Evaluation of Human Error Probabilities for Post-Initiating Events

by

Phillip E. Dawson

Submitted to the Department of Nuclear Science and Engineering
in Partial Fulfillment of the Requirements for the Degree of

Master of Science in Nuclear Science and Engineering

at the

MASSACHUSETTS INSTITUTE OF TECHNOLOGY
May 2007
[June 2007]

Signature of Author ______________________________
Phillip E. Dawson
Department of Nuclear Science and Engineering
9 May 2007

Certified by ______________________________
George E. Apostolakis
Professor of Nuclear Science and Engineering
Professor of Engineering Systems
Thesis Supervisor

Accepted by ______________________________
Michael W. Golay
Professor of Nuclear Science and Engineering
Thesis Reader

Accepted by ______________________________
Jeffrey A. Coderre
Associate Professor of Nuclear Science and Engineering
Chairman, Departmental Committee on Graduate Students

(This page intentionally left blank)

Evaluation of Human Error Probabilities for Post-Initiating Events

by

Phillip E. Dawson

Submitted to the Department of Nuclear Science and Engineering
on May 9, 2007, in partial fulfillment of the requirements for the
Degree of Master of Science in Nuclear Science and Engineering

ABSTRACT

The United States Nuclear Regulatory Commission is responsible for the safe operation of the United States nuclear power plant fleet, and human reliability analysis forms an important portion of the probabilistic risk assessment that demonstrates the safety of sites. Treatment of post-initiating event human error probabilities by three human reliability analysis methods are compared to determine the strengths and weaknesses of the methodologies and to identify how they may be best used. A Technique for Human Event Analysis (ATHEANA) has a unique approach because it searches and screens for deviation scenarios in addition to the nominal failure cases that most methodologies concentrate on. The quantification method of ATHEANA also differs from most methods because the quantification is dependent on expert elicitation to produce data instead of relying on a database or set of nominal values. The Standardized Plant Analysis Risk Human Reliability Analysis (SPAR-H) method uses eight performance shaping factors to modify nominal values in order to represent the quantification of the specifics of a situation. The Electric Power Research Institute Human Reliability Analysis Calculator is a software package that uses a combination of five methods to calculate human error probabilities. Each model is explained before comparing aspects such as the scope, treatment of time available, performance shaping factors, recovery and documentation. Recommendations for future work include creating a database of values based on the nuclear data and emphasizing the documentation of human reliability analysis methods in the future to improve traceability of the process.

Thesis Supervisor: George E. Apostolakis
Title: Professor of Nuclear Engineering

(This page intentionally left blank)

Acknowledgements

I would like to thank all of the people at MIT who helped me with my coursework and thesis. First and foremost, I would like to thank Professor Apostolakis for his patience and guidance through this process. The insights and advice he gave me will undoubtedly serve me throughout my career. I wish to thank my thesis reader Professor Golay, for his time and comments. I also want to recognize my fellow colleagues who willingly helped me at any hour: David Carpenter, Tyler Ellis, Erik Johnson, and Michael Stawicki.

Away from MIT, I would like to thank Jeffrey Julius at Scientech, Ken Canavan at EPRI, John Forester at Sandia National Laboratories, and Susan Cooper at the USNRC for answering my questions at the PRA subcommittee meeting of the ACRS in March 2007 and in subsequent emails or phone calls. I would also like to thank Mary Presley, an alumna of MIT, for her advice and gracious assistance. Finally, I would like to thank Annalise Gill and my family for their love, patience and support that have shaped me into the person I am today.

(This page intentionally left blank)

Table of Contents

(This page intentionally left blank)

Table of Figures

(This page intentionally left blank)

Table of Tables

(This page intentionally left blank)

Acronyms

ASEP	Accident Sequence Evaluation Program
ATHEANA	A Technique for Human Event Analysis
BWR	Boiling Water Reactor
CBDT	Cause-Based Decision Tree
CREAM	Cognitive Reliability and Error Analysis Method
EFC	Error Forcing Context
EOC	Error of Commission
EOO	Error of Omission
EOP	Emergency Operating Procedure
EPRI	Electric Power Research Institute
HCR	Human Cognitive Reliability
HEP	Human Error Probability
HFE	Human Failure Event
HI	Human Interaction
HRA	Human Reliability Analysis
IE	Initiating Event
MMI	Man-Machine Interface
NPP	Nuclear Power Plant
NRC	Nuclear Regulatory Commission
ORE	Operator Reliability Experiments
PRA	Probabilistic Risk Assessment
PSF	Performance Shaping Factors
PWR	Pressurized Water Reactor
SHARP	Systematic Human Action Reliability Procedure (EPRI HRA framework)
SHARP1	Revision of SHARP
SPAR-H	Standardized Plant Analysis Risk Human Reliability Analysis
THERP	Technique for Human Error Rate Prediction
TRC	Time Reliability Correlation
UA	Unsafe Action

(This page intentionally left blank)

1 Introduction

Human reliability plays an important role in the safety and reliability of the operation of complex technologies. Space exploration, large processing facilities, and nuclear power are all susceptible to mistakes committed by the human operators, and these errors need to be identified and analyzed in order to avoid loss of life, injury, and the engineering system itself. Mistakes can be costly in terms of both human life and monetarily. This paper is only concerned with the safety of nuclear reactors and specifically how the operators affect the probability of a failure event.

In order to better ensure the safety of nuclear reactors, the US Nuclear Regulatory Commission (NRC) requires probabilistic risk assessments (PRAs) for each reactor to determine that the nuclear power plant (NPP) is safe to operate. As part of the PRA, a human reliability analysis (HRA) is conducted to determine how the operators affect the safety of the plant. These analyses attempt to recognized and quantify how human error can lead to a failure of the NPP. Three models for the quantification of post-initiating events are investigated and compared to determine the relative strengths of the models and suggestions are made for future work in the HRA of NPP.

The HRA methods quantify human error probabilities (HEPs), and this is challenging for many reasons including the fact that human actions are unpredictable and influenced by many factors.[1] First, the scope of the HRA is identified in terms of the larger context of the needs of the PRA that the HRA results will be incorporated into. This involves an evaluation of the available resources and the type of human failure events that need to be addressed. From this information, appropriate models can be selected to perform the HRA. Next, the HEPs need to be identified through a rigorous search process. This involves the construction of logic structures and a screening process to identify the important human failure events. Many methods include iterative processes to help analysts with the screening process. Last, the HEPs are quantified to give numerical results to be included in the PRA. The quantification is the focus of this paper and the three models compared are A Technique for Human Event Analysis (ATHEANA),[2] Standardized Plant Analysis Risk Human Reliability Analysis (SPAR-H),[3] and the Electric Power Research Institute Human Reliability Calculator (EPRI HRA Calculator). Figure 1 shows an overview of the three main steps involved in an HRA:

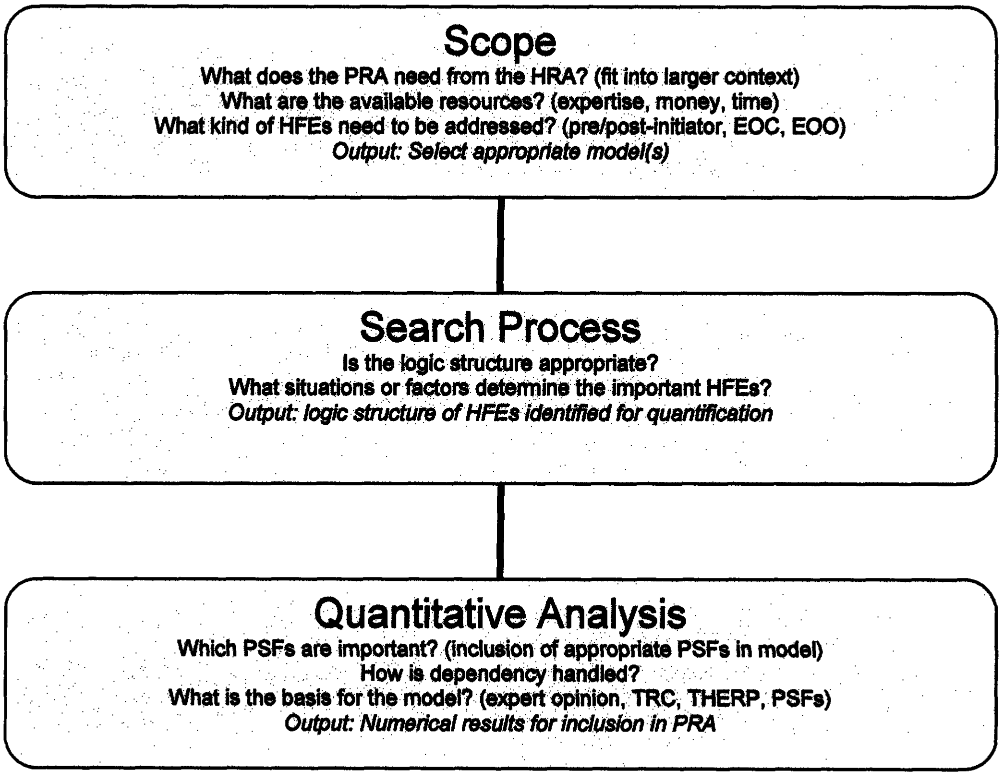


Figure 1: High Level Human Reliability Analysis Block Diagram

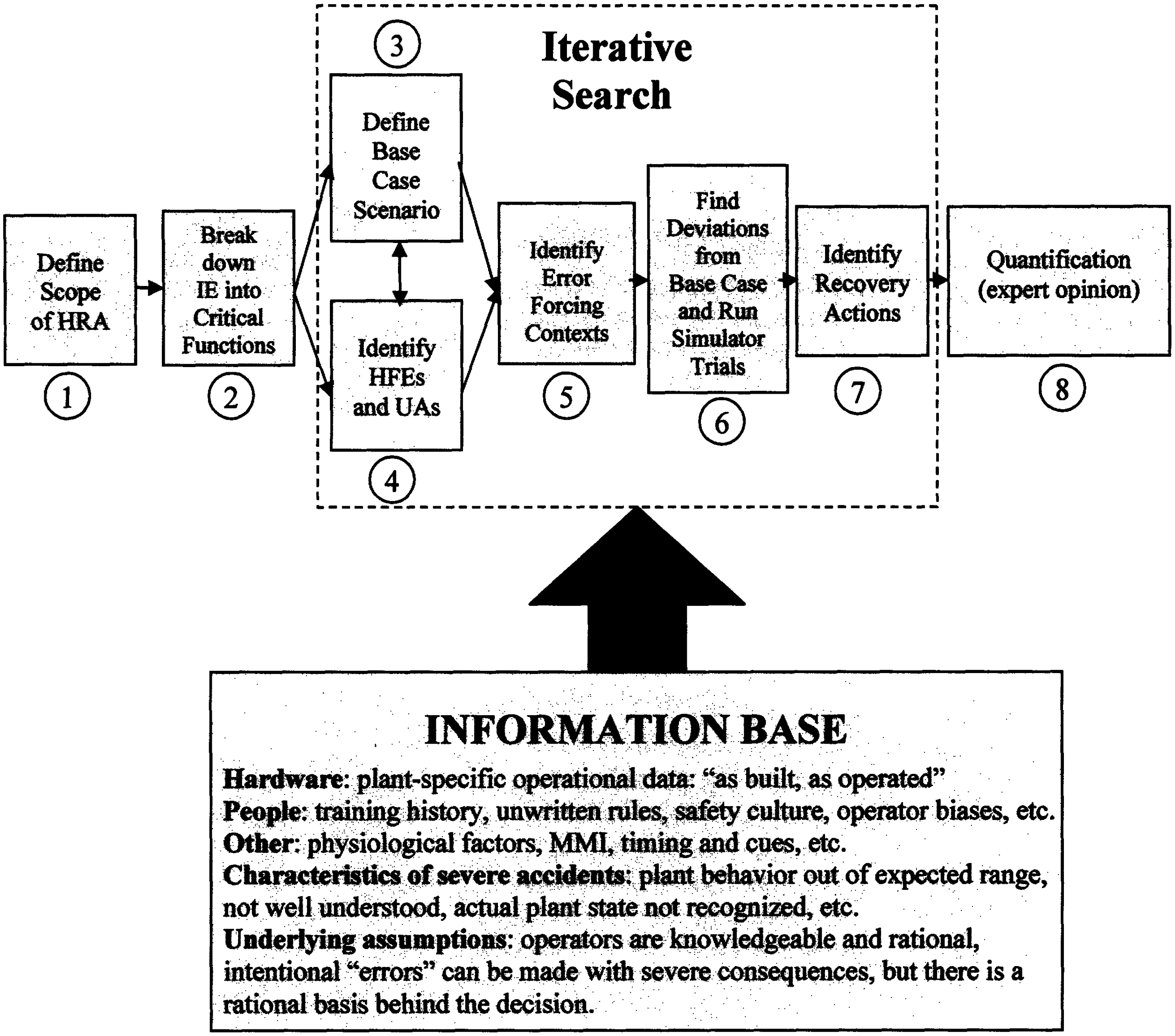
The next three sections provide overviews of how the three quantification methods work for ATHEANA, the EPRI HRA Calculator, and SPAR-H. Section five compares various aspects of the models including scope and the treatment of dependency, performance shaping factors, and time among others. Finally, the conclusions of the comparison are presented in section six along with suggestions for future work.

