National Aeronautics and Space Administration

Human Reliability Analysis Methods
Selection Guidance for NASA

Chandler, Faith, T.
NASA

Chang, Y.H, James
University of Maryland

Mosleh, Ali
University of Maryland

Marble, Julie, L.
Idaho National Laboratory
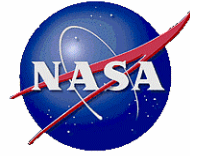
Boring, Ron, L.
Idaho National Laboratory

Gertman, David, I.
Idaho National Laboratory

NASA/OSMA Technical Report

July 2006

This page intentionally left blank.

National Aeronautics and Space Administration

Human Reliability Analysis Methods
Selection Guidance for NASA

Chandler, Faith, T.
NASA

Chang, Y.H, James
University of Maryland

Mosleh, Ali
University of Maryland

Marble, Julie, L.
Idaho National Laboratory

Boring, Ron, L.
Idaho National Laboratory

Gertman, David, I.
Idaho National Laboratory

NASA/OSMA Technical Report

July 2006

This page intentionally left blank.

ACKNOWLEDGEMENTS

The authors would like to thank the following individuals for their comments on an earlier draft of this document and for their attendance at the Human Reliability Analysis (HRA) method assessment workshop held in January 2006 at Kennedy Space Center, Florida: Dennis Bley (Buttonwood Consulting), Roger Boyer (NASA, Johnson Space Center), Andreas Bye (Institute for Energy Technology, Norway), Richard Cook (University of Chicago), Susan Cooper (United States Nuclear Regulatory Commission), Vinh Dang (Paul Scherrer Institute, Switzerland), Katrina Groth (University of Maryland), Bruce Hallbert (Idaho National Laboratory), Teri Hamlin (SAIC), Jeff Julius (Scientech), Gareth Parry (United States Nuclear Regulatory Commission), Pete Prassinos (NASA, Headquarters), Oliver Sträter (Eurocontrol, Belgium), and Michael Stewart (NASA, Johnson Space Center). Attendees of the HRA method assessment workshop are pictured in Figure i. We would also like to thank Barry Kirwan for his extensive comments and guidance on the earlier draft of this document. We would also like to thank Michael Stamatelatos, Bill Vesely, Homayoon Dezfuli and Lynne Loewy for their editorial comments.

Figure i. Attendees of the NASA HRA Workshop. Pictured (left to right): Front Row: Katrina Groth, Susan Cooper, Teri Hamlin, Faith Chandler, Oliver Sträeter, Ali Mosleh, Ron Boring, Andreas Bye, and Jeff Julius. Back Row: Dave Gertman, Roger Boyer, James Chang, Julie Marble, Gareth Parry, Richard Cook, Bruce Halbert, Dov Adelstein, Vinh Dang, and Dennis Bley. Several attendees are not shown in photograph.

TABLE OF CONTENTS

FIGURES

TABLES

ACRONYMS

AC	Action Characteristics
AHP	Analytic Hierarchy Process
ANN	Annunciator
APJ	Absolute Probability Judgment
ASEP	Accident Sequence Evaluation Program
ASP	Accident Sequence Precursor
ATHEANA	A Technique for Human Event ANAlysis
BNFL	British Nuclear Fuels LLC
CAHR	Connectionism Assessment of Human Reliability
CESA	Commission Errors Search and Assessment
CEV	Crew Exploration Vehicle
CNI	Constrained Non-Informative
COCOM	Contextual Control Model
CODA	Conclusions from Occurrences by Descriptions of Actions
CREAM	Cognitive Reliability and Error Analysis Method
CRT	Cathode Ray Tube
EF	Error Factor
EFC	Error Forcing Context
EOC	Error of Commission
EOO	Error of Omission
EOP	Emergency Operating Procedures
EPRI	Electric Power Research Institute
ESD	Event Sequence Diagram
ET	Event Tree
EVA	Extra Vehicular Activity
FFD	Functional Flow Diagram
FHEP	Final Human Error Probability
FLI	Failure Likelihood Index
F-V	Fussell-Vesely
GEMS	Generic Error Modeling System
HEA	Human Error Analysis
HCR	Human Cognitive Reliability

HFE	Human Factors Engineering
HEA	Human Error Analysis
HEART	Human Error Assessment and Reduction Technique
HEP	Human Error Probability
HRA	Human Reliability Analysis
HRR	Human-Rating Requirements
IDA	Information Decision Action
IE	Initiating Event
INL	Idaho National Laboratory
ISCT	Individual Simulator Critical Tasks
ISS	International Space Station
IVA	Intra Vehicular Activity
KSC	Kennedy Space Center
LCC	Launch Commit Criteria
MAPPS	Maintenance Personnel Performance Simulation
MAT	Maximum Allowable Time
MIDAS	Man-Machine Design Analysis System
MLD	Master Logic Diagram
MMS	Man-Machine System
NARA	Nuclear Action Reliability Assessment
NASA	National Aeronautics and Space Administration
NEA	Nuclear Energy Agency
NPP	Nuclear Power Plant
NPR	NASA Procedural Requirements
NTD	NASA Test Director
NUREG	Nuclear Regulation
OAT	Operator Action Tree
OIS	Operational Intercommunications System
OMI	Operations and Maintenance Instruction
ORE	Operator Reliability Experiment
OSHA	Occupational Safety and Health Administration
PIF	Performance Influencing Factor
PFMEA	Process Failure Mode and Effects Analysis
PRA	Probabilistic Risk Assessment

PRACA	Problem Reporting and Corrective Action
PSF	Performance Shaping Factor
PSI	Paul Scherrer Institute, Switzerland
RAW	Risk Achievement Worth
Rf	Recovery Factors
SHARP	Systematic Human Action Reliability Procedure
SHARP 1	Systematic Human Action Reliability Procedure (enhanced)
SLI	Success Likelihood Index
SLIM	Success Likelihood Index Methodology
SLIM-MAUD	Success Likelihood Index Methodology – Multi-Attribute Utility Decomposition
SNL	Sandia National Laboratory
SPAR-H	Standardized Plant Analysis Risk HRA Method
SRR	System Requirements Division
STA	Shuttle Training Aircraft
STAHR	Socio Technical Assessment of Human Reliability
THERP	Technique for Human Error Rate Prediction
TRC	Time Reliability Correlation
UMH	University of Maryland Hybrid
UCB	Uncertainty Bound

Human Reliability Analysis Methods: Selection Guidance for NASA

1.   INTRODUCTION

Human performance has played, and continues to play, a pivotal role in NASA missions. Effective human performance can lead to the accomplishment of NASA mission objectives. In contrast, human errors during system design, fabrication, testing, ground processing, launch control, mission control, and operations may place astronauts in danger, cause injuries or fatalities in operations on the ground, damage hardware and facilities, or cripple a payload or spacecraft before it completes its mission. The effects of human error have been evident in large-scale NASA mishaps, such as the Wide-Field Infrared Explorer, Mars Climate Orbiter, NOAA N Prime, and Genesis. Consequently, NASA has determined the need to include human performance analysis in the evaluation of risk to existing systems and future programs in order to cost-effectively improve safety and overall performance.

This report deals with the subject of Human Reliability Analysis (HRA) for NASA applications, with special emphasis on the selection of methods that can support Probabilistic Risk Assessment (PRA) being conducted on future systems, such as the Crew Exploration Vehicle (CEV), lunar lander, and lunar base. In this context, HRA is the use of systems engineering and behavioral science methods to evaluate the interaction between humans and the system, including the identification, qualitative analysis, and quantitative analysis of human actions, so that the impact of these actions on overall system reliability and their contribution to risk can be understood and managed.

In order for NASA to successfully employ HRA in PRA to support risk-based decision making and design trades, NASA must adopt a standard approach to evaluating and managing human performance related risks. To that end, the NASA Office of Safety and Mission Assurance initiated a study to evaluate existing HRA methods to determine their suitability in the aerospace domain and recommend adoption of methods for use on current and future NASA systems and missions. Although this study evaluated HRA methods that are applicable to human interactions for maintenance activities (ground processing) and flight operations (launch control, mission control, and space flight crew), it predominantly focused on providing recommendations for the quantitative analysis of space flight crew human performance in the support of PRA.

NASA prepared this report with support from the University of Maryland and Idaho National Laboratory. The work was reviewed and enhanced by HRA experts from organizations world-wide via a HRA workshop and individual comments. The guidance and recommendations provided here were developed by professionals experienced in conducting and evaluating HRAs, and in eight cases, the professionals were the authors of existing HRA methods. Consequently, the report reflects the perspectives of HRA experts gained from commercial and government efforts across a variety of domains and applications.

The purpose of this report is to describe what HRA methods can be used to identify, quantify, and evaluate Exploration Systems Mission Directorate program risk, and aid decision making from early conceptual design throughout the life cycle of the program. This report describes the HRA methods study and provides both HRA selection guidance and recommendations that NASA should consider for immediate implementation. This report emphasizes early life cycle integration of HRA and application of Human Factors Engineering (HFE) to provide cost-effective error management recommendations during concept development.

Section 1 provides a general overview of the HRA process, describes NASA's unique performance shaping factors, provides a set of recommended quantitative HRA methods for NASA use, and offers guidelines for selecting the appropriate method. The second section provides a description of the NASA HRA methods study, the criteria used to select methods for the evaluation, and the criteria used to compare the methods for the final selection. Section 3 provides a detailed look at the results of the study for each method evaluated. The appendices provide additional detail on NASA requirements and performance shaping factors

This report is not intended to instruct a novice on how to perform HRA. Rather, it is intended to provide the risk analyst with a familiarization of the HRA process, a list of recommended methods, and rationale used as the basis for the selection of these methods.

1.1   Background

Historically, NASA used HFE on the design of spacecraft and human-system interfaces to improve crew performance and mission success. In more recent years, NASA has applied HFE to reduce human error and improve the safety of public aviation and selected space systems. Over the last 10 years, NASA has applied Human Error Analysis (HEA) and, more specifically, HRA to ground-based operations, design processes, testing, and space system operations to evaluate risk and enhance safety and mission success.

As early as January 1969 NASA was considering the impact of human error on manned space flight systems. The Office of Space Flight System Safety Requirements for Manned Space Flight, Safety Program Directive No. 1 (NASA, 1969) stated that hazard analysis should be completed for both "human and equipment failure on the safety of the system." This directive also required the results of hazard analysis to be employed in eliminating and controlling critical hazards. NASA focused HFE efforts on space flight, developing state of the art interfaces, habitation systems, interior layouts, and information management to achieve NASA's goals. These goals have been realized, in part, through the development and application of Man System Integrations Standards for the Space Shuttle (NASA, 1995) and for the International Space Station, NASA-STD-50005 (NASA, 1995).

Quantification of human error occurred much later. NASA's PRA efforts were initiated in 1988. They began as result of two influential reports. First, in October 1986, the report "Investigation of the Challenger Accident," authored by the United States House of Representatives Committee on Science and Technology, indicated that NASA could not focus its Space Shuttle resources effectively without a means of estimating the probability of failure for Shuttle events. Second, in January of 1988, the report "Post-Challenger Evaluation of Space Shuttle Risk Assessment and Management," authored by the Slay Committee, reinforced this sentiment by recommending NASA apply PRA approaches to the Shuttle Risk Management Program. Together, these reports resulted in the birth of the NASA PRA efforts.

In 1991, the Magellan Mishap (caused by human error) at Kennedy Space Center (KSC) sparked interest at the NASA Center-level for applying HEA and HFE to reduce human errors in ground processing. In 1993, the KSC Shuttle Ground Processing Human Factors team was formed to generate human error evaluation techniques and apply them to accident investigations.

In 1998, the NASA Office of Safety and Mission Assurance established the first Human Reliability Program to focus HFE on the identification, reduction, and management of human error in the Space Shuttle Program ground processing activities at KSC. The goal was to focus on critical processes that exhibited high error rates and resulted in a significant number of problem reports (damage and delays). This effort led to the completion of three projects aimed at evaluating and reducing human error. First, a Human Factors Process Failure Mode and Effects Analysis (HF PFMEA) methodology was created to identify potential human errors (failure modes), factors that contribute to human error, and potential consequences of these errors. The HF PFMEA methodology expected to generate recommendations to reduce error and mitigate its effects. The methodology was used to evaluate the Space Shuttle dome heat shield installation and removal process, and resulting recommendations have led to increased safety and efficiency in Space Shuttle processing. The second and third projects were continuations of the first, both seeking to apply HFE methodology and principles to the reduction of error through error management and system redesign. These projects led to innovative new technologies that are applicable to NASA and commercial industry. From this spawned the further refinement of the HF PFMEA methodology, the development of HF PFMEA software (now also commercially available through RELEX corporation), and HF PFMEA training. Since then, this methodology has been applied to the Space Shuttle, International Space Station, and Payload processing activities at KSC and other locations.

Simultaneously, efforts were underway to generate a world class PRA tool (the Quantitative Risk Assessment System) and to incorporate HRA into PRAs. As a part of these efforts, the Shuttle Probabilistic

Risk Assessment Team (SPRAT) generated a Human Reliability Scope Study outlining an approach that was used to perform detailed Human Error Probability (HEP) calculations to identify relevant crew actions for four pilot systems. Over the years, as the Shuttle PRA was refined, more HRA was performed. The NASA Shuttle PRA used the Technique for Human Error Rate Prediction (THERP) as a screening tool and evaluated pre-initiating events (Shuttle ground processing errors), initiating events (crew errors), and post-initiating events (crew errors) using the Cognitive Reliability and Error Analysis Method (CREAM). Another major NASA program, the International Space Station program, chose to identify human errors in their accident scenarios and use 1 x 10^{-3} (also expressed as 1E-3) as the HEP, rather than explicitly quantifying the contribution of human error to risk.

In 2002, NASA approved its first set of Agency-level human-rating requirements. These requirements, called the NASA Procedural Requirements (NPR) 8705.2, *Human-Rating Requirements for Space Systems* (NASA, 2005), take a proactive approach to human error management. (For more detail see Appendix A). To satisfy these requirements, NASA human space flight programs must perform HEA and identify cuts set(s) that can contribute to loss of life or loss of vehicle, in order to determine where failure tolerance and other error management solutions are needed.

NASA has successfully developed sophisticated, reliable, and robust systems, and mishaps and incidents due to equipment failure have become more and more rare. Despite this success, the human component of systems remains fallible, and the number of mishaps attributed to human error is significant. Human error is a significant contributor to mishaps which result in loss of life; damage to one-of-a-kind hardware, government equipment, and government facilities; and loss of scientific knowledge. In a recent study of NASA mishaps, it was found that 57% of Type A mishaps[a] were caused by human error (1996-2005). This is consistent with general industry, in which human error contributes to approximately 80% of the accidents (Hollangel, 1993). This is also relatively consistent with civil and military aviation accident data which indicates that human error is a contributor in 70% to 80% of mishaps (Wiegmann and Shappell, 2001). These percentages further demonstrate the need for NASA's human-rating requirements and error management philosophy.

NOAA N Prime Mishap
September 2003
Proximate Cause: Human Error

As NASA expands its capability in PRA, striving to become a world leader in risk assessment for the aerospace domain, it will also expand HRA applications to other NASA systems and future applications. In part, HRA's growth is fueled by the fact that human error is a significant contributor to NASA mishaps, as is articulated in numerous NASA mishap reports. The goal is to apply HRA to all PRAs (at a minimum apply HEA to qualitative risk assessment) to provide relevant, practical, and timely contributions to NASA's management of risk. This application of HRA to PRA will influence future design decisions by identifying ways to reduce the likelihood of human error and by making systems safer and more effective.

---

[a] A Type A Mishap is a mishap resulting in one or more of the following: (1) an occupational injury or illness resulting in a fatality, a permanent total disability, or the hospitalization for inpatient care of 3 or more people within 30 workdays of the mishap; (2) a total direct cost of mission failure and property damage of $1 million or more; (3) a crewed aircraft hull loss; (4) an occurrence of an unexpected aircraft departure from controlled flight (except high performance jet/test aircraft such as F-15, F-16, F/A-18, T-38, OV-10, and T-34, when engaged in flight test activities). **Note that the study did not include auto accidents or death by natural causes in the analysis.**

1.1   General Overview of HRA

To optimize total system reliability, NASA must consider hardware, software, and human reliability in the design and analyses of systems (Figure 1). Human reliability refers to 1) the probability that the human elements will function as intended over a specified period of time under specified environmental conditions, and 2) the probability that no extraneous human actions detrimental to the system reliability or availability will be performed. Human reliability in a space system includes the reliability of the crew in space and the personnel on the ground.
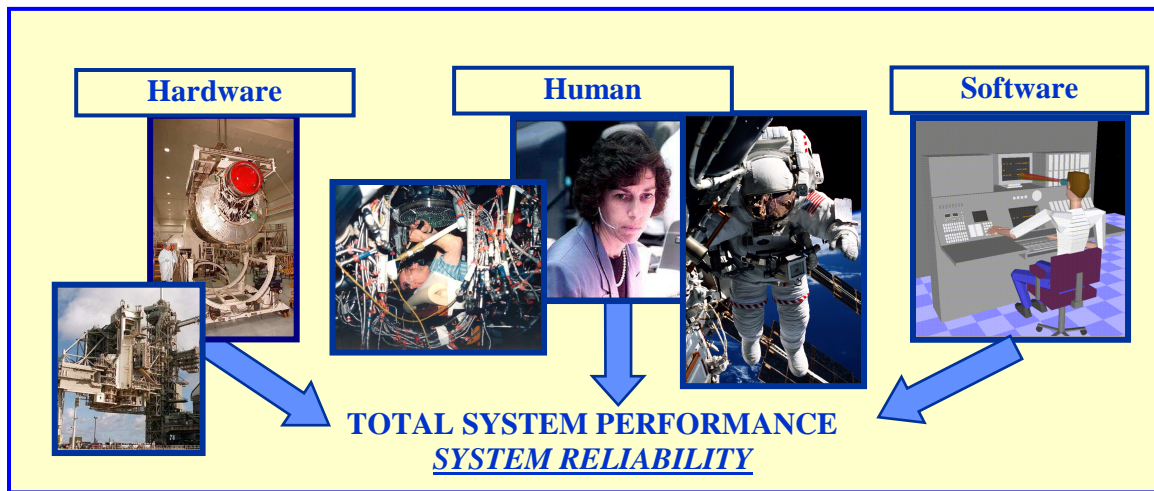
Figure 1 . Total System Performance

HRA is a comprehensive and structured methodology that can support NASA programs. Using HRA, NASA can evaluate existing or future systems to model what human actions[b] or errors can negatively impact the system, predict how often these will occur, and identify the consequences if they do occur.

Recommendation: Use HRA throughout the life cycle of the system, beginning early in the system design process.

HRA can have benefits at every phase of the system life cycle. During the design phase, HRA is a tool that can be used to support the evaluation of concept designs by quantitatively comparing two design solutions and determining which designs best achieve the program risk objectives. HRA can assist in the identification of human actions (and corresponding system interfaces) that pose the most significant risk to the system. It can be used to identify potential errors or accident scenarios so that design modifications can be made prior to system fabrication, leading to fewer and less costly design changes later. Additionally, HRA can be used during system operation to evaluate and compare proposed system upgrades or evaluate factors that are contributing to significant problems, such as anomalies, damaged property, and/or delays. To achieve the most effective results, the same HRA activities should support both the HFE interface design and the

---

[b] Human actions in this context refer to human errors and not violations. Human error is defined as either an action that is not intended or desired by the human or a failure on the part of the human to perform a prescribed action within specified limits of accuracy, sequence, or time such that the action or inaction fails to produce the expected result, and has led or has the potential to lead to an unwanted consequence. Violation is defined as an action that was intended and desired by the human that departs from rules (e.g., intentionally skipping a step in a procedure or taking a short cut) or breaks the law (e.g., speeding on the highway).

PRA, ensuring that the task analysis, accident scenarios, and mitigations are consistent, represent the actual system, and address the critical risks.

Recommendation: HFE design efforts should pay special attention to the human interactions and accident scenarios identified by the HRA as critical for overall system reliability and safety and generate solutions that mitigate risk.

HRA Context in PRA. NASA is utilizing PRA as a unifying process for the consideration of hardware, software, and human reliability in the design and analyses of systems (Figure 1). As noted in the *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners* (NASA, 2002), "The PRA ultimately presents a set of scenarios, frequencies, and associated consequences, developed in such a way as to inform decisions regarding the allocation of resources to accident prevention. This could be changes in design or operational practice, or could be a finding that the design is optimal as is." PRA scenarios typically consist of an initiating event (IE), one or more pivotal events, and resulting end states. Definitionally, an initiating event is an event that has the potential to cause loss of a system function leading to an undesired end state such as loss of life, damage to or loss of property or equipment, failure of a mission, unavailability of a system, or damage to the environment. A pivotal event is an event that is a success or a failure of a response, or an occurrence or non-occurrence of an external condition or key phenomenon, which occurs after the initiating event and mitigates or aggravates the severity of the consequence.

Pivotal events have at least two possible outcomes reflecting success or failure of the event. The pivotal event outcomes depend in turn on success or failure of hardware, software, or human interactions with the perturbation posed by the IE and conditions determined by any preceding pivotal events. The IE and pivotal events will have associated frequencies and probabilities, and the end states reflect the consequences of particular combinations of the IE with specific outcomes of the pivotal events. Experience has shown that a successful PRA process will result in a comprehensive, organized collection of scenarios that reflect a consideration of the IEs, all relevant outcomes of the pivotal events, and quantifiable likelihoods of occurrence of the end states. The complete PRA process is described in the *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners* (NASA, 2002)], Section 3.

Recommendation: HRA must be an integral part of the PRA development, from its earliest stage, in order to identify, analyze, and, if necessary, quantify, the points where humans interact with the hardware and software in each scenario.

Steps in a HRA: In general, the HRA process has a number of distinct steps, including the HRA problem definition, task analysis, human error identification, human error representation (including the modeling of dependencies between different human errors), and human error quantification (Figure 2). The way in which each step is conducted is dependent upon the HRA method used and the purpose of the analysis. After an HRA is completed, NASA engineering may incorporate error management techniques to reduce errors or mitigate their effects. Further quantification can be done to verify that the measures were effective in lowering the impact of human error on the overall system reliability.
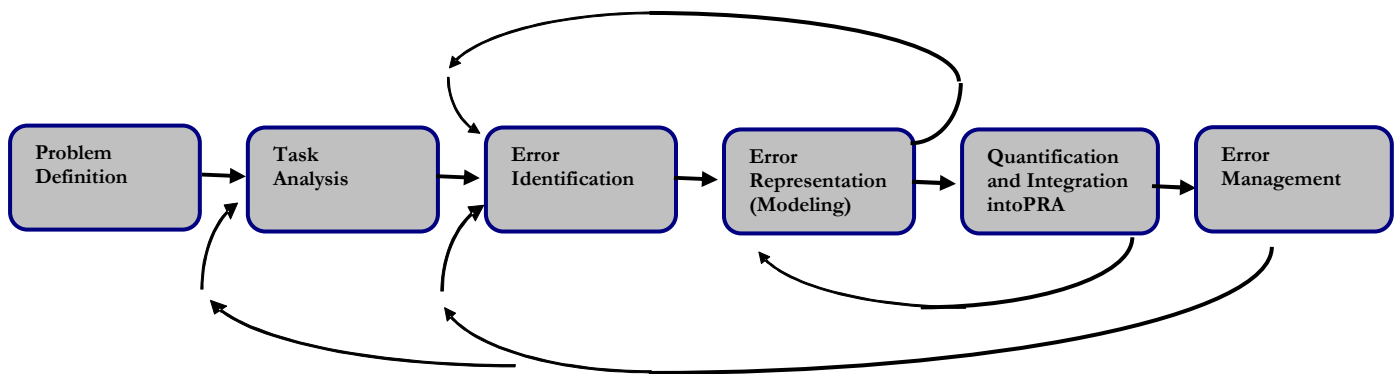
Figure 2 . Basic Steps in the HRA Process

Problem Definition.

The problem definition is the first step in the process and is used to determine the scope of the analysis, including what type of analysis (qualitative or quantitative) will be conducted, what tasks (normal, emergency) will be evaluated, and what human actions (pre-initiating actions, initiating actions, and post-initiating actions) will be assessed. There are two factors that impact the determination of scope of the analysis. The first is the system's vulnerability to human error.

A NASA space system's vulnerability to human error is dependent upon the complexity of the system (and how the NASA team understands this complexity), the amount that the human interacts with the system (either through maintenance, operation, and/or recovery), and how the human-system is coupled. (A tightly coupled system does not allow the user the flexibility to use alternatives or wait for a repair when there is a failure). An optimal system design would be error tolerant, have less complexity (or easily understood processes and design), provide the capability for the human to detect and correct errors, and be loosely coupled, allowing the human flexibility in operations if failures occur. In general, when a system is more vulnerable to human error, then a larger scope and comprehensive analysis is needed to fully understand and mitigate the human contribution to system risk.

The second factor that impacts the determination of the scope of the analysis is the purpose of the analysis. For NASA, the analysis could be completed to support accident investigation, anomaly and problem report evaluation, process improvement, a PRA for design trades, or operational improvements.

Once the purpose is identified, the analyst can determine what type of analysis (qualitative or quantitative) will be conducted. Typically, analyses completed for accident investigation, problem report evaluation, and general process improvement are more qualitative in nature, whereas those analysis that support design trades and major operational improvements are quantitative. A qualitative analysis may identify "what can go wrong," including what potential errors or violations could occur, the potential effect of these errors, and those errors that pose the most significant risk. Qualitative analysis is performed through modeling of the task or tasks with an emphasis on points at which the human acts. At those points, the model

identifies factors that could affect performance. A qualitative analysis such as the HF PFMEA (described in Section 3) is one method that has been adopted and applied to Space Shuttle processing, payload processing, and the evaluation of control rooms. Quantitative analysis, which was the focus of the NASA HRA study, will be discussed in more detail later in this report.

During the problem definition phase, determining what type of human actions will be evaluated is very important, because the number and type of errors included in the analysis can lead to an underestimation or overestimation of the impact of the human errors on the system risk. There are an infinite number of possible human actions, and it is impossible to predict every potential human error and how it will impact the system. However, it is possible to complete a comprehensive evaluation of the human system interfaces and processes and identify many potential human errors and their effects.

Recommendation: When NASA is conducting a HRA in support of a PRA, it is recommended that the scope include the assessment of human errors (both errors of omission and errors of commission) and take into account pre-initiating actions, initiating actions, and post-initiating actions.

This recommendation is supported by NASA's mishap investigation data which indicates that human error was the initiating event or post-initiating event in approximately 24% of NASA's Type A mishaps. The data also indicates that human error (e.g., an engineering design error, inspection error, test error, calibration error, or other maintenance error) was a pre-initiating event in approximately 57% of NASA's Type A mishaps in the last ten years. Both errors of omission (failing to do something) and errors of commission (doing something incorrectly) have been listed as causes or contributing factors to NASA mishaps and accidents in other industries. Consequently, to accurately reflect the risk contribution, it is best practice to include both errors of omission and errors of commission in the HRA.

Task Analysis (Task Decomposition).

The second step in the HRA process is task analysis, a systematic method to identify, list, and break down each task into the steps and substeps that describe the required human activities in terms of physical actions and/or cognitive processes (e.g., diagnosis, calculation, and decision making) necessary to achieve the system's goal. Over 25 variations of task analysis exist to accomplish different goals such as task data collection, task description, simulation, behavior assessment, and task requirement evaluation. Describing each variation of task analysis is beyond the scope of this report and the reader should refer to Meister (1985) and Kirwan and Ainsworth (1992) for additional detail.

Often a task analysis begins after a functional analysis has been completed to identify all system functions that must be performed to achieve the system's goals. A function is an action that can be accomplished by either a human or machine. It can be instantaneous (e.g., provide thrust or power up equipment) or occur over time (e.g., provide thermal control, monitor life support system). A function can be complex, requiring a series of steps to complete (e.g., provide propulsion or position the davit crane), or simplistic (e.g., attach the fan belt). As a part of function analysis, functional flow diagrams (FFDs) are developed to depict the chronological sequence of functions and illustrate the overall flow in the process. The functional analysis and FFD (or process flow diagram) provide the global sequence and framework for the task analysis.

The goal of task analysis is to decompose the high level functions into tasks, tasks into their constituent subtasks, and subtasks into human actions. A thorough task analysis ensures that the entire system or process has been completely evaluated and all potential actions have been identified. When considering the level of task decomposition, the analyst must consider the purpose of the task analysis and the resources available. If the task analysis is being performed to understand the risk contribution of human errors at a system's functional level (as needed in early system design), or the resources are limited, the task analysis should be kept at a higher level. (This is often called a screening analysis.) However, if the task analysis is being conducted to evaluate a specific process that is experiencing problems, or is being used to further understand a task that has been identified as a significant risk contributor, a more detailed task analysis (with decomposition to lower levels) is warranted.

Today, there are software tools that support task analysis and discrete-event simulation. Two such MicroSaint modeling tools, IMPRINT and C3TRACE (Kilduff, P. W., Swoboda, J. C., & Barnette, B. D., 2005), were developed by the Army Research Laboratory and are free for U.S. Government use. These tools allow the analyst to perform mission, function, and task decomposition, as well as measure human multi-channel workload and other human performance shaping factors such as employee education, time in the job, age, and more. The analyst can also define attributes such as the priority of the task, the situation awareness level required, the task mode (manual, automatic, or both), and task time, to name a few. All tasks are then represented in task level network diagrams. Task analysis completed with these tools can support the HRA, strengthen the system design process, and evaluate the effectiveness of manpower organizational design.

After the task analysis is complete, the HRA analyst should request system/process owners (e.g., engineers, technicians, and operators) to verify that all steps have been included and described accurately. A task can also be modeled or evaluated using a computer-aided design tool such as Man-Machine Design and Analysis System (MIDAS). MIDAS is a fully integrated human performance modeling tool that provides a graphic representation of the human, equipment, crew stations, and environments. It can be used to evaluate the task sequence, workload, anthropometrics, reach envelope, situation awareness, eye gaze behavior, and predict performance (Hart, et. al, 2001).

A comprehensive task analysis identifies all human actions and serves as the building block for understanding all the places where human error can occur in the process. The task analysis is the foundation of good human error analysis. (It should be noted that some HRA methods (e.g., Success Likelihood Index

Methodology (SLIM) and Absolute Probability Judgment (APJ)) do not perform this step, but instead quantify the task failures rather than the individual human actions.) NASA has not adopted a specific task analysis method, because each method has different advantages and disadvantages and should be selected based on purpose of the analysis and the HRA approach used.

Error Identification.

The third and the most important step in the HRA is human error identification, where basic human actions are evaluated to determine what human errors and violations can occur, have potential contributions to hazardous events, and should be included in the analysis. If the analyst fails to identify critical human errors, then their contribution to risk will not be included in the HRA/PRA, and the overall system risk will be underestimated.

The analyst must determine what type of human error will occur and the performance factors that could contribute to the error. To accomplish this, the analyst must identify and understand the different types of human errors that can impact the system. Human actions/interactions within a system can be broken down into two main types of elements, a cognitive response or a physical action, and their related errors. Within the NASA environment, many important cognitive human responses contribute to accidents, including the failure to perceive a signal or alarm, failure to interpret the information correctly, and the failure to make the correct decision. Human actions/errors that could adversely affect the system reliability, such as those that occur prior to the initiating event (pre-initiating actions), serve as the initiating event (initiating actions), or provide the ability to detect and correct the system failure or mitigate the effects (post-initiating actions), should be included in the analysis.

Recommendation: When a task has a significant cognitive aspect, such as human decision-making tasks during inspections, space vehicle launches, piloting, and manual docking, the HRA should include the evaluation of errors for both cognitive responses and physical actions.

Human actions and errors cannot be considered in isolation from the system and environment in which the human works. The system design (hardware, software, and crew habitable environment) affects the probability that the human operator will perform a task correctly or incorrectly for the context and specific situation. Consequently, it is important to evaluate the factor(s) that may increase or decrease the likelihood that these errors will occur. A Performance Shaping Factor (PSF) is anything that can affect the ability of the person to carry out the task. External PSFs are outside the individual's control (e.g., design of the task, tools and equipment, environmental factors, policies, and procedures). Internal PSFs are human attributes brought to the task by the person that, in some cases, can be influenced by the person (e.g., skills, knowledge, abilities, attitudes, fatigue, etc.). Once PSFs are identified, their influence on the potential human error is determined so that the basic human error rate can be modified (adjusted) per the specific situation.

When human error identification and PSF evaluation are being completed for an existing process where problems have occurred (e.g. hardware damage, personnel injury/illness, and schedule delays), the analyst can identify from problem reports, mishap reports, and other data logs the actual errors that did occur and the PSF that are present in the existing environment.

No one methodology or technique will be able to identify all human errors or list all conditions and circumstances under which human error will occur. However, in cases where the analysis is being completed for an existing system (e.g., payload or space flight hardware ground processing), this is a much easier task. The analyst can use a structured method to identify potential errors (errors of omission and commission) for each potential action and determine which are most credible and most hazardous by evaluating the task itself and the PSFs that are present in the environment. The analyst can gather supporting evidence by evaluating NASA data such as problem reports, mishap reports, safety logs, personnel records, and system data. Each type of record will provide a different kind of information. For example, Occupational Safety and Health

Administration (OSHA) logs will provide the incidence and severity information on the types of accidents and injuries that have occurred and a general statement of why they occurred, giving some indication if a human action or error was contributory or causal. NASA mishap reports will describe the human errors that caused or contributed to the mishap and the factors that influenced the error. NASA mishap reports also provide an event and causal factor tree graphically representing the chain of events that occurred. Other data sources, such as worker's compensation records, will furnish additional information on the incidence and severity of injuries and the costs associated with them. The review of Problem Reporting and Corrective Action (PRACA) and quality records can also assist the analyst in identifying what jobs have recurrent problems, and, in some cases, how these problems are linked to human actions such as workmanship. Together, this data will allow the analyst to confirm what errors have occurred, their effects, and PSFs so that, at a minimum, these errors are included in their analysis. It should be noted that when existing processes are evaluated, more PSFs may be included in the analysis than in those done for new systems, because their presence and effect can be confirmed through observation, interviews, and existing records.

The analyst evaluating tasks that will be performed on new space system designs will not have the luxury of evaluating the existing task and will only have data for related systems. However, human error identification and analysis are still possible. Each human action can still be evaluated to determine the most probable human errors using expert judgment and computer-aided design tools that simulate the human performing the task on the system.

Error Representation (Modeling). The fourth step in HRA is human error representation, also described as modeling. This step is conducted to help visualize the data, relationships, and inferences that cannot be as easily described with words. Human error modeling allows the analyst to gain insight into the causes, vulnerabilities, recoveries, and possible risk mitigation strategies associated with various accident scenarios. Modeling is done as a part of the PRA, where the HRA analyst provides input to PRA products. Human errors can be modeled and represented in a Master Logic Diagram (MLD), Event Sequence Diagram (ESD), Event Tree (ET), Fault Tree (FT), or a generic error model and influence diagram.

Each scenario begins with an Initiating Event (IE). An IE is generally developed and supplied to the model as a frequency from sources outside the scenario. Human errors can be represented in a MLD as a contributor to an IE; note, however, for a complex system, there are very few human errors that serve as single point failures that can cause a critical function to fail, resulting in an undesired end state such as loss of vehicle. When determining which human errors should be considered for analysis, the system should first be evaluated in its "normal" operating condition, where all human actions that can physically vary the system can be identified and modeled.

For each IE that is represented in a MLD, a corresponding ESD is developed. The ESD is a flow chart that depicts the sequence of events that occurred, or failed to occur, in a scenario and the resulting end state for each. Human errors can be represented in the ESD as the sub-event or as a pivotal event. Typically, human errors are not in themselves modeled as an initiating event; rather, the consequence of the human errors (e.g., subsequent system failure) serves as the initiating event. Figure 1A, (from the *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners*, Chapter 6) displays a MLD that represents a situation in which the failure of a demanded system to run (failure of a pivotal event) has some possibility of being recovered by the crew (success of a pivotal event), and restored to operation. There is not sufficient detail at this level of modeling to determine whether diagnosis is necessary, or how complex the recovery task might be. Operational history might be available to aid in task analysis and error identification, depending on the life-cycle phase of the problem being analyzed.
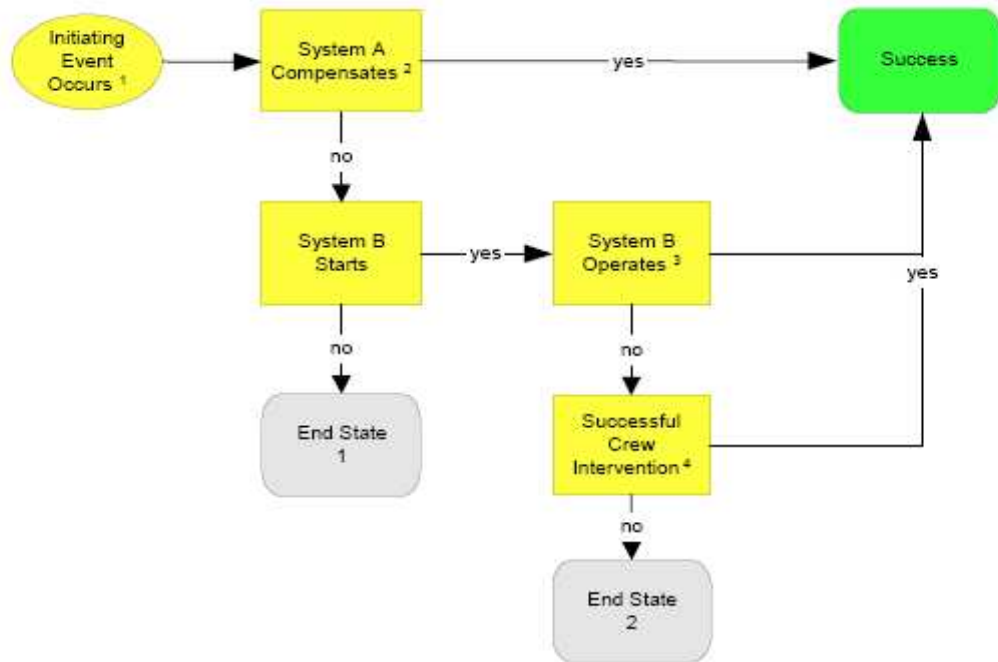
Figure 1A. Master Logic Diagram for a Scenario with Crew Intervention (from NASA, 2002).

Once an ESD is developed, it is possible for the analyst to develop an ET that represents the basic initiating event and the occurrence or non-occurrence of pivotal events that can lead through recovery to normal status of the system or to an undesired outcome. The ET illustrates the event sequence, progressing from left to right in the chronological order of events, providing end-to-end traceability of accident scenarios. The ET structure uses binary logic to split each basic event into two branches according to their consequences, with pivotal events leading to failure on lower branches, and pivotal events leading to success on corresponding and parallel upper branches. ETs are predominately used to identify the event sequences which result in some undesired outcome and serve as a basis for scenario quantification. Human actions that serve as pivotal events (aggravating or mitigating the initiating event) in the system failure must be modeled in sufficient detail such that the events can be quantified. For example, human errors/actions that represent the activation of equipment corresponding to different systems are included in ET for both normal and emergency procedures. Human recovery actions that could improve the system condition after a failure, or can be used to recover or replace a failed system, also may be included in ETs. Figure 1B, (from *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners,* Chapter 6) continues the example from the MLD figure above. The "Crew Intervention" pivotal event is shown in the ET as the last top event (R), where success of the intervention results in scenario success (End-State 3), while intervention failure results in End-State 4, representing some degree of overall failure of the system.
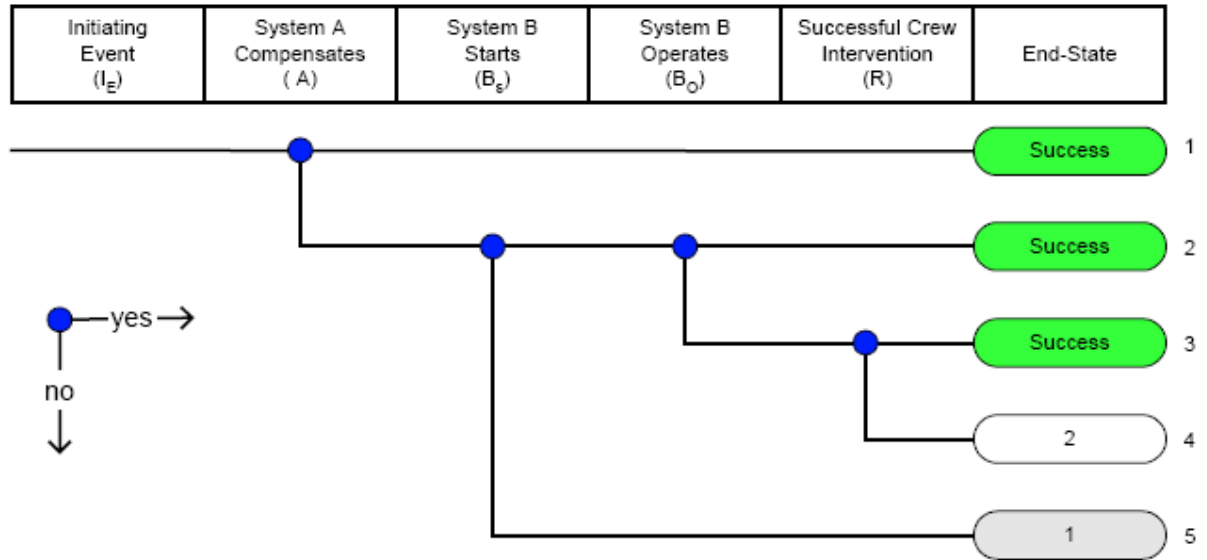
Figure 1B. Event Tree for a Scenario with Crew Intervention (from NASA, 2002).

Complex pivotal events can also be modeled using FTs. A FT represents the human and system actions and inactions in a logical top down structure, starting with the undesirable outcome at the top. The FT uses AND and OR logic to show what basic event (human error, hardware failure, software failure, or environmental change) or combination of events could have combined to produce the top event. The FT top event is defined and named to correspond to an ET top event, and is logically linked to it in the PRA modeling software. FTs are developed in "failure space," and the basic event probability is the probability of failure. The combinations of events in the FT that evaluate to failure at the top are supplied to the ET logic during evaluation, and those combinations are then found in the resulting cut sets.   FTs are developed for all non-trivial pivotal events depicted in the ET. If a human error causes an IE, then the human error is typically explicitly modeled in the FT. In general, individual human errors (pre-initiating event actions) are included in FT when the human action (test, maintenance, calibration, or operation) can change the system state or disable the system causing a system component failure or loss of system function. Human errors/actions that affect the development of the sequence and apply to a single system are modeled at this level. It should be noted that in some HRAs that support PRA, the human contribution to an individual component failure is not included in the analysis because it is assumed that the human error is represented as a component failure at a higher level. When FTs are used, it is important to determine how far down in the fault tree the errors will be represented. If human errors/actions are very low in the FT, they will not have much importance on the overall PRA (quantitatively).

