html, http://www.aero-space.nasa.gov/goals/index.htm

[26] An Aviation Safety Research Investment Strategy, April 23, 1997

[27] http://avsp.larc.nasa.gov/dfs/ASIST.pdf

[28] http://www.nlr.nl/public/hosted-sites/ariba (July 2002)

[29] Guide to Methods & Tools for airline flight safety analysis, Issue1 December 2001, Prepare by: GAIN Working Group B, Analytical Methods and Tools

[30] Hasan, Shahab, Robert Hemm and Scott Houser. Logistics Management Institute, *Preliminary Results of an Integrated Safety Analysis of NASA Aviation Safety Program Technologies: Synthetic Vision and Weather Accident Prevention*, March 2002

[31] Luxhøj, James T., Apichart Choopavang, and Donald N. Arendt, "Risk Assessment of Organizational Factors in Aviation Systems," Air Traffic Control Quarterly, Vol. 9, No. 3 (2001), pp. 135-174. (Special Issue on Flight Safety).