2 ATHEANA – A Brief Overview and Important Terminology

ATHEANA was developed to improve the capabilities of HRA, and in particular, the method was designed to realistically represent and quantify behavior observed in accidents and near-miss events at NPPs. ATHEANA targets specific sets of conditions that make up the context of a situation that can "trigger error mechanisms in plant personnel." [4] Both intentional and unintentional errors of commission (EOCs) and errors of omission (EOOs) are quantified by ATHEANA. For each HFE, this method attempts to identify important contexts, called error forcing contexts (EFCs), that may lead operators carrying out an inappropriate action. The quantification process uses expert opinion to define the probability of failure within these contexts. While quantification is the primary focus of this paper, the search process of ATHEANA is presented here as the context for the consensus-based expert opinion quantification method.

ATHEANA is unique in its ability to conform to specific scenarios instead of providing limited options that can be adjusted to fit a particular context. The method provides more adaptability to situations than more rigid models that rely on predetermined PSFs to differentiate between varying conditions. Coupled with the search process, ATHEANA can provide results tailored to the specific characteristics that are likely to drive human performance. NUREG-1624 provides guidance for post-initiating events and analysis of post-initiator HFEs has been the focus for the model, but there is no reason that the same process could be applied to pre-initiator HFEs. This flexibility is an advantage of the method over many current HRA techniques.

Figure 2 provides a graphical summary of the ATHEANA process; these eight steps will be reviewed.



Hardware: plant-specific operational data: "as built, as operated"
People: training history, unwritten rules, safety culture, operator biases, etc.
Other: physiological factors, MMI, timing and cues, etc.
Characteristics of severe accidents: plant behavior out of expected range, not well understood, actual plant state not recognized, etc.
Underlying assumptions: operators are knowledgeable and rational, intentional "errors" can be made with severe consequences, but there is a rational basis behind the decision.

Figure 2: Summary Flow Chart of the ATHEANA Process[5]

2.1 Steps 1-4: Identifying Human Failure Events[5]

The first portion of the ATHEANA framework is designed to identify the significant modes of human failure to include in the plant's PRA. The first step requires the HRA team (a multi-disciplinary team comprised of experts such as HRA analysts, PRA analysts, operators, trainers, and thermo-hydraulics specialists) to carefully define the scope of the HRA required to do an "adequate" job for the overall PRA. The HRA team must decide where the boundaries are for the analysis. Quantitative screening can be helpful in determining the necessary scope, and NUREG-1624 suggests a method for identifying candidate HFEs.

Once the important initiators are selected, and their corresponding event trees identified, the HRA team proceeds to prioritize the plant functions, systems, and equipment required to respond to the accident initiators. With these critical items identified, candidate HFEs can be more readily identified by examining the initiating event (IE) event tree and then systematically evaluating the event tree branch points for possible human-caused functional failures.

With the scope and issue clearly defined, the HRA team moves on to describe the base-case scenario, or the expected evolution of the accident scenario. Part of defining the base case includes describing and understanding the human performance context of the scenario. The following is a list of suggested components of the base case description, taken from the ATHEANA user's guide: [6]

- A list of possible causes of the initiating event(s)
- A brief, general description of the expected sequence of events (as in PRA event trees), starting before reactor trip
- A description of the assumed initial conditions of the plant
- A familiarization/description of the expected plant conditions for the accident sequence
- A specification of the *expected sequence timing* of plant status changes

- A description of the *expected trajectories, over time, of key parameters* indicating plant status and a specification of the status of indications and other *cues* that are expected as the sequence evolves
- Any assumptions of expected plant behavior, system/equipment/indicator responses, and operator response
- A discussion of the procedures expected to be used for the given situation
- A description of *key operator actions, and their timing*, expected during the scenario progression

Concurrently with the base case definition, the HRA team identifies and defines the possible human failure events and the corresponding unsafe actions (UAs, where one or more unsafe action makes up a HFE). In order to identify the HFEs, a systematic process, building on the IE event trees from step 1, is followed by determining: [6]

1. whether the function is necessary or undesired
2. the system(s) or equipment that perform the function
3. the pre-initiator status of the system(s) or equipment
4. the functional success criteria for the system(s) or equipment
5. the functional failure modes of the system(s) or equipment
6. how the operator interacts with the equipment and deciding if EOCs, EOOs, or both are relevant.

Both the ATHEANA user's guide[6] and NUREG-1642 provide tables of example UAs for generalized equipment functional failure modes to help the HRA team through this step. At this point the HRA team has developed a list of HFEs and associated UAs important to the scope of the PRA.

2.2 Steps 5-7: The Error Forcing Context

These next steps focus on defining EFCs that are most likely to lead to accident scenarios. ATHEANA attempts to define EFCs based on the following characteristics: many severe accidents share common attributes, plant behavior may be out of the expected range, plant behavior may not be well understood by operators, plant procedures may not helpful/appropriate, and the actual plant state may not being recognized by operators.[5] These error forcing contexts are comprised of a combination of plant state (hardware) and other performance shaping factors (PSFs). The goal of identifying these EFCs is to find regimes where operators believe an inappropriate action is the correct action, because these are the situations that cause EOCs.

Vulnerabilities based on the knowledge of an operator need to be identified because these may result in HFEs. In order to identify the PSFs, plant-specific background data must be considered, such as: formal procedures, crew characteristics/dynamics, ergonomics, informal rules and biases. An example of an informal rule is "beat the automatic system when practical," as opposed to "wait for the automatic system before taking action."[2] The former rule might carry with it a greater chance for an EOC to occur. Plant biases include frequency and recency biases that arise from plant operation and simulator training (operators may have an expectation of how a scenario will unfold).[6] The goal of this search is to find scenarios that might prove "troublesome" to operators and produce an error-forcing context. Simulator exercises may prove helpful in this step in that they allow the HRA analyst to observe how the operators behave and think. They can also serve to test theories of operator response.

With the base case defined, and having an idea of what the important PSFs might be, the HRA team searches for and defines potential deviations from the base case. These are credible scenarios that include the identified EFC, and nuclear records show that no serious accidents have developed from base case scenarios. Section 2.4 will provide an example of a deviation scenario and its accompanying EFC.

The analysts must take into account other complicating factors, as well as recovery factors, as part of the context of quantification. Complicating factors can be PSFs or physical conditions not yet considered for a particular EFC such as additional hardware failures, configurations problems, unavailabilities, or factors typically not considered in a PRA. Specifically, there are two groups of PSFs that can contribute to the EFC: those triggered by the defined EFC and additional PSFs not specific to the context. New plant conditions or PSFs need to be included in the scenario definition, and they can also activate different or more error mechanisms, creating an iterative process until the EFC is described completely by the PSFs and physical conditions identified.

To prevent an unrealistically conservative estimate, recovery factors are included in the analysis. In search for potential recovery actions and evaluation of their feasibility, there are five steps outlined in the ATHEANA user's guide: [6]

1. Define possible recovery actions given a HFE/UA has occurred
2. Consider *time available* for diagnosis and execution of potential recovery actions
3. Identify recovery cues
    a. Timing of recovery cues
    b. How compelling these cues are
        i. Do they strongly alert operators to a need for recovery?

        ii. Is there sufficient information to identify the most applicable recovery action?

4. Identify additional resources for aid in recovery (i.e., more staff) and associated timing
5. Assess the strength of recovery *cues and timing* with respect to EFCs – is there a "high" or "low" likelihood of successful recovery? Deviation scenarios with high likelihood of recovery need no further analysis or quantification. In this step, there are some suggested factors to consider when assessing feasibility of recovery: [6]
    a. Dependencies between the initial error and recovery actions that would make recovery unlikely
    b. Initial mindset (or diagnosis) of the situation may be hard to break
    c. Distractions or attention to other activities could cause new cues to be overlooked
    d. Operators may delay recovery action because there is a negative consequence to taking the action; this is especially relevant when plant hardware providing an alternative recovery is "almost" repaired.

NUREG-1624[2] gives further guidance on assessing recovery actions by type of failure: "thinking" (mistakes and circumventions) or "doing" (slips and lapses). For thinking failures, this involves assessing how the operator could persist in believing their UA is the "correct" action, using the same process and information from step 5. "Doing" failures, as suspected, are more straightforward. First the team must decide whether the slip/lapse is recoverable at all: was plant hardware irreparably damaged? Was it so damaged that the time for recovery is greater than the time available? If recoverable, then the team must determine whether the slip/lapse can induce a mistake. If so, then it should be further analyzed, if not, then it can be dropped from further analysis. At the end of this process, the final EFC for the HFE and UAs describes all of the foreseen scenarios.

2.3 Step 8: Quantification[5]

ATHEANA calculates the conditional likelihood of an UA given the occurrence of an EFC. This is a departure from the typical HRA methods that quantify the probability of human error under plant conditions specified in the event and fault trees. The quantification has three stages. The first assigns a probability to the EFC. Next, the conditional likelihood f the UAs capable of causing a HFE are determined, and lastly, the conditional likelihood of no recovery from each UA is calculated.

The search process for EFCs ends when the team feels assured that the EFC is sufficiently well defined and that both the frequency of the context and the conditional probability of the UA in that context can be estimated with an appropriate degree of confidence. To test the adequacy of EFCs, the HRA team can do simulator tests, compare EFCs with past operational experience and with human performance checklists.[7]

The rigorous way to quantify the probability of a human failure event (P(HFE)) is:

$$P(HFE|S) = \sum_{all\_i,j} P(EFC_i, S) * P(UA_j|EFC_i, S) * P(\overline{R}|EFC_i, UA_j, S) \qquad [1]$$

where S refers to the PRA accident scenario, and $P(\overline{R}|EFC_i, UA_j, S)$ is the probability of non-recovery given an unsafe action has occurred in an error forcing context for that scenario. Note that non-recovery is only modeled given an unsafe action because otherwise there is not an operator action or inaction to recover from. The probabilities are then summed over all UA/EFC combinations.

This rigorous method, however, requires too much resolution to be a feasible method of quantification, and so the HRA team may choose the following method: [6]

$$P(HFE|S) = \sum_{all\_i,j} P(EFC_i|S) * P(UA_j|EFC_i, S) \qquad [2]$$

In this case, recovery is factored implicitly into $P(UA_j|EFC_i, S)$, as explained below. For this formulation, the HRA team must use an expert opinion elicitation approach to quantification. [8]

An error forcing context is comprised of two parts: plant state (hardware) and performance shaping factors. For quantification purposes, $P(EFC_i|S)$ is taken directly from the PRA, and represents the plant state portion of the EFC. $P(PSF_i|S)$ is not included explicitly in the formulation because these are based on the scenario context, so these PSFs are implicitly taken into account in quantifying $P(UA_j|EFC_i, S)$ through expert opinion elicitation. [8]

The term "error *forcing* context" can be misleading because it implies that the conditional probability $P(UA_j|EFC_i,S)$ should be near unity. This, however, is not the case. EFCs are contexts which increase the likelihood of error, and in some cases "trigger" error, but do not generally *force* an error.

The probability of an UA for a specific EFC is taken from the consensus expert opinion elicitation process described below. Due to the way the expert opinion elicitation process is structured, aleatory uncertainty, recovery, and dependencies are all holistically incorporated into $P(UA_j|EFC_i,S)$. This judgment-based quantification consists of six steps:[8]

1. Discuss HFE and influences, identify specific EFC and "aleatory" PSFs
2. "Calibrate" experts
3. Elicit an estimated curve
4. Each expert presents his estimated curve to the group of experts
5. Open discussion amongst experts
6. Arrive at consensus curve

1. So that the experts understand exactly what they are quantifying, they should discuss and fully understand the scenario, including: the definition of the HFE in question, the plant state (part of the EFC), and the relevant PSFs (the other part of the EFC). To prevent the experts from being overwhelmed, only the most relevant PSFs to the scenario at hand should be taken into account. The importance of scenario specificity cannot be over emphasized. Most methods that incorporate PSFs do so by generically applying an adjustment factor (i.e., increase the failure probability by a factor of 2 if there

is time pressure). ATHEANA recognizes that in some situations these factors may not have a large impact; for example, time pressure may not impact operators in a very familiar situation on which they get trained frequently.

This is how ATHEANA incorporates recovery actions and the PSF portion of the EFC. Aspects such as scenario timing and relevant cues are woven into the description of the HFE and become the context within which the experts judge the probability of a UA.

2. The next step is to calibrate the experts so they have a more intuitive understanding of what a probability really is (i.e., 1 failure in 10 trials is 0.1). Here, they are encouraged to think about failures as a number of x failures in n trials instead of directly estimating a probability.