In the ET of Figure 1B, the last top event, "Crew Intervention," could be further decomposed using a FT linked to the ET. Figure 1C displays a FT that breaks the intervention into diagnosis and action components (basic events). Each of these components can be analyzed and quantified, and are modeled under a logical OR gate (represented by ᑌ) to reflect the fact that failure of either of these basic events will constitute failure of the "Crew Intervention" top (pivotal) event. In this example, the human errors are modeled as direct contributors to the top event, and one or the other will appear in all cut sets generated for End-State 4 of this scenario.

13
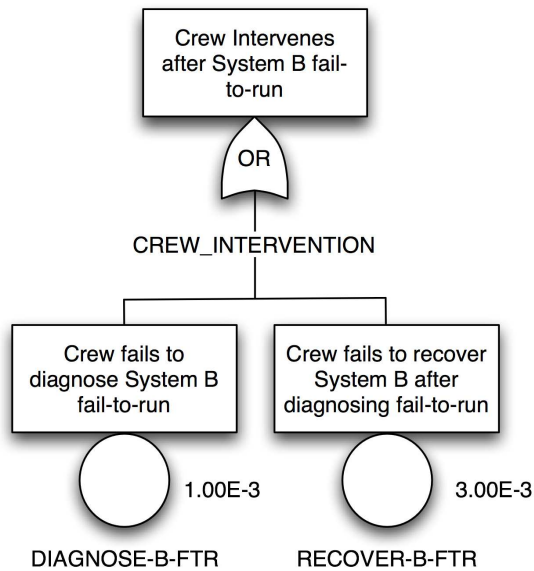
Figure 1C. Fault Tree Logic for "Crew Intervention" Top (Pivotal) Event.

Human errors can also be represented using a number of error modeling techniques that each have different advantages and disadvantages. For example, Generic Error Modeling System (GEMS) has been used to represent how error mechanisms work, the Socio Technical Assessment of Human Reliability (STAHR) has been used to illustrate the influences affecting decisions and actions, and the Maintenance Personnel Performance Simulation (MAPPS) has been used to identify maintenance activities and their sequence and interrelationships.

It should be noted that, during this step of the HRA, the analyst must also consider modeling dependencies between different types of human errors. There are a number of different types of dependencies including: 1) the likelihood that one human error contributes to or causes another, 2) the likelihood that one person has the same error repeatedly (reads all the displays incorrectly), or 3) that one person makes an error (calculates the pressure incorrectly), increasing the likelihood that a second person will make an error. Dependencies are difficult to model, and only a few HRA methods provide methods and/or guidance on this topic.

For NASA, the FT and ET structure allow visualization of the effects of combinations of failures and are the preferred methods for representing human actions, because these methods are consistent with the way the Agency models hardware and software failures prior to incorporating the data into MLD used in the PRAs.

Quantification and Integration into PRA.

Quantification is the process used to assign probabilities to the human errors. The HEPs are incorporated into the PRA to provide comprehensive accident-sequence quantification and allow the practitioner to determine which human errors were the most significant contributors to system risk. For example, in Figure 1C, the hypothetical HEP is 1E-3 for diagnosing failure to run and 3E-3 for recovering from the failure to run. These HEPs are incorporated into the overall PRA risk calculation.

14

The steps in quantification are dependent upon the method being used. (A variety of methods will be described in detail, compared, and contrasted later in this report). All HRA methods recommended in this report allow generation of an estimated HEP that may be incorporated into an FT as part of the PRA.

The method by which quantification is completed is dependent upon the resources available (time and money to perform quantification), the experience level of the analyst, and the relevant available data. The data must be sufficient to allow the analyst to estimate the frequency with which the errors may occur and the number of opportunities for these events. The data may come from databases, simulations, or expert judgment.

Quantification can be performed as a screening analysis or as a detailed HRA. The purpose of the screening analysis is to limit the number of human errors/actions that must be evaluated. (Usually this is done to save time and/or money). For example, if the analyst identifies a large number of human errors that require quantification, this can be time consuming and resource intensive. Rather than quantify all the human errors, it is more desirable to perform a conservative screening of these errors, so that the more significant contributors to overall system risk can be identified for detailed HRA.

Screening can be accomplished using a qualitative method or a quantitative method. When a qualitative screening method is used, the analyst excludes some human actions/errors from consideration based on a set of assumptions. For example, the analyst may exclude an error because it will be accounted for in the analysis in other ways or if the impact of the error is insignificant because the effects will be detected and corrected by other means (e.g., the analyst may screen out a maintenance calibration error because the system will automatically calibrate itself upon initialization). The analyst may also use a quantitative method to screen. In this case the analyst uses a very conservative HEP in the PRA logic models to determine if the error has a significant impact on the overall system risk. Those errors that appear in the minimum cut sets of the PRA models, thus indicating they have a significant impact on system risk, are further evaluated using detailed quantification methods. There are importance measures such as risk reduction and risk increase calculations that provide rankings for prioritization of events. Those errors that do not have a significant impact on system risk are not analyzed any further. Detailed HRA analysis is performed on all human errors that were determined to be significant by the screening analysis.

Recommendation: The analyst should carefully weigh the value of a screening analysis vs. a detailed HRA. Where time and resources allow, a screening analysis allows the analyst to establish risk significant events and conduct a detailed HRA only on those events.

Following quantification of the risk, the analyst must determine the relevance of the failure estimate and the uncertainty of probabilities[c]. Uncertainty is high when there is sparse data on the human initiating events, site conditions, and related human errors. Uncertainty is also influenced by the analyst's understanding of the influence of performance shaping factors and inability to identify and model all failure modes. The uncertainty surrounding the HEPs may be expressed via a distribution. (Typically the PRA uses a log-normal distribution to express these uncertainties). The HRA quantification techniques each treat uncertainty differently, and some methods do not provide a means to estimate uncertainty at all.

Recommendation: Risk significant human errors should be included in the PRA master logic diagram, event tree, and fault tree. HEPs and uncertainty information should be included according to the conventions of the HRA method in use.

---

[c] The analyst should perform uncertainty analyses to evaluate their degree of knowledge or confidence in the risk quantification. There are two types of uncertainty, "aleatory" (or stochastic-type) and "epistemic." The first is randomness-driven and the second is associated with errors in the models and insufficient knowledge of modeled processes. Monte Carlo or related simulation methods can be used to perform the uncertainty analysis (NASA, 2002).

Once the HRA has been completed, and the human errors have been modeled and quantified as part of a PRA, the risk calculations are performed to evaluate the overall system risk. The PRA team will perform sensitivity analyses to identify those human errors (or system inputs) that cause the greatest changes in partial or final risk results. This will help the HRA analysts focus on those human errors.

The analyst will want to determine which human errors are dominant contributors to system risk and make decisions about the design of human-machine interfaces. To do this, the HRA/PRA team will rank leading contributors to risk in decreasing order of importance (importance ranking). For places where human error has been identified as a dominant contributor to risk, the system owner may decide to take actions to make the system more error tolerant. This can be achieved by implementing a human error management process that prevents the error from occurring, through the inclusion of barriers such as physical guards and dissimilar adapters, provides a means to detect and correct the errors, reduces the frequency of the error through modification of performance shaping factors that have a negative impact on the human error rate, or through the mitigation of the negative effects of the error.

Human Error Management.

Human error management can be employed by HFE after a mishap, if a problem process has been identified, or if HEA/HRA has identified potential human errors that pose significant risk to the system. Human error management philosophy assumes that humans will remain fallible. Even well trained crew will make errors. However, the philosophy indicates that potential human errors can be identified, and good engineering can eliminate some errors, minimize others, and lessen the negative impact of most of those remaining. In general, human error management seeks to develop a system that minimizes errors and tolerates those that remain to provide the maximum assurance that the system will not experience a catastrophic failure or result in a major mishap, despite the errors that may occur.

NASA's human error management philosophy is represented by the human reliability curve and the recommended error management strategy shown on the right side of Figure 3. The human reliability curve is a very simple conceptual model that depicts the relationship between the system design characteristics and the rate of human error. As the system design characteristics (including physical hardware design and all other external performance shaping factors) are improved to prevent significant human errors[d], the human reliability improves, and the corresponding system reliability improves. In contrast, where the system design does not prevent human error, human reliability is lower, and the overall system reliability is lower. However, even when human reliability is not optimal, the system reliability can remain high if the system (hardware, software, or human) can detect and correct the human error or mitigate its negative consequences prior to the undesired outcome.
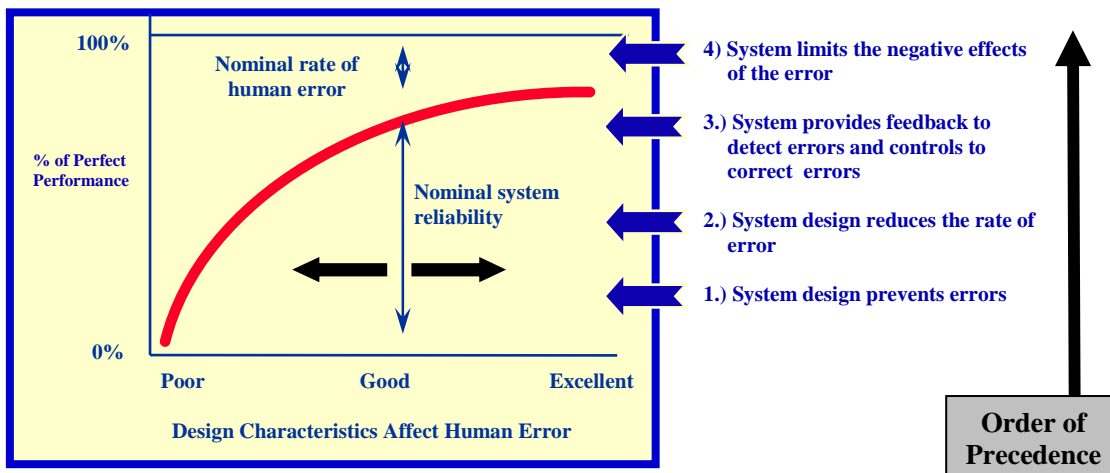
Figure 3 . Human Error Management Philosophy

NASA human error management philosophy is to manage the risk associated with human error in the following order of precedence: 1) system design prevents the error, 2) system design reduces the rate of error, 3) system provides feedback to detect errors and controls to correct error, and 4) system limits the negative effect of the errors.

---

[d] Significant human errors are human errors that have been shown by HEA or HRA to have a significant impact on overall system reliability.

The most effective way to ensure reliable human performance is to incorporate a design that prevents the error from occurring. So, the first step is to prevent the error from occurring. Table 1 provides examples of ways to prevent human errors.

Table 1. Possible methods to prevent human error (NASA, 2005).

Possible Methods to Prevent Error
• *Automatic Sequencer (prevents human's mis-sequencing)*
• *Automation (prevents human's calculation errors)*
• *Automation (prevents human's monitoring errors)*
• *Boundary/Barrier to Entry (prevents entry into area)*
• *Breakaway (prevents system overload errors)*
• *Button/Switch Cover (prevents inadvertent activation)*
• *Constraint (limits movement)*
• *Control Limit (prevents exceeding boundaries)*
• *Dead Man Switch (prevents use)*
• *Dissimilar Shape Connectors (prevents incorrect connection)*
• *Dissimilar Size Connectors (prevents incorrect connection)*
• *Exclusion Design (design makes it impossible to make error)*
• *Guards (prevents entry into an area)*
• *Guides (prevents going out of boundary)*
• *In-process Feedback (feedback embedded in task step)*
• *In-process Verification (self-check embedded in task step)*
• *Interlock (prevents action out of sequence)*
• *Keyed Connector (prevents incompatible connections)*
• *Limiters (limits human action)*
• *Load Limiting Fuses (prevents overloads)*
• *Lock-in (prevents premature stopping of process)*
• *Lockout (prevents access)*
• *Machine Guards (prevents entry into area)*
• *Rate Limiter (prevents excess rate)*
• *Safeguards (prevents use, will not operate under unexpected conditions)*
• *Selection Limits (prevent incorrect selection)*
• *Shields (prevents access)*
• *Speed Restrictor/Governor (prevents excess speed)*
• *Timer Lockouts (prevents activation of equipment at wrong time)*
• *Torque Limiter (prevents excess torque)*

If the error can not be prevented, then the error rate should be reduced. Typically, this is done by performing a detailed evaluation of the human-system interface(s) to determine what PSFs are increasing the likelihood of the error and (re)designing the interface to mitigate the effects of those PSFs. If the remaining potential human errors are still significant risk contributors to overall system reliability and safety, then the system must provide a means to automatically (via hardware and software) detect and correct the error or provide the human with the capability to detect and correct the error. Possible mechanisms that allow a human to detect a human error include performing a manual system audit, check, review, inspection, test, or

receiving feedback from the system via an alarm, alert, or warning signal that an error has occurred. A variety of different input devices and controls can be used to correct a human error.

If potential human errors cannot be corrected, then the system should mitigate the negative effects of the error. This can be accomplished by isolating the human error so that it does not cause another failure, designing a failure path to direct and control the effect of the error, or ensuring that the effect of the error is not catastrophic (e.g., providing redundancy to enable a continued function after a function has be lost due to human error).

HRA Methods.

Very few HRA methods provide instructions on how to perform all of the basic steps described in this section. For example, many of the quantification methods do not provide specific instructions for how to complete task analysis and error identification. Consequently, analysts often use a combination of methods to perform HRA. This will be discussed later in the report.

HEA[e] focuses on five steps of this process (problem definition, task analysis, error identification, modeling, and error reduction). HEA, often referred to as qualitative HRA, is a good alternative to quantification when quantification is too time consuming or cost prohibitive. For example, it would be impractical for NASA to perform a PRA on each and every process that is being performed in the Space Shuttle or International Space Station processing areas, and yet, these processes are vulnerable to, and affected by, human error. HEA is useful when human errors have contributed to poor quality products, high injury rates, hardware or property damage, or delays. Additionally, HEA is valuable during the detailed design phase when HFE is evaluating and designing specific human-system interfaces and generating error reduction strategies.

The HRA methods recommended in this document focused on quantification to support PRA. Most, therefore, are relatively weak in qualitative task analysis and context characterization. To remedy this, a strong qualitative method (currently NASA uses HF-PFMEA) can serve as a complementary tool to any of the four methods. This may require some modifications of the method and the HF-PFMEA approach for compatibility and interface consistency (e.g., in terms of error taxonomy, task characterization, and PSFs).

---

[e] HEA is a systematic approach to evaluate human actions, identify potential human error, model human performance, and qualitatively characterize how human error affects a system. HEA is often referred to as a qualitative HRA.

1.2   NASA's Unique PSFs

Human reliability analysis, for the most part, is performed first through the development of a model of the task or tasks in question, determination of areas in which human actions have a role, and then the determination of PSFs within those tasks. NASA missions are unique, both in the tasks that are performed, but also in the factors that can affect human performance. PSFs unique to space missions include the effects from zero gravity, microgravity, and isolation on crew performance. These effects are highly significant for long duration manned missions such as a one-year International Space Station mission, a month-long stay on the lunar surface, or a 30-month journey to Mars and back, but also impact performance on much shorter duration missions, such as the typical duration of an Orbiter mission.

Effects of microgravity and zero gravity can be grouped into three categories: tools and equipment, human health, and behavioral health and performance. In the first category, the analyst must consider how the change in gravity influences the human's use of tools and equipment. Because tools float and the human floats, the procedures to complete tasks include steps such as tethering or securing the person, tool, and/or object to be worked on. Methods to secure the tools and worker can impact the speed and precision of the human performance. In the second category, the analyst must consider how human health is impacted by microgravity and zero gravity. There are numerous health effects experienced in space such as bone loss, muscle alterations and atrophy, neurovestibular adaptation, cardiovascular alterations, altered wound healing, radiation exposure effects, and nutritional changes. As the humans' physiology is affected, their performance is affected. For example, as muscles atrophy, physical tasks requiring force (such as opening a hatch or latch) may be more difficult to do. In the third category, the analyst must consider how the human performance is affected by the behavioral health changes experienced by the human including changes in psychosocial adaptation, neuropsychological changes, sleep and circadian rhythm changes, and changes in the cognitive abilities (e.g., time and space distortions, difficulty concentrating, memory problems, and slowing of intellectual activities). All of the PSFs can impact the potential for human errors that are significantly affected by time, require team cohesion, or concentration for calculation, diagnosis, and problem solving. (For additional detail on NASA's unique PSFs, see Appendix C, and for a description of the types of human actions that may occur in NASA space and ground-based missions, see Appendix B.)

Recommendation: The HRA should specifically consider those PSFs that are unique to space missions and explicitly document the relationship of these PSFs to those PSFs that are included in specific HRA methods.

1.3   HRA Methods Recommended for NASA Use: Capabilities and Characteristics

This section lists the HRA methods recommended for immediate use in NASA PRAs conducted for Exploration Systems Mission Directorate space flight systems trade studies and design analysis. This section provides comparisons of the methods in four areas: screening and qualitative analysis capabilities, quantitative analysis capabilities, model attributes, and resource requirements. Additionally, guidelines for selecting one or a pair of the four methods are provided.

Although there are many HRA methods available for use, NASA constrained the method selection to those method(s) that are immediately available for use and were applicable for analysis of new aerospace designs. The first constraint was imposed on the selection process because the Exploration Systems Mission Directorate is using and will continue to use PRA this year in support of design and trade studies. If an HRA method was not available for immediate use because it required a significant amount of modification for use in NASA PRAs, or was not practiced by/being applied by U.S. citizens (requiring non-nationals' support for its use or significant training of U.S. HRA analysts)[f], the method was eliminated from consideration. It should be noted that this firm constraint eliminated some very good methods that, with modification or U.S. application experience, would be applicable to the analysis of aerospace systems. No HRA method (with the exception of the HF PFMEA) has been specifically designed for an aerospace application. All the HRA methods that provide quantification techniques have roots in, and were designed for, nuclear power plant PRAs. As a result, they should be used with caution regarding their assumptions, application scope, and HEP and data transferability. The limitations of these methods for space mission activities are particularly evident in three areas: 1) coverage of PSFs and task characteristics unique to space missions; 2) applicability of the underlying data, HEP estimates, and PSF weights to space environment; and 3) significant differences in human action time scales between nuclear plant operation and space missions. More detailed information on how the four HRA methods were selected is provided in Section 2, and additional detail on these methods and the other methods evaluated is provided in Section 3.

The following four HRA methods were identified by international HRA experts as most applicable to aerospace applications and appropriate for use in NASA PRAs conducted on new system designs of space flight vehicles: THERP, CREAM, Nuclear Action Reliability Assessment (NARA), and Standardized Plant Analysis Risk HRA Method (SPAR-H). These methods were found to be the most suitable for NASA use in various applications based on the NASA constraints and the criteria provided for the HRA method study (see Section 2).

Recommendation: NASA HRA practitioners should utilize THERP, CREAM, NARA, and SPAR-H for quantitative HRA. Each method offers unique strengths that suit particular NASA HRA needs.

THERP.

THERP is comprehensive HRA methodology that was developed by Swain & Guttmann for the purpose of analyzing human reliability in nuclear power plants. THERP can be used as a screening analysis or a detailed analysis. Unlike many of the quantification methodologies, THERP provides guidance on most steps in the HRA process including task analysis, error representation, and quantification. THERP begins with system familiarization and qualitative assessment (task analysis and error identification). THERP can be

---

[f] Exploration Systems Mission Directorate trade studies will involve analysis of technical information that is restricted to U.S. citizens due to International Traffic in Arms Regulations (ITAR) and Export Administration Regulations (EAR).

used to analyze typical errors of omission and commission. It requires the analyst to construct a HRA ET to model the human error. The analyst then identifies the PSFs that affect human performance. THERP provides a list of three specified PSFs (training level, stress, and experiences) and allows the user to add additional PSFs. THERP allows the analyst to explicitly treat task-error dependencies and human recovery actions. THERP has five levels of dependency that can impact the overall probability of the error. The primary source for THERP HEP estimates is generated by the method. THERP has a large base of experienced analysts in the U.S., and it has been applied to nuclear power plants, off-shore oil drilling, and the NASA Space Shuttle program.

CREAM.

CREAM was developed for general applications and is based on the Contextual Control Model (Hollnagel, 1993). CREAM can be used as a screening analysis or a detailed analysis. CREAM does not provide specific guidance on all steps of the HRA process described earlier. For example, CREAM does not provide guidance on task analysis and error identification; however, it provides an aid (a list of fifteen basic cognitive tasks and their definitions) to support this step. CREAM requires the analyst to perform task decomposition that breaks the task down into subtasks. Each subtask is matched to one of the pre-specified cognitive activities in the list. For each subtask, the activity is further classified as an observation, interpretation, planning, or execution activity. Each of these activities has pre-determined error modes from which the analyst can select (e.g., wrong object observed). CREAM specifies 13 specific error modes which includes both errors of omission and errors of commission. CREAM provides a basic HEP value and upper and lower uncertainty bounds for each generic error. CREAM provides a list of nine PSFs that can be used to modify the HEP. Given that the analyst is selecting tasks, errors, and HEPS from standard tables, the reproducibility is high. CREAM does not provide a specific procedure that explicitly handles task-error dependencies or human recover actions. CREAM has a relatively large U.S. experience base and has been applied to nuclear power plants, off-shore drilling, and the NASA Space Shuttle program.

NARA.

NARA is a refinement of the Human Error Assessment and Reduction Technique (HEART). NARA can be used as a detailed analysis method (and does not provide an explicit method for screening). NARA does not perform all steps described in the basic HRA process; for example, it does not provide guidance on how to perform task analysis or error identification. Instead, the analyst must best match the task being analyzed to one of 14 generic tasks. NARA does not provide specific HEPs for error modes; rather it provides basic HEP values that apply to these generic tasks. The HEPs are adjusted based on a list of 18 PSFS (called Error Producing Conditions (EPCs)). NARA covers both short duration and long duration activities by providing EPCs for longer duration tasks. NARA does not explicitly cover task dependencies or error recovery (these are included in the definition of the generic tasks). NARA has not been applied to any specific domains; however, its parent method HEART has been applied to a number of domains including the chemical industry and weapons manufacturing. Although NARA has not been applied, it was ranked as an acceptable method for NASA use because it is an enhancement of HEART (modifying the grouping of generic tasks and weighting of PSFs) and, most importantly, because of its use of the CORE-DATA[g] human error data base.

SPAR-H.

---

[g] The CORE-DATA human error database provides the foundation of NARA results. The data comes from a rather large HEP data set that was initially used as the source, and subsequently screened to include the least subjective numbers to form the distribution for each of the Generic Tasks. This data has not been independently assessed and is not publicly available.

SPAR-H is a revision of the Accident Sequence Precursor (ASP) HRA screening method. SPAR-H can be used as both a screening method and a detailed analysis method. SPAR-H does not provide specific guidance on how to perform task analysis and error identification, but does tell the analyst to decompose each task to either a diagnosis or an action subtask. The method provides worksheets that allow the analyst to provide complete descriptions of the tasks and capture task data in a standard format. SPAR-H requires the analyst to determine the system activity type (power/operation or low power/shutdown) and then provides HEPs for the four combinations of the error type and system activity type (e.g., one combination is diagnosis and power/operation). The HEP is adjusted based on eight basic PSFs. SPAR-H also adjusts the HEP based on the dependency. A dependency condition table is provided that allows the analyst to evaluate the same crew, time (close or not close in time), information cues (additional information or no cues), and location (same or different location). SPAR-H treats restoration and recovery tasks as a separate event, which is specified and analyzed. SPAR-H has a large U.S. experience base, has been applied to over 70 U.S. nuclear power plants, and has recently been used to help support the Nuclear Regulatory Commission's Office of Reactor Regulation (NRR) Reactor Oversight Process.

Method Selection Criteria.

The final step in assessment of the final four HRA methods for applicability to NASA needs was to determine if the methods met the criteria that defined NASA needs. NASA provided many of these criteria before assessment of methods began, while other criteria were revealed as NASA tasks were assessed and input was received at an HRA workshop used to evaluate the NASA study. Each HRA method was evaluated across 17 attributes which are discussed in detail in Section 2. These attributes provided the foundation for comparison tables that could be used to compare and contrast methods to determine which most closely met the NASA requirements and preferences. The selection criteria and attributes depicted in the HRA comparison tables are presented in Table 2, and the comparison Tables for the top four HRA methods are Tables 3-7.

Table 2. HRA Method Selection Criteria and Discussion.

TABLE	CRITERIA and ATTRIBUTES	DISCUSSION
3	Applicable to existing aerospace designs?	The HRA method must be applicable to existing aerospace designs. Existing designs refer to current NASA systems (e.g., the International Space Station and Space Shuttle) where tasks performed in support or as part of the mission are fully defined and some data exist from actual experience or other sources (e.g., simulators and engineering judgment).
3	Applicable to aerospace system designs in the conceptual phase?	The HRA method must be applicable to aerospace system designs in the early conceptual design phase. Because HRA for NASA is required for systems that have not yet been developed and for tasks that will be performed for the first time, the method must be applicable to aerospace system designs in the early conceptual design phase where the required human activities and specific tasks are not fully defined. This would include systems such as the Crew Exploration Vehicle (CEV), lunar base, or a vehicle that will travel to Mars.
3	Ability to determine human error probability (HEP) quantification?	The HRA method must include procedures for error modeling and result in human error probability (HEP) quantification. To be most adaptable to NASA missions, the method (or set of methods selected) must include procedures for error identification, error modeling, and result in human error probability (HEP) quantification, rather than only a qualitative assessment of error likely scenarios. This requirement is prompted by the increased needs within NASA to provide quantified measures of increased safety for use in selection between plans and designs.
3	Screening capability?	The method should provide a screening capability. Some HRA methods require significant details, effort, and time to perform an analysis. In some cases, applying such extensive analyses may not be feasible (e.g., for new designs) or necessary (unimportant events in PRA context). Because many NASA tasks will never have been previously performed, screening for new designs or tasks in which the entire process is currently unknown is necessary.

TABLE	CRITERIA and ATTRIBUTES	DISCUSSION
3	Ability to update the model?	The HRA method must allow for updating the model. As the system design is refined and the human activities are specified more clearly during the system's development phase, it must be possible to update the model to provide more detailed analysis.
3	Provides guidance on task decomposition?	Task decomposition is the key process by which an HRA method breaks down the human activities of interest into subtasks that match the method's "units of analysis" or "basic tasks." Some HRA methods require and provide guidelines for such task decomposition processes, others do not. The subtasks defined by a method must adapt or match the task types in NASA missions.
3	Provide a flexible PSF list?	Most of HRA methods have a fixed set of PSFs for predictive analysis (i.e., calculating HEPs) or retrospective analysis (i.e., identifying the root causes). Some methods allow the analyst to specify the set of PSFs based on the task of analysis. NASA missions are unique in the PSFs that affect them. Therefore, the HRA methods must be able to adapt to these unique PSFs, either through adaptation or expansion of the set of PSFs. The method must be able to account for performance-shaping factors that are specific to the NASA environment and space missions (including physical and cognitive adaptations to microgravity and zero-gravity).
3	Broad coverage of error sources considered?	Does the method provide estimates for specific error modes or for broadly defined errors (e.g., omission, failure to respond in time, failure to complete a task)? This is especially critical for NASA because data must reflect the types of errors of most interest to NASA. *(Note: Error sources may be associated with a variety of tasks. Three overarching categories are generally considered:* *Ergonomic factors refer the design of controls and interfaces, lighting, system design, physical workload, and/or physical work;* *Cognitive factors refer to decision making, mental*

TABLE	CRITERIA and ATTRIBUTES	DISCUSSION
		workload, mental fatigue; *Organizational factors refer to the impact of management on reliability and performance of human activities, work processes, task organization, safety culture, and team. Often this may include communications.*)
4	Procedures for error identification, error modeling, and HEP quantification for use in PRA models, and other predictive and retrospective analyses?	Does the method provided specific procedure and data used to calculate HEPs? Because NASA users will not primarily be HRA experts, the calculation procedure should be straight forward or well documented.
4	Address errors of omission <u>and</u> errors of commission?	The HRA method must address both errors of omission and errors of commission. NASA requires HRA for both nominal and emergency operations. Nominal operations refer to the human activities performed in situations where there are no system failures. Emergency operations refer to human activities performed with at least the perception of partial or full system failure. This might require the use of emergency operating procedures and could change the character of the mission or lead to a mission abort or a system escape.
4	Provides guidance on treatment of error recovery?	The HRA method should include explicit treatment of error recovery.
4	Provide explicit treatment of task/error dependencies?	Are dependencies (of multiple tasks or actions) and recovery (from error) explicitly modeled in calculating HEPs? Because analysts may have little experience, it is critical that the HRA method provide a way to identify and address task dependency and recovery within HEP estimates.
4	Provide a method for uncertainty bound estimation?	It is preferable that the method provides instructions to assess HEP uncertainty bounds.

TABLE	CRITERIA and ATTRIBUTES	DISCUSSION
5	Validated?	HRA methods have "face validity." NASA requires validations for "error identification" and "HEP estimation." Error identification refers to whether specific instructions are provided by the method to identify the risk-related human tasks to be modeled inside the PRA model. In current practice, error identification is typically performed when developing the PRA model.
5	Reliable and reproducible?	The error analysis, identification, and error probabilities in the HRA method should have good reliability and reproducibility. For a method to be most useful for NASA needs, different analysts should obtain same results by applying the method to the same problem, independently for a second time.
5	Low sensitivity?	Sensitivity or the amount of change in results (e.g., HEPs) as a function of changes in the input variables was critical for NASA. In order for the methods to be most transparent to the users, and for reliability and reproducibility to be maximized, it was critical that the HRA method not yield large changes in the HEP calculated when only small changes were made in the PSFs.
5	Use data from a variety of sources?	The HRA method should adapt to data from a wide variety of sources, including simulators, human performance studies, and potentially, expert opinion. Changes in the minor details of a context should not have a large effect on the error probability computed (i.e., low sensitivity). Data sources and assumptions should be transparent.
5	Broad experience base?	Experience base referred to the areas and industries in which the HRA method has been applied. This provided an indication of the degree to which the method can be applied to different NASA mission areas. For NASA, a U.S. experience base is required for methods that will be applied immediately because NASA technology can only be evaluated by U.S. citizens for export control reasons.
6	Ability to be operated by user who is not expert in HRA?	In HRA, people with three levels of HRA-related knowledge regularly perform analyses. These levels of knowledge are: • HRA Specialist: Many years of experience, capable of making judgement on key aspects of the analysis when the method does not provide

TABLE	CRITERIA and ATTRIBUTES	DISCUSSION
		specific instrucations. • HRA Analyst: About one year of experience in HRA practice. • PRA Analyst: Not necessarily familiar with HRA methods, but capable of performing general engineering analysis by following instructions. With respect to NASA HRA analysis, most persons performing analysis would not be HRA experts, but they may have extensive experience with PRA. Therefore, a lower required level of knowledge, that of someone who is informed about factors affecting human performance but not highly educated in the area, is expected. For the purpose of the assessment of HRA methods, this study assumed that analysts would most likely be PRA analysts.
6	Minimal expenditure of resources?	Another consideration in determining recommended HRA methods for NASA use is the required resources (including time to perform and software tools) of a method. It is expected that analysts performing HRA will not have significant amounts of time because of the need to meet deadlines. HRA methods can be divided into 3 required levels of effort: • Low: Requires a few hours of effort. • Medium: Requires a few days, up to one week of effort. • High: Requires a few weeks to a few months. In addition, HRA methods can further be differentiated by whether they require certain computer programs or software to perform the calculation. NASA stated a preference for HRA methods that require less level of effort and those which are self-contained or only use commonly available software programs.
7	Available for immediate use with reasonable cost?	The HRA method must be available for immediate use. HRA is currently being performed for NASA missions. Therefore, to the extent possible that relevant human error data can be obtained to inform the analysis, the method(s) selected should be available for immediate

TABLE	CRITERIA and ATTRIBUTES	DISCUSSION
		use by NASA as part of the CEV design process. A critical aspect of this is that the method should be usable by an analyst who is not expert in HRA, and the method should not require significant amounts of training to yield reliable, reproducible results between analysts. Most HRA methods are publicly available and free for use. However, some methods require certain tools for analysis. In some situations, additional data may be required to perform an analysis. Because NASA needs to immediately implement use of the HRA method(s), there was a preference for publicly available, free for use methods.

Table 3 . Screening and qualitative analysis capabilities.

Method	Screening[1]	Task Decomposition[2]	PSF List	Causal Model[3]	Coverage [4] 1: Ergonomics 2: Cognitive 3: Organizational
THERP	Yes	Screening, Diagnosis, and Action	3+[5]	Single layer[6]	1 and 3
CREAM	Yes	15 task types	9 for quantification; many for root causes	Multi-layer	1, 2, and 3
NARA	No	14 generic tasks	18	Single layer	1, 2, and 3
SPAR-H	No	Diagnosis, Action	8 for quantification; Many for root causes	Single layer	1, 2, and 3

[1] For preliminary screening and identification of PRA-significant human activities.
[2] Decomposing identified PRA-significant human activities into subtasks for more specific error analysis.
[3] Whether a causal chain is provided for the analyst to identify the "root causes" from the immediate causes.
[4] Method provides "relatively detailed" instructions for assessing the PSFs or factors' effect on the specific problem scope.
- Ergonomics – design of controls systems, machine aspects, lighting, system design, physical workload, physical fatigue; i.e. anything physical or physiological
- Cognitive – decision making, mental workload, cognitive fatigue, i.e. anything cognitive
- Organizational – design of tasks, management impact on reliability of human, work processes, task organizations / procedural alignment, safety culture, team, communications
[5] THERP allows user to add PSFs other than the three explicitly specified PSFs (i.e., Training Levels, Stress, and Experiences) to adjust the HEP values, however, it does not provide guidance on the effect of these PSFs to the analyst.
[6] Single means that only a list of PSFs is provided. Multi-layer means that an explicit causal chain is provided, considering interdependencies of causal factors. Such causal models can be used in retrospective analysis and root cause identification.

Table 4 . Quantitative analysis capabilities.

Method	Primary Source for HEP Estimates		Analysis Approach		HEPs for Specific Error Modes	Explicit Treatment of		Uncertainty Bounds Estimation[1]
	Number provided by method	Number produced by analyst	Manual	Software Aided		Task /Error Dependencies	Recovery (includes actions with feedback)	
THERP	√		√		Specifies typical Omission/Commission errors	√	√	√
CREAM	√		√		13 error modes			√
NARA	√		√		None specified			
SPAR-H	√		√		Diagnosis & Action	√	√	√

[1] None of the methods that provide uncertainty bound capability make a distinction between aleatory or epistemic nature of the uncertainties.

Table 5 . Other model attributes.

Method	Error Identification		HEP Estimation			Experience Base**
	Reproducibility[1]	Validation[2]	Validation[2]	Reproducibility[1]	Sensitivity[3]	
THERP	Does not do error identification	Does not do error identification	1 known validation; Widely referred	Medium	Low	Widely used and referenced (nuclear, off-shore oil drilling, and NASA)
CREAM	Does not do error identification	Does not do error identification	None	High	Low	Several nuclear, off-shore oil drilling, and NASA applications
NARA	Does not do error identification	Does not do error identification	An informal verification against HEART in a NPP PSA setting	Medium	High	Not applied
SPAR-H	Does not do error identification	Does not do error identification	Base Rate comparison[4]	High	Low	Many U.S. nuclear applications

** Refers to the number and variety of applications.
[1] Reproducibility refers to the level of consistency in results produced by different individuals for the same task.
[2] Refers to formal independent validation of the results.
[3] Refers to the amount of change in numerical results when input parameters are changed.
[4] Performed by the method developers.

Table 6 . Resource requirements.

Method	Knowledge Level			Tool		Level of Efforts Required For Error Identification	Level of Effort Required For Calculating HEPs[1]
	HRA specialist (knowledge based)	HRA analyst (about one year of experience)	PRA analyst (skilled based)	Manual	Computer Code Needed		
THERP		√		√		*	Medium
CREAM		√		√		*	Medium
NARA		√		√		*	Low
SPAR-H			√	√		*	Low

* Method does not include specific guidelines for error identification.

(1) Low = Look up in office, done in minutes to a day
Medium = Up to 2 weeks
High = Resource or time intensive (includes expert elicitation, more than 2 weeks of effort)

Table 7 . Cost and availability of method, tools, and data.

Method	Method/Software	Parameter Values[1]	Raw Data[2]
THERP	Free	Free	Not available
CREAM	Free	Free	Not available
NARA	Free	Free	Not available publicly
SPAR-H	Free	Free	Not available

[1] Refers to all parameter values needed by the method (e.g., median time, reference HEP).
[2] Refers to original "raw" data used in developing the HEP estimates.

1.4 Guidelines For Selecting a HRA Method or Pair of Methods

Compared to other HRA methods considered, the four HRA methods recommended are: (1) relatively easy to use; (2) provide an explicit procedure for HEP estimation; and (3) do not require extensive information from task analysis (when used in their respective screening modes). There are significant differences between the selected methods, which may make each better suited to a different type of analyst or question to be analyzed.

CREAM and SPAR-H use broader definitions of tasks making them easier to apply to a wider spectrum of space activities. CREAM's set of generic tasks is defined based on human information processing. The set of generic tasks provided in CREAM covers a range of activities while remaining at a level of specificity that is easily adaptable to NASA mission tasks. CREAM's generic tasks may assist an analyst when a well understood task is to be performed in a new setting and, therefore, is a useful tool for screening. The provision of activities could allow the analyst to focus on the incorporation of PSFs that are unique to the NASA mission.

The task types used in SPAR-H, while also based on human information processing, characterize tasks in a much simpler way, by dividing tasks by emphasis on cognitive workload or physical workload (or potentially a combined rate). SPAR-H characterizes tasks as Diagnosis or Action, which can be generalized to any task, but is not as specific as tasks defined in CREAM or THERP. Therefore, SPAR-H is not as effective for screening analyses, but can be extremely powerful to assess the effects of performance-shaping factors on a task. Because NASA may wish to assess the potential for error in new mission tasks and activities (activities that may not have been performed 'for real'), the ability to estimate probabilities without having to specify exact tasks may be useful. In addition, more so than other methods, SPAR-H does not require that the analyst be familiar with human performance, just the task. Through its standardized process, it assures reliability between analysts who have the same understanding of the task in question. Its results are easily reproduced, because it provides standard worksheets that ensure the analyst uses the same process each time.

NARA combines context characteristics and human tasks, and (like CREAM) defines a set of "generic tasks," also largely based on the human information processing model. These tasks can be generalized to match a subset of space activities. Among the four methods, NARA has the most extensive use of real data to support its HEP quantification. One of the appealing features of NARA is its use of actual human error data (i.e., CORE-DATA) in most cases. This contrasts with the other methods that are either totally or partially expert judgment based.

Finally THERP, when compared to the other three methods, is highly granular in its task decomposition (e.g., Opening a Valve). Treatment of human performance is much like treatment of the mechanical system, with significant emphasis on actions and much less emphasis (especially when compared to the other selected HRA methods) on cognitive aspects of performance. THERP relies on task granularity and a small number of PSFs. THERP is effective when the task is well understood, but the potential cognitive impacts on performance are not understood. Like CREAM, THERP can assist the analyst in the identification of potential human errors during screening analyses. The tasks used in THERP can be and have been generalized for use in non-nuclear power applications. (The THERP screening method has been used in the Space Shuttle PRA for about 80% of roughly two hundred human basic events in the PRA model.)

The four methods selected do not individually necessarily meet all NASA selection criteria and requirements described in Section 2. They are simply identified as better candidates relative to other existing methods. Therefore, the reviewers of this report strongly recommended pairing the methods to allow provision of data and coverage of all potential error types, as shown in Table 8. Four possible combinations are recommended and listed in order of preference by the reviewers. Each pair's relative ability to provide data support and coverage were rated, and the number of stars indicates the relative ranking.

Table 8. Ranking of the paired method.

Suggested combinations	Data Support	Coverage
THERP + CREAM	**	***
THERP + NARA	***	**
CREAM + NARA	***	***
CREAM + SPAR-H	**	***

Once the Exploration Systems Mission Directorate has selected a method or pair of methods for use in a program, the program should develop a set of guidelines regarding their effective use. Such guidelines should include clear statements on which domain(s) of human activities will be covered, the level of detail expected in the analysis, all assumptions, if and how additional microgravity-related PSFs should be included, if and how the model parameters and HEPs generated by each method should be modified for space applications, and how future human error data from simulations, usability studies, and actual use will be used to update the analysis. Additionally, the program should have a structured method to collect, process, catalogue, and store data used during the HRA. The method should provide the HRA users and Exploration Systems Mission Directorate program management with the capability to identify and locate data sources (e.g., human error probabilities, PSF impact studies) for the life of the program.

The HRA methods considered in this document focused on quantification. Most, therefore, are relatively weak in qualitative task analysis and context characterization. To remedy this, a strong qualitative method (currently NASA uses HF-PFMEA) can serve as a complementary tool to any of the four methods. This may require some modifications of the methods and the HF-PFMEA approach for compatibility and interface consistency (e.g., in terms of error taxonomy, task characterization and PSFs).

1.5   Special PRA Considerations

Section 1.1 provided a high-level overview of HRA with respect to incorporating human errors into a PRA model. PRA requirements for HRA may vary considerably across applications. The PRA analyst who incorporates HRA into PRA should therefore closely follow guidance in the *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners* (NASA, 2002), especially Chapter 9 on HRA. It should, however, be noted that Chapter 9 predates the discussion found in this document. In particular, the Chapter 9 focuses on THERP, without a discussion of the CREAM, NARA, and SPAR-H methods advocated here.

Special considerations may dominate the selection of a particular method for use in the PRA. Table ZZ below provides a summary of considerations for incorporating each method into an overall PRA model. THERP, CREAM, NARA, and SPAR-H are discussed in terms of four common PRA functions identified in this document:

- *Human Error Identification* refers to the method's capabilities regarding providing task analysis guidelines and support for qualitative human error analysis.

- *Screening* refers to the method's capabilities to assist with risk significance determination prior to conducting a detailed HRA.

- *HEP Generation* refers to the method's approach to estimating HEPs or providing quantitative risk information.

- *Uncertainty Calculation* refers to the method's approach at providing uncertainty bounds suitable for incorporation in a standard PRA model.

Further detailed discussion about each method is provided in Chapter 3 of this document.

Table 1D. Comparison of HRA Methods in Terms of PRA Considerations.