3. Now, each expert should come up with a 7-point estimation of $P(UA_j|EFC_i,S)$, including the 1st, 10th, 25th, 50th (median), 75th, 90th, and 99th percentiles. The experts should start by setting the 1st and 99th percentiles to be the probabilities for the best and worst case scenarios, respectively. This is where the "aleatory" PSFs come in – the best scenario is when there are no adverse PSFs, and the worst is when all the PSFs are in play. In exercising his/her judgment, an expert would then think about an effective crew and imagine them in the best circumstances for the 1st percentile and imagine a particularly ineffective crew with communication difficulties and imagine them in the worst circumstance for the 99th percentile estimate.

There have been extensive studies demonstrating that expert opinion is replete with potential biases, and that experts have difficulty consistently assessing the extremes of a range. Reference 8 describes some of these biases, and suggests methods to alleviate these effects.

4-6. After the experts have their curves, they present them to the group and the group deliberates until a consensus is reached. Part of the reason behind a consensus-based approach is to avoid unintentional bias. The epistemic uncertainty for this method would be a family of curves, one for each expert's opinion – see Table 1.

Table 1: Example Expert Opinion Elicitation Results for Failure to Isolate a Stuck-Open Atmospheric Dump Valve within 30 Minutes of the Initiating Event[9]

Expert	**Percentiles**						
	1st	**10th**	**25th**	**50th**	**75th**	**90th**	**99th**
#1	0.01	0.03	0.05	0.08	0.4	0.8	1.0
#2	0.001	0.003	0.008	0.02	0.07	0.1	0.8
#3	0.001	0.01	0.03	0.06	0.4	0.6	0.9
#4	0.005	0.01	0.02	0.033	0.1	0.6	0.8
Consensus	**0.004**	**0.01**	**0.03**	**0.05**	**0.2**	**0.5**	**0.9**

2.4 Example of the Search Process

The search process is presented as a series of steps that the HRA team took for this limited trial of ATHEANA.[2] Each step was accompanied by a set of guiding questions to aid the team in identifying important scenarios. For brevity, these questions are not presented here – most of them can be found in the Sections above, and the rest can be found in NUREG-1624. The HRA initially chose three initiating events to quantify analyze: MLOCA, LOSP and ATWS. Only the MLOCA IE will be examined here. Figure 3 is the simplified event tree for the MLOCA, and Table 2 is a list of top events and their descriptions.

1. First the HRA team selected the initiating event of interest (MLOCA) and prioritized the functional requirements as represented by the nodes of the event tree. These functions/priorities were:

- Makeup: Medium Priority
- Heat Removal: Medium Priority
- Long-Term Heat Removal: High Priority

2. The team then examined the safety functions required, defined their success criteria, and identified their failure modes. For each failure mode, the team asked "how can the operators produce the effects characterized by the failure modes identified?" From this process, they found two unsafe actions of particular significance, as described in NUREG-1624:

- Operators stop pump (function: makeup; system: high-pressure injection)
- Operators operate pump outside design parameters (function: long-term cooling; system: residual heat removal (RHR) system)

3. Addressing these unsafe acts, the team searched for and defined important error forcing contexts. Here the team identified a credible accident sequence for each EFC to simulate in order to test the strength of the EFC. Both sequences involve a MLOCA where system repressurization is not possible, and continuous high-head injection is required to keep the core cool and covered. Furthermore, in these sequences it is unclear whether the steam generators act as a heat source or provide a heat sink because the break is the primary method of heat removal, and the primary system pressure is less than the pressure of the secondary system.[2,5]

HFE #1 – Inappropriate Termination of Makeup

The error forcing context of this scenario is a deceptive failure of the RCS pressure indicators. In this simulation, one RCS pressure indicator was under repair and the second failed stuck during operation at 550 psig. This was intended to make the operators believe that the indicator is functioning normally, when in fact it indicated greater sub-cooling than reality. This misinformation would prompt the operators to shut off the pumps early (as directed by the procedure). In this case, core damage would ensue if recovery of injection was not restored in a timely manner.

HFE #2 – Inappropriate Depletion of Resources

In this scenario, there is increased RWST depletion due to containment spray system activation during the LOCA. If the RWST "empty" alarm sounds, high-head pumps should be stopped until reconfiguration is complete, or pump cavitation will occur leading to core damage unless the pressure can be reduced and low-pressure injection is initiated. To "trigger" the operator error of 'failing to stop the high-head pump,' the RWST alarm was made inoperable due to IRTU maintenance.

The conclusion of this limited test search was the identification of at least one strong EFC. The failed RCS pressure indicators indeed caused the simulation crew to prematurely stop the pumps even though the potential of a failed indicator was recognized (but not verbalized) by one operator. However, the inoperable RWST "empty" alarm did not prove to be a significant EFC at all – the crew paid sufficient attention to the RWST level throughout the simulation, and the Work Control Supervisor recognized that the IRTU maintenance could fail the alarm. It was also brought to the attention of the trainers that in unfamiliar or tricky situations, operators might not adhere to the strong tendencies developed through training like the "think it, say it" rule. Neither scenario was considered significant enough to retain in the demonstration plant's HRA.

Table 2: MLOCA Event Tree Top-Event Summary[9]

Top Event ID	Title	Description
RW	RWST	RWST failure - no inventory for RCS makeup.
ALT	Alternate Cooling	Alternate cooling to the charging pumps.
CSA	Charging Pump	Centrifugal Charging Pump Train A failure.
CSB	Charging Pump	Centrifugal Charging Pump Train B failure.
SIA	SI Pump	Safety Injection Pump Train A failure.
SIB	SI Pump	Safety Injection Pump Train B failure.
EF	EFW	EFW failure - motor and turbine-driven pumps.
OD	Operator - Depressurizes RCS	Operator failure to depressurize RCS for RHR injection, given failure of HPI.
RA	RWST Valve	RWST isolation valve Train A failure to remain open - RHR and CBS Train A suction path.
RB	RWST Valve	RWST isolation valve Train B failure to remain open - RHR and CBS Train B suction path.
L1	RHR Miniflow	RHR Train A failure in miniflow recirculation.
L2	RHR Miniflow	RHR Train B failure in miniflow recirculation.

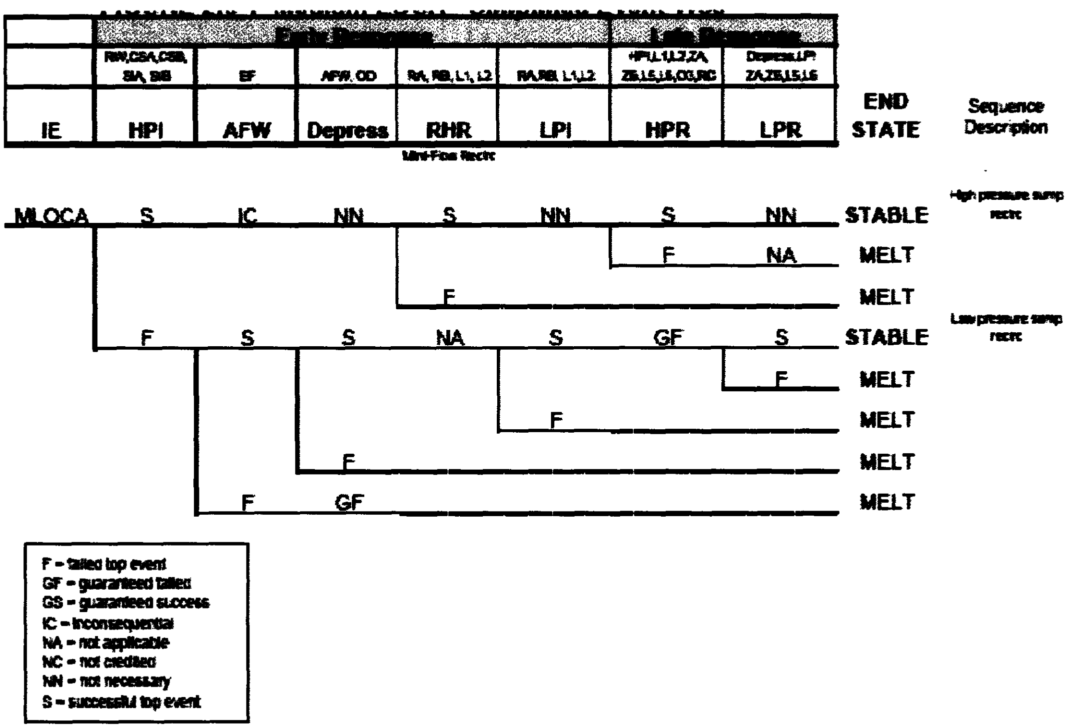


Figure 3: Simplified MLOCA Event Tree[9]

3 EPRI HRA Calculator – A Brief Overview and Important Terminology

EPRI produced the HRA Calculator to provide a tool that would produce documented and reproducible results that is less resource intensive than competing HRA methodologies. The HRA Calculator is a piece of software instead of being a methodology for completing the search process and/or quantification process for the HRA portion of a PRA. The EPRI HRA Calculator combines the SHARP1 framework with five quantification methods (HCR/ORE, CBDT, THERP execution analysis, THERP annunciator response and SPAR-H) to create an HRA tool that is easy to use, consistent, transparent, and non-resource intensive.[17] Unlike ATHEANA, the EPRI

Calculator does not attempt to break down human failure events into specific unsafe actions, or even into specific contexts. Rather, it takes a more generic approach, using time reliability correlations (TRCs) and generic decision trees. This method is specifically designed to be usable by a PRA expert with some HRA training and instruction on use of the EPRI Calculator in the case that an HRA expert is not available. Figure 4, below, provides a graphical summary of the EPRI HRA Calculator; these six steps will be reviewed briefly here.[5]

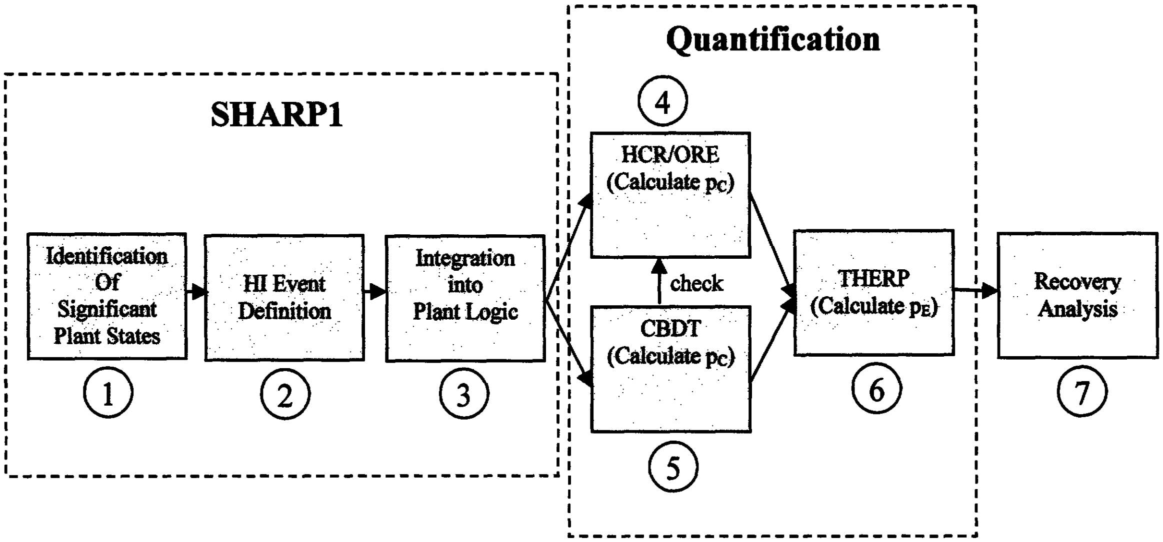


Figure 4: EPRI HRA Calculator Summary Flow Chart[5]

3.1 Basic Terminology: HI Types, Cue-Response Structures, and Timing

To aid in the definition of human interactions (HIs), SHARP1 defines three broad types of human interactions:[10]

Type A – Pre-Initiating event related HI

These HIs are associated with normal operation. Activities associated with maintenance, tests, calibrations, and evolutions of the plant are all Type A HIs. Before the initiating event, an HI can affect the availability of systems.

Type B – Initiating event related HI

In this case, a worker initiates an event by misaligning a system or through malfunction of equipment that trips or inserts false control signals. These HIs are not represented in the HRA because they are implicitly included in the PRA as part of the frequency of generic initiating events. The frequencies can be generated using plant operating histories.

Type C – Post-initiating event HI

These interactions are performed by plant staff after an initiating event, and the two kinds of Type C HIs are CP and CR, which are procedural actuation of systems and recovery actions, respectively. The former includes actions directed by procedures that will terminate the accident, and the latter includes recovery actions.

Only the post-initiating event HI quantification will be compared in this work. Type CR HIs are highly scenario specific and are not incorporated in the plant logic model. Instead, CR events are modeled as correction factors for individual scenarios. The following steps, then, are primarily concerned with type CP human interactions.

In 1989, EPRI carried out a set of simulator experiments, the Operator Reliability Experiments (ORE), in order to gather data to aide in HFE quantification.[11] This program was meant to validate and support the HCR TRCs which rested upon the cognitive categorizations of skill, rule, and knowledge based behavior.[12] The experiments, however, did not support this grouping, and cue-response structures became the focus for quantification. The procedure driven HIs can be modeled using cue-response structures based on five distinct scenarios. These cue-response structures for Type CP actions, taken from EPRI TR-100259 are presented here:[13]

CP1: Response following a change in the plant damage state that is indicated by an alarm or value monitored parameter (e.g., response to a spurious pressurizer spray operation in a PWR).

CP2: Response following an event that gives rise to a primary cue (as in CP1) that has to be achieved <u>when</u> a parameter is exceeded or can be seen not to be maintainable below a certain value (e.g., initiate RHR when the suppression pool (SP) temperature exceeds 95^{oF} in a BWR). These HIs involved a waiting period after the primary cue in order to reach a determined plant state.

CP3: Response following an event that gives rise to a primary cue (as in CP1) that has to achieved <u>before</u> some plant parameter reaches a critical value (e.g., initiating SLCS before SP temperature reaches 110^{oF} in a BWR). This critical value can be regarded as a soft prompt, or secondary cue.

CP4: Performing a step in a procedure that is being followed as a result of a plant disturbance (e.g., inhibiting ADS before lowering level in a BWR, in response to an ATWS). The cue in this case is generally associated with completing the previous step.

CP5: Maintaining a variable parameter below, at, or within specific limits (e.g., controlling the level in a steam generator to prevent overfill or dryout). This is a control action.

Only CP1 – CP3 HIs can be quantified using the HCR/ORE process. The timing information for those cue-response structures are described by the following: $T_{1/2}$ is the median crew response time to initiate the appropriate action, T_m is the time required to execute the appropriate action (the "manipulation" time), T_w is the time available to diagnose and initiate the appropriate action, and T_{sw} is the total time window between an initial cue (the time origin) and irreversible plant damage.

3.2 Steps 1-3: SHARP1 and Event Definition

SHARP1 is the result of improvements made on the SHARP model as recommended by the EPRI sponsored Benchmark of SHARP report.[14] The Nuclear Utility Services Corporation reviewed SHARP and created an accident scenario and benchmark process. The experts found that SHARP should emphasize the integration of the HRA into the overall PRA methodology. Instead of breaking the steps of the SHARP method apart as being completed by human reliability or systems analysts, another suggestion was to form an integrated team to follow the entire process through. The evolution to SHARP1 included a new approach, emphasizing the integration of the HRA with modeling the plant. The new method also includes only four steps, now called stages that are iterative instead of sequential as the original seven step SHARP.

The first step is the identification of significant plant states. In order to do this, the HRA team must limit the scope and context of the analysis. This includes defining initiating event groups and documenting possible plant responses to each group, including success criteria definitions for each function in the event tree and identification of proper EOPs After the scope and context of the analysis is set, the team can proceed to qualitatively screen the interactions.[15, 5] This is done by identifying significant plant states and functions that are crucial to accident mitigation. Because the EPRI HRA Calculator is software, the process is well defined by the inputs to the software. For example, the software will ask for specific information, and follow up with the next appropriate step.

Understanding the Emergency Operating Procedures is crucial when trying to understand the procedure based post-initiating events. This includes identifying failures in following the EOPs that can lead to unique and significant evolutions of the initial scenario. This step in the event definition stage is where the team can dive into the details of specific scenarios. While there is no specific guidance on how this breakdown and impact assessment should be done, SHARP1 refers to a variant of NUREG/CR-3177 as a possible procedure. This variant involves identifying critical values of key plant parameters associated with EOP response points and evaluating HIs via these parameters. However the team decides to carry out this step, they should be thorough and include such components as: [13]

- examining why the HI is required
- understanding how the HI is carried out

- identifying scenario-specific performance shaping factors
- identifying and understanding dependencies
- understanding failure consequences on the plant
- understanding failure consequences on subsequent operator actions
- identifying possible effects of training on operator actions (similar to what ATHEANA calls identifying "unwritten rules")
- defining time sequence of the accident progression
- defining the cue-response structure of the HI

The plant logic models need to be updated to incorporate the failure modes and dependencies. Failure modes should be modeled in only as much detail as necessary to capture the proper dependencies because too much detail at this level will make quantification significantly more difficult. Once this step is complete, the team should double check that the overall plant model is self-consistent, all assumptions are documented and well understood, and HI basic events are clearly defined and ready for quantification.[5]

3.3 Steps 4-6: Quantification

The EPRI quantification method is based on dividing the human failure event into a failure to initiate the proper action (p_C) and failure to execute (p_E). The probability p_E is quantified using THERP, where look-up tables for simple manipulation actions based on non-nuclear data, along with PSF correction factors, are used to find the probability of failure. The probability p_C, however, is more difficult to estimate. The first choice for estimating p_c is to simply look it up using a TRC – a curve, as in Figure 6, that correlates non-response probability to available time (T_W, $T_{W'}$).[13,5] This time-reliability approach is called the HCR/ORE method, and is further described in Section 3.2.1. Figure 5 is a simple event tree that shows how the two probabilities are related to success.

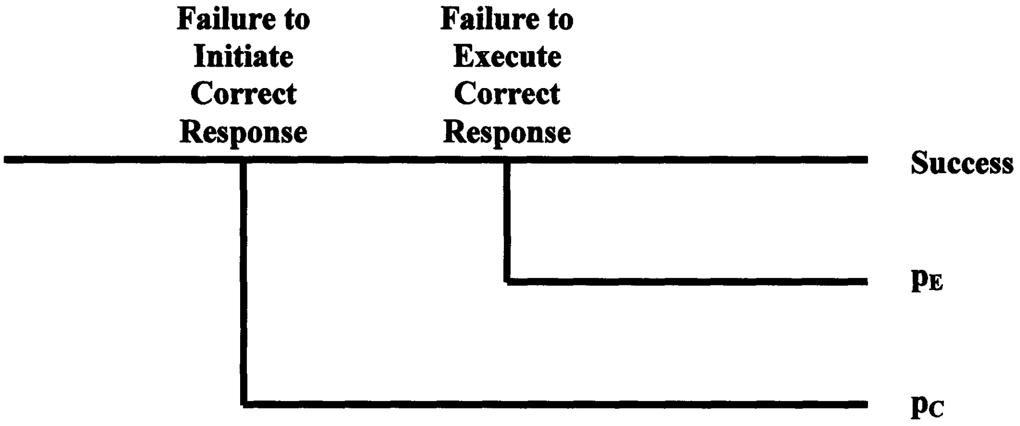


Figure 5: Generalized Event Tree for Calculating HEPs[13]

For a short available time (T_W), this method works quite well. However, the data used to create these curves fails to include those points where the operators misdiagnosed the situation and were on the "wrong path." For these cases, the extrapolated curve, seen as the dotted line, is not an accurate assessment of the HEP. For long times ($T_{W'}$), the actual probably of non-response would behave asymptotically, with a minimum that reflects a failure of the operator to properly diagnose the correct action, as seen in the figure below. Therefore, the HCR/ORE method is only useful for some situations. Other situations require an alternate method; that alternate approach is generally the cause-based decision tree (CBDT) method described in Section 3.2.2.

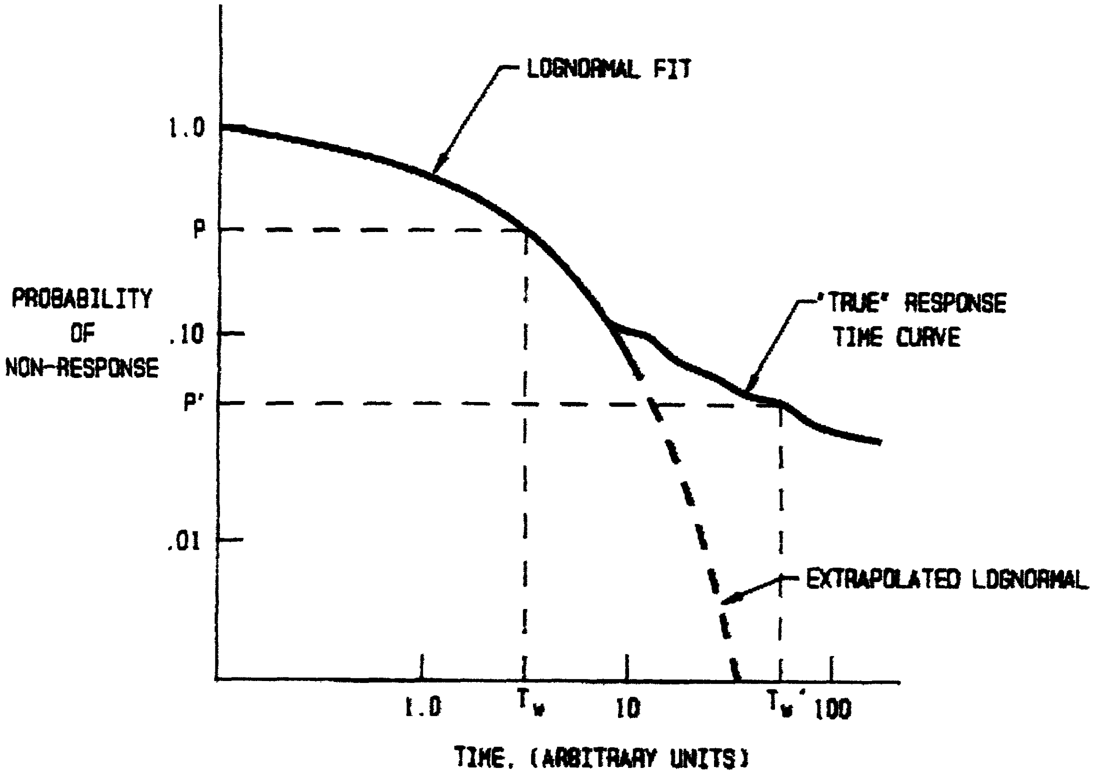


Figure 6: Conceptual Representation of the p_c Distribution as a Function of Available Time (T_w, $T_{w'}$).[13]

3.2.1 HCR/ORE

In addition to validating the cue-response structures, the ORE data also demonstrated that the lognormal distribution was a good approximation for HEP quantification, and so the HCR/ORE correlation was developed:[13]

$$p_c = \Pr(t_r > T_w) = 1 - \Phi\left[\frac{\ln(T_w / T_{1/2})}{\sigma}\right] \qquad [3]$$

where $T_{1/2}$ is the median response time, σ is the logarithmic standard deviation of normalized time, t_r is the response time, T_w is the available time, and $\Phi()$ is the standard normal cumulative distribution. This correlation is demonstrated in Figure 7 below: [13]

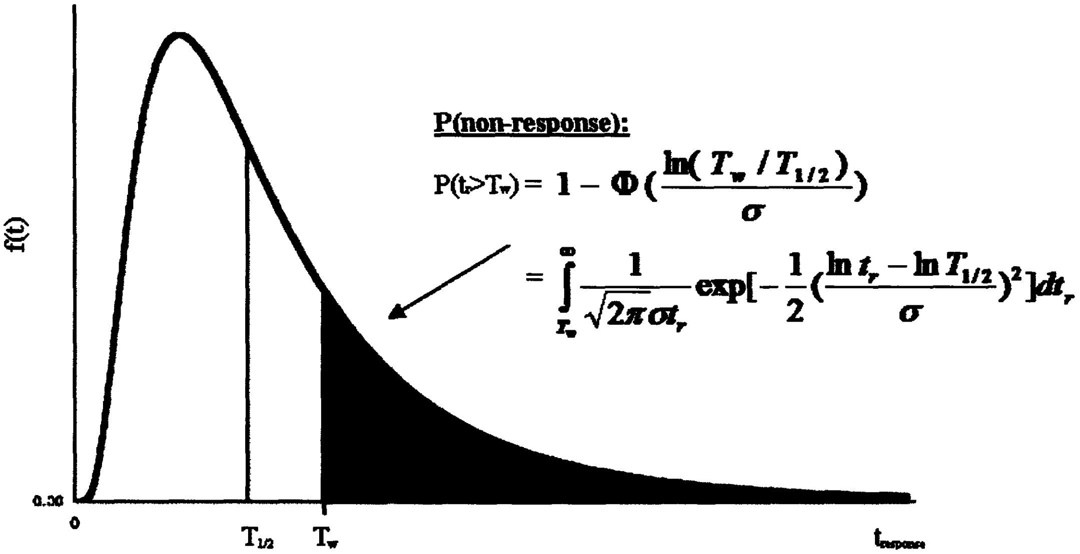


Figure 7: HCR/ORE Correlation, Lognormal Distribution of Response Time

The HRA team can apply these curves to simply estimate p_C. First they determine which cue-response structure is appropriate and find the $T_{1/2}$ and σ for that curve. Then, for a given scenario, they determine the time window of the total system ($T_{sw'}$) of a given HI from thermo-hydraulic codes (like MAAP). This system window must be adjusted (T_{sw}) to fit the HI event description (i.e., MAPP may give the total time until an event, but if the HI event begins with an alarm or parameter value, the time origin must also begin at the point that the alarm is sounded or the parameter value is reached). To use the curve the team must calculate the normalized time window: [13]

$$T_{w\_normalized} = (T_{sw} - T_m)/T_{1/2} = T_w/T_{1/2} \qquad [4]$$

again, T_m is the time needed to actually execute the necessary action, or the "manipulation time." Using the appropriate cue-response curve, the team can then just look up p_C on the graph for a given normalized time window (see a), or calculate it using Equation 3.