Method	Human Error Identification	Screening	HEP Generation	Uncertainty Calculation
THERP (Also see Section 3.1)	THERP includes two phases of error identification: familiarization and qualitative assessment. THERP is largely scenario or task based. The scenarios or tasks do not directly map to all NASA activities. Any mapping of THERP tasks to NASA activities should be thoroughly documented and justified. THERP uses an HRA event tree, which differs from the standard PRA event tree and may not seamlessly integrate into a PRA model. It is possible, in practice, to use PRA event trees for most THERP modeling.	THERP provides screening for diagnosis and rule-based action. Tables are provided to produce screening values based on available time to diagnose and on errors per critical step for actions. The screening criteria may not fit the full realm of NASA relevant activities.	Table 20 in THERP provides common tasks and subitems corresponding to the level of performance. For each subitem, a corresponding HEP and Error Factor (EF) are provided. Dependency between multiple tasks is corrected using a simple dependency calculation. The THERP tasks and subitems may not generalize to NASA activities, especially cognitive activities, which were mostly documented after the creation of THERP. Any mapping of THERP Table 20 items to NASA activities should be thoroughly documented and justified. Historically, THERP HEPs have tended to be more conservative than the HEPS generated through newer HRA methods.	The Error Factor (ratio of 95th to 5th percentile HEP estimates) is provided for all HEP calculations.

CREAM (Also see Section 3.4)	In CREAM, tasks are decomposed into subtasks, which are mapped to one of 15 pre-specified cognitive activity types. Because CREAM emphasizes the cognitive activity of the individual—not the work activity—all NASA activities can be mapped to the cognitive activity levels.	CREAM includes a "basic method" and an "extended method" of quantification. The basic method is designed for screening. By assessing the improvement or reduction in reliability according to nine screening PSFs, CREAM provides a simple table corresponding to four resultant screening values. The CREAM screening PSFs may not map to all NASA relevant activities. Any mapping should be thoroughly documented and justified.	For each subtask, the cognitive activity type is identified, along with the corresponding human function (Observation, Interpretation, Planning, or Execution). This produces a basic HEP, which is adjusted according to the nine CREAM PSFs. While the cognitive activity levels may be seen as universal and highly applicable to NASA applications, the PSF list is more restricted and may not map to all NASA relevant activities. Any mapping should be thoroughly documented and justified. No correction is provided for dependency between subtasks in CREAM.	A table is provided for uncertainty according to human function type. This table provides the lower and upper bound (5th and 95th percentile) values.
NARA (Also see Section 3.6)	NARA does not provide an explicit process for error identification. However, 14 generic tasks are provided. The limited number of generic tasks may not encompass all NASA relevant activities. The mapping of NASA activities to NARA generic tasks should be thoroughly documented and justified.	NARA does not provide an explicit process for screening. In practice, since NARA is a simple-to-apply method, it is possible to multiply the generic task values by the PSFs to quickly arrive at screening values.	Generic tasks are multiplied by PSFs (called Error Producing Conditions in NARA). The generic tasks and error producing conditions may not encompass all NASA relevant activities. The mapping of NASA activities to the generic tasks and error producing conditions in NARA should be thoroughly documented and justified.	NARA provides uncertainty information on a limited range of generic tasks and error producing conditions. This lack of consistent uncertainty information may limit the seamless integration of NARA into some PRA models.

SPAR-H (Also see Section 3.9)	SPAR-H does not provide an explicit process for error identification. SPAR-H analysts have often followed THERP.	SPAR-H does not provide an explicit process for screening. In practice, since SPAR-H is a simple-to-apply method, it is possible to multiply PSF weights by nominal HEP values quickly for screening purposes.	The nominal HEP for diagnosis (1E-2) or action (1E-3) is multiplied by assignment levels for each of eight PSFs. The PSF worksheets distinguish between at power and low power/shutdown, a distinction that may not apply to NASA. However, the PSFs are broad and encompass most NASA relevant activities. Any mapping of NASA's unique PSFs to SPAR-H should be thoroughly documented and justified. A simple correction factor (borrowed from THERP) is applied for dependency between multiple human events.	SPAR-H does not provide uncertainty bounds; instead, it uses the constrained non-informative prior, which allows the approximation of a beta distribution based on a single central point, the calculated HEP. This approach, in practice, requires PRA software capable of modeling the constrained non-informed prior when modeling uncertainty in SPAR-H.

1.6　Summary of HRA Recommendations

Chapter 1 has provided an overview of HRA processes, a list of recommended practices, and a discussion of considerations for incorporating HRA into PRA. We repeat the nine recommended practices below for the human reliability analyst and risk manager to use as a checklist when carrying out an HRA. While these recommendations are not meant as an exhaustive list of factors to consider in performing a high quality HRA, these recommendations provide anchor points for NASA's goal to adopt a standard approach to evaluating and managing human performance related risks. These recommendations should be employed as part of any NASA PRA and HRA process. Further, in broadening the application of HRA throughout NASA, these recommendations should guide an overall human error management strategy that serves to mitigate sources of human error from design through mission execution.

Recommendations

1. Use HRA throughout the life cycle of the system, beginning early in the system design process.

2. HFE design efforts should pay special attention to the human interactions and accident scenarios identified by the HRA as critical for overall system reliability and safety and generate solutions that mitigate risk.

3. HRA must be an integral part of the PRA development, from its earliest stage, in order to identify, analyze, and, if necessary, quantify, the points where humans interact with the hardware and software in each scenario.

4. When NASA is conducting a HRA in support of a PRA, it is recommended that the scope include the assessment of human errors (both errors of omission and errors of commission) and take into account pre-initiating actions, initiating actions, and post-initiating actions.

5. When a task has a significant cognitive aspect, such as human decision-making tasks during inspections, space vehicle launches, piloting, and manual docking, the HRA should include the evaluation of errors for both cognitive responses and physical actions.

6. The analyst should carefully weigh the value of a screening analysis vs. a detailed HRA. Where time and resources allow, a screening analysis allows the analyst to establish risk significant events and conduct a detailed HRA only on those events.

7. Risk significant human errors should be included in the PRA master logic diagram, event tree, and fault tree. HEPs and uncertainty information should be included according to the conventions of the HRA method in use.

8. The HRA should specifically consider those PSFs that are unique to space missions and explicitly document the relationship of these PSFs to those PSFs that are included in specific HRA methods.

9. NASA HRA practitioners should utilize THERP, CREAM, NARA, and SPAR-H for quantitative HRA. Each method offers unique strengths that suit particular NASA HRA needs.

2. HRA Methods Study

2.1 Purpose

In order for NASA to successfully employ HRA in PRA to support risk-based decision making and design trades, NASA must adopt a standard approach to evaluating and managing human performance related risks. There have been about fifty HRA methods published in terms of producing HEPs. Although there are a large number of HRA methods available for use, the majority were developed for use in the nuclear power industry and, in many cases, are used to evaluate existing nuclear power plants with well developed procedures. Some of these methods have been modified and used in petrochemical, automotive, and aviation industries and applied to newer system designs. Each of the methods varies slightly in the way it decomposes tasks, identifies errors, calculates HEPs, treats dependency and recovery, and represents uncertainty. As HRA methodology has evolved since the early 1960s, the approach has changed from attempts to create human-error databases, parallel to those created for hardware components, to use of expert judgment techniques, back to a renewed variation of data-base uses. Additionally, early methods evaluated errors, later methods evaluated tasks rather than behavioral elements, with some methods focusing on human actions, and others relying on information processing theories and focusing on cognitive behaviors. With all these variations, and no single quantitative method created for aerospace applications or designed to include the unique PSFs found in microgravity and zero gravity applications, a debate ensued amongst PRA analysts and HRA analysts concerning the selection of the best method for NASA's use on new system designs, such as the CEV.

To that end, the NASA Office of Safety and Mission Assurance conducted the HRA methods study to evaluate the suitability of existing and formally applied HRA methods and to recommend adoption of a set of methods for use in various aerospace applications. This study focused on the evaluation and selection of HRA methods that can support PRA being conducted on future systems, such as the CEV, lunar lander, and lunar base. The goal was to identify methods that could be used during early concept design, when little system information was available, and yet were flexible enough to allow growth of the HRA as more information became available when the system progressed through later phases of the system life cycle (fabrication, test, and use). Although there are many HRA methods available for use, NASA constrained the method selection to those method(s) that are immediately available for use because the Exploration Systems Mission Directorate is using and will continue to use PRA in support of design and trade studies conducted this year and in future years to support the development of the CEV as it prepares for its first operational flight in 2012. Consequently, the HRA method selected must be ready for immediate use to support Exploration Systems Mission Directorate PRAs. (Although this paper evaluates methods for their applicability to existing NASA systems, the intent is not to imply that treatments of HRA in existing NASA PRAs are inadequate or that they will be updated using these methods selected in this study.)

2.2   Method

The first step in assessment of HRA methods for applicability to NASA needs was to determine those criteria that defined NASA needs. A set of criteria was developed before the assessment of methods began, and other criteria were revealed and added as the NASA tasks were assessed and the study progressed. The second step was to select a set of candidate HRA methods for evaluation. HRA experts from NASA, University of Maryland, and Idaho National Laboratories evaluated existing HRA methods and selected a set of candidate methods for consideration based on the preliminary criteria. During the third step, a literature review was completed, the candidate methods were researched, and descriptions of the methods were generated. Fourth, the criteria were refined and a list of desired attributes was developed. Fifth, the descriptions of each method were updated with information on each attribute and the methods were compared. A draft report of this study was generated, providing an overview of the study purpose, a description of NASA's unique tasks and PSFs, NASA's criteria, detailed descriptions of the candidate methods, how they met or did not met the criteria, and method comparison tables. The sixth step and seventh steps were a peer review and the formulation of recommendations by HRA experts. To implement these steps, the report was circulated to a group of internationally recognized HRA experts (Appendix D) and practitioners for review, and a HRA workshop was held in January 2006 at Kennedy Space Center, Florida. The experts were asked to evaluate the methods and verify the set of candidate methods was appropriate, the comparison attributes were sufficient, that each method was adequately represented in the description, and the comparison tables accurately reflected the methods' capabilities and limitations. The experts were invited to provide comments and make corrections to the report. Those experts that participated in the workshop were also asked to help evaluate the selected methods and identify the best candidates for NASA applications. The feedback from the experts at the workshop, and via written comment before and after the workshop, was incorporated into the final draft of this report.

Summary of Steps in the HRA Methods Study:

(1)   Initial selection of criteria

(2)   Initial selection of a set of candidate methods

(3)   Literature review and description of history and manner of use of the selected methods

(4)   Refinement of selection criteria and HRA method attributes to be evaluated

(5)   Comparison of all candidate methods using criteria

(6)   Peer review of the study, including evaluation of methods selected, representation of methods and comparison of methods

(7)   Formulation of the recommendations to NASA regarding the most appropriate HRA methods for varying problem types

(8)   Integration of feedback and completion of final report

2.3   Criteria Used in Evaluation

The criteria were developed with the ultimate objective of assessing and comparing HRA methods and their suitability for use in risk and reliability studies of various NASA space systems and missions, with special emphasis on Exploration Systems that will be designed to travel to the moon and Mars. To accomplish this, NASA developed a set of loosely defined criteria to be used to evaluate the existing 50+ HRA methods and down select to a set of methods for detailed evaluation and consideration. The initial criteria required that the method 1) be flexible enough to be applicable to aerospace systems allowing it use on CEV, lunar landers, and lunar bases, 2) be applicable to early system designs where little information is known about the human tasks and potential errors allowing use on CEV trade studies, 3) provide the capability to perform HEP quantification in support of PRAs, 4) provide the capability to account for NASA's unique PSFs (described in detail in Appendix C), 5) be published and available to U.S. citizens allowing easy and immediate access for NASA analysts, and 6) be applied to an existing system where the application is published and available allowing NASA access to an example use. (This was also intended to eliminate from consideration any methods that are early in development and have no real world applications). Other characteristics that were desirable but were not met by all candidate HRA methods included the capability of the method to perform screening, evaluate both errors of omission and commission, provide a method to evaluate cognitive errors, have a data source that is applicable to NASA, and provide guidance on how to handle dependency, recovery errors, and model uncertainty. From these criteria, 12 methods were selected for further evaluation.

At the HRA workshop, after long discussions concerning HRA method availability for use, two other significant criteria evolved. First, the method must have a U.S. experience base, allowing NASA to immediately use the method. If the expertise only resides in other countries, and in most cases it will take a substantial amount of time (up to one year) to train US HRA analysts to reliably perform the method, the method is really not available for immediate use on Exploration Systems Mission Directorate programs. This is because most NASA technology that is being evaluated in PRAs has some type of export control restrictions, and only U.S. nationals are permitted to participate in the hardware review necessary for the HRA. The second criteria, although a somewhat less important deciding factor, was the knowledge or expertise level required by the analyst. As NASA considered its resource base of PRA and HRA experts, it recognized that the majority of NASA and NASA contract professionals are PRA experts, with very few HRA analysts having background and knowledge in detailed task analysis, error identification in field operations, cognition, and HFE. Consequently, some methods would be more difficult to employ, taking a longer time to train professionals to become proficient on the method. Although outside experts could be used to perform the analysis, NASA and/or NASA contractors would be required to review and approve the analysis, still requiring some level of knowledge and expertise about the correct application of the methods.

After the literature review was initiated, a list of attributes was developed so that each of the methods could be described and compared in a variety of different categories. The list was refined with input from the HRA workshop. In the end, 17 attributes were used to compare the methods, and they are:

1. Developmental Context
2. Screening
3. Task Decomposition
4. PSF List and Causal Model
5. Coverage
6. HEP Calculation Procedure
7. Error-Specific HEPs
8. Task Dependencies and Recovery
9. HEP Uncertainty Bounds
10. Level of Knowledge Required

11. Validation
12. Reproducibility
13. Sensitivity
14. Experience Base
15. Resource Requirements
16. Cost and Availability
17. Suitability for NASA Applications

A description of each attributes is provided below. The attributes could be considered selection criteria, although no one method met all NASA selection criteria and requirements described in this section. Consequently, the criteria simply allowed comparison of the HRA methods to determine which are better candidates.

2.3.1   Description of Each Attribute Used for the HRA Method Comparison

1. **Developmental Context**: The history and domain in which the method was developed. Such context often has a significant impact on the method's structure, assumptions, focuses, format, results, and extent of applicability to other domains and contexts.

2. **Screening**: Some HRA methods require significant details, effort, and time to perform an analysis. In some cases applying such extensive analyses may not be feasible (e.g., for new designs) or necessary (unimportant events in PRA context). To address this some methods also provide a screening analysis procedure which requires less information and effort.

3. **Task Decomposition**: Task decomposition is a key process by which an HRA method breaks down the human activities of interest into a list of subtasks that match the method's "units of analysis" or "basic tasks." Some HRA methods require and provide guidelines for such task decomposition process, others do not.

4. **PSF List and Causal Model**: Most of HRA methods have a fixed set of PSFs for predictive analysis (i.e., calculating HEPs) or retrospective analysis (i.e., identifying the root causes). Some methods allow the analyst to specify the set of PSFs based on the task of analysis. For retrospective analysis, a causal model is necessary for the analyst to identify the root causes. For the methods that provide only a set of PSFs without explicitly specifying their dependencies, these PSFs are considered as proximate causes. These methods are not credited for having a causal model. Only the methods that provide multiple layers and explicit PSFs dependencies are credited for having a causal model.

5. **Coverage**: This refers to the aspects of tasks and error sources covered by the method. Three categories are included:

   a.   Ergonomics: refers to design of controls systems, machine aspects, lighting, system design, physical workload, physical fatigue, i.e. anything physical or physiological

   b.   Cognitive: refers to decision making, mental workload, cognitive fatigue; i.e. anything cognitive

   c.   Organizational: refers to design of tasks, management impact on reliability of human activities, work processes, task organization, safety culture, team, possibly includes communications

6. **HEP Calculation Procedure**: Refers to the specifics of the procedure and data used by the method in calculating HEPs.

7. **Error-Specific HEPs**: Whether the method provides estimates for specific error modes or for broadly defined errors (e.g., omission, failure to respond in time, failure to complete a task).

8. **Task Dependencies and Recovery**: Whether dependencies (of multiple tasks or actions) and recovery (from error) are explicitly modeled in calculating HEPs.

9. **HEP Uncertainty Bounds**: Whether the method provides instructions to assess HEP uncertainty bounds.

10. **Level of Knowledge Required**: three levels of HRA-related knowledge are identified:

    a. HRA Specialist: Many years of experience, capable of making judgement on key aspects of the analysis when the method does not provide specific instrucations.

    b. HRA Analyst: About one year of experience in HRA practice.

    c. PRA Analyst: Not necessarily familiar with HRA methods, but capable of performing general engineering analysis by following instructions.

11. **Validation**: HRA methods normally pass the simple test of "face validity." In this review an attempt is made to further determine a method's level of empirical validity (the extent to which models, assumptions, procedures, and results have been validated experimentally). Validations for "error identification" and "HEP estimation" are discussed separately. (Error identification refers to whether specific instructions are provided by the method to identify the risk-related human tasks to be modeled inside the PRA model. In current practice, error identification is typically performed in developing the PRA model.)

12. **Reproducibility**: Whether different analysts would obtain same results by applying the method to the same problem. Reproducibility for "error identification" and "HEP estimation" is discussed separately.

13. **Sensitivity**: The amount of change in results (e.g., HEPs) as a function of changes in the input variables.

14. **Experience Base**: The areas and industries in which the HRA method has been applied.

15. **Resource Requirements**:

    a. Level of Effort to Perform a Typical Analysis: Three levels are identified:

        (1) Low: Requires a few hours of effort

        (2) Medium: Requires a few days, up to one week of effort

        (3) High: Requires few weeks to a few months

    b. Required tools (e.g., whether certain computer programs are required to perform the calculation)

16. **Cost and Availability**: Most HRA methods are publicly available and free for use. However, some methods require certain tools for analysis. In some situations, additional data may be required to perform an analysis. Acquiring such tools and data could involve additional cost.

17. **Suitability for NASA Use**:

   (1)   The method should include procedures for error identification, error modeling, and HEP quantification for use in Probabilistic Risk Assessment (PRA) models and other predictive and retrospective analyses.

   (2)   The method must be applicable to aerospace system designs in the early conceptual design phase.

   (3)   The method must be capable of being updated to provide a more detailed analysis as the system design is refined and the human activities are specified more clearly during the system's development phase.

   (4)   The method should be applicable to existing aerospace designs.

   (5)   The method must be able to account for performance shaping factors that are specific to the NASA environment and space missions (including physical and cognitive adaptations to microgravity and zero-gravity).

   (6)   The method must be applicable to both nominal and emergency operations.

   (7)   The method must be available for immediate use by NASA as part of the CEV design process. The CEV System Requirements Review (SRR) will begin in the spring of 2006. HRA methods will be used shortly thereafter to make system design trades.

3. HRA Methods Selected for Review and Comparison

Initially, 50+ HRA methods were evaluated and twelve methods were selected for consideration and further review. During the HRA workshop, at the advisement of the experts, one method was dropped from consideration because it lacked real-world experience, and others were added. In the end, 14 methods were selected to be evaluated in this study. It should be noted that the HF PFMEA method is included in this list of 14 because it is discussed in the report as a good qualitative method; however, it does not provide a method to quantify HEPs.

This section summarizes the results of the study for the 14 HRA methods that NASA selected for evaluation.

1. Technique for Human Error Rate Prediction (**THERP**, Swain & Guttman, 1983).

2. Accident Sequence Evaluation Program (**ASEP**, Swain, 1987).

3. Success Likelihood Index Methodology (**SLIM**, Embrey, Humphreys, Rosa, Kirwan & Rea, 1984).

4. Cognitive Reliability and Error Analysis Method (**CREAM**, Hollnagel, 1998).

5. Human Error Assessment and Reduction Technique (**HEART**, Williams, 1986; 1988).

6. Nuclear Action Reliability Assessment (**NARA**, Kirwan, Gibson, Kennedy, Edmunds, Cooksley & Umbers, 2004).

7. A Technique for Human Event Analysis (**ATHEANA**, Barriere, Bley, Cooper, Forester, Kolaczkowski, Luckas, Parry, Ramey-Smith, Thompson, Whitehead & Wreathall, 2000).

8. Connectionism Assessment of Human Reliability (**CAHR**, Sträeter, 2000; 2005).

9. Standardized Plant Analysis Risk HRA Method (**SPAR-H**, Gertman, Blackman, Marble, Byers, Haney &Smith, 2005).

10. University of Maryland Hybrid (**UMH**, Shen & Mosleh, 1996).

11. Commission Errors Search and Assessment (**CESA**, Reer, Dang & Hirschberg, 2004).

12. Human Factors Process Failure Modes & Effects Analysis (**HF PFMEA**, Broughton, Carter, Chandler, Holcomb, Humeniuk, Kerios, Bruce, Snyder, Strickland, Valentino, Wallace, Wallace & Zeiters, 1999).

13. Time Reliability Correlation (**TRC**, Dougherty & Fragola, 1987).

14. EPRI Caused Based Decision Tree (**CBDT**, Parry, Lydell, Spurgin, Moieni & Beare, 1992; Moieni, 1994; Moieni, Spurgin, & Singh, 1994).

3.1 Technique for Human Error Rate Prediction (THERP)

3.1.1 Developmental Context

THERP (Swain& Guttman, 1983) was initially developed and used by Sandia National Laboratories (SNL) in 1961 for defense-related HRA analyses. WASH-1400 (1975) used THERP to perform HRA in the PRA framework for two United States commercial Nuclear Power Plants (NPPs). The experience gained through such practice led to the development of the THERP handbook by the same group of experts (Swain, 1990).

"The Handbook addresses the kinds of tasks that have been studied in PRAs of NPPs to date. These are divided into tasks performed during normal operating conditions and tasks to be performed after an abnormal event has occurred. For normal operating conditions, the tasks addressed by the Handbook are calibration tasks, tests performed after calibration, maintenance, or other operations, and the restoration of important safety components to their available states after carrying out tests, calibration, or maintenance. For abnormal conditions, the Handbook provides information to estimate HEPs for tasks to be performed to cope with abnormal events, many of which have not occurred in commercial plant operation." (Swain & Guttman, 1983). "The Handbook does not provide estimated HEPs related to the use of new display and control technology that is computer-based. Neither does the Handbook provide HEPs for corrective maintenance such as repairing a pump." (Swain & Guttman, 1983).

3.1.2 Screening

THERP provides instructions for screening two types of activity: diagnosis and rule-based action. The screening of diagnosis activity is based on available time (Table 10 .). Screening for rule-based behavior is shown in Table 11 .
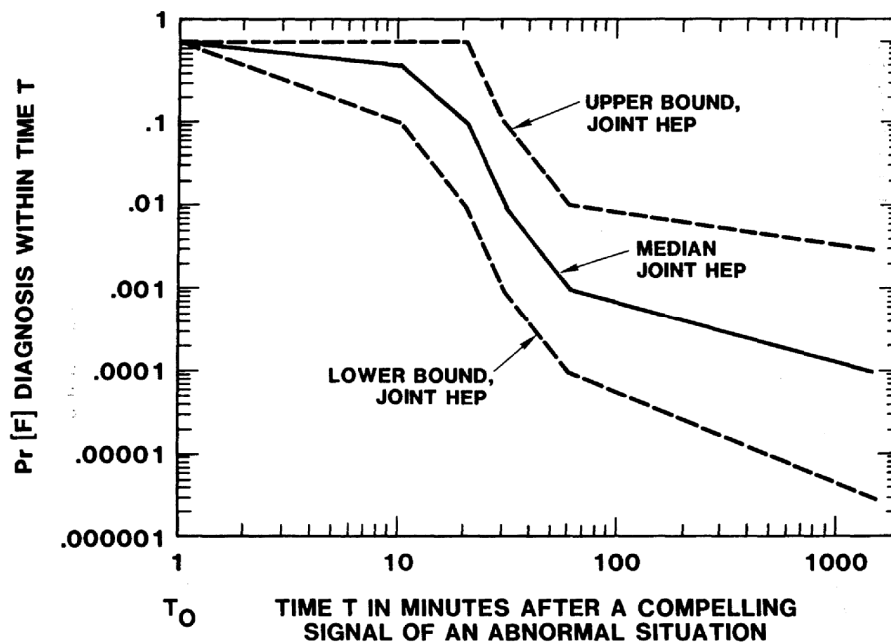
Figure 4 . Initial screening model of estimated human error probability and uncertainty bounds for diagnosis within time T of one abnormal event by control room personnel.

3.1.3   Task Decomposition

The following are the four phases of THERP:

1.   Familiarization.

    a.   Gather plant-specific and event-specific information.

2.   Qualitative assessment.

    a.   Perform preliminary task analyses including error identification.

3.   Quantitative assessment.

    a.   Estimate the HEPs.

4.   Incorporation into system risk and reliability model.

    a.   Perform sensitivity study to determine the impact on the system.

    b.   Incorporate results into system model (e.g., PRA or reliability model).

The following are the ten steps to performing qualitative and quantitative assessment through analyzing the man-machine system:

1.   Describe the system goals and functions of interest.

2.   Describe the situational characteristics.

3.   Describe the characteristics required of the personnel.

4.   Describe the jobs and tasks performed by the personnel.

5.   Analyze the jobs and tasks to identify error-likely situations and other problems.

6.   Estimate the likelihood of each potential error.

7.   Estimate the likelihood that each error will be undetected (or uncorrected).

8.   Estimate the consequences of each undetected (or uncorrected) error.

9.   Suggest changes to the system.

10.   Evaluate the suggested changes (repeat steps 1 through 9).

To calculate the HEP for a task, THERP provides a number of activities for the analyst to identify the HEP's existence in the tasks of analysis. Example activities are screening (including diagnosis-based or rule-based) and action (including check displays, perform control, and operating valves). See Figure 5(a) for the specific activities. Note, Figure 5(a) is the THERP search flow for human errors. Its objects of search include human activities and other factors that could affect HEPs.

3.1.4 PSF List and Causal Model

THERP provides a list of PSFs (see Table 9) but gives no specific rules to assess the states of these PSFs and their effects on HEPs.

Table 9 . The PSFs according to THERP.

External PSFs:	Internal PSFs:	
• Situational Characteristics - Control Room Architectural Feature - Quality of the Working Environment - Works Hours and Work Breaks - Shift Rotation and Night Work - Availability/Adequacy of Special Equipment/Tools and Supplies - Manning Parameters - Organizational Structure and Actions by Others - Rewards, Recognition, and Benefits • Task and Equipment Characteristics - Perceptual Requirements - Motor Requirements - Control-Display Relationships - Anticipatory Requirements - Interpretation - Decision-Making - Complexity/Information Load - Frequency and Repetitiveness - Task Criticality - Long- and Short-Term Memory - Calculation Requirements - Feedback - Dynamic Versus Step by Step Activities - Team Structure - Main-Machine Interface Factors • Job and Task Instructions - Operating Procedures - Oral Instructions	• Psychological Stressors - Suddenness of Onset - Duration of Stress - Task Speed - Task Load - High Jeopardy Risk - Threat of Failure, Loss of Job - Monotonous, Degrading, or Meaningless Work - Long, Uneventful Vigilance Periods - Conflicts of Motives About Job Performance - Reinforcement Absent or Negative - Sensory Deprivation - Distraction (Noise, Glare, Movement, Flicker, Color) - Inconsistent • Physiological Stressors - Duration of Stress - Fatigue - Pain or Discomfort - Hunger or Thirst - Temperature Extremes - Radiation - G-Force Extremes - Atmospheric Insufficiency - Vibration - Movement Constriction - Lack of Physical Exercise - Disruption of Circadian Rhythm	• Organizational Factors - Previous Training/Experience - State of Current Practice or Skill - Personality and Attitudes - Motivation and Attitudes - Knowledge of Required Performance Standards - Sex Differences - Physical Condition - Attitudes Based on Influence of Family and Other Outside Persons or Agencies - Group Identifications

Only three PSFs among the identified PSFs are used in HEP calculation. These are: Tagging levels (of components or controls), experience, and stress (Tables 20-15 and 20-16 of THERP).

THERP does not provide a causal model to describe the dependencies among the PSFs modeled.

Table 10 . Initial screening model of estimated human error probabilities and error factors for within time T by control room personnel of abnormal events annunciated closely in time.*

Item	T (Minutes** after $T_0^†$)	Median joint HEP for diagnosis of a single or the first event	EF	Item	T (Minutes** after $T_0^†$)	Median joint HEP for diagnosis of the second event[††]	EF
(1)	1	1.0	--	(7)	1	1.0	--
(2)	10	.5	5	(8)	10	1.0	--
(3)	20	.1	10	(9)	20	.5	5
(4)	30	.01	10	(10)	30	.1	10
				(11)	40	.01	10
(5)	60	.001	10				
				(12)	70	.001	10
(6)	1500 (≈ 1 day)	.0001	30				
				(13)	1510	.0001	30

* "Closely in time" refers to cases in which the annunciation of the second abnormal event occurs while control room personnel are still actively engaged in diagnosing and/or planning responses to cope with the first event. This is situation-specific, but for the initial analysis, use "within 10 minutes" as a working definition of "closely in time." Note that this model pertains to the control room crew rather than to one individual.

** For points between the times shown, the medians and EFs may be chosen from Figure 4 .

[+] T_0 is a compelling signal of an abnormal situation and is usually taken as a pattern of annunciators. A probability of 1.0 is assumed for observing that there is some abnormal situation.

[++] Assign HEP = 1.0 for the diagnosis of the third and subsequent abnormal events annunciated closely in time.

Table 11 . Initial screening model of estimated human error probabilities and error factors for rule-based actions by control room personnel after diagnosis of an abnormal event.*

Item	Potential Errors	HEP	EF
	Failure to perform rule-based actions correctly when written procedures are available and used:		
(1)	Errors per critical step without recovery factors	.05	10
(2)	Errors per critical step with recovery factors	.025	10
	Failure to perform rule-based actions correctly when written procedures are not available or used:		
(3)	Errors per critical step with or without recovery factors	1.0	--

* Note that this model pertains to the control room crew rather than to one individual.

51

3.1.5   Coverage

The THERP's error search scheme focuses on identifying ergonomic and organizational factors. Examples of ergonomic factors are "Procedural Items" (in Table 20-7 of THERP) and "Valve Selection" (in Table 20-13 of THERP). An example organizational factor is "Administrative Control" (in Table 20-6 of THERP). With respect to cognitive error modeling, THERP uses available time to determine the probabilities of diagnosing failure. The use of such time reliability correlation is not credited for having emphasis on cognitive analysis.

3.1.6   HEP Calculation Procedure

HEPs are calculated through a number of steps:

1.   Analyze the event:

a.   Construct the HRA Event Tree (ET). For each branching point of the HRA ET, use the HEP search scheme (see Figure 5(a)) to identify the likely human errors and the corresponding nominal HEPs as well as the uncertainty bounds.

b.   Identify factors and interactions affecting human performance: Assess the effect of the tagging levels, experience, and stress on the HEPs as well as the uncertainty bounds of the HEPs.

2.   Quantify effects of factors and interactions:

a.   Assess the levels of task dependencies based on the five-level dependency scale specified by THERP. Such dependencies would affect the task HEPs.

b.   Account for probabilities of recovery from errors: Assess the possible recovery branches in the HRA ET and assess the success probabilities.

3.   Calculate human error contribution to probability of system failure:

a.   Determine the success and failure consequences within the HRA ET and calculate the HEP of the HRA ET. The calculated HEP is used in the PRA model.
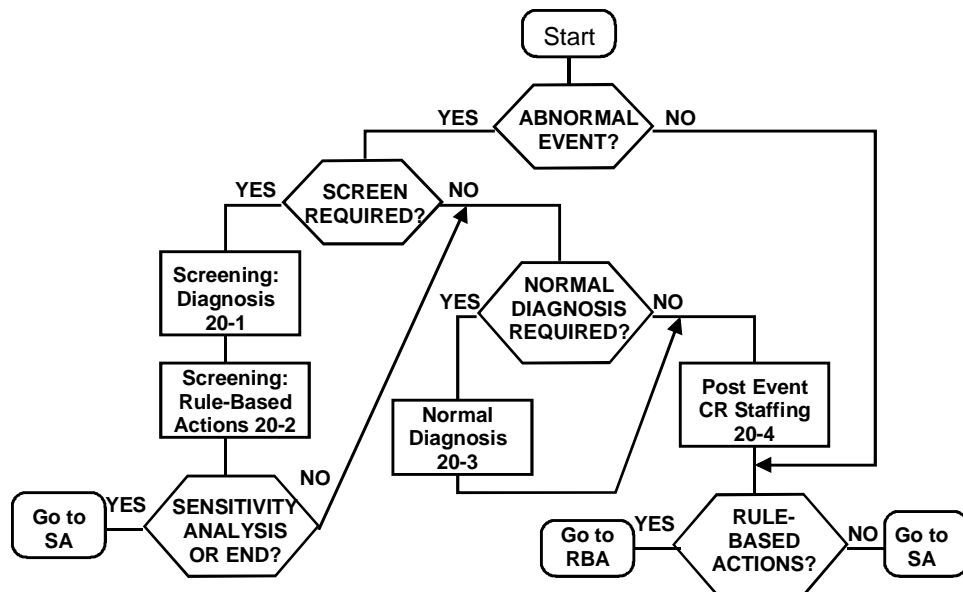
Figure 5(a) . THERP HEP Calculation Scheme (1/5).
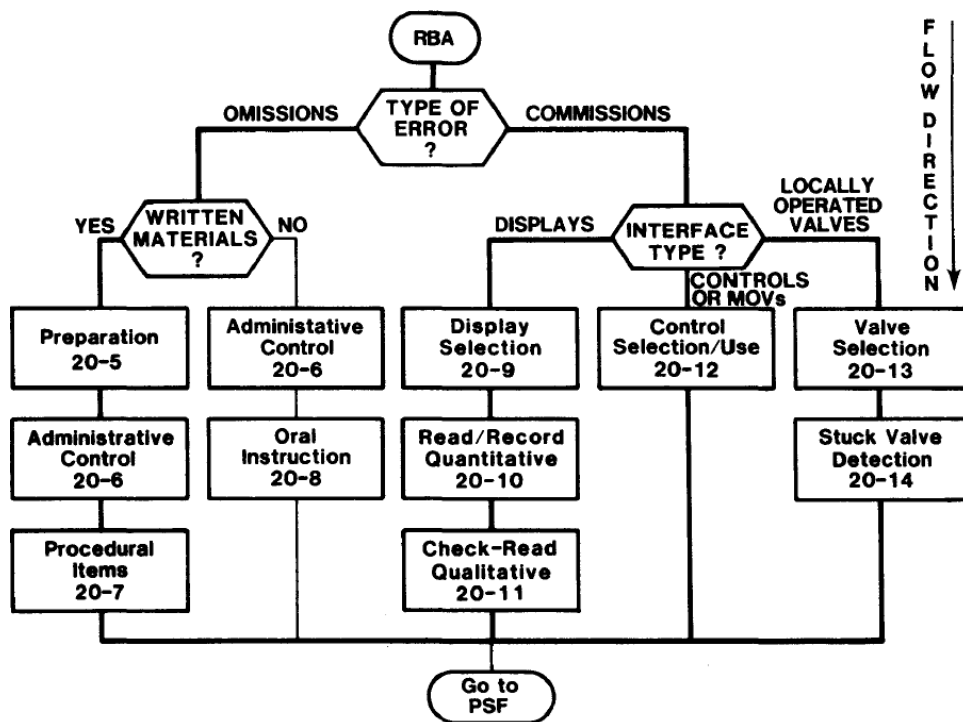
Figure 5(b) . THERP HEP Calculation Scheme (2/5).
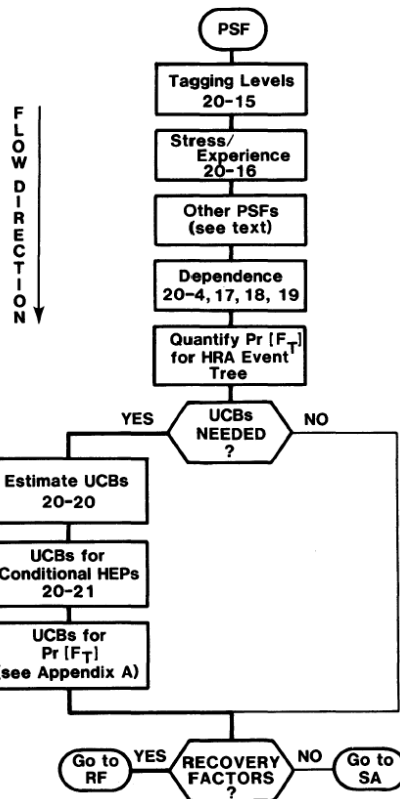
Figure 5(c) . THERP HEP Calculation Scheme (3/5).
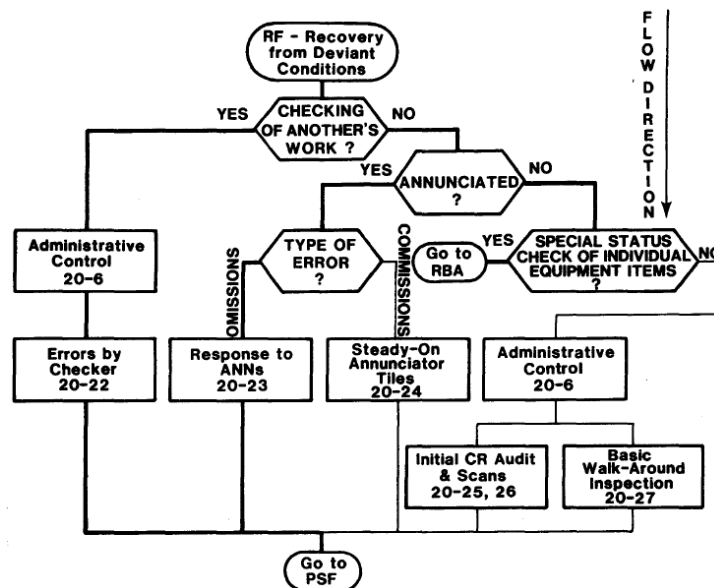
Figure 5(d) . THERP HEP Calculation Scheme (4/5).

54
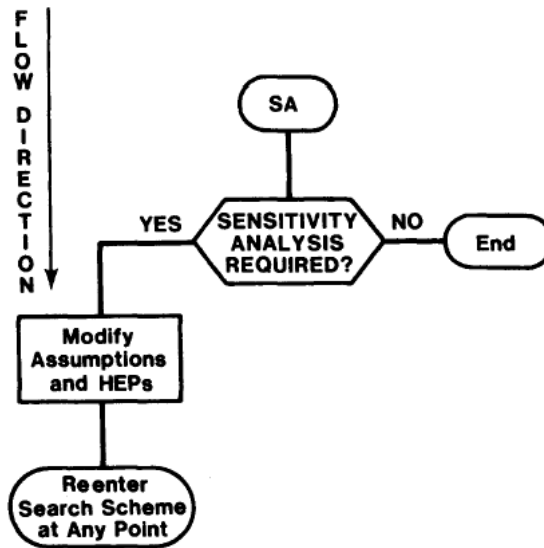
Figure 5(e) . THERP HEP Calculation Scheme (5/5).

3.1.7   Error-Specific HEPs

THERP calculates probabilities of the following types of errors:

- Screening and detection of system abnormalities

- Diagnosis and identification of the causes of system abnormalities

- Omitted actions, including actions in procedure preparation, use of a specified procedure (e.g., administrative control), execution of a procedure step, and providing an oral instruction

- Writing down incorrect information

- Acting on a wrong object; includes reading from an unintended display, acting at an unintended control, and unintended control (e.g., turn a control at wrong direction).

3.1.8   Task Dependencies and Recovery

THERP provides five levels of dependency between two consecutive operator activities. These activities are represented by branches of an HRA Event Tree. The five dependency levels are zero dependency (ZD), low dependency (LD), moderate dependency (MD), high dependency (HD), and complete dependency (CD). Although the authors state that "There are no hard and fast rules for deciding what level of dependency is appropriate for any situation; considerable judgment is required." The "time between tasks" is suggested as a key factor affecting the level of dependency.

Recovery, by other crew or by instrument, is explicitly covered. The framework therefore allows for explicit accounting of the impact of dependencies and recovery actions on the overall probability of error.

3.1.9   HEP Uncertainty Bounds

An error factor (EF) is used to represent the HEP uncertainty bounds. In THERP each activity is associated with a median HEP and an EF. The uncertainty bound of this activity is found by multiplying or

dividing the median HEP by the EF. For example, assume the median HEP and EF for a certain activity are 1E-3 and 10, respectively. The uncertainty bound of this activity is 1E-4, and 1E-2. A lognormal uncertainty distribution is often assumed.

3.1.10   Level of Knowledge Required

About one year of HRA experience is believed to be a sufficient level of experience for applying THERP. "The Handbook could certainly be used by persons who do not have a human factors background, particularly since it contains somewhat of a primer of human factors basics, and references to other human factors engineering handbooks. Given the complexity of the subject, however, users are likely to vary considerably in their ability to perform the analyses properly. The consistency and reliability of the Handbook users would certainly be improved if they participated in some human factors training, preferably emphasizing the use of the Handbook itself (this would probably be beneficial to most users with previous human factors training as well)." (Swain & Guttman, 1983).

Another reviewer, an authority in PRA, stated that the estimated HEPs in the handbook should be used only by human factors specialists. This PRA expert expressed doubts that analysts outside the human factors field would be able to make the proper evaluation of the role of PSFs (Swain & Guttman, 1983).

3.1.11   Validation

3.1.11.1   *Error Identification*

THERP does not provide explicit procedures for performing error identification.

3.1.11.2   *HEP Quantification*

The following are comments made by the authors and reviewers of THERP and other HRA/PRA experts:

- (by the handbook authors) The scarcity of objective and quantitative data on human performance in NPPs is a serious limitation. Most of the HEPs in this Handbook are what we call <u>derived data</u>. In some cases, they are extrapolations from performance measures, which may be only marginally related. In other cases the HEPs represent our best judgment based on our experience in complex systems and on our background in experimental and <u>engineering psychology</u>. The necessity to rely so heavily on judgment is a regrettable state of affairs, but a start needs to be made, and this Handbook is a first step toward what is really needed--a large data bank of human performance information directly related to NPP tasks.

- (by the handbook authors) "In general, our HEPs and models are based on studies and observations in the kinds of plants we have visited, which are listed in the Foreword (of THERP document). Some newer plants may incorporate human factors improvements that could make some of our estimates too pessimistic."

- One of the workshop participants who had participated in the Space Shuttle PRA HRA made the following comment on the quality of THERP quantification: "In general THERP is conceded by most people including the authors to yield a fairly conservative value. When we compared THERP derived values for HEPs that were fairly high contributors to the Shuttle LOCV with CREAM and simulator data (I believe we had four or five events that the Astronaut's office provided some simulator data for) we found that THERP consistently over predicted compared to the simulator data and it was also higher than the CREAM values."