This method is only valid in the ranges where operating and simulator data is available – *extrapolation of these curves may produce unrealistically low estimates*, and CBDTs should be used instead for these cases.[5] It is generally good practice to perform a CBDT analysis in addition to using HCR/ORE and use the highest (reasonable) HEP to be conservative. Furthermore, the probabilities taken from the HCR/ORE correlation are only as good as the inputs to the correlation: $T_{1/2}$ and σ. Sigma is generally taken from the ORE curves for a given cue-response structure. $T_{1/2}$ should be obtained from plant-specific, HI specific data, such as simulator experiments or operator/trainer judgment. For the latter case, it is recognized that operations personnel may not have a good grasp for the time required for more complicated actions. The time ranges can, in these cases, be indirectly estimated by having the personnel identify ranges of key parameters within which they operators might act – the times could then be obtained from the thermo-hydraulics code. The median time would then be the middle of the given range.

3.2.2 *Cause-Based Decision Tree (CBDT)* [13,5]

The CBDT method is used to find HEPs for situations where TRCs are not applicable, situations such as: CP-4 and CP-5 HIs, HIs where $T_w >> T_{1/2}$ (ample diagnostic time) and other areas where the HCR/ORE method is determined to be unrealistically low. This method is based on a decision tree decomposition of a HFE into situation specific failure mechanisms, associated PSFs and possible recovery modes. Each HI interaction is decomposed into two high-level failure modes, each of which are in turn broken down into four failure mechanisms. These modes and mechanisms are defined in EPRI TR-100259:

Mode 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.

Mode 2: Failure in the Procedure-Crew Interface

e) The relevant step in the procedure is skipped.
f) An error is made in interpreting the instructions.
g) An error is made in interpreting the diagnostic logic.
h) The crew decides to deliberately violate the procedure.

A decision tree is created for each failure mechanism (see Figure 8 a-h). The nodes for each tree are PSFs which were predetermined to be important. The definitions for each PSF and any addition guidance provided to analysts on how to assess each PSF is provided in EPRI TR-100259:

a) <u>Availability of Information</u>:

1. *Indicator Available in CR* – Is the indicator in the Control Room?

2. *CR Indicator Accurate* – Are the indications available accurate?
3. *Warn/Alt. Procedure* – If the displayed information is perceived to be unreliable, do the procedures direct the operator to alternate sources of information? Do they warn the operator the indication might be inaccurate?
4. *Training on Indicator* – Has the crew received training in interpreting or obtaining the required information under conditions similar to those prevailing in this scenario?

b) Failure of Attention:

1. *Low v. High Workload* – Do to the cues critical to the HI occur at a time of high workload or distraction? [Workload or distraction leading to a lapse of attention (omission of an intended check) is the basic failure mechanism for this mechanism.]
2. *Check v. Monitor* – Is the operator required to perform a one-time check of a parameter, or is he required to monitor it until some specified value? "Monitor" leads to a greater failure probability than "check" because the operator might miss (exceed) the specified value if he does not check the parameter frequently enough.
3. *Front v. Back Panel* – Is the indicator displayed on the front or back panel of the main control area? Does the operator have to leave the control area to read the indicator?
4. *Alarmed v. Not Alarmed* – Is the critical value of the cue signaled by an annunciator? If the alarm comes in long before the value of interest is reached, it will likely be silenced and therefore ineffective.

c) Misread/Miscommunicated Data:

1. *Indicator Easy to Locate* – Is layout, demarcation, and labeling of the control boards such that it is easy to locate the required indicator?
2. *Good/Bad Indicator* – Is the MMI good or bad? Is it conducive to errors in reading the display?

3. *Formal Comms.* – Is a formal or semi-formal communication protocol (i.e., 3-way communication) used for transmitting values? Is the value always identified with its associated parameter?

d) Information Misleading:

1. *All Cues as Stated* – Are cues/parameter values as stated in the procedure? For example, if high steamline radiation is given as one of the criteria for a decision or action, at the time for the given action, the steamline radiation indicator would read high, not normal. The "no" branch is used if an indicator is *not* obviously failed but would not give the anticipated value (i.e., if the steamline was isolated).
2. *Warning of Differences* – Does the procedure itself provide a warning that a cue may not be as expected, or provide instructions on how to proceed if the cue states are not as anticipated?
3. *Specific Training* – Have operators received specific training in which the cue configuration was the same as the situation of interest where the correct interpretation of the procedure for the degraded cue state was emphasized?
4. *General Training* – Have the operators received general training that should allow them to recognize that the cue information is not correct in the circumstances? That is, is it something that every licensed operator is expected to know? For the steamline example, the answer would be "yes" because isolations are common; for instrument abnormalities that only occur under a very special set of circumstances, the answer would be "no" unless the operators had received specific training. Operators cannot be expected to reason from their general knowledge of instrumentation to the behavior of a specific indicator in a situation where they are not forewarned and there are other demands for their time and attention.

e) Skip a Step in the Procedure :

1. *Obvious v. Hidden* – Is the relevant instruction a separate, stand-alone numbered step or is it easily overlooked? A "hidden" instruction might be on of several steps in a paragraph, in a note or caution, on the back of page, etc.
2. *Single v. Multiple* – At the time of the HI, is the procedure reader using more than one flowchart procedure?
3. *Graphically Distinct* – Does the step stand out on the page? This effect is diluted if there are several things on the page which stand out.
4. *Placekeeping Aid* – Are placekeeping aids, such as checking off completed steps, used by all crews?

f) <u>Misinterpret Instruction</u> :

1. *Standard Wording* – Does the step use unfamiliar or ambiguous nomenclature or grammatical structure? Does it require any explanation?
2. *All Required Information* – Does the step present all information required to identify the actions directed and their objectives?
3. *Training on Step* – Has the crew received training on the correct interpretation of this step under conditions similar to those in the given HI?

g) <u>Misinterpret the Decision Logic</u> :

1. *"NOT" Statement* – does the step contain the word "not"?
2. *AND or OR Statement* – Does the procedure step present diagnostic logic in which more than one condition is combined to determine the outcome?
3. *Both AND & OR* – Does the step contain a complex logic involving a combination of ANDed and ORed terms?
4. *Practiced Scenarios* – Has the crew practiced executing this step in a scenario similar to this one in a simulator?

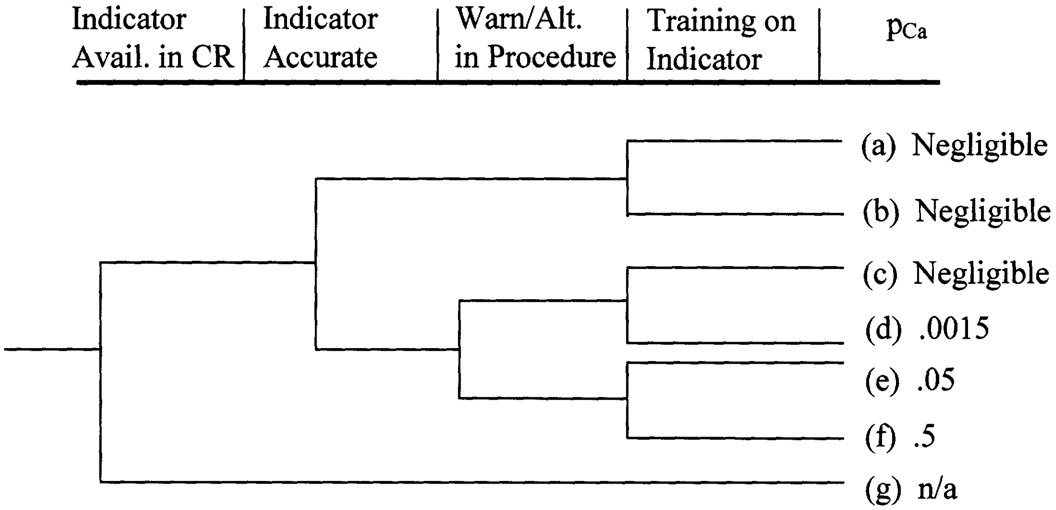
h) <u>Deliberate Violation</u> (*NOTE: this tree is rarely used in practice)

1. *Belief in Adequacy of Instruction* – Does the crew believe that the instructions presented are appropriate to the situation (even in spite of any potential adverse consequences)? Do they have confidence in the effectiveness of the

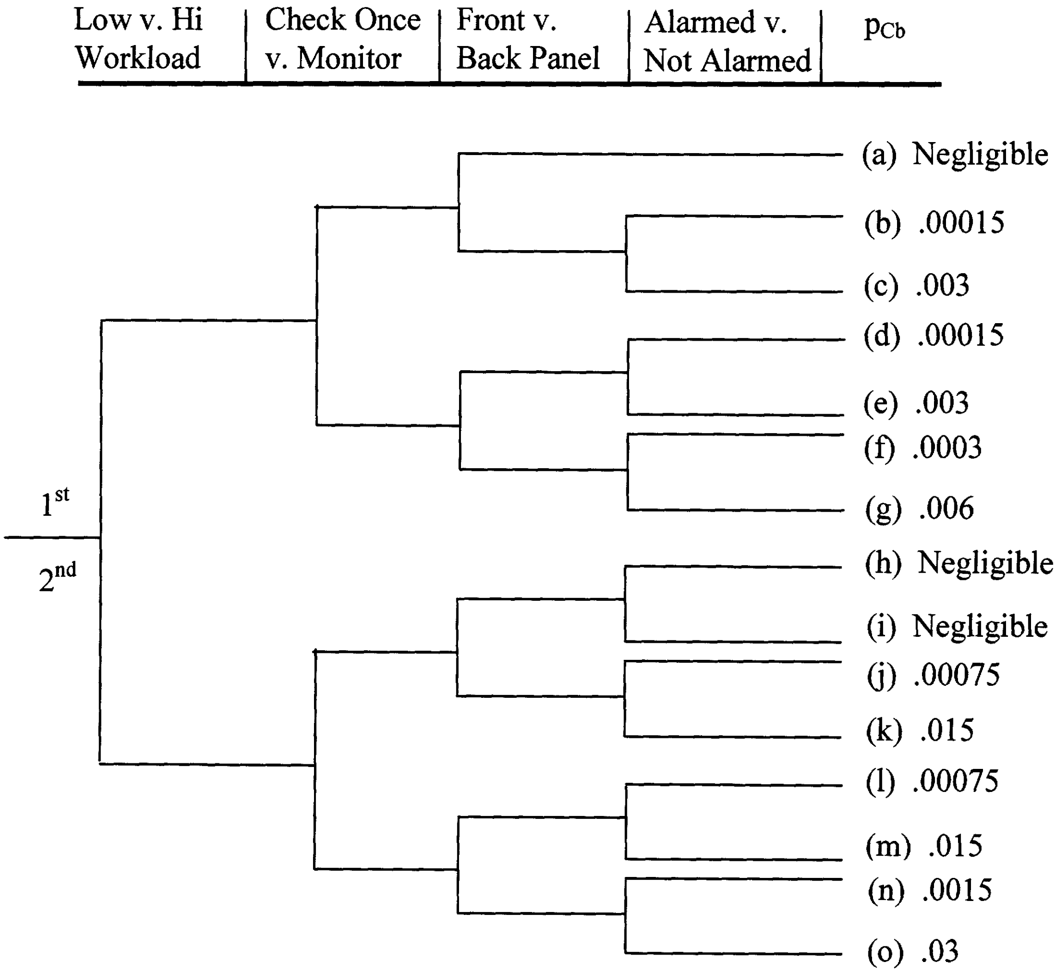
procedure for dealing with the current situation? In practice this may come down to: have they tried it in the simulator and found that it worked?

2. *Adverse Consequences if Comply* – Will literal compliance produce undesirable effects, such as release of radioactivity, damage to the plant, unavailability of needed systems or violation of standing orders? In the current regulatory climate, a crew must have *strong motivation* for deliberately violating a procedure.
3. *Reasonable Alternatives* – Are there any fairly obvious alternatives, such as partial compliance or use of different systems, that appear to accomplish some or all of the goals of the step <u>without</u> the adverse consequences?
4. *Policy of "Verbatim" Compliance* – Does the utility have and enforce a strict policy of verbatim compliance with EOPs and other procedures?