3.1.12   Reproducibility

3.1.12.1   Error Identification

THERP does not provide explicit procedures for performing error identification.

3.1.12.2   HEP Quantification

THERP provides a prescriptive HEP calculation procedure (see Section 3.1.6). Given clear context description and well-trained analysts, the reproducibility of the result is rated medium.

3.1.13   Sensitivity

Small changes in context specifications do not significantly change THERP HEP values. As a result, the sensitivity is rated low.

3.1.14   Experience Base

THERP is the product of the first large scale HRA study of complex systems. Its results have been used as a benchmark to develop other HRA methods (e.g., CAHR and SPAR-H), and its data have been adopted in other HRA methods (e.g., TRC and CREAM). Various studies in different industries (e.g., nuclear, chemical process, information technology, off shore platforms, and aviation) have used THERP for HRA studies.

3.1.15   Resource Requirements

The HEP calculations can be performed manually. Software would facilitate the HEP calculation, but is not a necessity. Applying THERP requires the analyst to identify the human tasks that match THERP's set of human activities to which the provided HEPs apply. Task dependencies and PSF influences must be considered in calculating the HEP of each subtask. Finally, an HRA event tree needs to be constructed with consideration of task recovery in order to determine the final HEP for the task of interest. THERP provides a fairly structured procedure for performing these steps. Therefore, the resource requirement for THERP is rated medium.

3.1.16   Cost and Availability

The THERP report is publicly available. There is no licensing fee associated with using THERP as a method or a source of HEP estimates.

3.1.17   Suitability for NASA Applications

THERP is partially applicable for assessing HEPs for existing aerospace designs in normal and emergency conditions. THERP task analysis and quantification schemes are best suited for routine tasks under normal conditions (e.g., proceduralized pre- and post-flight checks). Ground processing activities most closely match the situations for which THERP was developed. THERP does not address human performance in flight, zero gravity, or microgravity environments.

One workshop participant with experience in Space Shuttle PRA HRA stated: "I found THERP to be useful in our efforts on the Shuttle PRA to evaluate errors of omission. It was difficult to use on continuous feedback HRA types of errors and even the text of THERP (Swain & Guttman, 1983) indicated that THERP is not very useful in these types of error evaluations. Also it was not very helpful on the Errors of Commission. Our experience showed that the types of errors that THERP was useful to be used on were not

the types of errors that ended up contributing a lot to the Shuttle PRA LOCV. We did find THERP to be useful to help us define the uncertainty of some of our HEPs."

Another workshop participant stated that "For some applications, especially for routine tasks (e.g., assembly, maintenance, and possibly testing), ASEP and/or THERP may be considered. These methods are generally not suitable for tasks in 'Command and Control' or 'Space (during mission)'."

Other factors limiting the applicability of THERP include:

- Inconsistency in typical time windows: The available time windows for action between nuclear power plant operation and aerospace missions are often significantly different. Recovery time windows for nuclear accidents typically vary from hours to days. In comparison, some action time windows and system response times in aerospace scenarios are very short, particularly those in the dynamic phases of space vehicle flight such as ascent, docking, and descent. Such time window differences cast doubt on the applicability of the HEP estimates based mostly on nuclear reactor time scales.

- Required Information: THERP quantification relies on specific characteristics of tasks and activities. This limits the usefulness of THERP for application to new aerospace designs for which detailed system information is not available.

- Needed modifications: THERP has a fixed search scheme for identifying possible errors and calculating HEPs for activities in nuclear power plants. It requires significant revisions to the search scheme, parameter values, and addition of new PSFs for space missions (as compared with Ground Processing).

3.2 Accident Sequence Evaluation Program (ASEP)

3.2.1 Developmental Context

ASEP (Swain, 1987) is a simplified version of the THERP method developed by an author of THERP. ASEP is highly nuclear power oriented. The main goal of its development was to obtain order of magnitude estimates of HEPs without the level of effort required by THERP.

ASEP is one of the HRA methods that use time-reliability correlation as the basis for calculating cognitive/decision failure.

3.2.2 Screening

ASEP screening is by assuming HEP=1 for the action in the PRA model (See Step 1 in 3.2.6). This is a common practice in PRA screening of human failure events.

3.2.3 Task Decomposition

ASEP simply decomposes a task into subtasks of diagnosis and action.

3.2.4 PSF List and Causal Model

Time is the dominant factor for calculating HEPs. Its PSFs focus on training and knowledge (see Table 13). No causal model of human error is provided.

3.2.5 Coverage

HEP in ASEP is dominated by the time available for diagnosis. (See time reliability correlation shown in Figure 6 and Table 12 .)

3.2.6 HEP Calculation Procedure

For a given initiating event or multiple sequential initiating events, ASEP provides rules to calculate the HEPs. The HEPs for diagnosis and action errors are calculated separately. The basic HEP is calculated based on the time factor. The factors for adjusting the basic HEPs are based on task types.

ASEP provides a number of figures and tables to calculate HEPs:

- Step-by-step procedural instructions (ASEP Table 8-1)

- Normal diagnosis model (ASEP Figure 8-1 and Table 8-2)

- Rules for adjusting the diagnosis HEPs (ASEP Table 8-3)

- The annunciator response model (ASEP Table 8-4)

- The nominal rules for post-accident, post-diagnosis actions (ASEP Table 8-5).

Tables 8-2 to 8-5 and Figure 8-1 of ASEP are reproduced in this report (see Figure 6 and Tables 41 – 44).

The HEP calculation procedure (ASEP Table 8-1) can be summarized into the following steps:

1. Familiarize with the terminology and definitions (e.g., skill-, rule-, and knowledge- based behavior, action time, diagnosis time, maximum allowable time, annunciation time of an abnormal event, and cognition-related terms).

2. Screen the activities with HEPs assumed to be 1.0. Examples are cases where the required instrumentation fails to support diagnosis or post-diagnosis behavior, or the instrumentation is inaccurate or misleading.

3. Calculate HEP for *diagnosis error.*

    a. Estimate the Maximum Allowable Time for the abnormal event to be handled.

    b. Identify the required actions to handle the event.

    c. Measure or estimate the required time for the actions to be completed (some guidelines are provided for the estimation).

    d. Subtract the time estimated in Item "c" from the Maximum Allowable Time specified in step "a." This is the allowable time for diagnosis.

    e. Use Figure 6 and Table 12 (Figure 8-1 or Table 8-2 in ASEP) to calculate the basic HEPs for diagnosis error.

    f. Use Table 13 (Table 8-3 of ASEP) to adjust the basic HEP.

    g. Calculate HEP uncertainty bounds specified in the Normal Diagnosis Model (Figure 8-1 or Table 8-2 in ASEP).

    h. For multiple simultaneously occurring abnormal events, guidelines are provided to determine whether the subsequent events should be included in the analysis.

4. Calculate the HEP for *action error.*
    Use the nominal rules for post-accident, post-diagnosis actions (Tables 43 and 44; Tables 8-4 and 8-5 in ASEP) for calculation the action error.

5. Calculate the total-failure probability using an HRA tree.
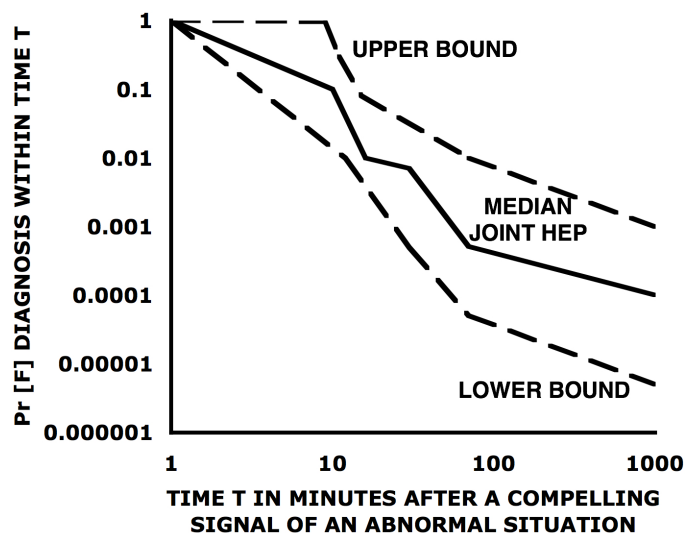
Figure 6 . Nominal model for estimating HEPs and uncertain bounds for diagnosis within time T of one abnormal event by control room personnel (Table 8-1 in ASEP, (After Swain, 1987)).

Table 12 . (Table 8-2 of ASEP) Nominal model of estimated HEPs and EFs for diagnosis within time T by control room personnel of abnormal events annunciated closely in time.*

Item	T (Minutes** after T_0^+)	Median joint HEP[++] for diagnosis of a single or the first event	EF	Item	T (Minutes** after T_0^+)	Median joint HEP[++] for diagnosis of the second event	EF	Item	T Minutes** after T_0^+)	Median joint HEP[++] for diagnosis of the third event	EF
(1)	1	1.0	--	(7)	1	1.0	--	(14)	1	1.0	--
(2)	10	.1	10	(8)	10	1.0	--	(15)	10	1.0	--
(3)	20	.01	10	(9)	20	.1	10	(16)	20	1.0	--
(4)	30	.001	10	(10)	30	.01	10	(17)	30	.1	10
				(11)	40	.001	10	(18)	40	.01	10
								(19)	50	.001	10
(5)	60	.0001	30								
				(12)	70	.0001	30				
								(20)	80	.001	30
(6)	1500	.00001	30								
				(13)	1510	.00001	30				
								(21)	1520	.00001	30

* "Closely in time" refers to cases in which the annunciation of the second abnormal event occurs while control room personnel are still actively engaged in diagnosing and/or planning responses to cope with the first event. This is situation-specific, but for the initial analysis, use "within 10 minutes" as a working definition of "closely in time."

Note that this model pertains to the control room crew rather than to one individual.

** For points between the times shown, use the medians and EFs from Table 8-1 for the first event, and interpolate between the tabled values for the second or third events.

+ T_0 is a compelling signal of an abnormal situation and is usually taken as a pattern of annunciators. A probability of 1.0 is assumed for observing that there is some abnormal situation.

++ ASEP Table 8-3 presents some guidelines to use in adjusting or retaining the nominal HEPs presented above.

Table 13 . (Table 8-3 of ASEP) Guidelines for Adjusting Nominal Diagnosis HEPs from Table 8-2 (of ASEP).

Item	General Rules
(1)	Use upper bound if:
	a. the event is not covered in training,
	<u>or</u>
	b. the event is covered but not practiced except in initial training of operators for becoming licensed,
	<u>or</u>
	c. the talk-through and interviews show that not all the operators know the pattern of stimuli associated with the event.
(2)	Use lower bound if:
	a. the event is a well-recognized classic (e.g., TMI-2 incident), and the operators have practiced the event in the simulator re-qualification exercises, and
	b. the talk through and interviews indicate that all the operators have a good verbal recognition of the relevant stimulus patterns and know what to do or which written procedures to follow.
(3)	Use nominal HEP if:
	a. the only practice of the event is in simulator requalification exercises and all operators have had this experience, or
	b. none of the rules for use of upper or lower bound apply.

Table 14 . (Table 8-4 of ASEP) The annunciator response model: estimated HEPs* for multiple annunciator alarming closely in time.**

Item	Number of ANNs	Pr[F_i] for each annunciator (ANN) (or completely dependent set of ANNs) successively addressed by the operator										$\overline{Pr[F_i]}^\dagger$
		1	2	3	4	5	6	7	8	9	10	
		(a)	(b)	(c)	(d)	(e)	(f)	(g)	(h)	(i)	(j)	(k)
(1)	1	.0001										.0001
(2)	2	.0001	.001									.0006
(3)	3	.0001	.001	.002								.001
(4)	4	.0001	.001	.002	.004							.002
(5)	5	.0001	.001	.002	.004	.008						.003
(6)	6	.0001	.001	.002	.004	.008	.016					.005
(7)	7	.0001	.001	.002	.004	.008	.016	.032				.009
(8)	8	.0001	.001	.002	.004	.008	.016	.032	.064			.02
(9)	9	.0001	.001	.002	.004	.008	.016	.032	.064	.13		.03
(10)	10	.0001	.001	.002	.004	.008	.016	.032	.064	.13	.25	.05
(11)	11-15	Pr[F_i] for each additional ANN beyond 10 = .25										.10
(12)	16-20											.15
(13)	21-40											.20
(14)	>40											.25

* The HEPs are for the failure to initiate some kind of intended corrective action as required. The action carried out may be correct or incorrect and is analyzed using other tables. The HEs include the effect of stress and should not be increased in consideration of stress effects.

An EF of 10 is assigned to each Pr*(F$_i$)* or $\overline{Pr(F_i)}$. Based on computer simulation, use of an EF of 10 for $\overline{Pr(F_i)}$ yields an approximately correct upper bound for the 95th percentile. The corresponding lower bounds are too high; they are roughly equivalent to 20th percentile rather than the usual 5th percentile bounds. Thus, use of an EF of 10 for the mean Pr*(F$_i$)* values provides a conservative estimate since the lower bounds are biased high.

** "Closely in time" refers to cases in which two or more annunciators alarm within several seconds or within a time period such that the operator perceives them as a group of signals to which he must selectively respond.

$\overline{Pr(F_i)}$ is the expected Pr*(F)* to initiate action in response to randomly selected annunciators (or completely dependent set of annunciators) in a group of annunciators competing for the operator's attention. It is the arithmetic mean of the Pr*(F$_i$)*s in a row, with an upper limit of .25. The $\overline{Pr(F_i)}$ column assumes that all of the annunciators (or completely dependent set of annunciators) are equal in terms of the probability of being noticed. See page 11-52 paragraph 2, in NUREG/CR-1278 (THERP) if this assumption does not hold.

Table 15 . (ASEP Table 8-5) Assessment of nominal HEPs for post-accident post-diagnosis action.

Item	HEP	EF	Action*
(1)	1.0	--	Perform a critical skill-based or rule-based action correctly when no written procedures are available. (Details of skill-based actions are not required to be written if they can be classified as "skill-of-the-craft"**.) This assessment is used even though it may be required for personnel to have memorized these actions. Instead, they would likely refer to the written procedures at a later time during the usual checking to see that all immediate emergency actions had been performed correctly. (See Table 2-1 of ASEP for definitions.)
(2)	var.	--	If sufficient information can be obtained per a task analysis, as described in Chapter 4 of NUREG/CR-1278, use the data tables in Chapter 20 of NUREG/CR-l278, adjusted for the effects of dependence, stress, and other performance shaping factors (PSFs) and error recovery factors (RFs) per the search scheme in Chapter 20. If this level of information cannot be obtained because of scheduling or other restrictions, use the remainder of this table.

Items (3), (4), and (5) present HEPs for the original performer of the action and must be adjusted for the effects of other operators and recovery factors (items 6 -9). These HEPs are for failure to correctly perform a critical post-diagnosis procedural action as part of a "step-by-step task"** or a "dynamic task"** done under "moderately high stress"** or "extremely high stress"**. See item 10 in Table 8-1 (of ASEP) for guidelines on how to apply these terms. It is assumed that "novice personnel" would be replaced by "skilled personnel" for critical actions.

Item	HEP	EF	Action*
(3)	0.02	5	Perform a critical action as part of a step-by-step task done under moderately high stress.
(4)	0.05	5	Perform a critical action as part of a dynamic task done under moderately high stress or a step-by-step task done under extremely high stress.
(5)	0.25	5	Perform a critical action as part of a dynamic task done under extremely high stress.

+If recovery of above errors made by the original performer is still possible at the point of error action, use following HEPs (6), (7), or (8) and related task and stress categories for a second person who checks the performance of the original performer.

Item	HEP	EF	Action*
(6)	0.2	5	Verify the correctness of a critical action as part of a step-by-step task under moderately high stress.
(7)	0.5	5	Verify the correctness of a critical action as part of a dynamic task done under moderately high stress or a step-by-step task done under extremely high stress.
(8)	0.5	5	Verify the correctness of a critical action as part of a dynamic task done under extremely high stress.*
(9)	var.	--	If there are error recovery factors (RFs) in addition to the use of human redundancy in items (6), (7), and (8), the influence of these RFs must be assessed separately. For annunciator RFs, use the Annunciator Response Model in Table 8-4 (of ASEP).

Item	HEP	EF	Action*
(10)	0.001	10	Perform a post-diagnosis immediate emergency action for the reactor vessel/containment critical parameters, when (a) it can be judged to have been committed to memory, (b) it can be classified as skill-based actions per Table 2-1 (of ASEP), <u>and</u> (c) there is a backup written procedure. Assume no immediate RF from a second person for each such action.

* The HEPs are for independent actions or independent sets of actions in which the actions making up the set can be judged to be completely dependent. Other levels of dependence among actions can be assessed by the analyst, using one or more methods for assessing dependence described in Chapter 10 of NUREG/CR-l278.

** See the prefatory section "Definitions of Technical Terms" for definitions of these frequently misunderstood terms.

⁺ Theoretically, if the HEP for item (7) is assessed as .5, the HEP for item (8) should be larger, e.g., .75. However, as .5 is already so large, any increase in the estimated HEP is judged to be unduly conservative.

3.2.7   Error-Specific HEPs

ASEP provides tables for calculating HEPs for failing to make correct diagnosis and perform required actions within a specified time.

3.2.8   Task Dependencies and Recovery

Task dependencies and recovery are implicitly accounted for at the level of ASEP modeling resolution (see Table 15).

3.2.9   HEP Uncertainty Bounds

Uncertainty bounds are provided for both diagnosis and action errors.

3.2.10   Level of Knowledge Required

ASEP is very prescriptive and can be performed by a PRA analyst with limited familiarity with HRA methods.

3.2.11   Validation

3.2.11.1   Error Identification

ASEP does not provide explicit procedures for error identification.

3.2.11.2   HEP Quantification

ASEP has been widely used in the U.S. nuclear industry. The ASEP document states that "The nominal values … are intended to err on the conservative side, when errors in estimation are made. However, the nominal values presumably avoid undue conservatism." (Swain, 1987).

An independent assessment of ASEP is documented in Gore, Dukelow, Mitts & Nicholson, 1995 where over 4000 individual simulator critical tasks (ISCTs) analyzed and compared the results with ASEP predictions. "Assessed within the context of the performance of critical tasks by individuals, the ASEP post-accident, post-diagnosis, nominal HRA procedure is found to be somewhat conservative on the average. For small values of HEP estimated by the ASEP procedure, there is little or no conservatism, but larger estimated HEP values exhibit significant conservatism." (Gore, Dukelow, Mitts & Nicholson, 1995).

The ASEP-based HEPs tend to be higher than those from THERP (Gertman et Blackman, 1993).

3.2.12   Reproducibility

3.2.12.1   Error Identification

ASEP does not provide explicit procedures for error identification.

3.2.12.2   HEP Quantification

Despite the fact that "lack of adequate documented guidance" has been cited to potentially "cause problems with traceability, consistency (especially if more than one analyst is involved), and repeatability" (US NRC, 2005), the ASEP quantitative reproducibility is rated as medium since the method is basically a time-reliability correlation method where, given the available response time, the HEP is highly reproducible.

3.2.13   Sensitivity

ASEP uses a time-reliability correlation for initial diagnosis errors. As a result, the HEPs could be very sensitive to the time factor. However, ASEP also provides some guidance and rules to adjust the final HEP values. This would reduce the sensitivity of the results. Sensitivity is rated medium.

3.2.14   Experience Base

ASEP has been widely used especially in the U.S. nuclear industry as documented in The NEA Committee on the Safety of Nuclear Installations (NEA Committee, 1998).

3.2.15   Resource Requirements

ASEP, by design, was intended for use by systems analysts who are not necessarily HRA specialists. The required resources are therefore rated low.

3.2.16   Cost and Availability

The method is publicly available. No specific software is required.

3.2.17   Suitability for NASA Applications

As stated earlier the primary factor in estimation of the basic HEPs in ASEP is time. As there are significant differences between time scales of interest in nuclear HRAs and those in many of the space mission activities, ASEP tables and figures need to be re-calibrated before they can be applied. With proper consideration of such differences ASEP could be used as a screening method in some cases. The overall suitability is rated low.

3.3 Success Likelihood Index Methodology (SLIM)

3.3.1 Developmental Context

SLIM (Embry, 1984) is not an HRA method per se, but rather a scaling technique. It has no fixed set of HEPs nor does it have a required set of PIFs/PSFs. It was developed under United States Nuclear Regulatory Commission sponsorship in the 1980's to formalize the use of expert judgment in estimating HEP values. It requires minimum data points (e.g., real event statistics) for HEP assessment. While the method has been extensively used in nuclear PRAs, as a computational framework, it can be easily applied to other domains.

3.3.2 Screening

SLIM does not provide a procedure for quantitative screening.

3.3.3 Task Decomposition

SLIM does not provides instructions for task decomposition. However, tasks need to be decomposed to a level consistent with the anchoring events in order to calculate HEPs.

3.3.4 PSF List and Causal Model

SLIM does not provide a fixed set of PSFs. It allows the analysts to identify them based on the situation being analyzed. It suggests some PSFs for the analyst to consider. These include:

- Quality of design

- Meaningfulness of procedures

- Roles in operation

- Teams

- Stress

- Morale/motivation

- Competence.

3.3.5 Coverage

Within the SLIM framework the analyst can define any type of error mechanism and cause, identify relevant PSFs, and obtain numerical anchor points for the corresponding error probabilities.

3.3.6 HEP Calculation Procedure

The quantification steps of SLIM are as follows:

1. Modeling of specifically relevant PSFs: Experts identify the PSFs relevant to the event of interest.

2.   Weighting the PSFs: Experts weigh the effect of each PSF.

3.   Rating the PSFs: Experts assess the state of each PSF.

4.   Calculating the Success Likelihood Indexes (SLIs): The values of SLIs are calculated using Equation 1.

$$SLI = \sum \left[Normalized \ \ Weight(PIF_i) \times State(PIF_i) \right]$$ (Eq. 1)

5.   Conversion of the SLIs to probabilities.

6.   Equation 2 is used to calculate the HEPs in SLIM.

$$Log(1 - HEP) = a \times SLI + b$$ (Eq. 2)

Using at least two sets of known HEPs and SLIs as reference points, the constants "a" and "b" can be obtained. Using the same equation (Equation 2) and replacing the SLI by the SLI of the task of interest, the HEP of the task can be calculated. Figure 7 shows an example of the linear relationship between Log (Success Probability) and SLI.

7.   Calculation of uncertainty bounds.

Performing a sensitivity study by changing PSFs' weights and ratings can determine the upper bound and lower bound of SLI, which in turn can determine the upper bound and lower bound of HEPs.
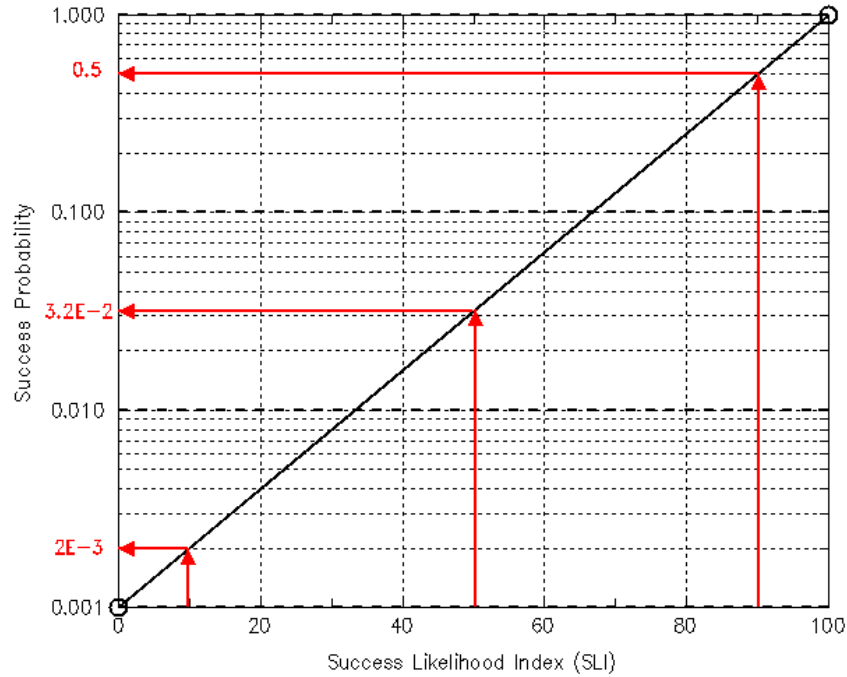
Figure 7 . An example of determining HEP by interpolating between two reference points (e.g., SLI = 0 and SLI = 100).

68

3.3.7   Error-Specific HEPs

No error-specific HEPs are provided. The computational form applies to any generic or specific error mode identified by the analyst.

3.3.8   Task Dependencies and Recovery

No procedures are provided for addressing task dependency or recovery.

3.3.9   HEP Uncertainty Bounds

HEP uncertainty bounds are determined by adjusting the weights and states of the PSFs.

3.3.10   Level of Knowledge Required

About one year of experience in the HRA field is believed to be a sufficient level for using SLIM.

3.3.11   Validation

3.3.11.1   Error Identification

SLIM does not provide explicit guidelines for error identification.

3.3.11.2   HEP Quantification

No known effort has been undertaken to empirically establish validity of the HEPs calculated by SLIM. A number of researchers have questioned the validity of the log-linear nature of the relation between PSFs and HEPs. SLIM assessment of HEPs requires a minimum of two HEP reference points that have similar task characteristics as the task being analyzed. Ideally it needs three calibration points due to possible 'inversion' of the calibration (Kirwan, 1994).

3.3.12   Reproducibility

3.3.12.1   Error Identification

SLIM does not provide explicit guidelines for error identification.

3.3.12.2   HEP Quantification

Reproducibility of analyses using SLIM is not very high since the method does not provide a standard or suggested list of PSFs. The reproducibility is also highly dependent on the anchor points used. The authors of the method maintain that the tasks for which HEPs are being generated should be from the same general task group. However, the determination of "general task group" is left for expert judgment. As a result, the reproducibility is rated low.

3.3.13   Sensitivity

The equation used in SLIM to calculate HEPs based on the SLI is highly sensitive at lower HEPs (e.g., success probabilities close to one). Therefore, an alternative index, the Failure Likelihood Index (FLI), has been used in some applications. Such practice moves the sensitive region from the lower bound HEP to the upper bound HEP. Since the HEPs of most tasks of interest are closer to zero than 1, the HEP values would not be very sensitive to the value of FLI.

3.3.14   Experience Base

SLIM and Failure Likelihood Index Methodology (FLIM) have been mostly used in nuclear plant PRAs by ABS Consulting™ (formerly PLG, Inc., an American consultancy firm). There are, however, some differences between PLG's™ implementation of SLIM and the 'original' SLIM implementation.

3.3.15   Resource Requirements

Applying SLIM requires a group of experts to identify the relevant set of PSFs, assess the weights and states of these factors, and find the appropriate anchor points for HEPs. The level of resources required is therefore rated high. Once this is done, the procedure for calculating HEPs is straight forward. A software pack, SLIM-MAUD, was developed to facilitate the process of quantitative calculation; however, it still requires experts' input on PSFs with corresponding weights.

3.3.16   Cost and Availability

Documents describing SLIM are publicly available. The method does not offer a data base. There is no licensing fee for the use of SLIM and the entire analysis can be conducted manually. However, the use of analytic hierarchy process (AHP) to determine PSF weights is typically easier with software. We note that when the number of factors to be considered is high, AHP analysis can be labor intensive.

3.3.17   Suitability for NASA Applications

The flexibility that SLIM provides for use of PSFs may be useful for adaptation to NASA missions; however, because SLIM does not provide guidance or definition of PSFs it may be more difficult to use for people who are not expert in HRA.

In general, since SLIM is an expert judgment-based HEP quantification framework, it is flexible for use in any industry including aerospace. However, SLIM requires the HEPs of some reference tasks in order to calculate the HEPs of new tasks. Identifying the reference tasks and obtaining credible HEPs for them may require significant work, both for existing and new aerospace designs. Development of such reference values may not be feasible within a short limit of time.

3.4 Cognitive Reliability and Error Analysis Method (CREAM)

3.4.1 Developmental Context

CREAM (Hollnagel, 1998) was developed for general applications and is based on the Contextual Control Model (COCOM, Hollnagel, 1993), which, from the information processing perspective, has emphasized the identification and quantification of so-called "genotype errors" (or cognitive errors).

3.4.2 Screening

CREAM provides a two-level approach to calculate HEPs: the *basic method* and the *extended method.* The basic method is designed for task screening. It provides simple rules to determine the HEP range for a task based on the combined PSFs states.

By applying the nine PSFs' values/states assessed in Table 16 , the type of "control mode" can be determined with use of Figure 8 . The HEP ranges for these four of control modes are:

- 5E-6 < HEP(Strategic) < 1E-2

- 1E-3 < HEP(Tactical) < 1E-1

- 1E-2 < HEP(Opportunistic) < 5E-1

- 1E-1 < HEP(Scrambled) < 1.

Table 16 . The CREAM PSFs and their influence on operators' performance.

CPCs	PSF State	Expected Effect on Performance Reliability
Adequacy of Organization	Very Efficient	Improved
	Efficient	Not significant
	Inefficient	Reduced
	Deficient	Reduced
Working Conditions	Advantageous	Improved
	Compatible	Not significant
	Incompatible	Reduced
Adequacy of MMI and operational support	Supportive	Improved
	Adequate	Not significant
	Tolerable	Not significant
	Inappropriate	Reduced
Availability of procedures/plans	Appropriate	Improved
	Acceptable	Not significant
	Inappropriate	Reduced
Number of simultaneous goals	Fewer than capacity	Not significant
	Matching current capacity	Not significant
	More than capacity	Reduced
Available time	Adequate	Improved
	Temporarily inadequate	Not significant
	Continuously inadequate	Reduced
Time of day	Day-time	Not significant
	Night time	Reduced
Adequacy of training and preparation	Adequate, high experience	Improved
	Adequate, limited experience	Not significant
	Inadequate	Reduced
Crew collaboration quality	Very efficient	Improved
	Efficient	Not significant
	Inefficient	Not significant
	Deficient	Reduced
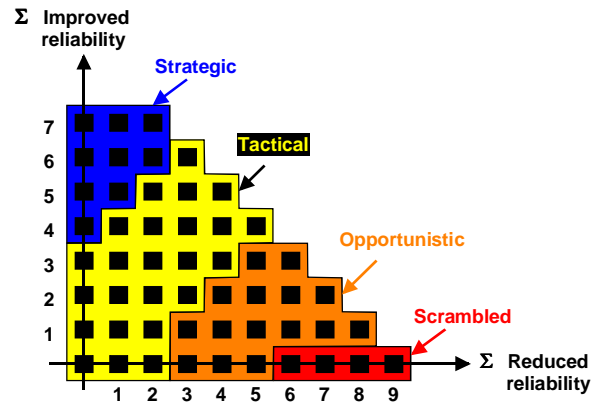
Figure 8 . The four control modes of CREAM.

3.4.3   Task Decomposition

CREAM identifies fifteen basic tasks (see Table 17) to decompose the human activities of interest.

3.4.4   PSF List and Causal Model

The CREAM method identifies a list of nine Common Performance Conditions (CPCs) (similar to PSFs) (Table 16) that could affect HEPs. These CPCs are:

- Adequacy of organization

- Working conditions

- Adequacy of MMI and operational support

- Availability of procedures/plans

- Number of simultaneous goals

- Available time

- Time of day

- Adequacy of training and preparation

- Crew collaboration quality

For retrospective analyses, a number of tables are provided that allow the analyst to trace back the root causes. The search scheme starts at the observable errors including actions:

- At wrong time,

- At wrong duration,

- Of wrong force,

- Of wrong distance/magnitude,

- Of wrong speed,

- Of wrong direction,

- At wrong object,

- In wrong sequence.

The analyst first identifies the type of error from the above list then follows the instructions provided to search the proximate causes and root causes. These causes cover the categories of:

- Observation (e.g., observation missed)

- Interpretation (e.g., faulty diagnosis)

- Planning (e.g., inadequate plan)

- Temporary person-related functions (e.g., memory failure and fear)

- Permanent person-related functions (e.g., cognitive style)

- Equipment failure

- Procedure

- Temporary interface problems (e.g., access limitation)

- Permanent interface problems (e.g., mislabeling)

- Communication

- Organization (e.g., maintenance failure and inadequate control)

- Training (e.g., insufficient skills)

- Ambient conditions (e.g., temperature)

- Working conditions (e.g., excessive demand).

3.4.5 Coverage

The CREAM PSFs for retrospective analysis (see Section 3.4.2) cover the areas of ergonomic, cognitive, and organizational factors.

3.4.6 HEP Calculation Procedure

The CREAM extended method is used for performing more detailed HEP assessments. The extended procedure includes the following steps:

1. Describe the task or task segments to be analyzed and perform task decomposition that breaks the task into a number of subtasks. Each subtask can be matched to one of fifteen pre-specified cognitive activities (see Table 17).

2. Identify the type of cognitive activity for each subtask.

3. Identify the associated human function of each subtask. Four types of human functions are identified: Observation, Interpretation, Planning, and Execution.

4. Determine the basic HEPs for all subtasks. A number of failure modes are identified. Each failure mode is associated with a basic HEP and uncertainty bounds (Table 18). The uncertainty bounds are shown in Table 20).

5. Determine the PSFs' effects on the subtasks' HEPs. Adjust the basic HEPs by multiplying by the adjustment factors based on the identified states of the PSFs (see Table 19).

6.   Calculate the task HEP based on the HEPs of subtasks.

Table 17 . The fifteen cognitive activities according to CREAM.

Cognitive Activity	General Definition
Coordinate	Bring system states and/or control configurations into the specific relation required to carry out a task or task step. Allocate or select resources in preparation for a task/job, calibrate equipment, etc.
Communicate	Pass on or receive person-to-person information needed for system operation by verbal, electronic, or mechanical means. Communication is an essential part of management.
Compare	Examine the qualities of two or more entities (measurements) with the aim of discovering similarities or differences. The comparison may require calculation.
Diagnosis	Recognize or determine the nature or cause of a condition by means of reasoning about signs or symptoms or by the performance of appropriate tests. "Diagnosis" is more thorough than "identify."
Evaluate	Appraise or assess an actual or hypothetical situation, based on available information without requiring special operations. Related terms are "inspect" and "check."
Execute	Perform a previously specified action or plan. Execution comprises actions such as open/close, start/stop, fill/drain, etc.
Identify	Establish the identity of a plant state or subsystem (component) state. This may involve specific operations to retrieve information and investigate details. "Identify" is more thorough than "evaluate."
Maintain	Sustain a specific operational state. (This is different from *maintenance* that is generally an off-line activity)
Monitor	Keep track of system states over time, or follow the development of a set of parameters
Observe	Look for or read specific measurement values of system indications.
Plan	Formulate or organize a set of actions by which a goal will be successfully achieved. Plan may be short-term or long-term.
Record	Write down or log system events, measurements, etc.
Regulate	Alter speed or direction of a control (system) in order to attain a goal. Adjust or position components or subsystems to reach a target state.
Scan	Quick or speedy review of displays or other information source(s) to obtain a general impression of the state of a system/subsystem.
Verify	Confirm the correctness of a system condition or measurement, either by inspection or test. This also includes the feedback from prior operations.

Table 18 . Matrix for determining the HEPs of CREAM cognitive activities.

Type of Functional Failure BHEP Type of HSI Activity	Type of Human Function												
	Observation			Interpretation			Planning		Execution				
	O1	O2	O3	I1	I2	I3	P1	P2	E1	E2	E3	E4	E5
	1E-3	3E-3	3E-3	2E-1	1E-2	1E-2	1E-2	1E-2	3E-3	3E-3	5E-4	3E-3	3E-2
Coordinate							▓	▓	▓	▓	▓	▓	▓
Communicate									▓	▓	▓	▓	▓
Compare				▓	▓	▓							
Diagnose				▓	▓	▓	▓	▓					
Evaluate				▓	▓	▓	▓	▓					
Execute									▓	▓	▓	▓	▓
Identify				▓	▓	▓							
Maintain							▓	▓	▓	▓	▓	▓	▓
Monitor	▓	▓	▓	▓	▓	▓							
Observe	▓	▓	▓										
Plan							▓	▓					
Record				▓	▓	▓			▓	▓	▓	▓	▓
Regulate	▓	▓	▓						▓	▓	▓	▓	▓
Scan	▓	▓	▓										
Verify	▓	▓	▓	▓	▓	▓							

*shaded cells are the possible types of human errors
BHEP: Basic human error probability

O1: Wrong object observed   O2: Wrong identification   O3: Observation not made
I1: Faulty diagnosis   I2: Decision error   I3: Delayed interpretation
P1: Priority error   P2: Inadequate plan
E1: Action of wrong type   E2: Action at wrong time   E3: Action on wrong object   E4: Action out of sequence
E5: Miss action

3.4.7   Error-Specific HEPs

The extended CREAM method identifies error modes of the four information processing phases: observation, interpretation, planning, and execution. The specific error modes within each of these four phases are:

1.   Observation:

    a.   O1: Wrong object observed.

    b.   O2: Wrong identification.

    c.   O3: Observation not made.

2.   Interpretation:

    a.   I1: Faulty diagnosis.

    b.   I2: Decision error.

    c.   I3: Delayed interpretation.

3.   Planning:

    a.   P1: Priority error.

    b.   P2: Inadequate plan.

4.   Execution:

    a.   E1: Action of wrong type.

    b.   E2: Action at wrong time.

    c.   E3: Action on wrong object.

    d.   E4: Action out of sequence.

    e.   E5: Miss action.

3.4.8   Task Dependencies and Recovery

CREAM does not provide a specific procedure for identifying and accounting for task or error dependencies. Similarly, error recovery is not explicitly discussed.

3.4.9   HEP Uncertainty Bounds

Uncertainty bounds are assigned to each basic error mode as shown in Table 20 ; however, no guidance is provided on how the uncertainty changes with the assignment of the Common Performance Conditions.

3.4.10   Level of Knowledge Required

About one year of experience in the HRA field is believed to be a sufficient level for an analyst to follow and apply the CREAM method in short period of time.

3.4.11   Validation

3.4.11.1   Error Identification

CREAM does not provide an explicit procedure for error identification.

3.4.11.2   HEP Quantification

The parameter values used in CREAM are derived from four other HRA methods (Williams, 1989; Swain & Guttman, 1983; Swain, 1987; Gertman & Blackman, 1993) and expert judgment. Thus, the validity of CREAM is strongly dependent on the credibility of those HRA methods and the quality of judgments made in their selection. No known empirical validation of the qualitative and quantitative results of CREAM has been conducted.

3.4.12   Reproducibility

3.4.12.1   *Error Identification*

CREAM does not provide an explicit procedure for error identification. (The generic task characterization and error taxonomy are indirect aids to the analyst for this purpose.)

3.4.12.2   *HEP Quantification*

CREAM provides a list of basic human activities. Decomposing the analysis into a limited set of subtasks defined by the basic human activities is relatively straightforward, and reproducibility is high. Other factors contributing to result reproducibility are that CREAM allows for a prescribed but more detailed specification of information processing phase and assessment of PSFs states.

Table 19 . The PSFs' coefficients for adjusting basic HEPs.

PSF	PSF State	Type of Human Function			
		OBS	INT	PLAN	EXE
Adequacy of Organization	Very Efficient	1.0	1.0	0.8	0.8
	Efficient	1.0	1.0	1.0	1.0
	Inefficient	1.0	1.0	1.2	1.2
	Deficient	1.0	1.0	2.0	2.0
Working Conditions	Advantageous	0.8	0.8	1.0	0.8
	Compatible	1.0	1.0	1.0	1.0
	Incompatible	2.0	2.0	1.0	2.0
Adequacy of MMI and operational support	Supportive	0.5	1.0	1.0	0.5
	Adequate	1.0	1.0	1.0	1.0
	Tolerable	1.0	1.0	1.0	1.0
	Inappropriate	5.0	1.0	1.0	2.0
Availability of procedures/plans	Appropriate	0.8	1.0	0.5	0.8
	Acceptable	1.0	1.0	1.0	1.0
	Inappropriate	2.0	1.0	5.0	
Number of simultaneous goals	Fewer than capacity	1.0	1.0	1.0	1.0
	Matching current capacity	1.0	1.0	1.0	1.0
	More than capacity	2.0	2.0	5.0	2.0
Available time	Adequate	0.5	0.5	0.5	0.5
	Temperately inadequate	1.0	1.0	1.0	1.0
	Continuously inadequate	5.0	5.0	5.0	5.0
Time of day	Day-time	1.0	1.0	1.0	1.0
	Night time	1.2	1.2	1.2	1.2
Adequacy of training and preparation	Adequate, high experience	0.8	0.5	0.5	0.8
	Adequate, low experience	1.0	1.0	1.0	1.0
	Inadequate	2.0	5.0	5.0	2.0
Crew collaboration quality	Very efficient	0.5	0.5	0.5	0.5
	Efficient	1.0	1.0	1.0	1.0
	Inefficient	1.0	1.0	1.0	1.0
	Deficient	2.0	2.0	2.0	5.0

Table 20 . Uncertainty bounds for HEPs according to CREAM.

Cognitive Function	Generic Failure Type	Lower Bound (5 percentile)	Basic Value	Upper Bound (95 percentile)
Observation	O1. Wrong object observed	3.0E-4	1.0E-3	3.0E-3
	O2. Wrong identification	1.0E-3	3.0E-3	9.0E-3
	O3. Observation not made	1.0E-3	3.0E-3	9.0E-3
Interpretation	I1. Faulty diagnosis	9.0E-2	2.0E-1	6.0E-1
	I2. Decision error	1.0E-3	1.0E-2	1.0E-1
	I3. Delayed interpretation	1.0E-3	1.0E-2	1.0E-1
Planning	P1. Priority error	1.0E-3	1.0E-2	1.0E-1
	P2. Inadequate plan	1.0E-3	1.0E-2	1.0E-1
Execution	E1. Action of wrong type	1.0E-3	3.0E-3	9.0E-3
	E2. Action at wrong time	1.0E-3	3.0E-3	9.0E-3
	E3. Action on wrong object	5.0E-5	5.0E-4	5.0E-3
	E4. Action out of sequence	1.0E-3	3.0E-3	9.0E-3
	E5. Missed action	2.5E-2	3.0E-2	4.0E-2

3.4.13   Sensitivity

The CREAM framework provides a relatively stable HEP output. The sensitivity to the change of context parameters is rated low.

3.4.14   Experience Base

CREAM has been applied in a number of HRAs in various industries including nuclear power, off-shore oil drilling, chemical process, and aerospace. The latter includes application in NASA Space Shuttle and International Space Station PRAs.