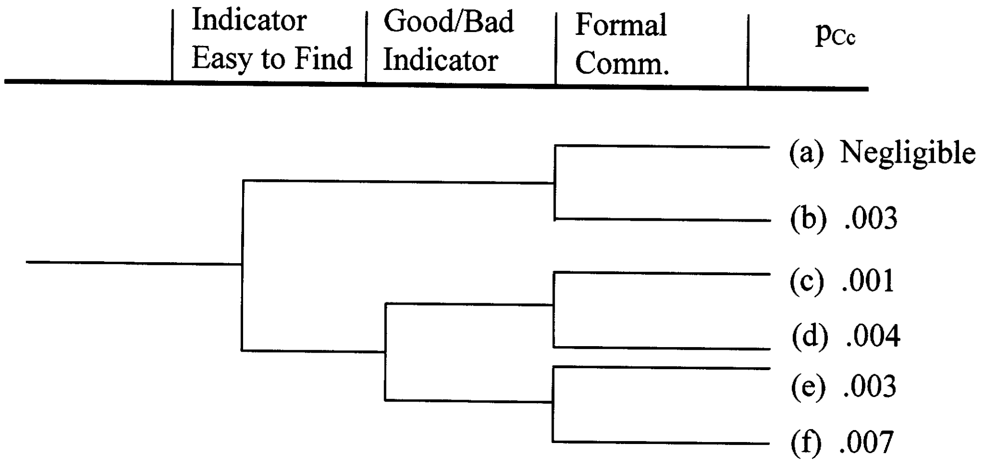
a) Data not Available:



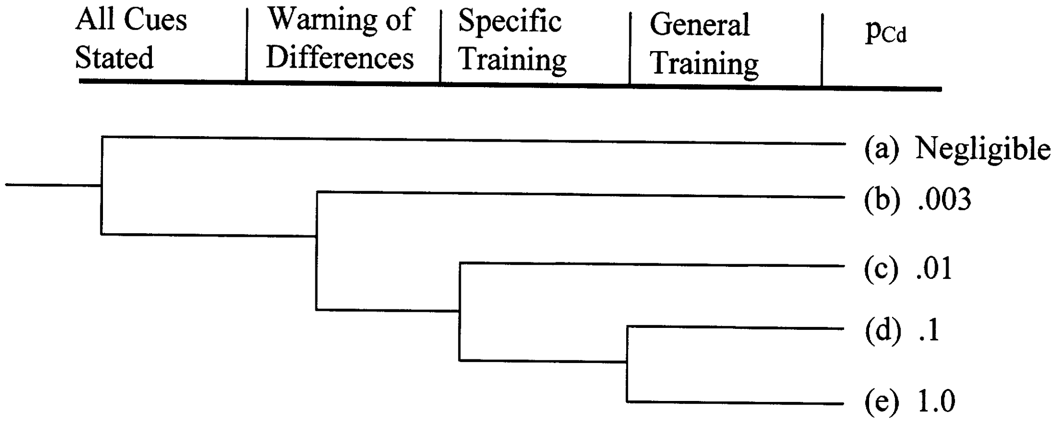
b) Failure of Attention:



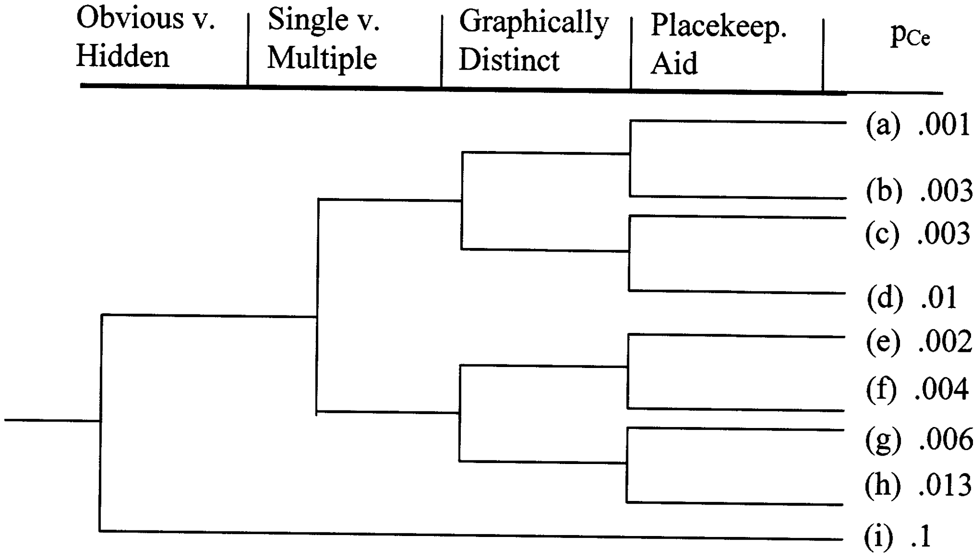
c) Misread/Miscommunicated Data:



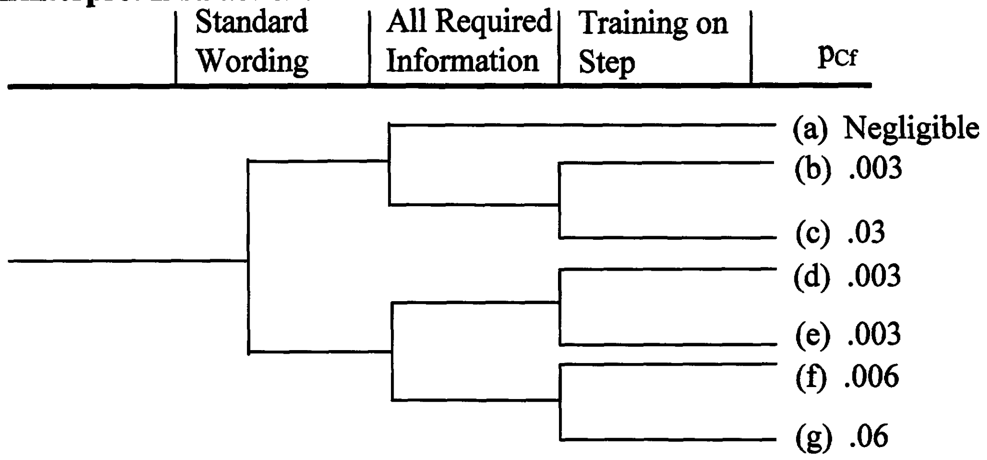
d) Information Misleading:



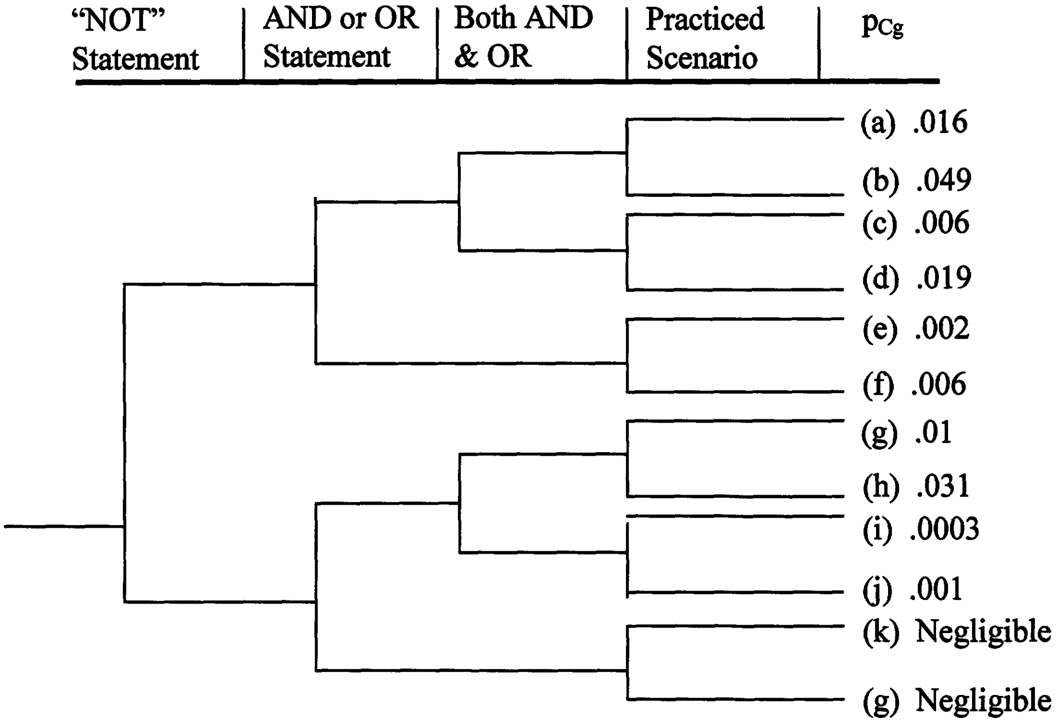
e) Skip a Step in the Procedure:



f) Misinterpret Instruction:



g) Misinterpret Decision Logic:



e) Deliberate Violation:

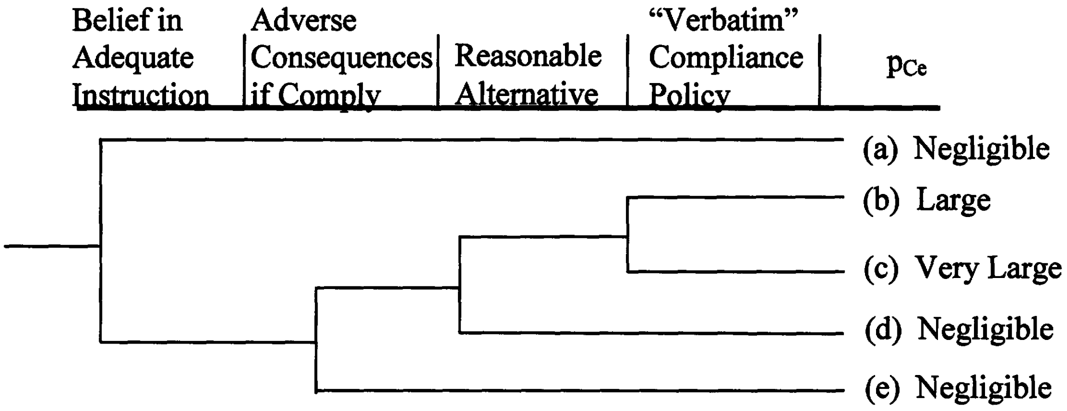


Figure 8: CBDT Failure Mode Decision Trees, a-h[13]

3.2.3 *Calculating the HEP*

To calculate the HEP, all of the applicable failure mechanisms need to be included, and for each failure mechanism, the analyst chooses the branch that most closely describes the HI being quantified. This calculation breaks the options into discrete probabilities using the failure mode decision trees, which helps eliminate some of the possible inconsistency between analysts. There is no way to eliminate the use of judgment when calculating human error, but the software attempts to provides a rigid structure for the calculations. The total HEP is then calculated according to the following equation:[13]

$$p_c = \sum_{i=1,2} \sum_j p_{ij} p_{nr}^{ji} \qquad [5]$$

where p_{ij} is the probability of mechanism j of mode i occurring, and p_{nr}^{ij} is the associated non-recovery probability for that mechanism.

The probabilities p_{ij} come from generic decision trees for a given mechanism. These default probabilities are generally taken from THERP, but can be adjusted by the HRA team. Some of the THERP values have been altered because of feedback from the ORE and other reviews of THERP.[13] The probability p_{ij} is then taken from the branch which represents the expected plant conditions. Similarly, recovery factors can also be taken from THERP or estimated by the HRA team, and are influenced primarily by opportunity for review by another operator and time available (Table 3 and Table 4). To avoid over-crediting recovery, credit is only given for one recovery mechanism; however, override values may be used if credit for multiple recoveries can be justified.[5]

Table 3: Available Recovery Factors for a Given Recovery Time[5]

Recovery Factor	Time Effective
Other (Extra) Crew	At any time that there are crew members over and above the minimum complement present in the CR and not assigned to other tasks
STA	10 to 15 minutes after reactor trip.
ERF/TSC	1 hour after reactor trip – if constituted
Shift Change	6 hours after reactor trip given 8 hour shifts 9 hours after reactor trip given 12 hour shifts

Table 4: Example Recovery Checklist with Probability for Recovery (modified) [13,5]

HI Failure Mode		**Review Type Available**				
Tree	**Branch**	**Self-Review**	**Extra Crew**	**STA Review**	**Shift Change**	**ERF Review**
a	a	No Credit	0.5	No Credit	0.5	0.5
b	d	X	No Credit	X	X	X
c	f	No Credit	No Credit	X	X	X

d	c	No Credit	0.5	X	X	0.1
e	a	X	0.5	No Credit	X	X
f	i	No Credit	0.5	X	X	X
g	b	No Credit	0.5	X	X	X
h	c	No Credit	X	X	No Credit	No Credit

4 SPAR-H– A Brief Overview and Important Terminology

The Standardized Plant Analysis Risk Human Reliability Analysis (SPAR-H) provides a simplified method for estimating HEPs at US nuclear power plants (NPPs).[3] SPAR-H provides steps for generating both pre-initiator and post-initiator HEPs, and this paper is only concerned with post-initiator HEPs for comparison with the other models.