3.4.15   Resource Requirements

Use of CREAM for HEP calculation requires identifying human-system interactions (see Table 18) and assessing their failure probabilities. The required effort is rated medium, requiring anywhere between a few days and a few weeks of training. No computer code is necessary.

3.4.16   Cost and Availability

The CREAM-related publications are available in open literature. The method provides the values for needed parameters, and no licensing fee is required.

3.4.17   Suitability for NASA Applications

The CREAM analysis units are "basic human activities" which are generic in nature. As a result, at the level of task description consistent with such basic human activities, the method can be applied to existing aerospace designs for both normal and emergency operations. For new aerospace design, since detailed task

information is not available, CREAM's basic method could be used for screening purposes. The CREAM basic HEP calculation method provides HEP ranges for four control modes (see Section 3.4.6). The nine common performance conditions identified in CREAM need to be expanded to include the PSFs experienced in the zero gravity and microgravity environments.

As stated earlier CREAM has been used in two recent NASA PRAs (Space Shuttle HRA and an earlier version of the International Space Station). Results of those HRAs have not yet been publicly released and the applicability of CREAM HEPs to NASA specific tasks is still under review.

One of the workshop participants with extensive HRA experience, including HRA analysis of the Space Shuttle PRA, stated that in her experience CREAM was relatively easy to use and that in many respects the method seemed to be a good complimentary/supplemental approach to THERP.

3.5   Human Error Assessment and Reduction Technique (HEART)

3.5.1   Developmental Context

HEART (Williams, 1986; 1988) was adopted for use in a number of PRAs performed in the United Kingdom nuclear power plants in the early 1990's. Its approach to HEP assessment differs from methods that require task decomposition. "Generic tasks" are defined with corresponding basic HEPs. Each generic task is described by a few sentences that specify the nature of the human action and its context. In order to determine a base HEP, the analyst must first identify the generic task that provides the closest match to the task of interest. Such an approach greatly reduces the effort required for calculating HEP.

3.5.2   Screening

HEART does not provide an explicit procedure for screening. However, since applying the method is relatively easy and does not require significant resources, the entire method can be used in screening.

3.5.3   Task Decomposition

HEART does not provide an explicit procedure for task decomposition. Instead, nine generic tasks (see Table 21) are specified for the analyst to identify the best-matched generic task for the task of interest.

3.5.4   PSF List and Causal Model

HEART provides a long list of PSFs (see Table 22) that can be used to modify the basic HEPs. These PSFs are treated as independent. No causal model is provided for the identification of "root causes" of the listed PSFs (proximate causes).

3.5.5   Coverage

HEART PSFs cover ergonomic, cognitive, and organizational factors.

3.5.6   HEP Calculation Procedure

The following summarize the HEART steps for HEP assessment:

1.   Identify the most appropriate task description (from a list of Generic Tasks; Table 21) for the task to be analyzed.

     Table 21 shows nine Generic Tasks with their corresponding basic HEPs and uncertainty bounds. The basic HEP values apply to these generic tasks when they are performed in "perfect" conditions. The HEPs are adjusted using steps 2 through 4 when the generic tasks are performed in less than perfect conditions.

2.   Identify all of the applicable Error-Producing Conditions (EPCs) from the provided list. Thirty-eight EPCs are identified with corresponding multipliers. These EPCs and their corresponding weights are shown in Table 22 .

3.   Assess the state of the EPCs by assigning a value ranging between 0 (best, positive) to 1 (worst, negative).

4. Calculate the final HEP using Equation 3:

$$Final\ HEP = Basic\ HEP \times \prod_{i=1}^{n}\left[(Effect_{EPC\_i} - 1) \times State_{EFC\_i} + 1\right] \qquad \text{(Eq. 3)}$$

Table 21 . HEART's nine generic tasks and corresponding basic HEPs and uncertainty bounds.

General Tasks	Basic HEP	5th – 95th Percentiles
(A) Totally unfamiliar, performed at speed with no real idea of likely consequences	0.55	0.35 – 0.97
(B) Shift or restore system to a new or original state on a single attempt without supervision or procedures	0.26	0.14 – 0.42
(C) Complex task requiring high level of comprehension and skill	0.16	0.12 – 0.28
(D) Fairly simple task performed rapidly or given scant attention	0.09	0.06 – 0.13
(E) Routine, highly-practiced, rapid task involving relatively low level of skill	0.02	7E-3 – 4.5E-2
(F) Restore or shift a system to original or new state following procedures, with some checking	3E-3	8E-4 – 7E-3
(G) Completely familiar, well-designed, highly practiced, routine task occurring several times per hour, performed to highest possible standards, by highly-motivated, highly-trained and experienced person, totally aware of implications of failure, with time to correct potential error, but without the benefit of significant job aids	4E-4	8E-5 – 9E-3
(H) Respond correctly to system command even when there is an augmented or automated supervisory system providing accurate interpretation of system state	2E-5	6E-6 – 9E-4
(I) Miscellaneous task for which no description can be found	3E-2	8E-3 - .11

Table 22 shows the PSFs used in HEART (termed Error Producing Conditions) and the corresponding "weight factors."

Table 22 . HEART Error Producing Conditions, weight factors, and remedial measures.

Error Producing Contexts (multiplicative weight factor shown in parentheses)		Remedial Measure
1	Unfamiliarity (×17)	Train operators are to be aware of infrequently-occurring conditions, simulate such situations, and teach an understanding of the consequences.
2	Time Shortage (×11)	Management must be aware that shortage of time is likely to impair the reliability of decisions, both their own and their staff's, and try to ensure that sensitive decisions are not made against the clock.
3	Low Signal/Noise ratio (×10)	Strenuous efforts must be made to ensure that such ratios do not fall to unreasonably low levels.

Error Producing Contexts (multiplicative weight factor shown in parentheses)		Remedial Measure
4	Features override allowed (×9)	If the consequence of placing a system in an inappropriate state is potentially damaging, suitable inter-locking and inhibition must be provided, together with any suitable time-outs to return features to their appropriate quiescent state.
5	Spatial and functional incompatibility (×8)	Such incompatibilities should not occur. Sufficient information is now known about human engineering and population stereotypes that the problem need not arise to any extent. Where information about functional compatibility is needed, advice should be obtained from trained ergonomists, who will either know how to arrange a design for spatial or functional compatibility, or how to run an appropriate experiment to find out what is required.
6	Model mismatch (×8)	Designers of systems and equipment are not always right. Operators sometimes often have better ideas and possess views about how a system should function which are contrary to those of system designers. Under pressure, particularly, operators will revert to their own perceptions of how a system should function, often with undesirable consequences. To protect against such mismatches, systems designers must try to find out what their users' expectations are, and then design these characteristics into the system, omitting their own prejudices, as they do so.
7	Irreversibility (×8)	Obvious means should be provided to ensure that errors can be reversed easily, with preference for means of reversing by the actions which created the error in the first place.
8	Channel overload (×6)	It should never be necessary to monitor more than one information channel at any one time. Single events should not occur at more than three per second.
9	Technique unlearning (×6)	The greatest possible care should be exercised when new techniques are being considered to achieve the same outcome. They should not involve adoption of opposing philosophies.
10	Knowledge transfer (×5.5)	Reliance should not be placed on operators' transferring their previous knowledge without loss of precision and meaning. If such perfect transfer is required, suitable job aid must be made available for reference.
11	Performance ambiguity (×5)	The required performance standards must be tested for comprehensibility on the user population to ensure that there is no ambiguity.
12	Misperception of risk (×4)	It must not be assumed that a user's perception of risk is the same as the actual level. If necessary a check should be made to ascertain where any mismatches might exist and the extent of the mismatches.
13	Poor feedback (×4)	A task analysis will show the points at which feedback must be available to operators. Ergonomists can advise on the best form of feedback if doubts should arise; what one is looking for is complete "system transparency."
14	Delayed/incomplete feedback (×4)	System response times should never exceed four seconds, and there must always be sufficient information to enable operators to step confidently on to the next part of a task. If doubt exists the feedback is incomplete.
15	Inexperienced (×3)	Personnel criteria should contain specified experience parameters thought relevant to the task. Chances must not be taken for the sake of expediency.
16	Impoverished information (×3)	Procedures should be human-engineered and treated for operability. It should be assumed that when personnel are required to communicate with each other that very considerable information loss will occur. Procedures must not rely on accurate verbal transmission of information for success.

Error Producing Contexts (multiplicative weight factor shown in parentheses)		Remedial Measure
17	Inadequate checking (×3)	When high reliability is paramount, independent checks on accuracy should be made, by people and systems that do not have any vested interest in the success or failure of an individual. Blame should not be attached to any inadequacies found at this level.
18	Objectives conflict (×2.5)	Objectives should be tested by management for mutual compatibility, and where potential conflicts are identified, these should either be resolved to make them harmonious or made prominent so that a program can be created to reconcile such conflicts in a rational fashion when they arise.
19	No diversity (×2.5)	It should not be assumed that operators will rely totally on a single information source for confirmation of accuracy, and inquiries should be made to ascertain what additional sources are referred to, so that these sources are not denied by operators, and, if possible are enhanced.
20	Educational mismatch (×2)	The job profile should identify any potential mismatch of recruits against requirements. Educational standards should be made explicit; there should be no ambiguity.
21	Dangerous Incentives (×2)	It is intuitively obvious that people work for rewards of various natures. If the reward for doing something quickly is greater than the reward for doing it accurately, or the reward for omitting an action is greater then the reward for performing it, we should not be surprised if that is what happens. The reward system must be evaluated carefully, therefore, to ensure that the desired behavior is emitted, rather than that which might be constructed as being appropriate simply because facets of the task are seen to conform to a partial criterion. If in doubt, seek advice from management scientists and/or psychologists.
22	Lack of exercise (×1.8)	Frequent rest breaks should be designed into the job, and the system made tolerant to personnel taking breaks as the need arises. Tuition should be given in techniques for maintaining high levels of arousal, such as postural change, personal ventilation, and recognition of fatigue symptoms. Encouragement should be given to engage in appropriate mild forms of physical exercise and relaxation and stress control. On-the-job refresher training and frequent exercises to maintain and enhance levels of competence and awareness of technical progress innovation should be given.
23	Unreliable Instruments (×1.5)	When instrumentation is found to be unreliable, operators will cease to trust its indications to the extent of ignoring valid information, preferring to believe their own interpretations, despite overwhelming evidence to the contrary. If instrumentation is thought likely to be unreliable it should be withdrawn from service and more reliable instrumentation substituted. No doubts should exist about its suitability.
24	Absolute judgments (×1.6)	Operators must not be placed in the position of having to make judgments about the meaning of data which are outside their span of comprehension or experience - a task analysis will reveal when such conditions are likely to arise, and management must plan for such contingencies by recognizing the circumstances and taking full responsibility for actions which might be taken on their behalf. "Brainstorming" and problem-solving workshops are helpful to identify some of the most bizarre situations in which staff and management can find themselves. It is likely that discussion of these 'gray areas' of organizational behavior will reinforce mutual respect and anticipate future conflict and/or issues of culpability at a time of zero threat.

Error Producing Contexts (multiplicative weight factor shown in parentheses)		Remedial Measure
25	Unclear allocation of function (×1.6)	As with the area above, doubt must not exist about responsibilities. Whilst they can, and should, be stated on paper, joint preparation of a functional specification will remove doubts and anxieties, and lead to the development of healthy attitudes towards the system design concepts. Organizational development specialists and/or behavioral scientists should be involved in facilitating the preparation of a satisfactory working protocol.
26	Progress tracking lack (×1.4)	Various job aids must be supplied in order to ensure that operators do not get out of step with the task in hand. These can range from checklists through mimics to electronic monitoring of progress against targets. If such aids are introduced they must be piloted to ensure that they are compatible with user needs and that there is an incentive to use them. Agronomists can advise on these job design aspects.
27	Physical capabilities (×1.4)	It should be self-evident that tasks must not exceed the operators' capabilities. Reference to human factors standards will ensure that these capabilities are not exceeded.
28	Low meaning (×1.4)	Meaning can be built into a job by preparing job descriptions with the staff concerned, showing them the significance of their contribution to corporate objectives, designing variety into their duties by arranging for job features such as task rotation to enhance system awareness, and holding periodic reviews of working practices to ensure that symptoms of alienation are not manifesting themselves. Behavioral scientists can advise on suitable precautions.
29	Emotional stress (×1.3)	Management and medical staff must be vigilant to recognize the onset of emotional problems which can manifest themselves via symptoms such as excessive absence, persistent lateness, obsessive behavior, lack of cooperation, and exceptional fatigue. Personal stress control training programs could be considered and potentially stressful decision-making circumstances identified so that the conditions can be modified to limit occurrence of extreme generalized stress.
30	Ill-health (×1.2)	Until it is pointed out, it is not apparent that ill-health can have such deleterious effects on performance. Often the effects of, say, a cold or flu do not manifest themselves until well into a shift. By now it should be obvious that operators and managers who are ill should not attempt to undertake work requiring high reliability, and out of respect for others, for system integrity, and peace of' mind they should stay away until recovered. A medical awareness program would be helpful.
31	Low morale (×1.2)	Apart from the more obvious ways of attempting to secure high morale by way of financial reward, for example, other methods involving participation, trust, and mutual respect often hold out at least as much promise. Building up morale is a painstaking process which involves a little luck and great sensitivity. Employees must be given reason to believe in their employer and themselves. This can be accomplished by a battery of activities, such as joint preparation of work plans and objectives, maximal delegation of authority, reward for effort end results, provision of subsidized fringe benefits, firmness of resolve and openness. It is not achieved to any great extent by appeals to workforces to stick by management. The respect necessary to make morale rise is earned, not enforced. A sensitive, caring management would be unlikely to encounter such problems.

Error Producing Contexts (multiplicative weight factor shown in parentheses)		Remedial Measure
32	Inconsistency of display (×1.2)	Even if the conventions adopted for display layout and procedure design are not human-engineered for ease of use, they must be consistent within themselves; e.g., if a display is showing an increasing value even though in an analogue sense the portion shown is decreasing, this convention must be adhered to throughout - even though such a principle is "wrong" (such an approach would not be encouraged, of course).
33	Poor environment (×1.15)	It should be self-evident that a poor environment is likely to impair performance. By and large this should not occur nowadays because of the introduction of legislation to control environments. To minimize any deleterious effects work physiologists, ergonomists, and/or architects should be consulted for details of appropriate parameters.
34	Low loading - (×1.1) 1st half hour - (×1.05) each hour	Prolonged inactivity or highly repetitive cycling of low mental workload tasks must be avoided. Generally when signal frequency falls below two per minute or involves little or no variability, vigilance in performance will degrade. To combat such effects the introduction of artificial signals has been found to be helpful, and job enrichment (with the introduction of different, more varied tasks) has been found to minimize boredom and better hold attention. Rather than combat these effects, it is better to ensure that such conditions do not arise in the first place; e.g., observation tasks demanding high human reliability should never require sessions of longer than one hour's concentration and tasks involving very low signal frequency should not be designed (if possible such tasks should be automated).
35	Sleep cycle disruption (×1.1)	Only extreme sleep deprivation will cause performance degradation. Our major interest, therefore, is in keeping small amounts of deprivation to a minimum. this can be achieved by keeping operators on a "stable" shift system such that there are no radical changes to either the patterns or the time of day over which such changes occur. The frequency with which changeovers occur should be as low as can reasonably be achieved. Advice should be sought from work physiologists.
36	Task pacing (×1.06)	Although all work ultimately involves some element of pacing, the unwitting or deliberate introduction of pacing will lead to a slight reduction in reliability. This can be avoided by checking work systems to ensure that there is sufficient 'buffering' such that operators are not subject to undue pressure and can work at their own preferred pace - the one which best matches their capability.
37	Supernumeraries	Where possible, limit gatherings of staff at workplaces to those necessary to perform tasks satisfactorily.
38	Age (×1.02)	Monitor perceptual capabilities of personnel required to perform task demanding high acuity and accurate information processing.

3.5.7   Error-Specific HEPs

HEART calculates HEPs for "errors" associated with the genetic tasks. There are no specific error modes attached to these HEPs.

3.5.8   Task Dependencies and Recovery

The effects of task dependencies and recoveries are implicitly embedded in the definitions of the generic tasks.

3.5.9 HEP Uncertainty Bounds

Uncertainty bounds are assigned to each generic task (see Table 21). HEART does not provide instructions on how the uncertainty bounds might change when the Error Producing Conditions are assigned.

3.5.10 Level of Knowledge Required

About one year of experience in the HRA field is believed to be a sufficient level for an analyst to learn the HEART method in a short period of time.

3.5.11 Validation

3.5.11.1 Error Identification

HEART does not provide an explicit procedure for error identification.

3.5.11.2 HEP Quantification

Two validation exercises have been reported, both in relation to nuclear power plant operation, one by the author of HEART, and one by British Nuclear Fuels LLC (BNFL). Both validations reached similar conclusions: "Three basic conclusions can be drawn from these studies. The first is that, as intended, assessments tend to be conservative, e.g., assessed probabilities of failure tend to be slightly higher than are observed in practice. The second is that the precision achieved may be judged as 'reasonable' with 70% of assessments falling within a factor of 10 of the measured value and 85%+ falling within a factor of 100. The third finding is that the longer the period of training, the more precise will be the assessments. For assessors given 2 hours of training the precision they may be expected to achieve can be of the order of 80% within a factor of 10 and 95+% with a factor of 100." (Williams, 1988).

"The proposed EPCs and 'data' cannot be regarded as definitive but they are derived from a number of sources, and their reliability is based on that which has been observed in experimental and epidemiological studies." (Williams, 1988)

3.5.12 Reproducibility

3.5.12.1 Error Identification

HEART does not provide explicit procedures for error identification.

3.5.12.2 HEP Quantification

The key factor affecting result reproducibility is the ability of the analyst to identify the most appropriate "generic task" for the task of interest. Generic task descriptions by definition require the analyst's assessment of applicability to the specific situation of interest. Therefore there is the possibility that different analysts will select a different generic task (and therefore different HEPs) for the same task. This problem is more serious in applications to the design of future NASA missions where tasks are not detailed. Thus, the reproducibility for HEART is rated low.

3.5.13 Sensitivity

One significant weakness of the HEART method arises from its dependence on accurate identification of generic tasks. Small changes in the context description are likely to result in identification of a different generic task for the task of interest. In addition, some error producing contexts carry a heavy weight in terms

of impact on HEP values. Small changes in the states of these error producing contexts could also change the HEP values significantly. As a result, the sensitivity of HEART method is rated high.

3.5.14   Experience Base

HEART has been applied in a number of domains. "Commercial evaluations of HEART have produced generally favorable comments especially from the chemical industry and weapons manufacturers. The nuclear industry by way of contrast has thus far reserved judgment, citing technique immaturity, non-publication of the data-base and the theoretical justification of the method as some of the reasons why endorsement should be withheld." (Williams, 1988).

3.5.15   Resource Requirements

Use of HEART does not require task decomposition. The HEP assessment process is straightforward once the task mapping is done. No specific computational tool is required. The required resource is rated low.

3.5.16   Cost and Availability

HEART methodology is well documented in publicly available literature.

3.5.17   Suitability for NASA Applications

HEART does not require detailed task-related information for calculating HEPs. This characteristic and the simplicity of use make HEART appealing for application to new aerospace designs. The analyst does need some situation-specific information in order to identify the Error Producing Conditions (i.e., PSFs) which are used to adjust the HEPs. The HEART approach is also suitable for existing aerospace designs if the level of detail offered by "generic tasks" adequately corresponds to the task being analyzed. Some issues are:

1. Only 9 generic tasks are identified. This is not enough to cover all aerospace human activities. The high specificity of some of the generic task types may make exact assignment of tasks difficult. Questions are:

    a. Can the generic tasks defined in HEART be adapted for NASA needs?

    b. If the generic tasks do not adapt well, can a good set of generic tasks and corresponding high fidelity data be defined for NASA use?

2. The PSFs need to be expanded to cover the space mission environments and tasks.

3. The weights of existing PSFs need to be calibrated to NASA data.

4. The relevance of the data behind HEART to space mission applications is a key concern.

3.6   Nuclear Action Reliability Assessment (NARA)

3.6.1   Developmental Context

NARA (Kirwan et al., 2005) is a refinement of the HEART method to (a) have better fit to nuclear contexts, (b) consider errors of commission, (c) have substantial data support, (d) consider long time scale scenarios, and (e) have better guidance on usage. NARA uses the same approach as HEART to calculate HEPs. The main differences between NARA and HEART are (a) the grouping of the generic tasks, (b) the weights of the error producing contexts, and (c) the use of the CORE-DATA human error database in NARA.

NARA uses different weights for some of the error producing conditions than HEART. This suggests that the PSFs' weights and perhaps the basic HEPs of the general tasks of HEART and NARA need to be revisited carefully for NASA applications.

3.6.2   Screening

NARA does not provide an explicit procedure for screening. However, since applying NARA is relatively easy and does not require significant resources, the entire method can be used in screening.

3.6.3   Task Decomposition

NARA does not provide explicit guidance on task decomposition. Instead, fourteen generic tasks (Table 23(a)) are specified for the analyst to identify the best-match generic task for the task of analysis.

3.6.4   PSF List and Causal Model

NARA provides a list of Error Producing Conditions (Table 23(a)). These are equivalent to PSFs. No causal model in terms of PSFs, their interdependencies, and other causal factors is provided. The list in Table 24 is not a complete set of NARA Error Producing Conditions because some of EPCs are still under review. It is expected that the complete set of Error Producing Conditions with corresponding weights will be available in March 2006. According to the developer of NARA, the method has an error reduction module (not publicly available yet).

3.6.5   Coverage

The NARA generic tasks and EPCs cover aspects of ergonomics, cognitive, and organizational factors explicitly in some cases and implicitly in others.

3.6.6   HEP Calculation Procedure

NARA uses the same general procedure as HEART to calculate HEPs, which was discussed in the previous section and is not repeated here. The calculations are based on NARA generic tasks with corresponding basic HEPs (Table 23(a)).

Table 24 provides a partial list of NARA Error Producing Conditions (EPCs) with maximum HEP multipliers. Long duration activities (up to 24 hrs) are covered in NARA.

Table 23(a) . The generic tasks of NARA (partial list).

	Generic Task	Basic HEP
A1	Carry out simple single manual action with feedback. Skill-based and therefore not necessarily with procedure.	0.005
A2	Start or reconfigure a system from the Main Control Room following procedures, with feedback.	0.001
A3	Start or reconfigure a system from a local control panel following procedures, with feedback.	0.003
A4	Reconfigure a system locally using special equipment, with feedback; e.g., closing stuck open boiler SRV using gagging equipment. Full or partial assembly may be required.	0.03
A5	Judgment needed for appropriate procedure to be followed, based on interpretation of alarms/indications, situation covered by training at appropriate intervals.	0.01
A6	Completely familiar, well designed highly practiced, routine task performed to highest possible standards by highly motivated, highly trained, and experienced person, totally aware of implications of failure, with time to correct potential error. Note that this is a special case.	0.0001

Table 23(b) . The generic tasks of NARA for checking correct plant status and availability of plant resources.

	Generic Task	Basic HEP
B1	Routine check of plant status.	0.03
B2	Restore a single train of a system to correct operational status after test following procedures.	0.007
B3	Set system status as part of routine operations using strict administratively controlled procedures	0.0007
B4	Calibrate plant equipment using procedures; e.g. adjust set-point.	0.003
B5	Carry out analysis.	0.03

Table 23(c) . The generic tasks of NARA for alarm/indication response.

	Generic Task	Basic HEP
C1	Simple response to a key alarm within a range of alarms/indications providing clear indication of situation (simple diagnosis required). Response might be direct execution of simple actions or initiating other actions separately assessed.	0.0004
C2	Identification of situation requiring interpretation of complex pattern of alarms/indications. (Note that the response component should be evaluated as a separate GTT)	0.2

Table 23(d) . The generic tasks of NARA for communication.

	Generic Task	Basic HEP
D1	Verbal Communication of Safety-Critical Data.	0.006

The PIFs (error producing contexts) and their corresponding HEP multipliers are shown in Table 24 .

In NARA, an HEP is calculated by the following equation (Equation 4):

$$HEP_f = HEP_i \times \prod_{j=1}^{N} \left\{ \left[Weight\left(PIF_j\right) - 1 \right] \times State\left(PIF_j\right) + 1 \right\}$$

(Eq. 4)

Where N is the number of applicable PIFs, and $0 \leq State\left(PIF_j\right) \leq 1$

3.6.7 Error-Specific HEPs

Similar to HEART, NARA calculates HEPs for "errors" associated with the genetic tasks. There are no specific error modes attached to these HEPs.

3.6.8 Task Dependencies and Recovery

NARA takes a holistic approach to calculating HEPs. The task dependencies, recovery, and other factors are covered in the definition of generic tasks. In its application, it is possible to adopt the dependency modeling of other HRA methods (e.g., THERP).

3.6.9 HEP Uncertainty Bounds

NARA provides uncertainty boundaries for some tasks.

3.6.10 Level of Knowledge Required

HRA analysts with about one year of HRA experience are expected to be able to learn how to apply the method with little or no training.

Table 24 . NARA PSFs and corresponding weight factors (partial list).

NARA EPC ID	NARA EPC Description	NARA EPC Effect
1	A need to unlearn a technique and apply one which requires the application of an opposing philosophy.	24
2	Unfamiliarity; e.g., a potentially important situation which only occurs infrequently or is novel.	20
3	Time pressure.	11
4	Low signal to noise ratio.	10
5	Difficulties caused by poor shift hand-over practices and/or team coordination problems or friction between team members.	10
6	A means of suppressing or overriding information or features which is too easily accessible.	9
7	No obvious means of reversing an unintended action.	9
8	Operator inexperience.	8
9	Information overload, particularly one caused by simultaneous presentation of non-redundant information.	6
10	Poor, ambiguous, or ill-matched system feedback.	4
11	Shortfalls in the quality of information conveyed by procedures.	3

NARA EPC ID	NARA EPC Description	NARA EPC Effect
12	Operator under-load/boredom.	3
13	A conflict between immediate and long-term objectives.	2.5
14	An incentive to use other more dangerous procedures.	2
15	Poor environment.	8
16	No obvious way of keeping track of progress during an activity.	2
17	High emotional stress and effects of ill health.	2
18	Low workforce morale or adverse organizational environment.	2

3.6.11 Validation

3.6.11.1 Error Identification

NARA does not provide explicit guidance on error identification.

3.6.11.2 HEP Quantification

No validation study has been conducted to assess the quality of the HEP numbers produced by NARA. The CORE-DATA human error database provides the foundation of NARA results. However, the CORE-DATA is not publicly available. Also, the quality of the CORE-DATA has not been independently assessed. Other data from published sources were also used in determining the HEPs associated with various generic tasks. According to the developers of the method, a rather large HEP data set was initially used as the source, and subsequently screened to include the least subjective numbers to form the distribution for each of the generic tasks.

3.6.12 Reproducibility

3.6.12.1 Error Identification

NARA does not provide explicit guidance on error identification.

3.6.12.2 HEP Quantification

Like HEART, the reproducibility of NARA results is strongly dependent on whether the most appropriate generic task can be found easily. In this respect, since the organization and specification of the generic tasks in NARA is better than HEART, reproducibility is rated medium.

3.6.13 Sensitivity

As in HEART, small changes in the context description are likely to result in identification of a different generic task for the task of interest. In addition, some EPCs carry a heavy weight in terms of impact on HEP values. Small changes in the state of those EPCs could change the HEP values significantly. Therefore, the HEP estimates from NARA are very sensitive to the identification of the generic task and error producing conditions and, thus, the sensitivity of method is rated high.

3.6.14   Experience Base

No field application of NARA has been documented.

3.6.15   Resource Requirements

Use of NARA does not require task decomposition. No specific computation tool is required. The required level of effort is rated low.

3.6.16   Cost and Availability

The NARA method is available for public use. The raw data (e.g., CORE-DATA and other data gathered from the British Nuclear Industry and other sources) are proprietary and available by fee.

3.6.17   Suitability for NASA Applications

Similar to HEART, NARA does not require detailed task-related information for HEP estimation. This characteristic and the simplicity of use make NARA appealing for application to new aerospace designs. The NARA approach is also suitable for existing aerospace designs if the level of detail offered by generic tasks adequately corresponds to the task being analyzed. However, the number of NARA generic tasks is limited and most likely inadequate to cover all space mission activities. As in the case of HEART, the challenge is in adapting and extending the generic tasks for NASA applications. Similarly EPCs and weight factors need to be calibrated for space applications, and relevance of the data behind the method has to be established and/or new data and estimates need to be developed.

3.7   A Technique for Human Event Analysis (ATHEANA)

3.7.1   Developmental Context

ATHEANA is the product of a multi-phase research sponsored by the U.S. Nuclear Regulatory Commission. The initial effort started in 1992, aiming for more comprehensive coverage of operator response in the PRAs of nuclear power plants, particularly EOCs. It contains a detailed search process that promises to determine cognitive vulnerabilities in crews that may not be discovered when applying other HRA methods. The publications covering results of this research include [Barriere, Luckas et al. 1994; Barriere, Wreathall et al. 1995; Cooper, Luckas et al. 1995; Cooper, Ramey-Smith et al. 1996; and Barriere, Bley et al. 2000].

ATHEANA was designed to be a full scope HRA method including capability for performing predictive task analysis (or error identification) and retrospective event analysis. It offers a procedure to search for and identify errors based on context analysis.

3.7.2   Screening

ATHEANA does not provide rules for screening analysis.

3.7.3   Task Decomposition

ATHEANA analysis focuses on the formation and effects of Error Forcing Contexts (EFC). There is no explicit procedure to guide the analyst in task decomposition.

3.7.4   PSF List and Causal Model

ATHEANA uses the concept of EFC to characterize types of scenarios in which human errors are most likely. ATHEANA breaks EFCs into two groups -- those that are characteristic of the initiator or accident sequence and those that are characteristic of the system or function.

EFCs related to the initiator or accident sequence include:

- Short time to damage;

- Unfamiliarity;

- Single functional failure that can yield damage;

- Distraction of control room team;

- Forced independent action by one member of team;

- Potential for complex, hidden, and/or unfamiliar conditions;

- Multiple (possibly conflicting) priorities;

- Wide range of accident responses, plant dynamics/conditions represented, and relatively low-frequency events.

EFCs related to the system state or function includes

- Short time to damage;

- Single functional failure that can yield damage;

- Function needed early in accident response;

- Little or no redundancy of systems or equipment that perform a plant function;

- Dependencies between redundant systems and equipment;

- Scarcity of action cues that creates high potential for confusion and complications;

- Functional failure including irreversible plant or equipment damage with no easy recovery options; and

- Human-intensive accident response, important principally for EOCs.

The two sets of EFCs are not meant to be all encompassing; rather they serve to ensure that the expert analyst at a minimum considers those factors relative to the task of interest.

ATHEANA provides the following PSFs to guide the experts to identify the EFCs:

- Procedures

- Training

- Communication

- Supervision

- Staffing

- Human-system interface

- Organizational factors

- Stress

- Environmental conditions

- Strategic factors such as multiple conflicting goals, time pressure, limited resources.

ATHEANA relies on experts to determine the weights of various PSFs. The method does not provide a model of dependencies among these PSFs, and its causal model is the notion of error forcing context as identified by the analyst for a given analysis. ATHEANA does provide instruction for the analyst to identify dependencies between tasks.

3.7.5 Coverage

ATHEANA analysis emphasizes the identification of the situations which would cause operator unsafe actions (error forcing contexts). In principle, these should cover cognitive, organizational, and ergonomic factors (see the suggested PSFs shown in Section 3.7.4).

3.7.6 HEP Calculation Procedure

ATHEANA is an expert judgment based method. The following are the steps that guide the experts in their analysis:

1. Define and interpret the issue being analyzed.

2. Define the resulting scope of the analysis.

3. Describe the base case scenario (e.g., the nominal/expect operator's response).

4. Define human failure events (HFEs) and unsafe actions (UAs) of concern.

5. Identify potential vulnerabilities.

6. Search for deviations from base case scenarios:

    a. Search by keyword to consider types of physical deviations (e.g., larger, smaller, faster, and slower),

    b. Examine the key decision points in related procedures to see if deviation from the base scenario could lead to inappropriate actions,

    c. Search dependencies between equipment faults and support system failures which could create cognitive challenge.

7. Identify and evaluate complicating factors.

8. Evaluate the potential for recovery.

9. Interpret the results (including quantification if necessary).

10. Incorporate into the PRA (if necessary).
    Specifically for quantifying HEPs, ATHEANA uses the following three steps:

    a. Assess the probability of EFCs in the particular accident scenario of analysis; i.e., P(EFC).

    b. Assess the conditional likelihood of the UAs that can cause the human failure event; i.e., P(UA|EFC).

    c. Assess the conditional likelihood that the unsafe actions is not recovered prior to the catastrophic failure of concern; i.e., P(fail recovery|UA, EFC).

    The HEP can be calculated by Equation 5:

$$HEP = \sum P(EFC_i)P(UA_{ij} \mid EFC_i)P(\text{fail recovery} \mid UA_{ij}, EFC_i)$$ (Eq. 5)

3.7.7 Error-Specific HEPs

ATHEANA attributes errors identified in retrospective event analysis to three causes:

- Information processing,

- PSFs, and

- Significant plant conditions.

The typical failure modes of these three categories are listed below:

1.   Information processing failure:

    a.   Monitoring and detection.

        (1)   Operators unaware of actual plant state

        (2)   Operators unaware of the severity of the plant conditions

        (3)   Operators unaware of continued degradation in-plant conditions.

    b.   Situation assessment.

        (1)   Information is erroneous or misleading

        (2)   Plant indicators are misinterpreted

        (3)   Plant or equipment behavior is misunderstood

        (4)   Similarity of the event to other better-known events leads operator to form an incorrect situation model.

    c.   Response planning.

        (1)   Operators select inapplicable plans

        (2)   Operators follow prepared plans that are wrong or incomplete

        (3)   Operators do not follow prepared plans

        (4)   Prepared plans do not exist, so operators rely upon knowledge-based behavior

        (5)   Operators inappropriately give priority to one plant function over another.

    d.   Response implementation.

        (1)   Important procedural steps are missed

        (2)   Miscommunication

        (3)   Equipment failures hinder operators' ability to respond.

2.   Performance influencing factors:

   a.   Human performance capabilities at a low point

   b.   Time constraints

   c.   Excessive workload

   d.   Unfamiliar plant conditions and/or situation

   e.   Inexperience

   f.   Non-optimal use of human resources

   g.   Environmental factors and ergonomics.

3.   Significant plant conditions:

   a.   Extreme or unusual conditions

   b.   Contributing preexisting conditions

   c.   Multiple hardware failures

   d.   Transitions in progress.

The probabilities of these error modes and conditions are again estimates from expert judgment.

3.7.8   Task Dependencies and Recovery

Task and error dependencies are not explicitly addressed in ATHEANA; however, due to the flexibility of the framework, experts can always define the scenarios in terms of possibly inter-dependent tasks and consider the impact on HEP assessment. Error recovery is explicitly called out in Step 8 of the procedure (see Section 3.7.6).

3.7.9   HEP Uncertainty Bounds

Uncertainty is not specifically addressed in ATHEANA; however, the experts can assess uncertainty bounds using the same process used for assigning the nominal HEPs. A procedure for characterizing uncertainty within ATHEANA is now undergoing review by the U.S. Nuclear Regulatory Commission.

3.7.10   Level of Knowledge Required

Highly experienced HRA specialists are needed for performing analysis with use of ATHEANA.

3.7.11   Validation

The assessment of HEP values mainly relies on expert judgment. It lacks a database of suggested values to support the assessment. No known empirical validation has been conducted to assess the quality of the ATHEANA results, for both error identification and HEP quantification.

3.7.12   Reproducibility

3.7.12.1   Error Identification

ATHEANA has many ambiguous steps that make it hard to follow (see also Reer, Sträeter, Dang, & Hirschberg, 1999). The result means that reproducibility of ATHEANA is highly expert-dependent. It is expected that consensus reached by a group of experts is likely to be consistent with the conclusions reached by another group of experts as long as their judgments are based on the same information. In the absence of supporting evidence for such consistency, the reproducibility of ATHENA for error identification is rated medium.

3.7.12.2   HEP Quantification

As stated earlier the ATHEANA HEP quantification is expert opinion based, with a general form for decomposition of the HEP in terms of its key ingredients according to Equation 5. All the probabilities of Equation 5 have to be estimated by experts with no specific guidelines offered other than a quantitative scale for qualitative expressions of likelihood (e.g., 'infrequent' event translated into 1E-3 per year). The reproducibility of ATHENA HEPs is therefore rated low.

3.7.13   Sensitivity

ATHEANA provides guidelines for experts to convert qualitative frequency descriptions (e.g., 'infrequent' event) into quantitative frequencies (e.g., 1E-3 per year). Obviously different perceptions of event frequencies or applicability of PSFs could result in significant differences in the predicted numerical values for the HEPs. The sensitivity is rated medium.

3.7.14   Experience Base

The method has been used in a number of trial applications sponsored by the US NRC for nuclear power plant PRAs. ATHEANA has also been used by several teams involved in the Task 97-2 of Working Group on Risk Assessment (RISK) of Nuclear Energy Agency (NEA) (Grant, et al., 2000) including teams from United States, Japan, and The Netherlands. The lessons learned about the ATHEANA process include:

- Guidance was unclear about the relationship between unsafe actions and error forcing contexts;

- ATHEANA is resource intensive;

- Use of retrospective analysis for prospective analysis unclear;

- Quantification of the probability (Unsafe Action | Error Forcing Context) and Probability (recovery | Unsafe Action | Error Forcing Context) is not clear;

- Poorly defined terms for errors of commission;

- The method is not a toolbox. ATHEANA is currently a set of concepts and a vague procedure for how to apply them;

- The method can be made into a toolbox. It may be necessary to develop new representations to work on (e.g., error mechanism); and

- ATHEANA provided a good basis for discussions with management of the chemical facility.

The most recent extensive application of ATHEANA has been the HRA analysis of Pressurized Thermal Shock scenarios for four U.S. nuclear power plants sponsored by the U.S. Nuclear Regulatory Commission (Kirk, Malik, Santos, Dickson, Pugh, Bass, Williams, Woods, Siu, Kim, Kolaczkowski, Whitehead, Bessette, Arcieri, Fletcher, Mosleh, & Chang, 2006).

3.7.15 Resource Requirements

Performing ATHEANA requires a significant amount of joint effort and time of system experts and HRA experts. No specific computation aids are needed. The history of ATHEANA is that it has always been applied by a team; it is therefore difficult to speculate whether a single analyst could apply ATHEANA. As a result, the level of effort for applying ATHEANA is rated high.

3.7.16 Cost and Availability

ATHEANA is publicly available through the U.S. Nuclear Regulatory Commission. In practice, the greatest cost associated in carrying out an ATHEANA analysis is the cost associated with assembling a panel or committee of experts to conduct the analysis.

3.7.17 Suitability for NASA Applications

ATHEANA is an expert judgment-based HRA method. Even though it was primarily developed for the nuclear industry, its framework is suitable for application to aerospace tasks. However, its application would require significant expertise in aerospace-related tasks and system-related information. For new aerospace designs, the available information is likely not detailed enough for experts to make credible judgments. Since the ATHEANA guidelines for searching for error forcing contexts were developed for nuclear operations, new guidelines would need to be developed for aerospace tasks. This is expected to require significant effort.

A workshop participant, one of the ATHEANA developers, commented that "ATHEANA is weak in lack of a broad pool of practitioners familiar with the method and the lack of immediately available supporting documentation."

The suitability of ATHEANA for NASA applications is therefore rated low.

3.8   Connectionism Assessment of Human Reliability (CAHR)

3.8.1   Developmental Context

CAHR (Sträeter, 2000; 2004) was developed to be "a method for a systematic evaluation of events with a view to human errors which at the same time will make it possible to build up an empirical database for reliability parameters." (Sträeter, 2000). The original purpose was to improve the basis for HRA process and to provide a knowledge base for human failure events, failure modes, and quantitative assessment. Recently a CAHR-based prospective Human Reliability Assessment process (Sträeter, 2005) was also developed.

CAHR was initially developed and populated with failure event data obtained from operating events that occurred in German nuclear power plants. Later it was also applied to the German automotive industry, the maritime environment, aviation, and air traffic management. Recently it was applied for the assessment in the early conceptual development phase of the European operational concept for 2020 (Trucco, Leva, & Sträeter, 2006) (provided by the author of CAHR).

3.8.2   Screening

CAHR does not provide explicit guidelines for screening. However, since the use of the method is computerized (CAHR uses key word search to obtain HEPs), the required effort is minor. Its quantification method can be also used in a screening process.

3.8.3   Task Decomposition

CAHR uses a structured Man-Machine System (MMS; Figure 9) as the analysis block to represent a task. A MMS contains the possible interaction paths: within an operator (e.g., cognitive activities), between operators, between an operator and the system, and between an operator and the environment. Each human activity is represented by a highlighted activity specified in a MMS. Thus, the sequence of a task can be represented by a number of MMSs.

3.8.4   PSF List and Causal Model

CAHR identifies 30 PSFs distributed into 6 major groupings (Table 25). No explicit causal model is provided.

3.8.5   Coverage

CAHR event analysis based on the MMS framework covers ergonomic, cognitive, and organizational factors.
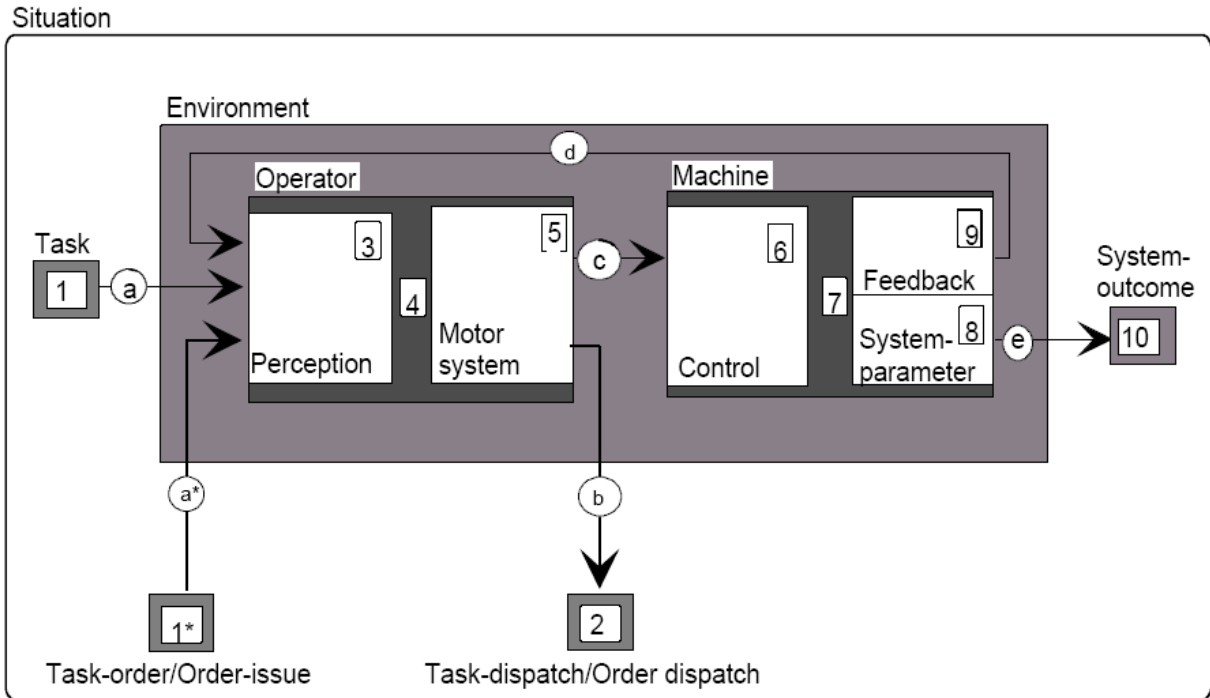
1. Task Order: the oral or written orders for the operator to perform the task
2. Task Dispatch: inform others about his task
3. Perception: awareness of the state of the external world
4. Operator: the individual involved in the task
5. Motor System: carrying out oral or physical actions
6. Control: the design of the system controls
7. Machine: the system
8. System Parameters: indicating the state of the system
9. System Feedback: system response to actions
10. System Outcome: system faults, if any
11. Environment: the situation related PSFs
12. Situation: overall description of the Human-System Interaction

Figure 9 . The MMS of CAHR.