The origin of this methodology is the Accident Sequence Precursor (ASP) program established by the NRC in 1979 as a response to the Risk Assessment Review Group report in 1978. The original PRA models developed as part of ASP needed further refinement, and the ASP HRA method was updated in 1994 by the Idaho National Laboratory. In 1999, a second revision was made and this became known as the SPAR-H method, which also included LP/SD scenarios as a separate class of events.[3]

SPAR-H is strictly a quantification method and does not include a unique approach to search and/or screen HFEs. NUREG/CR-6883 does refer to ATHEANA and SHARP1, and includes a section comparing the results from the SPAR-H quantification process with THERP, ASEP, HEART, CREAM, and SHARP1. This section will focus on providing an understanding of the basis for SPAR-H and how it handles varying situations.

The essential parts of the framework can be summarized as follows:[3]

- Probabilities separated into contributions from diagnosis failure and action failure
- Accounts for the context associated with HFEs by using PSFs and dependency assignment to adjust base-case HEPs
- Pre-defined base-case HEPs and PSFs with guidance to improve consistency and simplify the quantification process
- Uncertainty analysis uses a beta distribution
- Provides worksheets to further facilitate consistency and reproducibility

4.1 Model of Human Performance

The cognitive model used by the SPAR-H method is based on the information processing and stimulus-response models. The former models behavior as a combination of perceptual elements, memory, and decision making, while the latter largely ignores any thought processes and treats human behavior as reflexive to a situation based on associations between actions and either rewards or punishments. These two types psychology are used separately in the SPAR-H model: information processing is used for the diagnosis and the stimulus-response is used for the action of the operator. Operational factors that need to be accounted for by the human performance model are accounted for in SPAR-H by the eight performance shaping factors. These operational factors come in four main groups: inflow and perception; working memory/short-term memory; processing and long-term memory; and response.[3]

Inflow of information can be visual, auditory, or kinesthetic. The information flows through filters, which change how the information is perceived and interpreted. For example, noise and auditory distractions in an environment can mask the strength of a message such as an annunciator. In the context of SPAR-H, perception is detection, and the perception of an operator is effected by experience, learning, and beliefs. These are included by the assignment of the PSFs.

Working and short-term memory form the second operational factor covered by SPAR-H and processing and long-term memory are the third. The working memory is treated as the ability of the operator to keep information in an active mental state, which is differentiated from long-term memory that needs to be activated and retrieved. The capacity of short-term memory is not set, but it varies with the way information can be grouped. To aid the memory, SPAR-H models procedures and similar elements as external memory that aid with both short and long term memory. This factor is modeled as the procedures and job aid PSF. Other PSFs account for variable environments and mental demands affecting the memory.[4]

The fourth operational factor is the response of the personnel, and similarly, SPAR-H accounts for varying scenarios with PSFs. All except two of the PSFs that will be explained in detail in a later section affect the response of the operator. Generally, SPAR-H provides guiding analysis as opposed to a mathematical model of human information processing. The information processing model reflects psychological principles that can be linked to human performance. Table 4.1[3] describes which PSFs are associated with each of the four operational factors.

Table 5: Operational Factors in SPAR-H, including how PSFs are incorporated

(The numbers after each entry refer to the PSF list at the bottom of the table.)

Inflow and Perception	Working Memory/ Short-term Memory	Processing and Long-term Memory	Response
Presence[6,5] (is the signal there?) and opportunity (is anyone present to receive the signal?) Human sensory limits[2,5,7] Modality [6,5] (verbal, graphic/symbol, text) • echoic • iconic • kinesthetic Interference [6,5,4,7] (signal, noise)	Limited capacity[3] *Serial processing *Good only for a short time [2,3,5,4] (20 seconds) Right amount of attention [2,3,4,5,7] required Rehearsal [2,3,5,7] Physical and mental health [7]	Training[4] (models, problem solving, behaviors) • learning Experience[4] (models, problem solving, behaviors) • learning Culture [8] (societal, organizational, interpersonal, (crew)) • learning Intelligence/cognitive skills[3,4,1,5,7] (decision making, problem solving) Interference factors [6,2,3,7] (distraction) Available time [1,3] Physical and mental health [7]	Training (actions) [4] *Existing models of behavior *Practice and skill Experience [4] (actions) • practice and skill • existing models of behavior Proper controls available[6] Human action limits[6,7] (physical strength and sensory acuity) Ergonomics of controls[6,3] complexity Environmental degradation[2,3,6] Time to react versus time available [1]

Performance Shaping Factors

Available time[1]	Procedures (including job aids)[5]
Stress and stressors [2]	Ergonomics and human-machine interface[6]
Complexity [3]	Fitness for Duty[7]
Experience and training [4]	Work processes[8]

Note. Available time, from the operator's perspective, is influenced by information complexity, which can take more processing and reduce the time available to act.

4.2 Task and Error Types

Diagnosis and action are two kinds of tasks completed by personnel as specified by SPAR-H. Diagnosis actions require thinking to observe and interpret the information to find the cause and choose a course of action. These actions rely on knowledge and experience to plan an appropriate course of action. Action tasks are activities dictated by rules, procedures, or diagnosis. Each task is assigned a probability, and the probabilities of a diagnosis task and its action task are summed to yield the joint HEP. Dependence between the two tasks is calculated to prevent underestimating an HEP.

SPAR-H does not differentiate between errors of commission and errors of omission, because the documentation states that experience shows that no more accurate prediction of error can be made by distinguishing between the two. Instead, base failure rates use a composite rate for omissions and commissions. It is suggested that HRA analysis follow a systematic search process to identify errors likely to result in unsafe acts.[3]

4.3 Treatment of Dependency

SPAR-H treats dependency similarly to THERP.[18] There are five levels of dependence ranging from no dependence to complete dependence. Guidance is provided for determining the level of dependence based on the time separation of actions, location of actions, additional clues, and if the same crew is involved. There is no explicit treatment of dependency across accident sequences, but the same concepts could be generalized to fit multiple HFEs in an accident sequence.

PSFs are treated as independent by SPAR-H; although, it is recognized that for specific scenarios one PSF may be chosen based on how which level of another PSF was selected. For example, when there is ample time, but also less than thirty minutes of time beyond the nominal time, it is expected that the obvious diagnosis complexity PSF level is chosen. To prevent PSFs from providing exaggerated results due to their independent treatment, SPAR-H provides a method for making the HEPs less conservative. This is accomplished by using an adjustment factor if more than three PSFs were determined to be negative.[4]

4.4 Performance Shaping Factors

The definitions of the eight PSFs are discussed below, and there is some overlap that could not be avoided between some PSFs. The PSFs are typically stratified into levels that are assigned probabilities, and diagnosis and action events are typically evaluated for each PSF individually. For all of the PSFs, there is an insufficient information level that can be used as a default.[3]

1. Available Time- This is the time a crew has to diagnose and respond to abnormal events. For the diagnosis stage, the available time is broken down into groups ranging from inadequate time with a P(fail)=1, to nominal time and extra time. Extra and expansive time both require greater than thirty minutes beyond the nominal time required to respond to the event. The categories that have sufficient time, but less than thirty minutes are corrected by having the analyst use the "obvious diagnosis" PSF for the complexity. That eliminates needing another range for the available time. The time range for actions is greater because actions can be much faster, such as just pressing a button or turning a switch. The groups for the action probabilities range from inadequate time to greater than fifty times that required. The nominal time for actions is defined as having "some extra time" in addition to the required time to complete the appropriate action, unlike the diagnosis stage definition.

2. Stress/Stressors- This PSF models the negative impacts of stress. Even though stress has been recognized as a positive motivating force in some situations, the stress on a NPP crew modeled by SPAR-H assumes that it is detrimental. Stressors are environmental factors such as intense noise, heat, or increased levels of radiation that can affect performance. The three levels of

stress identified are extreme, high, and nominal. There is also an insufficient information option in the case that a stress level cannot be assigned. Extreme stress is caused by a sudden stressor and lasts for an extended period. For example, the potential for radioactive release would be modeled as an extreme stress. High stress is defined as above nominal levels, and it degrades performance. This level can be caused by distractions and unexpected events. The nominal level of stress is conducive to good performance.

3. Complexity- This PSF has more overlap with others and is hard to quantify independently from other factors. The complexity is a measure of how difficult a task is to perform in context. More demanding tasks with multiple parts that are ambiguous are highly complex, while nominally complex tasks are not difficult to perform due to little ambiguity and/or few variables. Figure 4.2 shows many factors that contribute to the complexity of a situation.

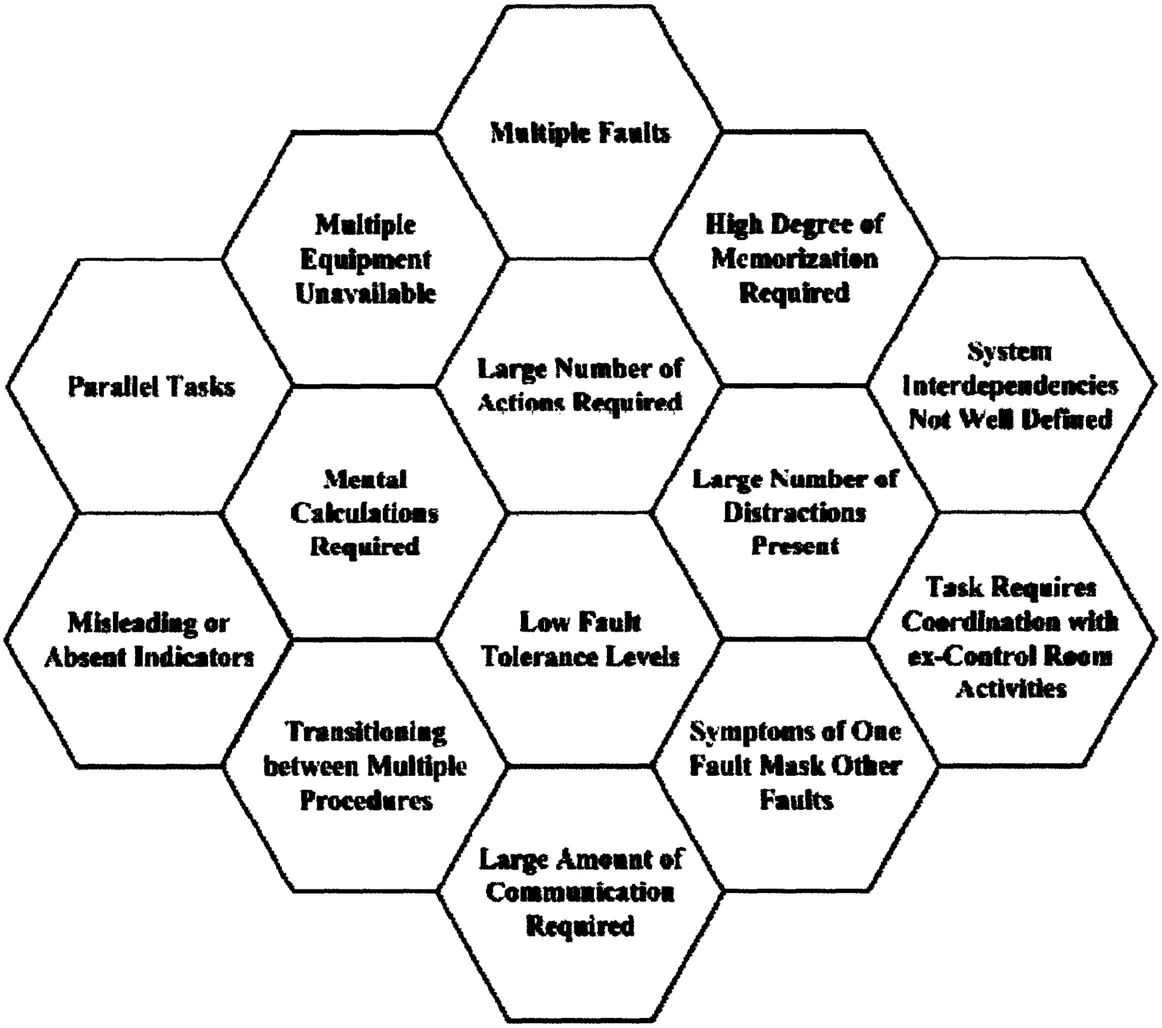


Figure 9: Diagram of Factors that Contribute to the Complexity PSF in SPAR-H[3]

There is no procedure for determining the influence of complexity based on the factors in Figure 4.2 that contribute. As mentioned earlier while discussing the available time PSF, the obvious diagnosis level of complexity is for tasks that are unlikely to be misdiagnosed. There is no need for an obvious action complexity PSF because the nominal complexity covers simple actions.