Table 25 . The PSFs modeled in CAHR classified based on the subject of their influence.

Task	Person	Feedback
a. Task preparation	a. Processing	a. Arrangement of equipment
b. Simplicity of task	b. Information	b. Display range
c. Complexity of task	c. Goal reduction	c. Accuracy of display/Display precision
d. Precision	**Activity**	d. Labeling
e. Time-pressure	a. Usability of control	e. Marking
Order Issue	b. Handling/Usability of equipment	f. Reliability
a. Clarity/Precision of procedures	c. Monotony	**System**
b. Design of procedures	d. Position/-ability	a. Technical layout
c. Content	e. Quality assurance	b. External event
d. Completeness	f. Equivocation of equipment	c. Construction
e. Presence		d. Redundancy
		e. Coupled equipment

3.8.6   HEP Calculation Procedure

CAHR calculates HEPs based on keyword searches within the CAHR database. The database is constructed from analysis of actual events. Each event is represented by a number of MMS characteristics. A MMS consists of the following nine elements which model the human-human and human-system interactions:

- Situation

- Task

- Person

- Activity

- Feedback

- Order dispatch

- Order issue

- Environment

- System.

The activities of each of the above elements are represented by the following five attributes:

- Object

- Verb

- Indication

- Property

- Element.

CAHR provides a list of keywords for the analyst to use to search the database. Searching for the above attributes would generate the frequencies of certain activities. For example, searching keywords by typing "valve AND open AND omit," the analyst would obtain the number of the activities that involve omitting opening a valve (No. 1). Searching keywords "valve AND open" would generate the number of activities of opening a valve (No. 2). CAHR provides equations that calculate the HEP of omitting opening a valve based on the values of No. 1 and No. 2.

3.8.7   Error-Specific HEPs

Each HEP is calculated for a specific error mode. This is a function of how the analyst defines the key words for the search.

3.8.8　Task Dependencies and Recovery

Task dependencies and error recovery are implicitly accounted for in CAHR through the interdependencies and recoveries embedded within the events in the CAHR database.

3.8.9　HEP Uncertainty Bounds

The method does not provide uncertainty bounds on the generated estimates.

3.8.10　Level of Knowledge Required

An analyst with about one year of HRA experience is expected to learn how to use CAHR in a very short time.

3.8.11　Validation

3.8.11.1　Error Identification

CAHR does not provide explicit guidelines for error identification.

3.8.11.2　HEP Quantification

Some empirical validation studies have been conducted to assess the quality of the HEPs generated by CAHR ranging from inter-domain validation of Boiling Water Reactors and Pressurized Water Reactors, transfer of data between nuclear and automobile industries, and comparison of nuclear data with data from air traffic management (Sträeter, 2005).

Currently, about 220 analyzed events are coded in the CAHR database. The opportunities for error underlying these failures could vary widely. Acceptance of the HEP estimate may be determined as a function of the means by which denominators were estimated. In the case of maintenance-based failures, the denominators may have been constructed from maintenance records. Two factors could potentially affect the credibility of CAHR output: (a) potential bias in database construction and (b) insufficiency of data quantity and quality. The algorithm for quantifying absolute error rates without success data merits closer review.

3.8.12　Reproducibility

3.8.12.1　Error Identification

CAHR does not provide explicit guidelines for error identification.

3.8.12.2　HEP Quantification

CAHR provides a list of keywords to be used for searches. Since the same computerized database is used it is expected that the result reproducibility is rated high.

3.8.13　Sensitivity

The sensitivity of results to changes in input variables is a function of correlations that are internal to the CAHR database. The authors of this report could not make an objective assessment of the numerical sensitivity, which would have required a large scale set of sensitivity runs using the CAHR adapt base. It is clear that the internal correlations may change when new events are added to the database. Therefore the HEPs generated by CAHR are dependent on the quantity and nature of events in its database. Also, the

volume of data (number of events) in the database affects the stability of the results. According to the author of CAHR, the result becomes stable when more than 50 events are analyzed and built into the database. The sensitivity is rated low based on the assumption that the database contains more than 50 events.

3.8.14   Experience Base

CAHR has been used in several German nuclear power PRAs and also applied in automotive industry in that country. It has also been used in the early conceptual phase for the appraisal of human interventions in the safety assessment of air traffic management changes for the year 2020. Lessons learned from applying CAHR in Task 97-2 by the Working Group on Risk Assessment (RISK) of the Nuclear Energy Agency (NEA) (Grant, Holy et al., 2000) and follow up activities (Sträeter, 2005) include:

- Reliability of development of database queries is not yet measured;

- Quality of event descriptions must be improved through the use of a multidisciplinary team and improved data from the plant;

- Currently high levels of expertise are required to apply method;

- Method needs a search scheme for errors of commission;

- Psychological scaling model has analogies to other established logic models and mathematics; and

- Limited practical guidance is available at the moment.

3.8.15   Resource Requirements

A CAHR software package is necessary to develop HEPs. This assumes that the database is populated with at least 50 events that are judged to be applicable to the domain of interest. Given the proper database the analysis process is straightforward and resource requirement is rated low.

3.8.16   Cost and Availability

A software package was developed to construct a database and conduct analysis. Documents are publicly available. The author of CAHR can be contacted at oliver.straeter@eurocontrol.int for code availability.

3.8.17   Suitability for NASA Applications

CAHR requires an established database for performing a keyword search to calculate HEPs. Construction of such a database requires detailed event analysis (or analysis of critical tasks or simulator runs) following CAHR specifications. The current database is nuclear-oriented and may not be directly applicable to NASA tasks. "Building a NASA-relevant database would require an effort of about one day per event, given that the descriptions of events/accidents are available and no reanalysis is necessary" according to the author of CAHR. CAHR provides an approach to transfer data from entirely different applications (currently nuclear, automotive, air traffic management.) This is certainly an appealing feature.

3.9   Standard Plant Analysis Risk HRA Method (SPAR-H)

3.9.1   Developmental Context

SPAR-H (Gertman et al., 2005) was a revision to, and a replacement of, the U.S. Nuclear Regulatory Commission's Accident Sequence Precursor (ASP) HRA screening method. The revisions were intended to make the characterization of human performance in SPAR more realistic and to reflect new trends in HRA methods and data. Some of the goals of SPAR-H include ease of use and better representation of uncertainty and dependency information for use in SPAR PRA models of U.S. nuclear power plants. SPAR-H has been applied to over 70 U.S. nuclear power plants. SPAR-H was originally developing as a screening methodology, but later the method was extended for full HEP quantification.

3.9.2   Screening

SPAR-H does not provide a procedure for screening.

3.9.3   Task Decomposition

SPAR-H decomposes a task into subtask of "diagnosis" and/or "action."

3.9.4   PSF List and Causal Model

SPAR-H is based on an information-processing model of human cognition, yielding a causal model of human error. SPAR-H also provides discussion of the interdependencies of PSFs, which are often ignored in other HRA methods. This being said, the interdependencies are not available to the reader in terms of correlation coefficients. The eight PSFs used by the method are:

- Available time;

- Stress/Stressors;

- Complexity;

- Experience/Training;

- Procedures;

- Ergonomics/Human-machine interface;

- Fitness for duty; and

- Work processes.

The authors consider this to be a set of universal PSFs that will fit most applications for which a simple HRA method is required. Each factor represents the effects of a number of subfactors. For example, the PSF "complexity" contains the following subfactors:

- Multiple faults

- Multiple equipment unavailable

- High degree of memorization required

- Parallel tasks

- Large number of actions required

- System interdependencies not well defined

- Mental calculations required

- Large number of distractions present

- Misleading or absent indicators

- Low fault tolerance levels

- Task requires coordination with ex-control room activities

- Transitioning between multiple procedures

- Symptoms of one fault mask other faults

- Large amount of communication required.

SPAR-H discusses dependencies among the eight identified PSFs in qualitative terms but the quantitative impacts are not addressed.

3.9.5   Coverage

The eight PSFs of SPAR-H cover ergonomics, cognitive, and organizational issues in a broad sense.

3.9.6   HEP Calculation Procedure

The SPAR-H HEP quantification for a specific activity includes the following steps:

1. Determine the plant operation state and type of activity:

    a. Two distinctive plant states, at-power and low power/shutdown, and two types of activities, diagnosis and action, are modeled. Four HEP worksheets are provided for use in calculating the HEPs of the following four different combinations:

        (1) At-power operation and diagnosis activity

        (2) At-power operation and action activity

        (3) Low power/shutdown operation and diagnosis activity

        (4) Low power/shutdown operation and action activity.

2. Evaluate PSFs' states to determine the multipliers:

a. Tables are provided within the HEP worksheet for the analysts to check the most likely states of PSFs. For each worksheet, the analysis needs to identify the type of activity. Three types of activities are specified: diagnosis, action, and diagnosis-and-action. The base failure rates for these types of activities are identical for all worksheets. A HEP multiplier is assigned to each PSF's state. The HEP multiplier could have different values in different worksheets.

3. Two exclusive equations are provided to calculate the final HEP. The choice of one equation over another is dependent on the number of negative PSFs.

3.9.7 Error-Specific HEPs

HEPs are calculated for "diagnosis" and "action" failures. The document (Gertman et al., 2005) does not explicitly define what is meant by "diagnosis failure" and "action failure."

3.9.8 Task Dependencies and Recovery

SPAR-H provides guidelines to assess the level of dependency of actions. Factors yielding dependency include same operating crew, time proximity, same work location, and same information cues. Error recovery (of the error itself) is not modeled in SPAR-H. Functional restoration and recovery of systems is treated as a separate event which needs to be specified by the analyst.

3.9.9 HEP Uncertainty Bounds

Uncertainty is performed for the final HEP (FHEP), adjusted for PSF influence and dependency. The HEP is assumed to be the best estimate of the mean. A beta distribution is assumed for purposes of uncertainty assessment. A so-called "constrained non-informative prior" (CNI) distribution (Atwood, 1996), is used to characterize the uncertainty around the mean. In model development, the Monte Carlo capability of the SAPHIRE workstation software is used to propagate the uncertainty for human failure sub-events, much the same as it is performed for other components. A typical assessment will include either 1,000 or 5,000 passes. As in most HRA methods, no adjustment or uncertainty for the PSFs separate from the base level HEP is performed.

3.9.10 Level of Knowledge Required

A PRA analyst with general engineering and system knowledge is expected to be able to apply SPAR-H with minimum training.

3.9.11 Validation

3.9.11.1 Error Identification

SPAR-H does not provide guidelines for error identification.

3.9.11.2 HEP Quantification

No independent validation of the method has been documented. The SPAR-H authors have provided a comparison of the base failure rates with other HRA methods (Gertman et al., 2005). These comparisons are shown in Tables 26-28.

Table 26 . Action error type base rate comparison (Gertman et al., 2005).

Method	Error Type Description	Base Rate (5th – 95th percentile bounds)
SPAR-H	Action Task	0.001
HEART	D. Fairly simple task performed rapidly or given scant attention	0.09
	F. Restore or shift a system to original or new state following procedures, with some checking	0.003
CREAM	Tactical	0.001–0.1
ASEP	Table 7-3. Screening critical action, assuming moderate stress, and no recovery	0.05
THERP	Table 20-2 Rule based actions of control room personnel after diagnosis, with recovery. EF=10	0.025

Table 27 . Mixed-task base rate comparison (Gertman et al., 2005).

Method	Error Type Description	Base Rate
SPAR-H	Task involving both diagnosis and action	0.011
HEART	A. Totally unfamiliar, performed at speed with no real idea of likely consequences	0.55
	B. Shifts or restores system to a new or original state on a single attempt, without supervision or procedures	0.26
	C- Complex task requiring high level of comprehension and skill	0.16
	E. Routine, highly practiced, rapid task, involving a relatively low level of skill	0.02
	G. Completely familiar, well-designed, highly practiced, routine task occurring several times per hour, performed to highest possible standards by a highly motivated, highly trained and experienced person, totally aware of implications of failure, with time to correct potential error, but without the benefit of significant job aids	0.0004
	H. Responds correctly to system command, even when there is an augmented or automated supervisory system providing accurate interpretation of system state	0.00002
	M. Miscellaneous task for which no description can be found (Nominal 5th to 95th percentile data spreads were chosen on the basis of experience available suggesting log normality)	0.03
FRANCIE (5th-95th percentile)	1. Procedural Omission	0.0059
	2. Error of Intent	0.085
	3. Selection Error	0.015
	4. Awareness and Task Execution Related to Hazards/Damage	0.016
	5. Cognitive Complexity or Task Complexity Related	0.033
	6. Inspection/Verification	0.097
	7. Values/Units/Scales/Indicators Related	0.022
	8. Maintenance/Repair Execution	0.041

Table 28 . Diagnosis error type base rate comparison (Gertman et al., 2005).

Method	Error Type Description	Base Rate
SPAR-H	Diagnosis Task	0.01
CREAM	Tactical Control Mode	0.001–0.1
	Opportunistic Control Mode	0.01–0.5
ASEP	Table 7-2. Screening diagnosis, assumed to be under moderate stress, given 30 minutes. EF=10.	0.01
THERP	Table 20.1 Screening diagnosis. EF=10.	0.01
HEART	Miscellaneous task category "M," no description in other tasks (A-H) fits diagnosis tasking as well.	0.03
INTENT	Misdiagnose given like symptoms. Capture sequence based on stimuli.	0.057
	Competing goal states lead to wrong conclusion.	0.048
	Symptoms noticed, but wrong interpretation.	0.026

3.9.12 Reproducibility

3.9.12.1 Error Identification

SPAR-H does not provide guidelines for error identification.

3.9.12.2 HEP Quantification

The reproducibility is high due to the simplicity of the SPAR-H model and its clarity in defining the scope. The authors of SPAR-H report relatively high inter-rater reliability for the first update to SPAR-H in the mid 1990s, but there is no recent effort to reassert the inter-rater reliability. A key factor affecting the reproducibility is the task decomposition. The analyst needs to decompose a task into a number of action or diagnosis activities. The final HEP of the tasks is the result of a HRA tree constructed from these activities. These steps are simple and clear. The reproducibility is therefore rated high.

3.9.13 Sensitivity

The main factor contributing to sensitivity in SPAR-H is specifying the values of the PSFs. Given a clear state description provided in the worksheets, the sensitivity is rated low.

3.9.14 Experience Base

SPAR-H has been applied to over 70 PRA analyses of U.S. commercial nuclear power plants, in ASP event analysis, by the U.S. Nuclear Regulatory Commission inspectors as part of the Reactor Oversight Process and in other industries. U.S. utilities, through EPRI, also have access to the SPAR-H method in the form of an HRA calculator under development. This function is primarily used by members to gage their HRA responses against expected regulator assessment of the same activities through application of SPAR-H.

3.9.15 Resource Requirements

SPAR-H only requires decomposing tasks into a number of cognitive or physical activities. The HEP for each activity can be calculated by using the appropriate worksheet. The process is easy to follow. No specific software is required for calculating HEPs.

3.9.16   Cost and Availability

Documents describing the method are publicly available (Gertman et al., 2005). No licensing fee is required for applying the method.

3.9.17   Suitability for NASA Applications

SPAR-H classifies tasks into only two types: diagnosis and action. Such a simple classification makes SPAR-H suitable for new designs. SPAR-H can also be easily applied to existing aerospace designs including both nominal and emergency situations. Before such application, the following concerns need to be addressed:

1.   SPAR-H worksheets are designed for nuclear power operations, the worksheets need to be revised regarding the appropriate task description, operating conditions, and scope of PSFs and their corresponding weights. If the current PSFs are to be used, then the assignment of many factors such as habitat factors, muscle wasting and bone density factors, cardiovascular factors, and other types of illness and their effects to the appropriate PSF category must be well defined for the analyst.

2.   Since SPAR-H does not provide guidelines for task decomposition, the analyst has the responsibility to identify how many diagnosis and/or action activities should be considered for a given task. This consequently affects the HEP of the task. The issue becomes more significant for new aerospace designs, where the allocation of tasks may be in development.

3.10   University of Maryland Hybrid (UMH)

3.10.1   Developmental Context

The University of Maryland Hybrid (UMH) HRA method [Shen and Mosleh 1996] was developed to estimate the HEPs for the Calvert Cliffs nuclear power plant PRA. The method utilizes certain features of SLIM, HCR, Influence Diagrams Approach (IDA) [Shen and Mosleh 1996; Shen, Smidts et al. 1997; Smidts, Shen et al. 1997], and some empirical results (e.g., EPRI Operator Reliability Experiment (ORE) [Spurgin, Moieni et al. 1990]).

The HCR method is used to estimate the HEPs of the reference points required by the SLIM. Equations of SLIM for calculating the Success Likelihood Index (SLI) are significantly revised to account for non-linearity of the effect of some PSFs on human performance.

3.10.2   Screening

UMH does not provide a screening procedure.

3.10.3   Task Decomposition

The UMH does not provide explicit procedures for task decomposition. Similar to SLIM, the UMH method requires the task of analysis to be "similar" to the anchoring tasks.

3.10.4   PSF List and Causal Model

The method specifies 23 PSFs grouped into eight broad categories:

1.   Rush Perceived by Operator

 a.   (VT1) Level of stress due to perceived lack of time; e.g., the "rush index"

2.   Operator Training and Experience

 a.   (VE1) Training and experience in identifying the need for the required action

 b.   (VE2) Training and experience in diagnosing what needs to be done

 c.   (VE3) Training and experience in carrying out (performing) the required action

3.   Procedural Direction Available to the Operator

 a.   (VP1) Quality and adequacy of the procedural direction available for the required action in the given scenario

 b.   (VP2) Non-scenario related procedures, such operating procedures and annunciator response procedures, available to direct the required response

4. Personnel Availability and Communications

    a. (VA1) Adequacy of initial manning in the control room, relative to performing the required action in time

    b. (VA2) Whether the number of personnel who eventually show up in the control room become a distraction to the operators

    c. (VA3) Adequacy of the initial, as well as the eventual, manning outside the control room, relative to performing the required action in time

    d. (VA4) Barriers to communications and coordination between the control room and the equipment operators to perform actions outside the control room

    e. (VA5) Barriers to communications and coordination between the control room operators to perform actions inside the control room

    f. (VA6) Barriers to communications and coordination between the equipment operators to perform actions outside the control room

5. Plant Indications

    a. (VI1) Initial indications that inform the operator of the action to be performed

    b. (VI2) Later indications received in time to complete the action, assuming that the initial indications went unnoticed

6. Consequences Associated with the Action

    a. (VC1) Consequences of performing the required action - to the plant (detriment to performance)

    b. (VC2) Consequences of performing the required action - to the operators (detriment to performance)

    c. (VC3) Consequences of failing to perform the required action - to the plant (stimulant to performance)

    d. (VC4) Consequences of failing to perform the required action - to the operator (stimulant to performance)

7. Operator Confusion

    a. (VD1) Preceding related successful actions

    b. (VD2) Preceding related unsuccessful actions

    c. (VD3) Number of preceding and concurrent unrelated actions in progress while the operators are trying to cope

8. Equipment Location

    a. (VL1) Difficulties of access, quality, and location of local instrumentation and controls in the control room required to perform the action

    b. (VL2) Difficulty of gaining access to any locations required to perform the action including airlock and security doors, as well as the distances that must be traveled

The UMH method does not offer an explicit causal model.

3.10.5 Coverage

The ergonomic, cognitive, and organizational factors are covered by the UMH method. The ergonomics and organizational factors are considered in the PSFs (see Section 3.10.4). The cognitive aspects are covered by the task classification (Table 29 - cognitive complexity and phase of information processing) and some of PSFs listed in Section 3.10.4.

3.10.6 HEP Calculation Procedure

UMH method revises SLIM to obtain the Success Likelihood Index (SLI) of the task. The required reference points for use in SLIM are obtained through the HCR method.

Instead of using a fixed set of PSFs to assess the SLI value, the UMH method classifies tasks into seven categories (see Table 29). Each category has its own set of PSFs (subset of the 23 PSFs identified in Section 3.10.4). The task classification is based on the combination of cognitive demand (e.g., Skill-based, Rule-based, and Knowledge-based) and phase of information processing (e.g., Identification, Planning, and Response). These seven task categories are shown in Table 29 .

Table 29 . The task classification of UM Hybrid method.

	Identification	Planning	Response
Skill-Based	S-Id	S-P	S-R
Rule-Based	R-Id	R-P	--
Knowledge-Based	K-Id	K-P	--

The UMH method also revised the way that SLIs are calculated. Instead of Equation 1 to calculate SLI, Equation 6 is used.

$$SLI \left(Success\ Liklihood\ Index\right) = \left[\prod_{i=1}^{N}\left(PIF_i\right)^{W_i^*}\right]\left[\sum_{j=1}^{M}W_j\left(PIF_j\right)\right]$$
(Eq. 6)

Where N represents the total number of PSF "switches" in the SLI equation; M represents the total number of "non-switch" PSF in the SLI equation; and W_i^* and W_j are weighting factors.

Equation 6 shows three types of PSF influences:

1. General Switch.

    - A single PSF is capable of setting the SLI score to zero.

2.	Joint Switch.

	-	Multiple PSFs joined together are capable of setting the SLI score to zero.

3.	Non-Switch.

	-	PSFs could influence the score of SLI but are not capable of setting the SLI value to zero.

The relationships between the seven task types, their corresponding PSFs, and the PSF influences are shown in Table 30 . The PSFs shown in Table 30 are represented by abbreviations. The descriptions of these PSFs are found in Section 3.10.4.

Using the above information, the steps for calculating HEPs are:

1.	Identify the type of task (See Table 29).

2.	Assess the states of the PSFs/PIFs related to the task). The UMH hybrid method provides questionnaires to assist in the assessment of the PSFs' states. The relevant PSFs to a task are specified in Table 30 .

3.	Calculate SLI value (use of Equation 6). The values of weighting factors, W^*_i and W_j, are determined by experts.

4.	Select similar tasks with HEPs that can be calculated by HCR. These HEPs are used as the reference points to obtain the constants 'a' and 'b' in the SLIM method.

5.	With known constants 'a' and 'b', the HEPs of the tasks of analysis can be calculated using Equation 2 (Equation 7 is a copy of Equation 2 for convenience).

$$Log(1-HEP) = a \times SL1 + b \qquad \text{(Eq. 7)}$$

Table 30 . The UMH seven types of tasks and their corresponding PSFs and influences.

	S-Id	S-P	S-R	R-Id	R-P	K-Id	K-P
VT1	Switch(g)	Switch(g)	Switch(g)	Switch(g)	Switch(g)	Switch(g)	Switch(g)
VE1	X					X	
VE2		X					X
VE3			X				
VP1				Switch(j)	Switch(j)	X	X
VP2				Switch(j)	Switch(j)	X	X
VA1			X				
VA2	X	X	X	X	X	X	X
VA3			X				
VA4	Switch(j)	X	X	X	X	X	X
VA5		X	X	X	X	X	X
VA6			X				
VI1	Switch(j)	X		X	X	X	X
VI2		X		X	X	X	X
VC1		X			X		X
VC2		X			X		X
VC3		X			X		X
VC4		X			X		X
VD1	X	X		X	X	X	X
VD2	Switch(g)	Switch(g)	Switch(g)	Switch(g)	Switch(g)	Switch(g)	Switch(g)
VD3	X	X	X	X	X	X	X
VL1			X				
VL2			X				

S-Id: Skill-based Identification
S-P: Skill-based Planning
S-R: Skill-based Response
R-Id: Rule-based Identification
R-P: Rule-based Planning
K-Id: Knowledge-based Identification
K-P: Knowledge-based Planning
g   = General type of switch
j   = Joint type of switch

3.10.7   Error-Specific HEPs

The UMH method calculates HEPs for error modes specified by the analysts.

3.10.8   Task Dependencies and Recovery

Certain error dependencies are explicitly considered in the PSFs. Error recovery is not explicitly modeled.

3.10.9   HEP Uncertainty Bounds

Uncertainly bounds for the HEPs are produced by propagating the uncertainties of the input parameters (e.g., values of PSFs).

3.10.10   Level of Knowledge Required

Analysts with about one year of experience in the HRA field are expected to be able to learn how to use UMH with moderate amount on training.

3.10.11   Validation

3.10.11.1   Error Identification

UMH does not provide guidance on error identification.

3.10.11.2   HEP Quantification

No independent validation has been conducted on the quality of the HEP produced by UMH. In the only application of UMH in a nuclear plant PRA, the needed HEP anchor points (for calculating the parameters of the relation between SLIs and HEPs) were based on the Operator Reliability Experiment (ORE) database.

3.10.12   Reproducibility

3.10.12.1   Error Identification

UMH does not provide guidance on error identification.

3.10.12.2   HEP Quantification

The reproducibility inherits the weaknesses of SLIM and HCR models. It is rated medium.

3.10.13   Sensitivity

UMH is sensitive in one tail of the distribution. (This is the same as SLIM).

3.10.14   Experience Base

This method was used in the Calvert Cliffs nuclear power station PRA that was submitted to the U.S. NRC.

3.10.15   Resource Requirements

The UMH method requires the same resources as in SLIM and HCR. A computer code has been developed to facilitate HEP calculations and uncertainty assessment.

3.10.16 Cost and Availability

The method is documented in the Calvert Cliff Human Error Probability Methodology Report (Shen & Mosleh, 1996).

3.10.17 Suitability for NASA Applications

The UMH method combines appealing features of HCR, SLIM, and some of the cognitive methods. In doing so, it also inherits the limitations of these methods regarding resources and applicability to NASA missions. An advantage of UMH over SLIM is that it provides a list of PSFs. HCR is only used to obtain reference HEPs based on (nuclear) experimental data; therefore, for NASA applications such reference HEPs would need to be based on space mission activities. UMH's suitability for short term application is limited; however, its framework provides a relatively generic task classification that can be very useful in developing improved methods in the future.

3.11 Commission Errors Search and Assessment (CESA)

3.11.1 Developmental Context

The CESA method was developed at the Paul Scherrer Institute (PSI), Switzerland. The focus of the method is on the identification, characterization, and assessment of potential errors of commission. As in the earlier Borssele EOC screening study (Versteeg, 1998; Julius, Jorgenson, Parry, & Mosleh, 1995), one of the inputs to a CESA analysis is information from an existing PRA study. The method has been applied for a Swiss nuclear power plant, as reported in Reer, Dang &Hirschberg, 2004.

3.11.2 Screening

As a screening tool CESA selects the tasks to be analyzed and the PRA scenarios to be examined by prioritizing the systems, components, and scenarios for which an EOC contribution would have the largest impact. The Risk Achievement Worth (RAW) and Fussell–Vesely importance techniques are used for such selection.

3.11.3 Task Decomposition

The identification of potential EOCs is performed in CESA by analyzing task performance as guided mainly by the applicable operating procedures in a range of PRA scenarios. The tasks to be analyzed and the PRA scenarios to be examined are selected by prioritizing the systems, components, and scenarios for which an EOC contribution would have the largest impact.

In the qualitative analysis, the CESA method distinguishes among decision-making tasks, execution tasks, and error correction. Decision-making and execution tasks are analyzed to identify potential errors associated with these tasks and the consequence of these errors as they relate to an EOC opportunity.

3.11.4 PSF List and Causal Model

A list of 5 high-level PSFs specific to EOCs is provided; additional PSFs associated with the HRA method(s) used for quantification are also considered. These five PSFs are *training*, *procedures*, *indications*, *error's attraction or attractiveness*, and *operator's attention*.

Retrospective event analysis is not a feature of the method. No causal model is provided by CESA.

3.11.5 Coverage

Currently CESA relies on other HRA quantification method for HEP quantifications. Using the THERP method is suggested. THERP error quantification mostly covers the ergonomic and organizational factors. The CESA method has added emphasis on cognitive failure in identifying EOC opportunities. The cognitive relevant PSFs are added to the THERP quantification scheme.

3.11.6 HEP Calculation Procedure

The steps followed by CESA to identify and quantify EOCs.

1. Catalog required operator actions:
    The objective is to define and categorize possible operator actions on components (e.g., starts boiler) which could affect PRA results and could be considered as potential causes of system failures. This is

done by searching through the operating procedures to identify the operator's actions on these components.

2. Identify EOC events linked to important systems:
The EOC events are defined as operator actions that may contribute to the failures of PRA top events. The Risk Achievement Worth (RAW) is used as a threshold index to identify the important EOCs.

3. Identify specific EOC scenarios (EOC opportunities):
Search EOC opportunities in the top PRA sequences, specify the EOC scenarios, and prioritize the opportunities, based on the core damage frequency (Fussell–Vesely importance).

4. Characterize the EOC scenarios in detail and quantify:
Perform qualitative and quantitative analyses to determine the risk impact of the identified EOC scenarios and to provide insights for reducing the risk contributions from these EOCs.

An EOC scenario, as in Step 3, refers to a specific EOC (inappropriate action) in a specific PRA scenario. Within an EOC scenario, however, there may be multiple paths leading to the EOC. In Step 4, an important aspect of the detailed characterization is to define the various scenario evolutions that may result in the EOC. Each of these paths includes combinations of system conditions, human errors (at the sub-task level), and nominal actions (again at the sub-task level). Many elements of these tasks are quantifiable with existing HRA methods. For example, the failure to detect an out-of-tolerance system parameter may cause the personnel to perform the EOC; e.g. to trip the back-up system in this case. In the pilot study, the THERP method was used to quantify many of the EOC path elements.

3.11.7 Error-Specific HEPs

HEPs are calculated for the error modes specified by the analyst.

3.11.8 Task Dependencies and Recovery

Task dependencies are modeled using the THERP model. A checklist of eight dependency-related factors is provided to support the assessment of the conditional HEPs, including:

- Personnel that perform the tasks

- Performance locations

- Time when the errors are possible

- Procedure parts that call up the tasks

- Indications of plant conditions required for task performances

- Possible goals associated with the errors

- Procedure-related implication of the first error

- Equipment-related implications of the first error

Recovery is mentioned but no explicit procedure for analysis is offered.

3.11.9 HEP Uncertainty Bounds

The THERP uncertainty bounds are used.

3.11.10 Level of Knowledge Required

HRA analysts with about one year experience in the HRA field are expected to be able to learn how to use CESA with a moderate amount of training.

3.11.11 Validation

3.11.11.1 Error Identification

No independent validation has been conducted to assess the quality of the CESA results.

3.11.11.2 HEP Quantification

No known validation on HEPs generated by CESA has been conducted.

3.11.12 Reproducibility

3.11.12.1 Error Identification

CESA searches for error opportunities based on the PRA events and emergency operating procedures. The analyst identifies the errors through applying general guidelines. For a given PRA model and procedures, CESA is expected to have high reproducibility for identifying risk relevant tasks and errors.

3.11.12.2 HEP Quantification

Since CESA suggests use of THERP for quantification, the reproducibility for HEP quantification is rated the same as THERP (medium).

3.11.13 Sensitivity

Since CESA suggests use of THERP for quantification, the sensitivity in HEP quantification is rated the same as THERP (low).

3.11.14 Experience Base

The CESA method has been used in one Swiss nuclear power plant HRA study.

3.11.15 Resource Requirements

CESA requires the analyst to search through the PRA model and operating procedures to identify the risk significant EOCs. This could be resource intensive depending on the complexity of the models and procedures. The level of effort required is therefore rated medium. However, if electronic procedures exist that allow the analyst to perform a keyword search to identify the human actions of interest, the required effort would be significantly reduced.

3.11.16 Cost and Availability

The methodology document is publicly available. No licensing fee is required.

3.11.17   Suitability for NASA Applications

CESA requires a PRA model and operating procedure to identify the EOCs. Once the EOCs are identified, it uses THERP for calculating the HEPs. As such, the approach is not very effective for new aerospace designs. For existing designs, CESA identifies the risk tasks, but again uses THERP for error quantification. Thus, it has the same limitations as THERP. The suitability of CESA for NASA application is rated low.

3.12   Time Reliability Correlation (TRC)

3.12.1   Developmental Context

TRC was developed during the time period where HRA methods were moving from a procedure-oriented approach to a more task-oriented approach in response to the TMI accident. Prior to the development of the TRC method, correlating time and reliability had been used in several PRAs. The Oak Ridge National Laboratory (ORNL), within the same time frame (the early 80's), also performed simulator exercises to generate human performance data with a focus on the time-reliability relationship for nuclear plant operators.

TRC was developed by incorporating the time-reliability relationships from various data sources including Bott, Kozinsky, Crowe & Haas, 1981; Greene, 1969; and field data. The TRC application scope is for the *abnormal* situations where nuclear plant operator action is required, and the situation:

- demands the operators to respond to unintended conditions

- forces the operators to diagnose the situation at hand, interpret its implications on future plant operation, and decide on a plan to respond--all in a time window dictated by the unfolding events

- forces a response time that is uncertain in its details and can only be inferred from the pace of the change in critical plant parameters or anticipated by analysis

- demands that the operators succeed in their actions, since failure risks loss of property or even lives.

3.12.2   Screening

TRC offers a simple screening process for estimating HEPs. Four tables are provided by TRC for such purposes (see Tables 32-35).

3.12.3   Task Decomposition

TRC does not provide rules for task decomposition. The analyst specifies the task scope for HEP calculation.

3.12.4   PSF List and Causal Model

The TRC method uses the Success Likelihood Indicator (SLI) of the SLIM method to account for the effect of PSFs. The TRC method provides a list of PSFs for assessing the SLI value, including:

1.   Time constraint related.

    a.   One action with a short available time.

    b.   Multiple activities over a single duration.

2.   Diagnosis related.

    a.   Confusing indications.

b.   Credibility of events.

c.   Complexity of events or system.

3.   Decision making related.

a.   Planning or decision making required.

b.   Conflict between an option and a normal intention.

c.   Competing resources.

4.   Command and control related.

a.   Remoteness of people who need to coordinate.

b.   Remoteness of actions from control room.

c.   Distance between indications and controls.

5.   Physiology related.

a.   Hostile environment.

No dependency among the PSFs is discussed in TRC.

3.12.5   Coverage

As a time-reliability correlation approach, the available time is the dominant factor in diagnosis error in TRC. Even though several other PSFs are identified (as listed in Section 3.12.4), their aggregate effect on HEP values is limited to a small range between 0.5 and 2.

3.12.6   HEP Calculation Procedure

The TRC method uses a multivariate lognormal distribution (see Equation 8) to calculate the probability of an operator successfully responding to a situation within a given time.

$$T = M_R \times M_U \qquad \text{(Eq. 8)}$$

Where

T   is a random variable that accounts for the time needed for an operator to successfully respond to the situation

M_R is a lognormal random variable accounting for the uncertainty of the process

M_U is a lognormal random variable accounting for uncertainty in the model

The equations representing M_R and M_U are discussed below:

3.12.6.1   *M_R Component*

Equation 9 shows the form of lognormal probability density function:

$$f(t; \mu, \sigma) = \frac{1}{t\sigma\sqrt{2\pi}} e^{\frac{-(\ln t - \mu)^2}{2\sigma^2}}$$ (Eq. 9)

where μ and σ are the logarithmic mean and standard deviation.

The TRC method uses median (*M*) and error factor (*EF*) to specify the lognormal distribution for time-reliability correlation. These are related to μ, and σ through:

$$M = e^{\mu}$$ (Eq. 10)

$$M = e^{\mu}$$ (Eq. 11)

$$M = e^{\mu}$$ (Eq. 12)

$$EF = e^{1.645\sigma}$$ (Eq. 13)

In order to incorporate the situational effects, the M is written as a function of other factors:

$$M = K_C K_I M_{REF}$$ (Eq. 14)

where

M_{REF} is the reference median response time

K_C adjusts M_{REF} by as much as 2 and as little as 0.5 to account for taxonomic considerations. $K_C = 1$ if no rule is available, and $K_C = 0.5$ if a rule is available.

K_I adjusts M_{REF} by as much as 2 and as little as 0.5 to account for influences of performance shaping factors

The value of K_I is a function of the success likelihood index (SLI) of the SLIM method.

$$K_I = 2^{(1-2\,SLI)} \quad where \; SLI \in [0, 1]$$ (Eq. 15)

The values of the reference median response time (M_{REF}) and error factor (EF) can be obtained by fitting to existing data. TRC uses THERP values as "data." For diagnosis-dominant response not aided by rules, the values for M_{REF} and EF are 4 minutes and 3.2 minutes, respectively. These two numbers are obtained from THERP numbers by anchoring at values at 10 and 60 minutes.

3.12.6.2   *M_U Component*

The median of M_U is 1. The EF of M_U is obtained by assuming that:

- A "good" plant typically has a SLI of about 0.7

- The HEP at sixty-minutes, for rule-based response of a "good" plant is assumed to be 1E-6.

Based on the above assumptions, the EF of M_U is calculated as 1.68.

A software pack called Operator Reliability Calculation and Assessment (ORCA) was developed for TRC calculation.

For HEP assessment without using the ORCA code, the TRC method provides four tables. Table 31 shows the types of situations covered by these four tables.

Table 31 . TRC tables calculating HEPs manually.

Action Type	With Hesitancy	Without Hesitancy
Rule-Based	Table 32	Table 33
Knowledge-Based	Table 34	Table 35

Table 32 . Time-reliability correlation values for Rule-Based Action, without hesitancy.

Time (Min)	Success Likelihood Index (SLI)				
	0.1	0.3	0.5	0.7	0.9
5	3E-1	2E-1	1E-1	6E-2	3E-2
10	9E-2	4E-2	2E-2	8E-3	3E-3
20	1E-2	5E-3	2E-3	5E-4	1E-4
30	3E-3	9E-4	3E-4	6E-5	1E-5
60	1E-4	3E-5	6E-6	1E-6	2E-7

Table 33 . Time-reliability correlation values for Rule-Based Action, with hesitancy.

Time (Min)	Success Likelihood Index (SLI)				
	0.1	0.3	0.5	0.7	0.9
5	4E-1	3E-1	2E-1	2E-1	1E-1
10	2E-1	1E-1	9E-2	5E-2	3E-2
20	7E-2	4E-2	3E-2	1E-2	8E-3
30	3E-2	2E-2	1E-2	6E-3	3E-3
60	8E-3	4E-3	2E-3	9E-4	4E-4

Table 34 . Time-reliability correlation values for Knowledge-Based Action, without hesitancy.

Time (Min)	Success Likelihood Index (SLI)				
	0.1	0.3	0.5	0.7	0.9
5	7E-1	5E-1	4E-1	3E-1	2E-1
10	3E-1	2E-1	1E-1	6E-2	3E-2
20	9E-2	4E-2	2E-2	8E-3	3E-3
30	3E-2	1E-3	5E-3	2E-3	5E-4
60	3E-3	9E-4	3E-4	6E-5	1E-5

Table 35 . Time-reliability correlation values for Knowledge-Based Action, with hesitancy.

Time (Min)	Success Likelihood Index (SLI)				
	0.1	0.3	0.5	0.7	0.9
5	6E-1	5E-1	4E-1	3E-1	3E-1
10	4E-1	3E-1	2E-1	2E-1	1E-1
20	2E-1	1E-1	9E-2	5E-2	3E-2
30	1E-1	7E-2	4E-2	3E-2	1E-2
60	3E-2	2E-2	1E-2	6E-3	3E-3

3.12.7 Error-Specific HEPs

TRC only calculates the probability that the successful action is not taken by a specified time.

3.12.8 Task Dependencies and Recovery

Task dependency is not explicitly addressed by the TRC method. The data used by TRC are mainly from THERP numbers in which the recovery factor is covered. Therefore, recovery is implicitly covered in the TRC method.

3.12.9 HEP Uncertainty Bounds

The TRC method addresses uncertainties through the variable M_U (See section 3.12.6).

3.12.10 Level of Knowledge Required

An HRA analyst with about one year of experience in the HRA field is expected to learn how to use TRC in a short amount of time.

3.12.11 Validation

3.12.11.1 Error Identification

TRC does not provide guidance on error identification.

3.12.11.2   HEP Quantification

TRC calibrates the time-reliability curve mainly based on THERP values. There is no known validation of the quality of HEPs produced by TRC.

3.12.12   Reproducibility

3.12.12.1   Error Identification

TRC does not provide guidance on error identification.

3.12.12.2   HEP Quantification

Two key parameters affecting the reproducibility are the SLI and available time for response. Even though a list of PSFs is provided, SLI reproducibility is dependent on the analysts who determine the PSFs' states. With respect to the impact of the available time, while for existing systems the available time might be easy to determine, for new systems the uncertainty of the available time could be significant and subject to analyst judgment error. Given these issues the overall reproducibility of TRC HEPs is rated medium. The rating in part reflects the fact that TRC is a time-reliability correlation method where, given the available response time, the HEP is highly reproducible.

3.12.13   Sensitivity

TRC anchors its time-reliability curve to THERP HEPs at operator response to situations at 10 and 60 minutes. The HEPs after 60 minutes are based on assumed minimum HEPs, and are not very sensitive to time assessment. Thus, the HEPs (beyond 60 minutes) are relatively stable. On the contrary, for time values under 10 minutes, TRC values are sensitive to time variation. This is the region that is of interest for many NASA tasks.

3.12.14   Experience Base

The TRC method has been used in a number of U.S. nuclear power plant PRAs.

3.12.15   Resource Requirements

Using TRC requires assessing SLI and the time available for successful action. The effort required is rated medium. The reason is that while the estimation procedure is straightforward, the analyst needs to identify PSFs and calculate SLI through rating and weighting of the PSFs. HEPs can be calculated by using the Operator Reliability Calculation and Assessment (ORCA) software or by using Tables 32 to 35.

3.12.16   Cost and Availability

The method is publicly available (at the time this report was being prepared; the main reference book for the method was out of print). The cost or availability of the ORCA software is unknown.

3.12.17   Suitability for NASA Applications

TRC uses time as the dominant factor for calculating HEPs. The theoretical basis for this aspect of the method has been questioned by some. Besides this we note that the response time scales are different between aerospace and nuclear power applications. According to TRC, in the nuclear power context most actions involving complex cognition take place between 10 and 60 minutes into an accident. While this might also be the case for a small number of space mission activities, the vast majority will have a different time

scale, some much shorter, and some much longer. In addition, before any application to NASA tasks, TRC needs to be calibrated to NASA-specific performance data. But such data are not expected to be available for new space system designs. Based on these limitations, the suitability of TRC for NASA applications is rated low.

3.13   Human Factors Process Failure Modes and Effects Analysis (HF PFMEA)

3.13.1   Developmental Context

The Human Factors Process Failure Modes & Effects Analysis (HF PFMEA; Broughton, Carter, Chandler, Holcomb, Humeniuk, Kerios, Bruce, Snyder, Strickland, Valentino, Wallace, Wallace & Zeiters, 1999) was developed by Boeing for NASA. The method, an extension of the standard PFMEA framework, is essentially a qualitative analysis approach designed to help in identification of potential human errors (failure modes), factors that contribute to potential errors, and potential consequences of the errors (effects) and provide a qualitative means to evaluate those effects and rank risks.

3.13.2   Screening

The HF PFMEA document does not mention screening analysis. However, the framework is flexible and can be used as a screening tool. Screening can be done based on both the *likelihood* of error and *severity* of its consequence. In assessing these two factors, the HF PFMEA offers a scale (three levels for error likelihood and five levels for consequence severity).

3.13.3   Task Decomposition

The HF PFMEA provides a procedure to identify human activities, potential error modes, causes of the errors, and their thorough function and task analyses. The basic components of the HF PFMEA are:

1.   Task Description: Accomplished via Mission Description, Functional Analysis, Functional Flow Diagrams, Identification of Human-System Interfaces, and Task Analysis.

2.   Identification of a Behavioral Function Performed by Operator: Action Verbs from the Modified Berliner's Taxonomy are incorporated into the Hierarchical Task Analysis.

3.   Identification of Potential Human Error: The Potential Human Error List is used to assist the analyst in identifying and evaluating errors of omission and commissions.

4.   Identification of Performance Shaping Factors: Direct observation, operator/technician interviews, and evaluation of all relevant human-machine interfaces using the IAT-M and the PSF checklist to assist the analyst in identifying any conditions that would influence the workers' performance and increase the potential for human error.

5.   Estimation of the Frequency of Human Error: The performance shaping factors and barriers, past performance data (PRACA and mishap reports), and errors noted during the direct observation are used to estimate the frequency of each human error.

6.   Error consequences: Equipment FMEAs, past performance data (PRACA and mishap reports), and subject matter experts' opinions are used to determine the consequences of individual human errors. This data is also used to determine the frequency of those consequences.

7.   Consequences are ranked based on their likelihood of effect and severity of effect using NASA's 5x5 risk matrices. The NASA risk matrices identify three categories of risk: 1) high risk = action required, 2) moderate risk = action optional, and 3) low risk = action not required.

8. Recommendations to prevent error: Information gathered at team meetings with technicians, engineers, and subject matter experts are combined with data from human factors standards and guidelines to generate recommendations.

In one implementation (Broughton et al., 1999), the functional flow diagram (FFD) was used for functional analysis, and the Hierarchical Task Analysis method was used for task analysis. The functional analysis identified the nature and sequence of activities required to take place. Task analysis decomposed the individual tasks identified in the functional analysis into subtasks until human activities according to the following Berliner taxonomy could be identified:

1. PERCEPTUAL PROCESSES.

   a. Searches for and receives information.

   b. Detects, inspects, observes, reads, receives, scans, and surveys.

   c. Identifies objects, actions, and events.

   d. Discriminates, identifies, locates, and categorizes.

2. MEDIATIONAL PROCESSES.

   a. Analyzes, interprets, calculates, chooses, compares, computes, estimates, plans, and verifies.

3. COMMUNICATION PROCESSES.

   a. Advises, answers, authorizes, communicates, directs, indicates, informs, instructs, requests, transmits, and receives.

4. MOTOR PROCESSES.

   a. Simple/discrete
      Activates, adjusts, aligns, attaches, bends, carries, closes, connects, detaches, disconnects, folds, follows procedures, gives, goes, holds, insert, joins, lifts, lowers, moves, opens, places, positions, pours, presses, pulls, pushes, puts, releases, removes, rotates, sets, stamps, and writes.

   b. Complex/continuous
      Regulates, synchronizes, and tracks.

NASA partnered with the Relex Corporation to develop a software tool that assisted the analyst in the completion of a HF PFMEA. As part of the software development, the list of possible human actions was expanded. This allowed NASA and other users to generate a task statement and then take the verb in the task statement and compare it to the human action list. The action on the list that most closely resembles the user action is selected and a corresponding error list is provided.

3.13.4   PSF List and Causal Model

In the HF PFMEA approach (Broughton et al., 1999), failure modes are identified for each of the basic human activities identified (Section 3.13.3). Examples of failure modes associated with "Detect" are:

1. Fails to detect signal (frequency - error of commission or error of omission).

2.   Detects incomplete and/or partial information (accuracy - error of commission).

3.   Detects unnecessary information that hinders the task (accuracy - error of commission).

4.   Detects information at inappropriate time (timing - error of commission).

5.   Fixates on wrong information (accuracy - error of commission).

6.   Misinterprets readings/displays during detection (accuracy - error of commission).

HF PFMEA also provides an extensive list of factors (including PSFs) so the analyst can check the factors contributing to the error modes (see Section 3.13.3). This is especially useful when the analyst is evaluating existing process and trying to determine the specific PSFs that are causing errors, so that the PSFs can be modified and/or eliminated. An expanded PSF list is available in the updated version of this method taught by NASA and provided in the Relex Corporation HF PFMEA software package. The list covers factors in various categories including:

- Equipment/tool/part

- Environmental/facilities

- Job/task

- Technical knowledge

- Written information

- Verbal information (communication)

- Factors affecting individual performance

- Team factors

- Leadership/supervision

- Organizational issues

- Other contributing factors.

Each item listed above contains a list of subfactors. The list is extensive and detailed. For example, the factors affecting individual performance include the following seven categories:

- Physical health

- Fatigue

- Time constraints

- Peer pressure

- Body size/strength

- Personal event

- Workplace distraction or interruption during task performance.

     Each of the above seven categories contains a number of specific PSFs.

3.13.5   Coverage

     The HF PFMEA provides an extensive and detailed checklist for the analyst to identify the factors contributing to failure of the human activities (Section 3.13.3). They cover ergonomic, cognitive, and organizational factors.

3.13.6   HEP Calculation Procedure

     HF PFMEA is a qualitative process. Past performance data and expert judgment are used to rank the likelihood of a given human error as improbable, possible, or highly likely.

3.13.7   Error-Specifics HEPs

     Potential error modes are listed for each type of human activity listed in Section 3.13.3. The analyst has to identify the likely error modes from the list. The analyst uses information gathered from interviews, simulation (such as MIDAS), past performance data, and expert opinion to determine the specific error modes for a given process.

3.13.8   Task Dependencies and Recovery

     Task dependencies and recovery are not explicitly covered by PFMEA.

3.13.9   HEP Uncertainty Bounds

     Quantification on human error is not the focus of PFMEA. Human error likelihood is assessed on a qualitative scale based on past performance data and expert judgment.

3.13.10   Level of Knowledge Required

     An HRA or human factors specialist with about one year of experience in the field is expected to be able to learn the method in a short period of time. NASA has a web-based and classroom course that provides instruction on the method

3.13.11   Validation

3.13.11.1   Error Identification

     No validation of the method has been cited. The method has been applied to Space Shuttle processing, payload processing, and control room evaluations. It has also being used in the medical industry.

3.13.11.2   HEP Quantification

     HF PFMEA does not produce quantitative HEPs.

3.13.12 Reproducibility

3.13.12.1 Error Identification

Due to the fact that the documentation provides a very detailed check list, and training and software are available, the reproducibility is rated high.

3.13.12.2 HEP Quantification

HF PFMEA does not produce quantitative values of HEPs. The qualitative assignment of error likelihood and severity is made by the analyst. Reproducibility of a qualitative assessment is usually dependent on the granularity of the scale used. In the case of HF PFMEA the number of levels is small (for both likelihood and severity), and the reproducibility is expected to be moderate to high for a given error.

3.13.13 Sensitivity

HF PFMEA does not produce quantitative HEPs. Since the qualitative assignment of error likelihood and severity is made by the analyst it is difficult to determine how such assessments vary as a function of the factors (e.g., applicable PSFs) identified by the analyst.

3.13.14 Experience Base

The PFMEA framework is widely used in various industries. The method has been applied to Space Shuttle processing, payload processing, and control room evaluations. It has also being used in the medical industry.

3.13.15 Resource Requirements

Ideally HF PFMEA should be performed by a team consisting of system designers, engineers, and HF or HRA analysts. The method "can be time-consuming if the process is long, complex, or involves a lot of team members."(Broughton et al., 1999) The effort required for performing HF PFMEA is therefore rated high.

3.13.16 Cost and Availability

The HF PFMEA project was sponsored by NASA, and the methodology is free to the public. Web-based training is available for government employees via NASA's online learning system, SATERN. A PowerPoint screen capture of this course is available at the NASA Human Reliability website: http://humanreliability-pbma-kms.webexone.com/default.asp?link=. Software was developed by NASA in partnership with the Relex Corporation to perform the HF PFMEA method. The software provides step-by-step assistance to the analyst who is not familiar with the process or needs assistance with the analysis of human error. This software can be found at: http://www.relex.com/products/humanfactors.asp (NASA employees should contact Faith Chandler at faith.t.chandler@nasa.gov).

3.13.17 Suitability for NASA Applications

Performing the analysis to the level of detail described in the HF PFMEA method requires detailed system information which is normally not available at the early design stage. For existing systems, however, the HF PFMEA is a good tool to identify the important human actions, context characteristics, and applicable failure modes to generate recommendations for process improvements.

3.14   EPRI Cause Based Decision Tree (CBDT)

3.14.1   Developmental Context

Between 1986 and 1990, the Electric Power Research Institute (EPRI) conducted the Operator Reliability Experiments (ORE) project performing full-scale nuclear power plant control room simulator exercises. The aim was to collect operating crew response data to test the hypotheses of the Human Cognitive Reliability (HCR) method (Spurgin, Moiene, Gaddy, Parry, Orvis, Spurgin, Joksimovich, Gaver & Hannaman, 1990).

The EPRI Cause Based Decision Tree (CBDT) (Parry, Lydell, Spurgin, Moieni, & Beare, 1992; Moieni, Spurgin, & Singh, 1994a; Moieni, Spurgin & Singh, 1994b) has it roots in the ORE data and experiments. The Parry et al., 1992 report states two main objectives:

1.   Provide a guideline for the use of simulator data. The HCR method requires frequent use of extrapolation to assess HEPs. The assumptions for such extrapolation are not supported by the HRA community. The Parry et al., 1992 report provides a worksheet for calculating HEPs based on simulator data. The procedure provided in the accompanying worksheets would compensate for the weakness of HCR in the assumptions made for extrapolation.

2.   Provide an independent procedure to calculate HEPs (i.e., the CBDT) method that uses insights drawn from the ORE. The CBDT method provides eight decision trees and a table for the analyst to assess HEPs.

Only the CBDT method (Item 2 above) is evaluated here.

3.14.2   Screening

CBDT provides a table (Table 36) and a time-reliability curve (Figure 10) for screening.

Table 36 . Screening rules provided by CBDT.

Value	Conditions
0.1 - 0.5	Known problems with parts of procedure, procedural routes normally exposed during training, combination of time limited actions with little training, competing key actions, subtle faults or indications disguised by well-known transients
1E-6 - 1E-4	Well practiced actions both in the plant and at the simulator. Slowly changing transients with multiple chances for recovery by crew and others, clear indications and little chance of confusion, simple challenges not multiple failures
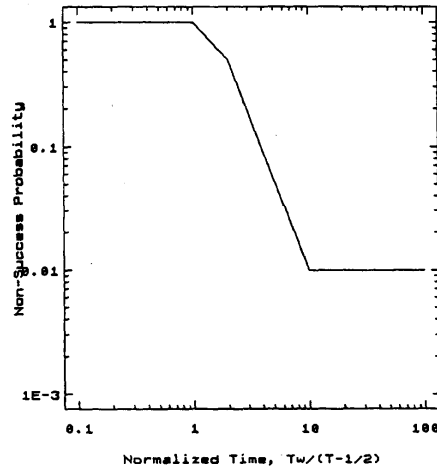
Figure 10 . The CBDT Screening Curve

3.14.3   Task Decomposition

As the ORE project mainly focused on the situations in which the operating crew were instructed to follow abnormal or emergency operating procedures to handle the problem, the CBDT method was developed to quantify human errors in such context. The human-system interactions are assumed to be clearly specified in the operating procedures. The identification and definition of the human-system interaction events are assumed to be done with the use of other techniques such as the EPRI SHARP 1 methodology (Wakefield, Parry & Spurgin, 1992).

3.14.4   PSF List and Causal Model

In the context of following operating procedures, two types of errors are identified:

- Failure in initiating correct response (P_C): the probability of failure to initiate timely and correct response due to failure (or delay) of detection, diagnosis, or decision.

- Failure to carry out required action (P_E): the probability of failure to execute the required response (i.e., slip type of error)

Two failure modes with corresponding failure mechanisms, contributing to P_C, are identified:

1.   Failures of the plant information-operator interface.

   a.   The required data are physically not available to the control room operators.

   b.   The data are available, but are not attended to.

   c.   The data are available, but are misread or miscommunicated.

   d.   The available information is misleading.

137

2. Failure of the procedure-crew interface.

    a. The relevant step in the procedure is skipped.

    b. An error is made in interpreting the instructions.

    c. An error is made in interpreting the diagnostic logic (this is a subset of Item f immediately above, but is treated separately for convenience).

    d. The crew decides to deliberately violate the procedure.

For each failure mechanism, a set of key PSFs are identified. These are:

1. (P_C^a) The required data are physically not available to the control room operators.

    a. Indication of availability in the control room.

    b. Control room indication accuracy.

    c. Warning/Alternate in procedure.

    d. Training on indicators.

2. (P_C^b) The data are available, but are not attended to.

    a. Low vs. High workload.

    b. Check vs. monitor.

    c. Front vs. back panel.

    d. Alarmed vs. not alarmed.

3. (P_C^c) The data are available, but are misread or miscommunicated.

    a. Indicators easy to locate.

    b. Good/bad indicator.

    c. Formal communication.

4. (P_C^d) The available information is misleading.

    a. All cues as stated.

    b. Warning of differences.

    c. Specific training.

    d. General training.

5. $(P_C{}^e)$The relevant step in the procedure is skipped.

    a. Obvious vs. hidden (instruction).

    b. Single vs. multiple (text procedure).

    c. Graphically distinct.

    d. Place-keeping aids.

6. $(P_C{}^f)$An error is made in interpreting the instructions.

    a. Standard unambiguous wording.

    b. All required information (are presented).

    c. Training on step.

7. $(P_C{}^g)$An error is made in interpreting the diagnostic logic.

    a. "NOT" statement (existed?).

    b. AND or OR statements (co-existed?).

    c. Both AND & OR (complex combinations existed?).

    d. Practiced scenario.

8. $(P_C{}^h)$The crew decides to deliberately violate the procedure.

    a. Belief in adequacy of instruction.

    b. Adverse consequence in compliance.

    c. Reasonable alternatives.

    d. Policy of "Verbatim" compliance.

3.14.5   Coverage

The CBDT method is designed to be used in the context of operators following a procedure to handle nuclear plant abnormal or emergency situations. The PSFs covered are dominated by the ergonomic and organizational factors and some cognitive factors.

3.14.6   HEP Calculation Procedure

The CBDT method uses decision trees to guide the assessment of HEPs by the analyst. Procedures are provided for estimating P_C and P_E.

3.14.6.1   Procedure for Assessing P_C

The CBDT method identifies eight key error mechanisms contributing to P_C. To assess the effect of each failure mechanism, a decision tree is developed. The analyst, guided by the decision tree, decides the

states of PSFs by which the HEP, due to the specific failure mechanism, is determined (see Section 3.14.4). The HEPs obtained through such decision trees do not include the recovery factors (e.g., error recovered by other crew members). The P_C-without-recovery is the sum of the eight non-recovery HEPs. Figure 11 shows an example decision tree (for assessing P_C^a-without-recovery).

The effects of recovery factors are explicitly modeled in CBDT by the use of a table (Table 37). Within this table, the possible recovery sources as well as their effects for each failure mechanism are specified. The HEP-with-recovery for each error mechanism is the product of its non-recovery HEP (obtained from the decision tree) and its "recovery factor" (obtained from Table 37). The P_C-with-recovery is the sum of the HEP-with-recovery of the eight error mechanisms.

3.14.6.2   *Procedure for Assessing PE*

The procedural steps for analyzing parameter P_E are as follows:

1. Define in detail the interaction to be performed by operating crews in terms of control board actions as described by the procedural steps (task analysis).

2. Develop a representation model for multi-step actions outlined in the procedure. Review carefully the steps in the procedure along with the control board layout to examine the need for breakdown of the interaction into subtasks. Also, review the applicable system fault tree(s) or process and instrument diagrams(P&IDs) to establish the functional requirements; e.g., if there are parallel trains a functional failure requires misalignment of both trains. This is important when considering dependencies between subtasks.

3. Include hardware failure events (or system unavailabilities) into representation model developed in step II.

4. Quantify probability of the manipulative error event represented in step III. The recommendations are as follows:

    a.   If there are no known problems with the control board labeling or other human factors, use estimates based on General Physics data (i.e., 0.01- 0.03) (Beare, Dorris, Bovell, Crowe & Kozinsky, 1983). A reduction in these values by a factor of five (i.e., 0.002 - 0.006) is recommended for the reasons stated earlier in this section (i.e., improved control boards, more training, more simulator time, etc.)

    b.   If it is shown that there are labeling or other human factors problems associated with certain control actions, then higher estimates for probability of slips should be used. In these situations, estimates can be made using information obtained during interviews with training instructors. A preliminary look at the ORE data shows that in cases where poor or unfamiliar labeling exists, slips can occur and, therefore, higher estimates for probability of slips are expected.

Ind. Avail in CR	CR Ind. Accurate	Warn/Alt. in Proc	Training on Ind.	$P_C{}^a$

(a) Neg.

(b) Neg.

(c) Neg.

(d) .0015

(e)   .05

(f)   .5

(g)   *
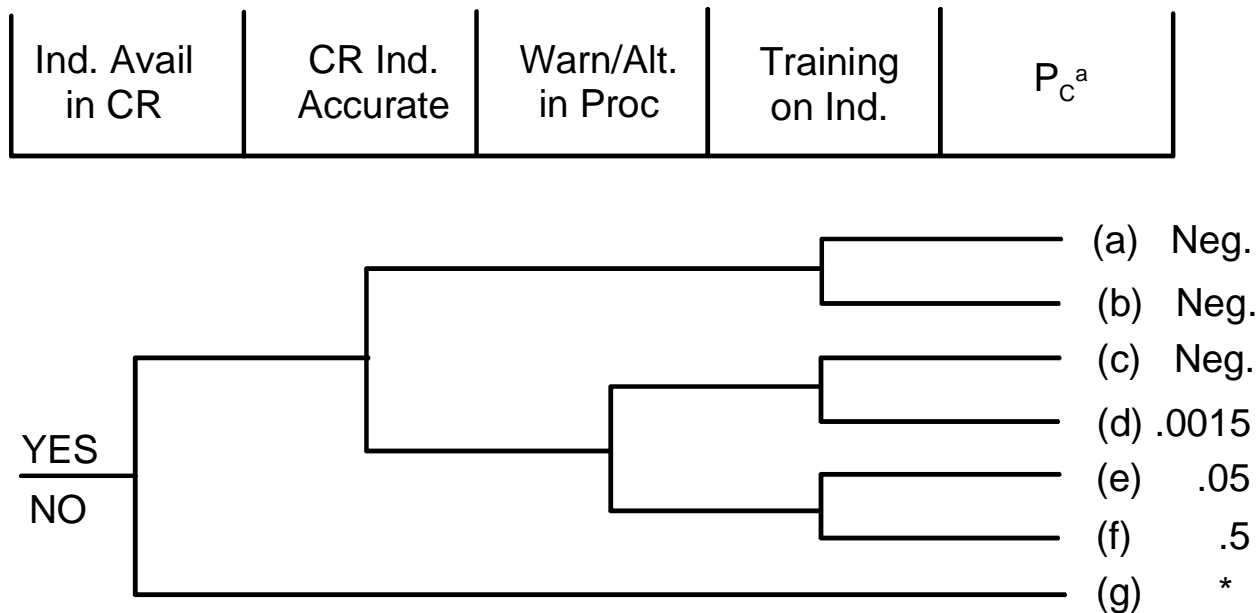
* In situations where the procedure or training specifies a course of action when the preferred information source is not available or the value of a parameter cannot be determined, the analyst must determine that the alternatives specified will lead to the same actions as the procedure would have directed, had the information been available. For situations where the crew must obtain information from ex-control room sources via a second-party report, the same analysis should be performed for the plant operator, who may have different procedures (or none) and very different training than members of the control room crew. The time for the second party to obtain the information should be subtracted from the available time window.

Explanation of Headings:
1. <u>Ind. Avail in CR.</u> Is the required indication available or functioning in the control room?
2. <u>CR Ind. Accurate.</u> Are the indications in the control room that are available accurate, or are they known to be inaccurate (e.g., due to degradation because of local extreme environmental conditions or isolation of the instrumentation)?
3. <u>Warn/Alt in Proc.</u> If the normally displayed information is expected to be unreliable, is a warning or a note directing to alternate information sources provided in the procedures?
4. <u>Training on Ind.</u> Has the crew received training in interpreting or obtaining the required information under conditions similar to those prevailing in this scenario?

Figure 11 . Decision Tree Representation of $P_C{}^a$, Availability of Information

Table 37 . Summary of values allowed for recovery.

Tree	Branch	Self-Review	Extra Crew	STA[a] Review	Shift[b,c] Change	ERF[d] Review
$P_C{}^a$	All	NC[f]	.5	NC	.5	.5
$P_C{}^b$	All	x[f]	NC	X	X	X
$P_C{}^c$	All	NC	NC	X	X	X
$P_C{}^d$	All	NC	.5	X	X	.1
$P_C{}^e$	a-h	X	.5	NC	X	X
$P_C{}^e$	i	.5	.5	X	X	X
$P_C{}^f$	All	NC	.5	X	X	X
$P_C{}^g$	All	NC	.5	X	X	X
$P_C{}^h$	All	NC	X	X	??	??

a. "STA (Shift Technical Advisor) review" is a review guided by a separate formal procedure such as the Westinghouse Safety Function Status Trees. For cue situations not covered in such documents, the STA's effects are credited in the "Extra Crew" column. Credit is not allowed for STA review until 15 minutes after the initial cue. Credit is only allowed for the initial STA review (for each shift) due to complete dependence --he will see it the first time or not at all.

b. Allow credit for shift change after 6 hours and every 8 hours thereafter.

c. Do not take extra credit for extra personnel on incoming shift.

d. Allow 30 minutes for first ERF (Emergency Response Facility) to be established, and take credit for only one ERF. The ERF may be credited with only one recovery opportunity per shift.

e. NC: no credit is allowed.

f. X: While the text suggests that revisiting these cases can be regarded as independent, the analyst may not feel confident that there are not in fact some underlying, yet unidentified, mechanisms that might induce dependency. Therefore, a higher value of X might be used instead of the HEP itself which would be more appropriate for complete independence.

Note that Swain and Guttman's nominal diagnosis model (Figure 12-4 of NUREG/CR1278) is also a TRC, but it exhibits several changes in slope, while empirically derived TRCs do not.

3.14.7   Error-Specific HEPs

CBDT method provides procedure for assessing the probabilities of eight error mechanisms. These error mechanisms contribute to failure to initiate correct response in timely manner due to failure (including delay) of detection, diagnosis, or decision. The eight error mechanisms are:

1.   The required data are physically not available to the control room operators.

2.   The data are available, but are not attended to.

3.   The data are available, but are misread or miscommunicated.

4.   The available information is misleading.

5.   The relevant step in the procedure is skipped.

6.   An error is made in interpreting the instructions.

7.   An error is made in interpreting the diagnostic logic (this is a subset of Item 6, but is treated separately for convenience).

8.   The crew decides to deliberately violate the procedure.

3.14.8   Task Dependencies and Recovery

Time-limitation dependency and cognitive dependency are discussed. However, their impacts on HEPs are not specified. Time limitation dependency would affect the available time assigned in the procedure related to the use of ORE simulator data (not covered in this evaluation, see discussion in Section 3.14.1). Cognitive dependency can potentially be handled in the task analysis (e.g., by SHARP 1) prior to use of CBDT for HEP quantification.

Recovery is explicitly modeled (see Table 37).

3.14.9   HEP Uncertainty Bounds

The CBDT suggests the use of the THERP uncertainty assessment technique or expert judgment to assign a range of uncertainty to the final HEP.

3.14.10   Level of Knowledge Required

HRA analysts with about a year working experience in the HRA field are expected to be capable of using CBDT.

3.14.11   Validation

3.14.11.1   Error Identification

CBDT does provide explicit guidelines for error identification.

3.14.11.2   HEP Quantification

The HEPs assigned in the decision trees are mainly based on THERP tables. Some values are also based on the judgment of the authors of the method. The source of each HEP used in the decision trees is explicitly specified. There is no known independent validation of CBDT results.

3.14.12   Reproducibility

3.14.12.1   Error Identification

CBDT does not do error identification.

3.14.12.2   HEP Quantification

The reproducibility mainly depends on the reproducibility of assessing the states of the thirty PSFs of the method (see 3.14.4). The process of assessing the states of these PSFs is mainly tied to the nature of the instructions of the operating procedures and observable plant and crew characteristics. As such, reproducibility is rated high.

3.14.13 Sensitivity

The CBDT HEPs are produced mainly based on THERP tables. Some values are based on expert judgment. The final P_C is a function of the states of thirty PSFs and the recovery factors. None of the PSFs can singly cause a significant change in the value of P_C. On the other hand some of the recovery factors could have a significant impact. However, since assessing the possibility and the nature of recovery actions is often not very difficult, recovery factors are not expected to be a major source of variability and sensitivity of the HEP results. The sensitivity is therefore rated low.

3.14.14 Experience Base

The CBDT method has been used in a number of U.S. nuclear power plant PRAs.

3.14.15 Resource Requirements

The CBDT procedures are very prescriptive and clearly stated. Their correct implementation, however, requires a good level of understanding of the system, possible accident environments, and the governing of emergency operating procedures. This might require significant effort (which can be done by one person instead of a group of analysts). The required level of effort is therefore rated medium.

3.14.16 Cost and Availability

The method is described in a proprietary document (Parry, Lydell, Spurgin, Moiene, & Beare, 1992) available through licensing from EPRI. EPRI can be contacted directly for the license fees.

3.14.17 Suitability for NASA Applications

While it can be argued that the conceptual framework of CBDT is general, the specifics of the method are closely tied to the nuclear plant environment. (This, of course, is the case with many other HRA methods reviewed in the study). In addition the CBDT procedure for HEP assessment is structured for a context mainly characterized by operators following operating and/or emergency procedures. Such procedures are typically not available for new designs. Even in the case of existing aerospace systems where such procedures do exist, they are very different than those used in nuclear power plants in most cases. For example, for the Space Shuttle the procedures are short and the decision time window, during abnormal situations, could be as short as one minute or a few seconds. Such differences would significantly limit the applicability of CBDT in its current form for NASA use.

3.15   Summary of HRA Method Attributes

To make comparison between the 14 HRA methods easier, Tables 38-44 were created and are shown below. Based on the selection criteria found in Section 1 and the input from the HRA experts during the peer review, four methods were recommended for NASA use. Similar comparison tables for those four methods are provided in Section 1.

Table 38 . Methods' features and capabilities.

Method	Screening	Task Decomposition	PSF List (number)	Causal Model	Coverage[1] 1: Ergonomics 2: Cognitive 3: Organizational
ASEP	No	Diagnosis, Action	Based on THERP	Single layer	1(limited), 2(limited), 3(limited)
SLIM	No	Not specified	User Defined (9 suggested)	Single layer	Analysts can define domain to be covered, suggested list covers 1,2,3
HEART	No	9 generic tasks	38	Single layer	1, 2, and 3
ATHEANA	No	Not specified	User-Defined	Implicit[8]	1, 2, and 3
CAHR	No	12 types of activities specified in the "Man-Machine System" diagram	35	Single layer	1, 2, and 3
UMH	No	7 task types	23	Two layers	1, 2, and 3
CESA	No	Decision error, error correction	5	Single Layer	1, 2 and 3
TRC	Yes	None specified	At least 1		Analysts can define domain to be covered, suggested list covers 1,2,3 (limited)
EPRI – Cause Based Decision Tree	No	Diagnosis, Action	30	Two layers	1, 2 (limited), 3
HP PFMEA	No	Functional analysis & Hierarchical task analysis	Many	Three layers	1, 2 and 3

[1] Method provides "relatively detailed" instructions for assessing the PSFs or factors' effect on the specific problem scope.

–   Ergonomics – design of controls systems, machine aspects, lighting, system design, physical workload, physical fatigue; i.e., anything physical or physiological

–   Cognitive – decision making, mental workload, cognitive fatigue; i.e., anything cognitive
–   Organizational – design of tasks, management impact on reliability of human, work processes, task organizations/procedural alignment, safety culture, team, communications

Table 39 . Methods' source, approach, and treatment of dependencies and recovery.

Method	Primary Source for HEP Estimates		Analysis Approach		HEPs for Specific Error Modes	Explicit Treatment of		Uncertainty Bounds Estimation[1]
	Number provided by method	Number produced by analyst	Manual	Software-Aided		Task/Error Dependencies	Recovery (includes actions with feedback)	
ASEP	√		√		Diagnosis & Action	√	√	√
SLIM		√	√	√	None specified			
HCR	√		√		Omission[2]			
HEART	√		√		None specified			
ATHEANA		√	√		Expert judgment	√	√	√
CAHR	√			√	Based on THERP			
UMH	√	√	√	√	Omission[2]	√		√
CESA	√	√	√		None specified	THERP-Based	THERP-Based	
TRC	√		√		Diagnosis only			√
EPRI CBDT	√		√	√	Diagnosis only		√	
HF PFMEA	--[3]	--[3]	√	Under development	Many error modes			

[1] None of the methods that provide uncertainty bound capability make a distinction between aleatory or epistemic nature of the uncertainties
[2] Fail to perform required action within specified time
[3] The HF PFMEA method does not calculate HEPs

Table 40 . Methods' error identification and HEP estimation.

Method	Error Identification		HEP Estimation			Experience Base**
	Reproducibility[1]	Validation[2]	Validation[2]	Reproducibility[1]	Sensitivity[3]	
ASEP	Does not do error identification	Does not do error identification	1 known validation	Medium	Medium	Many U.S. nuclear applications
SLIM	Does not do error identification	Does not do error identification	None	Low	Sensitive in one tail of Prob. Dist.	Many U.S. nuclear applications
HEART	Does not do error identification	Does not do error identification	2 known validations[5]	Low	High	U.K. nuclear, chemical and defense industries
ATHEANA	Medium	None	None	Low	Medium	U.S. nuclear, U.S. chemical weapons demilitarization, and railroad
CAHR	Does not do error identification	Does not do error identification	Benchmarked with THERP	High	Low	German nuclear, automotive, and air traffic management
UMH	Does not do error identification	Does not do error identification	None	Medium	Sensitive in one tail of Prob. Dist.	One U.S. nuclear application
CESA	High[4]	None	None	Medium	Low	One Swiss nuclear application
TRC	Does not do error identification	Does not do error identification	None	Medium	High in high probability l region	Several U.S. nuclear applications
EPRI CBDT	Does not do error identification	Does not do error identification	None	High	Low	Several U.S. nuclear applications
HF PFMEA	High	None	Does not do HEP quantification	Does not do HEP quantification	Does not do HEP quantification	NASA

**Refers to the number and variety of applications
[1] Reproducibility refers to the level of consistency in results produced by different individuals for the same task
[2] Refers to formal independent validation of the results
[3] Refers to the amount of change in numerical results when input parameters are changed
[4] Requires operating/emergency procedures and PRA models of the system
[5] One was performed by the method developer; another was performed by an independent expert

Table 41 . Resource requirements.

| Method | Knowledge Level | | | Tool | | Level of Effort Required For Error Identification | Level of Effort Required For Calculating HEPs[1] |
	HRA specialist (knowledge based)	HRA analyst (about one year of experience)	PRA analyst (skilled based)	Manual	Computer Code Needed		
ASEP			√	√		*	Low
SLIM		√		√		*	High
HCR		√		√		*	Medium
HEART		√		√		*	Low
ATHEANA	√			√		High	High
CAHR		√			√	*	Low[3]
UMH		√		√	√	*	Medium
CESA		√		√		Low	Medium
TRC		√		√		*	Medium
EPRI CBDT		√		√	√[2]	*	Low
HF PFMEA		√		√		High	**

*   Method does not include specific guidelines for error identification
**   Method does not include guidelines for HEP calculation

(1) Low = Look up in office, done in minutes to a day
    Medium = Up to 2 weeks
    High = Resource or time intensive (includes expert elicitation, more than 2 weeks of effort)
(2) A computer code is available to assist in the analysis, but is not essential.
(3) Given the proper database the analysis process is straightforward and resource requirement is rated low.

Table 42 . Resource requirements.

Method	Knowledge Level			Tool		Level of Effort Required For Error Identification	Level of Effort Required For Calculating HEPs[1]
	HRA specialist (knowledge based)	HRA analyst (about one year of experience)	PRA analyst (skilled based)	Manual	Computer Code Needed		
THERP		√		√		*	Medium
ASEP			√	√		*	Low
SLIM		√		√		*	High
HCR		√		√		*	Medium
CREAM		√		√		*	Medium
HEART		√		√		*	Low
NARA		√		√		*	Low
ATHEANA	√			√		High	High
CAHR		√			√	*	Low[3]
SPAR-H			√	√		*	Low
UMH		√		√	√	*	Medium
CESA		√		√		Low	Medium
TRC		√		√		*	Medium
EPRI CBDT		√		√	√[2]	*	Low
HF PFMEA		√		√		High	**

*   Method does not include specific guidelines for error identification
**   Method does not include guidelines for HEP calculation

(1) Low = Look up in office, done in minutes to a day
    Medium = Up to 2 weeks
    High = Resource or time intensive (includes expert elicitation, more than 2 weeks of effort)
(2) A computer code is available to assist in the analysis, but is not essential.
(3) Given the proper database the analysis process is straightforward and resource requirement is rated low.

Table 43 . Cost and availability of method, tools, and data.

Method	Method/Software	Parameter Values[1]	Raw Data[2]
THERP	Free	Free	Not available
ASEP	Free	Free	Not available
SLIM	Free	Not provided [3]	Not available
HCR	Free	Not provided [4]	Not available
CREAM	Free	Free	Not available
HEART	Free	Free	Not available publicly
NARA	Free	Free	Not available publicly
ATHEANA	Free	Not provided	Not provided
CAHR	License Fee	Free with code	Not available publicly
SPAR-H	Free	Free	Not available
UMH	Free	Free	Not available
CESA	Free	Not provided	Not available
TRC	Free	Free	Not available
EPRI CBDT	License Fee	Free with code	Free with code, primarily comes from THERP
HF PFMEA	Free	Not applicable (a qualitative method)	Not applicable (a qualitative method)

[1] Refers to all parameter values needed by the method (e.g., median time, reference HEP)
[2] Refers to original "raw" data used in developing the HEP estimates
[3] Refers to reference HEPs
[4] Refers to median response, $T_{1/2}$

Table 44 . Results of voting on methods suitability as (1) screening and (2) more detailed quantification.

Method	No. of Votes for Screening Method	Method	No. of Votes for Detailed Method	Total Number of Votes
THERP	3	THERP	5	8
ASEP	2	ASEP	0	2
SLIM	0	SLIM	1	1
CREAM	6	CREAM	7	13
HEART	2	HEART	1	3
NARA	7	NARA	2	9
ATHEANA	1	ATHEANA	0	1
CAHR	4	CAHR	0	4
SPAR-H	6	SPAR-H	4	10
UMH	1	UMH	0	1
CESA	0	CESA	0	0
HF PFMEA	2	HFP FMEA	0	2
EPRI CBDT	1	EPRI CBDT	1	2
TRC	0	TRC	0	0

150

3.16   Study Conclusion

After extensive discussion of NASA HRA needs, HRA methods selection criteria, and capabilities of the various methods, the most appropriate HRA method(s) for space mission (excluding ground processing and command and control) were identified for screening methodology (early concept design) and detailed analysis. THERP, CREAM, NARA, and SPAR-H were identified as most applicable to NASA HRA needs, particularly when used in combination. Further details on this conclusion are found in Section 1 and will not be repeated here.

Appendix A: HUMAN-RATING REQUIREMENTS SUPPORTING HUMAN ERROR MANAGEMENT

The NASA Office of Safety and Mission Assurance carefully crafted a set of requirements called the NASA Procedural Requirements (NPR) 8705.2, *Human-Rating Requirements for Space Systems* (NASA, 2005), to provide an extra level of safety for systems that will be operated and used by humans in space. These requirements incorporate design features into the system and implement safety processes to provide the maximum assurance that the crew and passengers will not sustain a permanent disability or fatality during the use of the system. These requirements cover all program activities including design, development, test, verification, management, and sustaining maintenance. As a part of this initiative, NASA approved its first Agency-level requirements that take a proactive approach to human error management (Figure A-1). NPR 8705.2 also contains notes that further explain the intent of these requirements.

Error Management for Single Point Failures	Error Management That Addresses Multiple Failures
The space system shall provide human error management in the following order of precedence (Requirement 34426): a. The system design prevents human error. b. The system reduces the likelihood of human error and provides the capability for the human to detect and correct the error through the incorporation of systems, controls, and associated monitoring. c. The system provides a method to limit the negative effects of errors so that the error does not result in a fatality or permanent disability.	The system shall be designed and operated so that neither two inadvertent actions during operation or in-flight maintenance nor a combination of one inadvertent action and one failure result in crew or passenger fatality or permanent disability (Requirement 34422). **Note**: *Inadvertent action includes, but is not limited to, out-of-sequence actions, wrong keystrokes, or inadvertent switch throws.*

Figure A-1. Human Error Management Requirements in NPR 8705.2A, Human-Rating Requirements for Space Systems.

The philosophy behind these requirements is that they:

1. Serve as a forcing function, ensuring the design community consciously considers human performance as a critical component in overall system performance and then designs the system accordingly.

2. Ensure that the design community predicts potential critical human errors and implements measures that prevent these errors or mitigates their effects early during the design process, rather than waiting until the design is complete.

3. Allow flexibility in error management by providing an order of precedence similar to the method to manage hazards documented in NPR 8715.3, NASA Safety Manual (NASA, 2004).

Often human error is only one of a multiple chain of events that together produce an accident. Because of this, it is not sufficient to address and manage each individual error (single point failure) independently. NASA must also evaluate the chain of events (combinations of human, equipment, and software failures) that will lead to accidents (Requirement 34422 in Figure A-1). These "minimal cut sets" can be identified using modeling techniques such as fault trees. Once identified, the cut sets allow analysts to identify scenarios that pose significant risk so that the risks can be managed.

One method to manage human-error's contribution to risk is to design an error-tolerant system, a system that tolerates multiple human errors without catastrophic failure or human injury. Hardware-software system design is the preferred method to produce an error-tolerant system. However, meeting the two-inadvertent action requirement may not always be possible in system design. In many phases of dynamic space flight, human error prevention may not be technically feasible. Consequently, the Human-Rating Requirements provide flexibility using paragraph 3.1.4 of NPR 8705.2 (Figure A-2). Although this paragraph provides the program manager with rationale for exceptions, deviations, and waivers from the error management requirement, it still requires human error analysis to be performed and requires the program to acknowledge and accept the risk posed by the potential human errors that have not been eliminated.

3.1.4 The Program Manager shall provide evidence and rationale that one or more of the following are met when requesting an exception, deviation, or waiver to the two-inadvertent action requirement (Requirement 34424).

a. Meeting the two-inadvertent action requirement is technically not feasible.

b. The program manager demonstrates through analysis that redundancy does not reduce the critical system contribution to cumulative risk, or the contribution of common cause failures to that critical system's failure.

c. The Program Manager has demonstrated by test data and comprehensive risk analyses that the system shall provide personnel with the capability to detect and recover from the inadvertent actions in time to prevent crew or passenger fatality or permanent disability.

Figure A-2. Paragraph 3.1.4 of NPR 8705.2.

For each new NASA crewed system, the Human-Rating Requirements will be translated by the program into lower level system requirements and functional design specifications. Naturally, during the system life cycle, the level of detail available with respect to potential errors will evolve from very generic function-related errors (when the system is in early conceptual design) to very specific cognitive and action execution error (when the system is fully developed). The HRA model will be refined and enhanced as the system is developed and becomes operational, and error management strategies will be incorporated as required.

Appendix B: SPACE MISSION HUMAN ACTIVITIES

None of the HRA methods that exist today were developed specifically for human activities related to space missions. Rather, the body of expertise and majority of HRA studies relate to commercial nuclear power plants. As an essential step towards the assessment of the applicability of current HRA methods for NASA use, an overview of the types of human activities related to space missions was created. It is provided here as a reference for those unfamiliar with the type and variety of activities related to space flight. This section provides a general description of the various activities undertaken by NASA that are related to space missions and then comments on the some of the differences that exist between factors affecting human performance in space compared to ground activities.

NASA human activities directly related to space flights can be classified into the following six categories:

1. Ground processing;
2. Space flight dynamic phases;
3. IVA or Intra Vehicular Activities (including maintenance, re-supply, communications, and science research);
4. EVA or Extra Vehicular Activities (including connecting and disconnecting cooling cables, etc);
5. Destination and surface operations; and
6. Earth landing.

Within these categories, NASA professionals perform a wide variety of activities, some similar to those seen in other complex industries, others unique to space flight. These categories are described in more detail in the following sections.

Ground processing

Ground processing includes a wide variety of human activities, such as system design, manufacturing and systems acquisition, vehicle assembly, preparation of science payloads, payload assembly, integrated vehicle and payload processing and test (e.g., integrating the crew exploration vehicle with the crew exploration launch vehicle), vehicle maintenance and repair, transport of the vehicle, and crew launch day preparation. To illustrate, below are some photographs of typical Space Shuttle ground processing activities (Figures B-1 through B-6). NASA has found that 78% of the Space Shuttle ground support operations incidents resulted from human error (Perry et al., 1993).
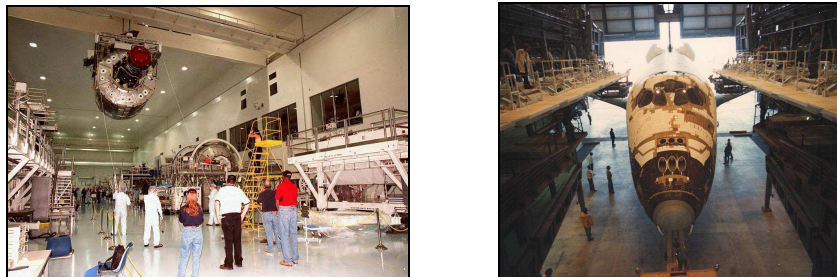
Figure B-2. Processing the International Space Station Node 1 in the Space Station Processing Facility, Kennedy Space Center, Florida (left), and preparing the orbiter in the Orbiter Processing Facility (OPF).

Figure B-2. Maintenance, repair, and payload processing: Technicians working on the Space Shuttle dome heat shields (left), technicians inspecting windows on the Space Shuttle in preparation for flight (middle), and technicians and scientists preparing the Genesis payload for its mission (right).

Figure B-3. Vehicle Transport: Shuttle near the OPF at KSC (left). Shuttle moves to the Vehicle Assembly Building at KSC (middle). Crew preparing for launch (right).

Space flight dynamic phases of flight

Space flight dynamic phases of flight include all activities during launch and lift-off, docking with other space flight vehicles, and descent to a planetary surface. Many of these activities are similar to those found in aviation and air traffic control. However, the impact of excessive gravitational forces and the physical, psychosocial, and cognitive aspects of space flight can have a significant impact on the flight crew's performance. Consequently, although the tasks themselves are not unique to NASA, the conditions in which the crew operates are novel and can impact human performance. During ascent, potential human errors that can cause significant risk to the crew can be made by ground processing during activities such as propellant tanking, by the launch control team during decision making and trouble shooting, or by the crew.

Figure B-4. Launch of STS 71 (left), launch control room at KSC during a launch (middle and right).

Intra Vehicular Activity (IVA)

IVA includes vehicle and/or system assembly; system maintenance (preventative and corrective); science research; command, control, and communications; resupply; planning and scheduling; and habitability (housekeeping, environmental control, exercise, health maintenance, food preparation, stowage, and waste management) (Figure B-5).
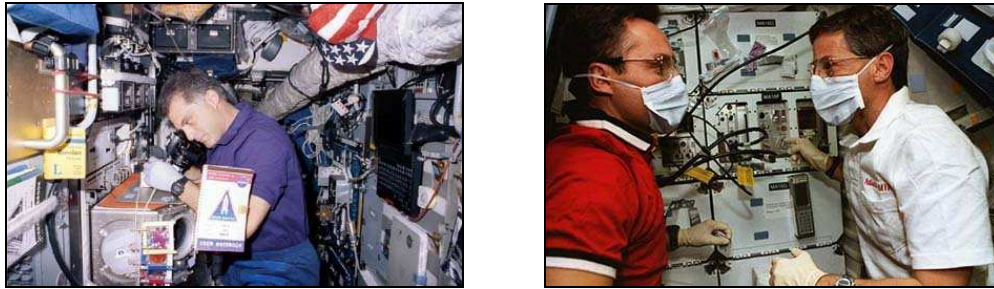
Figure B-5. IVA: Astronauts working on various science experiments.

Figure B-5 represents some of the types of In-Flight/En Route Activities that NASA professionals perform. Flight crew activities differ significantly from those performed on the ground. For example, simple science research activities, such as staining three slides takes 19 seconds longer in microgravity than on the Earth because the crew have to secure each item they use in microgravity, or the item will float away. Additionally, the flight crew activities are different than ground-based activities because the crew's physiology is affected while in space, and, consequently, the crew's performance is affected. Research indicates that two of the most significant factors affecting crew performance in space are microgravity and fatigue.

Extra Vehicular activities (EVA)

EVA include maintenance (preventative and corrective), science research activities outside the space vehicle (in zero-gravity or microgravity), and resupply (Figure B-6).
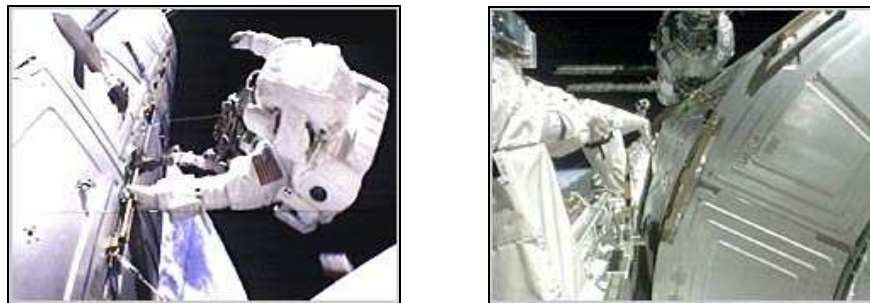
Figure B-6. Robert Curbeam disconnects power and cooling cables between Destiny and Atlantis on STS-98 (left). Curbeam EVA in the STS-98 Space Shuttle payload bay (right).

Destination surface operations and support

These include maintenance (preventative and corrective); science research activities in microgravity (1/3 or 1/6 of the Earth's gravity); mission planning; and command, control, and communications (Figure B-6).

Earth landing, egress, and recovery

This last set of activities related to space flight includes Earth landing, crew egress, crew recovery, and vehicle recovery (Figure B-7).

Figure B-7. Space Shuttle STS-71 Landing (left) and crew of Apollo 11 egressing the crew module (right).

Initially, the study described in this report evaluated HRA methods that are applicable to human interactions for maintenance activities (ground processing) and flight operations (launch control, mission control, and space flight crew). Consequently, Tables B1-B6 were created to help the HRA analysts determine 1) the specific types of activities that would be performed. 2) which activities would require HEA to provide error management as required by the NASA Human-Rating Requirements (illustrated by shading), 3) which activities must be completed in a short amount of time and thus required time is to be considered as a possible PSF (illustrated by TC = time critical). These tables are provided for reference for those who are unfamiliar with NASA space activities.

Table B-0.1. NASA human activities in ground processing and personnel involved.

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
Design	X	x								
Manufacturing	X	x	X							
Element Receipt & Acceptance	X	x		X						
Assembly & Integration	X	x	X							
Test	X	x		X						
Maintenance	X	x		X						
Repair & Refurbishment	X	x		X						
Logistics	X	x		X						
Spaceport Services	X	x		X						
Payload Processing	X	x		X						
Element Processing & Turnaround (Orbiter, SRB, Tank, ELV)	X	x		X						
Payload, Element, System Transport	X	x		X						
Launch Processing (Integration Element & Payload Processing)	X	x		X	X					
Crew (Launch Day Preparation)				X	X	X		x		

Key

X   =   Personnel Supporting the Activity
x TC   =   Personnel Perform The Activity. Activity Has Some Time Critical Elements

| X | Box Shaded As Below = Human Error Identification and Analysis Required to Meet Human-Rating Requirement

Table B-0.2. NASA human activities in EVA and personnel involved.

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
Preparation (Pre-Breathe, donning suit)						X	x	x		
Assembly						X	x	x		
Maintenance & Repair						X	x	x		
Science						X	x	x		
Operating Robotic Systems						x TC	x TC	x		

Key

X   =   Personnel Supporting the Activity
x TC   =   Personnel Perform The Activity. Activity Has Some Time Critical Elements
Box Shaded =   Human Error Identification and Analysis Required to Meet Human-Rating Requirements

Table B-0.3. NASA human activities in IVA and personnel involved.

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
Habitability (e.g., station cleaning, environmental control, food preparation, exercise, personal hygiene, stowage)						X	X	x		
Planning and Scheduling						X	X	x		
Science Research						X	X	x		
System Reconfiguration (e.g. ISS reconfiguration)										
Maintenance						X	X	x		

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
(preventative and corrective) including trouble shooting & diagnosis)										
Command, Control, and Communication						X	X	x		
Medical diagnosis and care						X	X	x		
Practice emergency procedures (Fires, Micro-Meteoroid penetrations, system failures)						X	X	x		
Operating Robotic Systems						x TC	X	x		
Vehicle and/or System Assembly						X	X	x		

Key

X	=	Personnel Supporting the Activity
x TC	=	Personnel Perform the Activity. Activity Has Some Time Critical Elements
		Box Shaded = Human Error Identification and Analysis Required to Meet Human-Rating Requirements

Table B-0.4. NASA human activities in EVA and personnel involved.

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
Preparation (Pre-Breathe, donning suit)						X	x	x		
Assembly						X	x	x		
Maintenance & Repair						X	x	x		
Science						X	x	x		
Operating Robotic Systems						x TC	xTC	x		

Table B-0.5. NASA human activities in destination surface operations & support and personnel involved.

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
Mission Planning						X	x	x		
Habitability						X	x	x		
Assembly						X	x	x		
Maintenance & Repair						X	x	x		
Science						X	x	x		
Ground Travel						X	x	x		
Operating Robotic Systems						X	x	x		

Table B-0.6. NASA human activities in Earth landing, egress, and recovery and the personnel performing these tasks.

Activities	Personnel									
Types of Human Activities	Administrative & Management	Design & Engineering	Fabrication & Manufacturing	Ground Processing	Launch Control	Flight Crew	Mission Control	Mission Management Team	Landing Site & Recovery	Search & Rescue
Landing Operations						X	x	x	x	
Search						x TC	x TC	x TC	x TC	x TC
Crew Recovery Operations						x TC	x TC	x TC	x TC	x TC
Vehicle Recovery Operations						X	x	x	x	x

<u>Key</u>

X = Personnel Supporting the Activity
x TC = Personnel Perform The Activity. Activity Has Some Time Critical Elements
Box Shaded = Human Error Identification and Analysis Required to Meet Human-Rating Requirements

Appendix C: UNIQUE PERFORMANCE SHAPING FACTORS

To further assist the NASA analyst in the application of HRA methods, the following section provides a discussion of the unique aspects of PSFs in NASA human-space-flight missions. This Appendix provides more detailed description of these unique PSFs and a list of key "bioastronautics risks" identified by the science community as needing additional research and/or countermeasures to ensure safe human space travel.

Zero gravity and micro gravity, constant radiation, extreme temperatures and a novel environment produce physiological changes and psychological changes in the human body. These changes also manifest themselves in pathological alterations that later affect the crewmember's return to life on Earth. The space vehicle is a biosphere, a home away from home that must maintain specific characteristics such as air quality, gas pressure, temperature, and humidity to keep the crew healthy and performing normally. Physiological changes begin to occur when the crew enters zero gravity. Just as everything else in a weightless environment, the crewmember and his tools float. Without the gravity, the fluids in the human body move from the lower extremities towards the head. The shape of the eyes change, and vision changes slightly. Although cardiovascular function is adjusted for the changes in workload on the body, red blood cell counts fall and the human's capability to respond to pathogens is altered. The human begins to loose body fluids (about 1 liter is lost). Along with this, many crewmembers experience nausea and vomiting in the first few days of flight. From the time the human reaches zero gravity, the adaptation process has the potential to affect human performance. As the central nervous system responds, the person may experience this motion sickness, which can last up to 2-3 weeks. **During this time, serious anomalies or significant task workload could potentially cause the crew to exhibit high error rates that could impact critical system functions.**

In addition to the immediate effects of the central nervous system adaptation to zero gravity, there are other physiological changes. The human no longer has the pull of gravity on the muscles, which is a requirement in maintenance of balance, or the normal requirement to use muscles to push, pull, and lift items in gravity. **Consequently, muscles lose strength, and bone mass is lost.** Vehicle systems that were easily operated on Earth (e.g., hatches) become more difficult to operate in space (e.g., crew may not have the strength to open the hatch). This degradation of muscle and bone continues throughout the human's time in space. For example, the human looses approximately 1-2% of bone mass every month (Nicogossian, 2003).

One area of greatest uncertainty regards the amount of cosmic radiation the humans will receive during their space mission. Crewmembers are exposed to a constant dose of cosmic and solar radiation. If they perform space walks, their exposure is even higher. Anomalous events (such as solar flares) have the potential to cause the crew to reach and/or exceed the lifetime radiation exposure limit.

In addition to the physiological changes that occur, psychosocial adaptation problems occur (Flynn, 2005). Both group cohesion and crewmember motivation are reduced. This has been documented in numerous settings from polar exploration teams to submarine crews. Isolation has other impacts on crew performance. The crew cognitive capabilities are lessened. Working memory, attention, and concentration have increased instability (Gushin, 1996). Additionally, interpersonal tension and the development of subgroups appear. This may restrict communication with team members in space and/or with team members on the ground. The effects of isolation have the potential to negatively impact crew performance, especially where extensive memory of task steps and/or group coordination is required. The crew also experiences sleep and circadian rhythm problems because normal day and night cues are lost. This can cause fatigue and impact crew performance.

Research into bioastronautics has identified a number of physiological and psychological issues unique to the space environment. Most of these factors fall outside of what are typically considered performance shaping factors by HRA methods; however, they may impact the potential for human error or interact with

more typical PSFs to increase the types of human errors that are seen. Table C-1 provides a sample set of the risks, their relative risk ranking by the science community, and the discipline area requiring research and/or mitigation. Up-to-date itemized risks can be seen in http://bioastroroadmap.nasa.gov/index.jsp. Some of these risk factors can be viewed as PSFs in the HRA context and the way that various methods try capture the effects of physical and psychological factors in the HEP assessment.

Table C-1. Sample set of risks and relative ranking.

ID[1]	Risk Title[2]	Rank[3]	Discipline Area[4]
1	Inability to Maintain Acceptable Atmosphere in Habitable Areas	1	Advanced Life Support
2	Inability to Provide and Recover Potable Water	2	Advanced Life Support
3	Inadequate Supplies (including maintenance, emergency provisions, and edible food)	2	Advanced Life Support
4	Inability to Maintain Thermal Balance in Habitable Areas	3	Advanced Life Support
5	Inability to Adequately Process Solid Wastes	3	Advanced Life Support
6	Inadequate Stowage and Disposal Facilities for Solid and Liquid Trash Generated During Mission	4	Advanced Life Support
7	Inadequate Nutrition (Malnutrition)	1	Food & Nutrition
8	Unsafe Food Systems	2	Food & Nutrition
9	Acceleration of Age-Related Osteoporosis	1	Bone Loss
10	Fracture & Impaired Fracture Healing	2	Bone Loss
11	Injury to Soft Connective Tissue, Joint Cartilage, & Intervertebral Disc Rupture w/ or w/o Neurological Complications	3	Bone Loss
12	Renal Stone Formation	4	Bone Loss
13	Occurrence of Serious Cardiac Dysrhythmias	1	Cardiovascular Alterations
14	Impaired Response to Orthostatic Stress	1	Cardiovascular Alterations
15	Diminished Cardiac Function	2	Cardiovascular Alterations
16	Manifestation of Previously Asymptomatic Cardiovascular Disease	3	Cardiovascular Alterations
17	Impaired Cardiovascular Response to Exercise Stress	4	Cardiovascular Alterations
18	Human Performance Failure Because of Poor Psychosocial Adaptation	1	Human Behavior & Performance
19	Human Performance Failure Because of Sleep and Circadian Rhythm Problems	2	Human Behavior & Performance
20	Human Performance Failure Because of Human System Interface Problems & Ineffective Habitat, Equipment, Design, Workload, or Inflight Information and Training Systems	3	Human Behavior & Performance
21	Human Performance Failure Because of Neurobehavioral Dysfunction	4	Human Behavior & Performance
22	Immunodeficiency/Infections	1	Immunology, Infection & Hematology (I, I & H)
23	Carcinogenesis Caused by Immune System Changes	1	I, I & H
24	Altered Hemodynamic and Cardiovascular Dynamics caused by Altered Blood Components	1	I, I & H
25	Altered Wound Healing	2	I, I & H
26	Altered Host-Microbial Interactions	3	I, I & H
27	Allergies and Hypersensitivity Reactions	2	I, I & H

Table C-1. (Continued).

ID[1]	Risk Title[2]	Rank[3]	Discipline Area[4]
28	Loss of Skeletal Muscle Mass, Strength, and/or Endurance	1	Muscle Alterations & Atrophy
29	Inability to Adequately Perform Tasks Due to Motor Performance, Muscle Endurance, and Disruption in Structural and Functional Properties of Soft & Hard Connective Tissues of the Axial Skeleton	1	Muscle Alterations & Atrophy
30	Inability to Sustain Muscle Performance Levels to Meet Demands of Performing Activities of Varying Intensities	2	Muscle Alterations & Atrophy
31	Propensity to Develop Muscle Injury, Connective Tissue Dysfunction, and Bone Fractures Due to Deficiencies in Motor Skill, Muscle Strength and Muscular Fatigue	3	Muscle Alterations & Atrophy
32	Impact of Deficits in Skeletal Muscle Structure and Function on Other Systems	not ranked	Muscle Alterations & Atrophy
33	Disorientation and Inability to Perform Landing, Egress, or Other Physical Tasks, Especially During/After G-Level Changes (Acute spontaneous & provoked vertigo, nystagmus, oscillopsia, poor dynamic visual acuity)	1	Neurovestibular Adaptation
34	Impaired Neuromuscular Coordination and/or Strength (Gait ataxia, postural instability)	2	Neurovestibular Adaptation
35	Impaired Cognitive &/or Physical Performance Due to Motion Sickness Symptoms or Treatments, Especially During/After G-Level Changes (Short-term memory loss, reaction time increase, drowsiness, fatigue, torpor, irritability, ketosis)	3	Neurovestibular Adaptation
36	Vestibular Contribution to Cardioregulatory Dysfunction (Postlanding orthostatic intolerance, sleep & mood changes)	4	Neurovestibular Adaptation
37	Possible Chronic Impairment of Orientation or Balance Function Due to Microgravity or Radiation (Imbalance, gait ataxia, vertigo, chronic vestibular insufficiency, poor dynamic visual acuity)	5	Neurovestibular Adaptation
38	Carcinogenesis Caused by Radiation	1	Radiation Effects
39	Late Degenerative Tissue Effects Including Non-Cancer Mortality, Cataracts, and Central Nervous System (CNS) Effects	2	Radiation Effects
40	Synergistic Effects from Exposure to Radiation, Microgravity & Other Spacecraft Environmental Factors	3	Radiation Effects
41	Early or Acute Effects from Radiation Exposure	4	Radiation Effects
42	Radiation Effects on Fertility, Sterility, and Heredity	5	Radiation Effects
43	Trauma and Acute Medical Problems	1	Clinical Capabilities
44	Toxic Exposure	2	Clinical Capabilities
45	Altered Pharmacodynamics & Adverse Drug Reactions	3	Clinical Capabilities
46	Illness and Ambulatory Health Problems	4	Clinical Capabilities
47	Prevention and Treatment of Decompression Sickness in NASA Operations	5	Clinical Capabilities
48	Difficulty of Rehabilitation Following Landing	6	Clinical Capabilities
49	Postlanding Alterations in Various Systems Resulting in	1	Multisystem (Cross-Risk)

Table C-1. (Continued).

ID[1]	Risk Title[2]	Rank[3]	Discipline Area[4]
	Severe Performance Decrements and Injuries		Alterations
50	Allergies and Hypersensitivity Reactions from Exposure to the Enclosed Spacecraft & Other Environmental Factors	3	Environmental Health
51	Inability to Maintain Acceptable Atmosphere in Habitable Areas Due to Environmental Health Contaminants	1	Environmental Health
52	Inability to Provide and Recover Potable Water Due to Environmental Health Contaminants	2	Environmental Health
53	Inadequate Nutrition (Malnutrition) Due to Inability to Provide and Maintain a Bioregenerative System	3	Advanced Life Support
54	Difficulty of Rehabilitation Following Landing Due to Nutritional Deficiencies	4	Food & Nutrition
55	Human Performance Failure Due to Nutritional Deficiencies	3	Food & Nutrition

[1] Risk Identification number: unique number assigned to each risk (1-55) used to track/identify each risk

[2] Risk: the title of each risk

[3] Rank: assigned to each risk within its discipline; a discipline may have more than one risk with the same risk ranking (risk order)

[4] Discipline area: there are 12 discipline areas in the BCPR representing the 55 risks

Appendix D: PEER REVIEW

D-1: Peer Review Team

A group of internationally recognized HRA experts and practitioners were invited to review the white paper and participate in a workshop to help evaluate the selected methods and identify the best candidates for NASA applications. The workshop was held January 16-18, 2006, at the NASA Kennedy Space Center in Florida. Among the participants were developers of 9 of the methods selected for review. Others were HRA analysts with extensive experience in various industries and NASA HRA/PRA task managers. Table D-1 provides the list the workshop participants and their affiliations.

Table D-1. NASA HRA workshop participants.

Name	Affiliation	Method(s) Developed
Adelstein, Bernard "Dov"	NASA Ames Research Center	
Bley, Dennis	Buttonwood Consulting, Inc.	ATHEANA
Boring, Ronald	Idaho National Laboratory	
Boyer, Roger	NASA Johnson Space Center	
Bye, Andreas	Institute for Energy Technology, Norway	
Chandler, Faith	NASA Headquarters	HF PFMEA
Chang, Yung Hsien James	University of Maryland/Paul Scherrer Institute, Switzerland	University of Maryland Hybrid
Cook, Richard	University of Chicago	
Cooper, Susan	Nuclear Regulatory Commission	ATHEANA
Dang, Vinh	Paul Scherrer Institute, Switzerland	
Gertman, Dave	Idaho National Laboratory	SPAR-H, ASP
Groth, Katrina	University of Maryland	
Hallbert, Bruce	Idaho National Laboratory	
Hamlin, Teri	NASA Johnson Space Center	
Julius, Jeffrey A.	Scientech	
Kirwan, Barry (provided written comments)	Eurocontrol, France	
Marble, Julie	Idaho National Laboratory	SPAR-H
Mosleh, Ali	University of Maryland	University of Maryland Hybrid
Parry, Gareth	Nuclear Regulatory Commission	
Prassinos, Pete	NASA Headquarters	
Sträeter, Oliver	Eurocontrol, Belgium	CAHR
Stewart, Michael	NASA Johnson Space Center	

D-2: Peer Review Process

Prior to the workshop the experts were provided with an earlier draft of this report and the set of 12 questions listed in Table D-2. These questions served to structure the workshop discussions of NASA HRA needs and assessment of HRA methods relative to those needs. During the workshop experts answered the questions and evaluated the report. All feedback received from the workshop has been incorporated into this document.

Table D-2. Questions posed in the white paper for HRA experts.

Do you agree with the pool of 12 HRA methods selected for review? If not, are there any methods that should be considered in addition to the 12 methods?
In addition to those described in the white paper, do you believe any other attributes should have been used to compare the HRA methods? If yes, what attributes?
Do you agree with the ranking of HRA methods presented in the white paper? Please provide comments and justification for repositioning a method ranking in listed in the white paper.
Of those HRA methods discussed, what existing method do you believe is the best to support early concept design? Please provide justification.
Of those HRA methods discussed, what existing method do you believe is the best to support unmanned and manned systems? Please provide justification.
What are the weaknesses of the existing approaches (that you identified in No. 4)? What improvements are needed?
Do the set of existing PSFs associated with the above method you have selected describe NASA's environment sufficiently, or should NASA's unique PSFs be incorporated into the NASA HRA?
If NASA should incorporate new PSFs into the analysis, what basis should be used for their inclusion?
What types of generic tasks best fit the task types encountered in NASA space missions?
Is the generic data derived from various industries potentially applicable to NASA space flight tasks? If yes, what data should be used?
Should NASA consider the dependencies between PSFs in the HEP calculations? If so, what data or evidence is available to assess PSFs' dependencies and level or degree of impact on HEPs?
Should the uncertainties in HEPs be estimated with consideration of both the epistemic and the aleatory aspects? If yes, what would be the aleatory source for HEPs? What would be the basis for epistemic characterization?
What future research in HRA do you believe is needed to support NASA's Exploration Mission Objectives?

Appendix E: DEFINITIONS

Error of Omission (EOO): The failure to perform the required action or complete the required task.

Error of Commission (EOC): Performing the required task or action incorrectly (wrong time, sequence, quantity, location), performing the require task or action on the wrong system interface/control, or performing an undesired action while attempting the desired action.

Functional Analysis: A systematic method used to break down the system or process into major components (functions) so that each can be described in terms of a system subgoal statement.

Functional Flow Diagram (FFD): A method of graphically representing a system's process flow in chronological order by generating functional blocks that describe "what is to be done" rather than the means to do it and organizing these blocks in order from process initiation to completion.

Human Error: Either an action that is not intended or desired by the human <u>or</u> a failure on the part of the human to perform a prescribed action within specified limits of accuracy, sequence, or time such that the action or inaction fails to produce the expected result and has led or has the potential to lead to an unwanted consequence.

Human Error Analysis (HEA): A systematic approach to evaluate human actions that identifies potential human error, models human performance, and qualitatively characterizes how human error affects a system. HEA provides an evaluation of human actions and error in an effort to generate system improvements that reduce the frequency of error and minimize the negative effects on the system. HEA is often referred to as a qualitative HRA.

Human Error Probability (HEP): A measure of the likelihood that a human will fail to imitate the correct, required, or specified action or response in a given situation or by commission will perform the wrong action. The HEP is the probability of the human failure event (ASME RA-S-2002).

Human Factors Engineering (HFE): The application of knowledge about human capabilities and limitations to system, equipment, job, or environment design and development to achieve efficient, effective, comfortable, and safe performance with minimum cost, manpower, skill, and training. Human engineering assures that the system, equipment design, required human tasks, and work environment are compatible with the sensory, perceptual, mental, and physical attributes of the personnel who will operate, maintain, control, and support it.

Human Failure Event: A basic event that represents a failure or unavailability of a component, system, or function that is caused by human action or an inappropriate action (ASME, RA-S-2002).

Human Performance: The physical and mental activity required of the crew and other participants to accomplish mission goals. This includes the interaction with equipment, computers, procedures, training material, the environment, and other humans.

Human Reliability Analysis (HRA): A method by which human reliability is estimated. In most cases, HRA is restricted to quantitative analysis**.**

Human Reliability: The probability of successful performance of the human activities necessary for either a reliable or an available system, specifically, the probability that a system-required human action, task, or job

will be completed successfully within a required time period, as well as the probability that no extraneous human actions detrimental to the system reliability or availability will be performed.

Initiating Event (IE): An event that has the potential to cause loss of a system function leading to an undesired end state such as loss of life, damage to or loss of property or equipment, failure of a mission, unavailability of a system, or damage to the environment.

Minimal Cut Set: A cut set containing the minimum subset of primary elements whose occurrence guarantees the occurrence of the top event (accident).

Performance Shaping Factor. A factor that influences human performance and human error probabilities.

Pivotal Event. An event that is a success or a failure of a response, or an occurrence or non-occurrence of an external condition or key phenomena after the initiating event that mitigates or aggravates the severity of the consequence.

Process Failure Mode Effect Analysis (PFMEA). A procedure used to analyze each component of the process for each possible failure mode and the "worst case" effect. It differs from the standard FMEA in that it analyzes the system's processes rather than a specific piece of equipment.

Tailoring: A process where a written authorization is given to the program from the Independent Technical Authority or designees prior to the approval of the Human-Rating Plan, allowing the program to exclude or modify a requirement in NPR 8705.2, Human-Rating Requirements for Space Systems, from the Human-Rating Plan, because the system does not have the component/subsystem described in that requirement, and consequently the requirement does not apply as written. For example, the system is not a flight vehicle; therefore, it is not required to perform flight tests.

Task Analysis. A systematic method to identify, list, and break down each task into the steps and substeps that describe the required human activities in terms of physical actions and/or cognitive processes (e.g., diagnosis, calculation, and decision making) necessary to achieve the system's goal.

Violation. An action that was intended and desired by the human that departs from rules (e.g., intentionally skipping a step in a procedure or taking a short cut) or laws (e.g., speeding).

Appendix F: REFERENCES

American Society of Mechanical Engineers. (2002). *Standard for Risk Assessment for Nuclear Power Plant Operations (ASME-RA-S-2002)*. New York, New York.

Atwood, C.L. (1996). *Constrained non-informative priors in risk assessment*. Reliability Engineering & System Safety. 53(1): p. 37-46.

Barriere, M., D. Bley, S. Cooper, J. Forester, A. Kolaczkowski, W. Luckas, G. Parry, A. Ramey-Smith, C. Thompson, D. Whitehead, and J. Wreathall (2000). *Technical Basis and Implementation Guideline for A Technique for Human Event Analysis (ATHEANA)*. NUREG-1624, Rev. 1, Washington D.C.: U.S. Nuclear Regulatory Commission.

Barriere, M.T., J. Wreathall, S.E. Cooper, D.C. Bley, W.J. Luckas, and A. Ramey-Smith (1995). *Multidisciplinary Framework for Human Reliability Analysis with an Application to Errors of Commission and Dependencies*. NUREG/CR-6265, Washington D.C.: U.S. Nuclear Regulatory Commission.

Barriere, M.T., W.J. Luckas, D. Whitehead, and A. Ramey-Smith (1994). *An Analysis of Operational Experience During LP&S and a Plan for Addressing Human Reliability Assessment Issues*, NUREG/CR-6093, Washington DC: U.S. Nuclear Regulatory Commission.

Beare, A.N., R.E. Dorris, C.R. Bovell, D.S. Crowe, and E.J. Kozinsky (1983). *A Simulator-Based Study of Human Errors in Nuclear Power Plant Control Room Tasks*, NUREG/CR-3309, Washington DC: U.S. Nuclear Regulatory Commission.

Bott, T.F., E.J. Kozinsky, C. Crowe, and P.M. Haas (1981). *Criteria for Safety-Related NPP Operator Actions: Initial PWR Simulator Exercises*. 1981, NUREG/CR-1908, Washington DC: U.S. Nuclear Regulatory Commission.

Broughton, J., R. Carter, F. Chandler, G. Holcomb, B. Humeniuk, B. Kerios, P. Bruce, P. Snyder, S. Strickland, B. Valentino, L. Wallace, T. Wallace, and D. Zeiters (1999). *Human Factors Process Failure Modes & Effects Analysis*. PGOC91-F050-JMS-99286: Boeing.

Cooper, S.E., A.M. Ramey-Smith, J. Wreathall, G.W. Parry, D.C. Bley, W.J. Luckas, J.H. Taylor, and M.T. Barriere (1996). *A Technique for Human Error Analysis (ATHEANA)*. NUREG/CR-6350; BNL-NUREG-52467, Washington D.C.: U.S. Nuclear Regulatory Commission.

Cooper, S.E., W.J. Luckas, and J. Wreathall (1995). *Human-System Event Classification Scheme (HSECS) Database Description*. 1995, BNL Technical Report No. L2415/95-1, New York: Brookhaven National Laboratory.

Dougherty, E. and J. Fragola (1987). Human Reliability Analysis: A Systems Engineering Approach with Nuclear Power Plant Applications, Canada: John Wiley & Sons Ltd.

Embrey, D.E., P. Humphreys, E.A. Rosa, B. Kirwan, and K. Rea (1984). *SLIM-MAUD: An Approach to Assessing Human Error Probabilities Using Expert Judgment*. NUREG/CR-3518, Washington, D.C.: Nuclear Regulatory Commission.

Flynn, C. (2005). *Four Primary Factors of Human Performance*. Aviation, Space, and Environmental Medicine, Official Journal of the Aerospace Medical Association, 76(6 Section II Supplemental).

Gertman, D., H.S. Blackman, J. Marble, J. Byers, L.N. Haney, and C. Smith (2005). *The SPAR-H Human Reliability Analysis Method*, NUREH/CR-6883, Washington DC: U.S. Nuclear Regulatory Commission.

Gertman, D.I. and H.S. Blackman (1993). *Human Reliability & Safety Analysis Data Handbook*. New York: John Wiley & Sons Ltd.

Gore, B.R., J.S.J. Dukelow, T.M. Mitts, and W.L. Nicholson (1995). *A Limited Assessment of the ASEP Human Reliability Analysis Procedure Using Simulator Examination Results*, NUREG/CR-6355, Washington DC: U.S. Nuclear Regulatory Commission, http://www.osti.gov/bridge/servlets/purl/119431-NBwPqV/webviewable/119431.pdf.

Grant, A., J. Holy, P. Pyy, R. Virolainen, G. Baumont, J.-M. Lanore, P. Le-Bot, K. Muramatsu, J. Mertens, O. Sträeter, A. Bareith, S. Czakó, E. Hollo, M. Fukuda, K. Muramatsu, G. Heslinga, M.F. Versteeg, J. Yllera, B. Liwång, S. Hirschberg, B. Reer, P. Meyer, D. Hamblen, J. Williams, M.A. Cunningham, J. Murphy, M.P. Rubin, and A. Ramey-Smith (2000). *Errors of Commission in Probabilistic Safety Assessment*, Nuclear Safety NEA/CSNI/R(2000)17, Le Seine St-Germain: the Organisation for Economic Co-operation and Development (OECD) Nuclear Energy Agency.

Greene, A.E. (1969). *Safety Assessment of Automatic and Manual Protection Systems in Reactors*, AHSB(S) R-172, England: UKEA Health and Safety Branch.

Gushin, V.I., A. Efimov, and T.M. Smirova (1996). *Work Capability During Isolation*, in *Advances in Space Biology and Medicine*, B. S.L., Editor, JAI Press Inc.: Greenwich, CT. p. Chapter 17.

Hart, S. G., Dahn, D., Atencio, A., Dalal, K.M. (2001). *Evaluation and Application of MIDAS V2.0*. Society of Automotive Engineers, Inc. 2001-01-2648.

Hollnagel, E. (1998). Cognitive Reliability and Error Analysis Method (CREAM). 1 ed., Elsevier.

Hollnagel, E. (1993). *Human Reliability Analysis: Context and Control*. Computers and People Series, ed. B.R. Gaines and A. Monk, London: Academic Press.

Julius, J.A., E.J. Jorgenson, G.W. Parry, and A. Mosleh (1995). A Procedure for the Analysis of Errors of Commission in a Probabilistic Safety Assessment of a Nuclear Power Plant at Full Power. Reliability Engineering & System Safety. 50(2): p. 189-201.

Kilduff, P. W., Swoboda, J. C., & Barnette, B. D. (2005). Command, control, and communications: techniques for the reliable assessment of concept execution (C3TRACE) modeling environment: The tool. (Report No. ARL-MR-0617). Army Research Lab, Human Research and Engineering Directorate, Aberdeen Proving Ground, MD.

Kirk, M., S. Malik, T. Santos, T. Dickson, C.E. Pugh, R. Bass, P. Williams, R. Woods, N. Siu, L. Kim, A. Kolaczkowski, D. Whitehead, D. Bessette, B. Arcieri, D. Fletcher, A. Mosleh, and Y.H. Chang (2006). *Technical Basis for Revision of the Pressurizer Thermal Shock (PTS) Screening Criteria in the PTS Rule (10CFR50.61)*, NUREG-1806, Washington DC: U.S. Nuclear Regulatory Commission.

Kirwan, B. (1994). A Guide to Practical Human Reliability Assessment. Taylor & Francis.
Kirwan, B., H. Gibson, R. Kennedy, J. Edmunds, G. Cooksley, and I. Umbers. *Nuclear Action Reliability Assessment (NARA): A Data-Based HRA Tool*. in *The 7th Probabilistic Safety Assessment and Management*. 2004. Berlin, Germany: June 14-18, 2004: Springer.

Kirwan, B., H. Gibson, R. Kennedy, J. Edmunds, G. Cooksley, and I. Umbers (2005). *Nuclear Action Reliability Assessment (NARA): A Data-Based HRA Tool.* Safety & Reliability Journal, Safety and Reliability Society, Manchester, 25(2).

Moieni, P., A.J. Spurgin, and A. Singh (1994). *Advances in Human Reliability Analysis Methodology. Part I: Frameworks, Models, and Data.* Reliability Engineering and System Safety, 44(1): p. 27-55.

Moieni, P., A.J. Spurgin, and A. Singh (1994). *Advances in Human Reliability Analysis Methodology. Part II: PC-based HRA Software.* Reliability Engineering and System Safety, 44(1): p. 57-66.

Mosleh, A., Y.H.J. Chang, F. Chandler, J. Marble, R. Boring, and D.I. Gertman (2006). *Evaluation of Current HRA Methods for NASA use (White Paper).* In *NASA HRA Workshop.* Kennedy Space Center, FL: January 16-18, 2006.

NASA (2005). *NASA Procedural Requirements 8705.2: Human-Rating Requirements for Space Systems.* Washington DC.

NASA (2004). *NASA Procedural Requirements 8715.3: NASA Safety Manual.* NASA, Washington, DC.

NASA (2002), *Probabilistic Risk Assessment Procedures Guide for NASA Managers and Practitioners.* Found at: http://www.hq.nasa.gov/office/codeq/doctree/praguide.pdf.

NASA (1995). International Space Station Flight Crew Integration Standard (NASA-STD-3000/T) Rev. C. SSP 50005.

NASA (1995). *NASA Standard 3000. Man System Integration Standard.* http://msis.jsc.nasa.gov

NASA (1969). Office of Space Flight System Safety Requirements for Manned Space Flight, Safety Program Directive No. 1.

NASA/JSC (2004). *Bioastronautics Roadmap*, NASA/SP-2004-6113: NASA, http://bioastroroadmap.nasa.gov/index.jsp.

National Research Council (1988). *Post-Challenger Evaluation of Space Shuttle Risk Assessment and Management.* (Slay Committee Report). National Academy Press, Washington, DC.

Nicogossian, A. (2003). *Medicine and Space Exploration.* The Lancet, 362(Supplement 1): p. s8-s9.

Parry, G.W., B.O.Y. Lydell, A.J. Spurgin, P. Moieni, and A.N. Beare (1992). *An Approach to Analysis of Operator Actions in Probabilistic Risk Assessment*, TR-100259, Palo Alto, California: Electric Power Research institute.

Perry, R.U. and committee (1993). *Report of the Shuttle Process Review Team.* NASA. Kennedy Space Center.

Reer, B., O. Sträeter, V.N. Dang, and S. Hirschberg (1999), *A Comparative Evaluation of Emerging Methods for Errors of Commission Based on Applications to the Davis-Besse (1985) Event.* PSI Bericht Nr. 99-11 ISSN: 1019-0643, Villigen PSI, Switzerland: Paul Scherrer Institut.

Reer, B., V.N. Dang, and S. Hirschberg (2004). *The CESA Method and its Application in a Plant-Specific Pilot Study on Errors of Commission.* Reliability Engineering & System Safety, 83(2): p. 187-205.

Shen, S.H. and A. Mosleh (1996). *Human Error Probability Methodology Report* RAN: 96-002: Calvert Cliffs Nuclear Power Plant: BGE.

Shen, S.H., C. Smidts, and A. Mosleh (1997). *A Methodology for Collection and Analysis of Human Data Based on a Cognitive Model: IDA.* Nuclear Engineering and Design, 172(1-2): p. 157-186.

Smidts, C., S.H. Shen, and A. Mosleh (1997). The IDA Cognitive Model for the Analysis of Nuclear Power Plant Operator Response Under Accident Condition. Part I: Problem Solving and Decision Making Model. Reliability Engineering and System Safety, (55): p. 51-71.

Spurgin, A.J., P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, and G.W. Hannaman (1990). *Operator Reliability Experiments Using Power Plant Simulators (Vol. 1-3 Technical Report)*, EPRINP-6937, Palo Alto, California: Electric Power Research institute.

Sträeter, O. (2005). Cognition and safety - An Integrated Approach to Systems Design and Performance Assessment: Ashgate. Aldershot.

Sträeter, O. (2000). Evaluation of Human Reliability on the Basis of Operational Experience, in Economics and Social Sciences. The Munich Technical University.

Swain, A.D. and H.E. Guttmann (1983). Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications. NUREG/CR-1278: Nuclear Regulatory Commission.

Swain, A.D. (1987). *Accident Sequence Evaluation Program Human Reliability Analysis Procedure.* NUREG/CR-4772, Washington DC: U.S. Nuclear Regulatory Commission.

Swain, A.D. (1990). *Human Reliability Analysis: Need, Status, Trends, and Limitations.* Reliability Engineering and System Safety, 29: p. 301-313.

The NEA Committee on the Safety of Nuclear Installations (1998). Critical Operator Actions: Human Reliability Modeling and Data Issues. Principle Working Group No. 5 - Task 94-1. NEA/CSNI/R (98)1: Nuclear Energy Agency.

Trucco, P., C. Leva, and O. Sträter (2006). Human Error Prediction in ATM via Cognitive Simulation: Preliminary Study. in PSAM8. New Orleans.

U.S. Nuclear Regulatory Commission (2005). HRA Method Review - ASEP. in Expert Workshop on the Evaluation of HRA Methods with Respect to HRA Good Practices (NUREG-1792). Rockville, Maryland: June 27-28, 2005.

Versteeg, M.F. (1998). Responses to PWG5 Task 94-1 Questionnaire on Critical Operator Actions, Netherlands, Part I, PWR (PSA Borssele NPP). NEA/CSNI/R(98)1/Add1, Appendix F: OECD, Nuclear Energy Agency.

Wakefield, D.J., G.W. Parry, and A.J. Spurgin (1992). *Systematic Human Action Reliability Procedure (SHARP) enhancement project, SHARP1 Methodology Report*, EPRI TR-101711, Palo Alto, CA: Electric Power Research Institute.

Wiegmann, D.A, Shappell, S.A. (2001). *Human Error Analysis of Commercial Aviation Accidents: Application of the Human Factors Analysis and Classification System (HFACS).*

Williams, J. (1988). A Data-Based Method for Assessing and Reducing Human Error to Improve Operational Performance. in IEEE Conference on Human Factors in Power Plants. Monterey California: June 5-9, 1988.

Williams, J. (1986). HEART: A Proposed Method for Assessing and Reducing Human Error. in The 9th Advances in Reliability Technology Symposium. University of Bradford.

Williams, J. (1989). Human Reliability Data - The State of the Art and the Probabilities. in Proceedings Reliability '89. United Kingdom: June 14-16.

U.S. House of Representatives Committee on Science and Technology. (1986). *Investigation of the Challenger Accident.* Washington, DC.

This page intentionally left blank.

http://www.hq.nasa.gov/office/codeq/rm/reference.htm