4. Experience/Training- The training and experience of the crew can affect their response to a scenario. This PSF accounts for the crew being trained for the particular scenario or similar scenarios and the years of experience. The low PSF level is for less than six months of experience, which at a US NPP would rarely apply to a licensed operator. The nominal level of 6 months of training and adequate schooling would be typical of a newly licensed operator, but most US NPP licensed operators and supervisors should qualify at the high level for this PSF, which requires extensive knowledge in a range of scenarios. Bypassing safety systems or operating in an abnormal configuration may decrease the level of the crew for a given situation.

5. Procedures- Even though all NPP documents are carefully formulated to provide the correct information, there may be situations where the information provided by the procedure is inadequate or incorrect. The diagnosis and action activities need to be evaluated separately, and the SPAR-H documentation warns against using the PSF for procedures to adjust for a complex task when the procedures are adequate. The diagnosis activity has a nominal level if the procedure is available and enhances performance as intended. Next is the available, but poor level, which impedes performance because it is hard to use, but all of the necessary information remains included. An incomplete procedure does not include the necessary information and the lowest level for a procedure is not available. At the other end of the spectrum, a diagnostic/symptom oriented procedure assists the crew

in diagnosis and helps keep the plant safe. Using a diagnostic oriented procedure should make it less likely that human error will have a negative impact on the state of the plant, but this is only true if the procedure is appropriate for the condition of the plant and is easy for the operating crew to follow. The action procedures have all of the same PSF levels except for the diagnostic/symptom oriented procedures.

6. Ergonomics/Human Machine Interaction- The ergonomics of the control room are accounted for by this PSF. The quality and quantity of information that the instrumentation, controls, and computers communicate to the crew effects its response to a scenario. As control rooms undergo updates, user friendly software becomes more important to the human machine interaction (HMI) as digital control and instrumentation becomes more prevalent. Another aspect of HMI is the set point of an alarm relative to a dangerous level. It can be detrimental to performance to have an alarm set point close enough to the limit that there is not enough time for the crew to react and correct the situation before surpassing the limit. The nominal level of this PSF is for HMI design that supports performance. Above nominal is the good level, in which case the HMI design improves task performance, and below are the poor and misleading levels, which negatively impact performance. The lowest level accounts for times when required information is not available to the NPP personel.

7. Fitness for Duty- This PSF describes the state of the crew both physically and mentally. Physical factors include fatigue, illness, drug use, and other medical problems that effect performance. Mental factors include overconfidence, personal problems, and other distractions that may similarly degrade performance of the operator. The levels for the fitness for duty PSF are unfit, degraded fitness, and nominal. The nominal in this instance is for the case where nothing degrades the performance of the individual. Degraded fitness refers to distractions and minor physical problems that negatively affect the employee's ability to perform such as headaches, fever, cold, or bad news. The unfit level is reserved for cases where the operator is unable to fulfill his or her duties.

8. Work Processes- This PSF measures the impact of management, organization, and supervision on the crew. SPAR-H maintains that supervisors need to be figures of authority, so a supervisor becoming involved in the specifics of an event instead of leading can be considered a breakdown in the work process and detrimental to the performance of the crew. Conflict within a crew, indecisiveness, and uncoordinated approaches to safety also negatively affect performance. The three levels of work process are poor, nominal, and good based on whether the various work processes at the NPP are detrimental, not affecting performance, or helpful. Examples of what would be poor PSF levels are inadequate turnover information, unclear performance expectations,

and conflict among personnel. The good level can be achieved by good communication and supportive policies enhancing the crew performance.

4.5 Uncertainty and Recovery

The SPAR-H method describes a way to calculate the uncertainty of an HEP. Epistemic and aleatory uncertainties are handled, but they are not separated and there is no mathematic process for separating the uncertainty into the two categories. SPAR-H uses a beta distribution to describe the uncertainty of an HEP, and the constraints on the distribution yield the greatest uncertainty for the HEP.[4]

All calculated HEPs are treated as mean values, so in order to determine the uncertainty, a second constraint is required. The fact that an HEP can only fall into the range between zero and one provides this constraint, and this constrained non-informative (CNI) prior maximizes the uncertainty for the given mean value. One potential problem with this method for calculating uncertainty is that by basing the uncertainty solely on the HEP, analyst variability is ignored. In other words, if two analysts find two different HEPs, then the two HEPs will have different uncertainties since the uncertainties are based only on the mean value. The uncertainty does not represent information based on the specifics of the accident scenario or analyst bias due to the use of the beta distribution.

Recovery actions in SPAR-H are modeled by the analyst in event and fault trees. This places the burden modeling the possibility of recovery on the analyst when developing the logic structures. An alternative way for the analyst to represent the influence of recovery on an HEP is to adjust the PSF values. Work practices, procedures, and ergonomics can be used to positively influence the HEP if a misdiagnosis is likely to be recovered.[3]

5 Comparative Analysis of ATHEANA, SPAR-H, and the EPRI HRA Calculator

The three HRA methods presented have various strengths and weakness that make them better suited to specific applications. The models are recapped with an overview before comparing particular aspects of the models to determine how they differ and what the implications of the differences may be. The goal of this chapter is to understand how the quantification methods vary when calculating post-initiating event HEPs.

<u>Summary:</u>

- Differences in Objectives:
    - EPRI: To quantify HEPs in a consistent, transparent and reproducible fashion. The method also aims to be less resource intensive and can be used by a PRA analyst without significant HRA expertise. A specialized HRA team is not required to use the EPRI HRA Calculator.

  - SPAR-H: This method was designed to provide a quantification methodology adequate for the ASP program that supports plant-specific PRA models for the US NRC. The method provides limited guidance for identifying or modeling HFEs.
  - ATHEANA: To find contexts where operators are likely to fail without recovery, and quantify the associated HEP. This includes inquiring into how operators can further degrade the plant condition while still believing their actions are correct. The search process of this model identifies PSFs and includes a screening process.

- Contexts Considered in Quantifying HEPs:
  - The three methods all incorporate aleatory and epistemic PSFs to some degree, and will be further discussed in the next section.
  - ATHEANA was not written to address pre-initiating HFEs and even though there are no technical problems using the method, there is no guidance provided by the ATHEANA documentation to guide quantification of pre-initiator events. This paper is concerned with only post-initiating events for the three models.
  - ATHEANA takes a look at a broader set of PSFs and contexts than the other two models discussed. The accident sequence includes the consequences of a misdiagnosis beyond simple failure of the procedure.
  - The EPRI Calculator is prescriptive and limits the PSFs and cognitive factors the analyst must consider, thus enhancing consistency among analysts and reducing the level of resources needed. However, CBDT and ATHEANA still have the same approach to evaluating cognitive context.
  - SPAR-H similarly limits the factors, this time to eight PSFs; this is only a quantification method, so there is no search process and little guidance provided to find HFEs, which would be necessary to perform the entire HRA portion of a PRA

- Quantification:
  - ATHEANA uses expert elicitation for quantification after the error forcing context is described. Agreement between the experts is addressed as part of the quantification method.
  - Depending on the type of HI – short or long available time, significant diagnosis – the EPRI Calculator defines the context based on the appropriate method. Quantification is completed using a time reliability curve or cause based decision trees.
  - SPAR-H uses nominal HEPs that are varied by using 8 fixed PSFs to account for the uniqueness of a situation. It deals with dependency between PSFs with a correction that takes the specifics of the situation into account only by using the number of PSFs included, ignoring that interactions between PSFs may vary with the combination of PSFs involved.

5.1 Terminology

Adjustment Factor (NUREG//CR-6883): This term describes the factor that is used to adjust PSFs for cases where more than two PSFs are determined to be negative. This factor is designed to prevent the model from producing HEPs that are overly conservative.

Human Failure Event (HFE): This term describes an event where a malfunction of a part of the plant is caused by human action. An error of commission or error of omission can be the source of an HFE. This term is widely used in ATHEANA and SPAR-H, but not used in the EPRI HRA Calculator. Instead, the calculator uses the term human interaction, explained below.[2]

Human Interaction (HI): The EPRI HRA Calculator defines human actions as any expected or actual action by a plant operator. Similarly to HFE, this includes both errors of commission and errors of omission. The HIs are broken down into Type A, B, and C depending on how the HI is related to the initiator of the scenario.

Error Forcing Context (EFC) and Failure Mechanisms a-h (NUREG-1624): An error forcing context describes the aggregate effect of the PSFs and plant conditions that make human error more likely. These are the scenarios targeted by ATHEANA. SPAR-H does use the term in NUREG/CR-6883, but only in reference to the ATHEANA model. The CBDT methodology that is part of the EPRI HRA Calculator does reference EFCs, but it does so in a more limited context. ATHEANA requires a more in-depth search process for EFCs.

Aleatory and Epistemic Uncertainty: Aleatory uncertainty describes irregularity that cannot be predicted. This uncertainty is due to random events that cannot be accounted for by having better equipment or performing more rigorous calculations. There is no way to lessen it because it is part of the system and cannot be removed. Epistemic uncertainty, on the other hand, can be reduced because it comes from a lack of familiarity. This type of error can be reduced by improving a method of measurement or quantification because the uncertainty from a lack of knowledge that can be improved through research and studies of a system.[19]

5.2 General Approach and Scope

This section compares the motivations and general approaches of the three models. The focus of this paper is the quantification of post-initiating events, but the entire scope of the model will also be addressed.

SPAR-H and the EPRI HRA Calculator were both developed to focus on the quantification of HEPs. This is why both models lack detailed guidance for the search and screening processes, but both models do suggest methods that can be used to identify and screen the HEPs. The HRA Calculator suggests following the SHARP1 framework for iteratively identifying and screening HFEs. For SPAR-H, there is not a specific method covered in great detail, but an example using ATHEANA is provided in section 4.2.2 of NUREG/CR-6883. The EPRI HRA Calculator does reference SHARP1 in enough detail for the model to function as an HRA tool that can stand alone and perform an entire HRA required for a PRA.

Unlike the other models, ATHEANA was designed to find specific scenarios in which operators are more likely to err. The methodology provides a detailed and rigorous search process that allows for a better representation of the important factors associated with an EFC. This is done by allowing the situation to drive development of the PSFs. Regarding the structure of PSFs, the opposite end would have models like SPAR-H with a fixed set of eight PSFs that can be altered to fit the context.

The resources required for the models vary with the level of detailed analysis that the model is capable of. Given this, the most resource intense model is ATHEANA, which requires an entire team of experts. The team consists of PRA, HRA, plant experts (operators and staff), human factors, and thermo hydraulic experts. This is in contrast to the expectations of the skill required to apply the EPRI HRA Calculator. Ideally, an HRA expert would use the EPRI software, but it can be used by a PRA expert with significant training in using the HRA Calculator.[4] SPAR-H falls in the middle between the two other models in terms of the skill requirements of the personnel. To fill out the work sheets of the SPAR-H model, at least one experienced HRA analyst would need to be part of the team performing the HRA.

The team of experts is also required for ATHEANA because not only is the search and screening detailed and rigorous relative to most HRA methods, but the quantification is based on expert opinion. As described in Chapter 2, the experts on the HRA team estimate seven percentiles for each HEP and after deliberation to work out differences in opinion, the average of the team is used to quantify the HEP. The EPRI HRA Calculator is designed to quantify the HEPs while minimizing the amount of expert judgment required. This allows the tool to be applied by non-HRA experts if an expert is not on the staff and/or there are not resources available to hire. SPAR-H relies more on expert judgment than the HRA Calculator. The eight PSFs need to be applied to the HFE by the expert, but to minimize the demand on the analyst, the various levels of the PSFs are well defined to decrease the impact of opinion on the process. The de-emphasis of expert opinion by the EPRI Calculator and SPAR-H help the repeatability of the HEPs quantified using the methods.

5.3 Available Time

When quantifying HEPs, time is a critical factor that is taken into account in different ways by all three methods. It is recognized that time effects the operators by applying pressure that can cause high stress situations, increasing the HEP. SPAR-H takes the simplest approach to incorporating time into the HEP of the three methods. This is done by dedicating a single PSF to available time. The levels are determined by how much time is available compared to how much is required to perform the necessary action. The method does recognize that the available time can affect other PSFs, but other PSFs are only affected if the available time meets certain criteria. The EPRI HRA Calculator incorporates available time into the HCR/ORE curves.[12] For this model, the time available to perform an action is based on the total time window for the disturbance minus the time to execute the correct action. ATHEANA incorporates the available time as part of the EFC. There are many recognized situations with little available time that operators are well trained for to ensure that they will perform the correct actions, but ATHEANA specifically searches for EFC that are unfamiliar, and for these, the available time is considered specifically when calculating the probability of an unsafe action given a particular EFC.[2]

In comparison, SPAR-H incorporates the simplest and least comprehensive method of incorporating the available time into the HEP. This puts the model at a disadvantage when compared to the HRA Calculator and ATHEANA because those models build a scenario that incorporates time, while SPAR-H separates it as one of eight PSFs that are essential independent. Available time is modeled in only a few discrete levels that do not differentiate between how critical time may be for a particular scenario in any way other than how the available time compares to the nominal time required to complete an action or diagnosis.

5.4 Performance Shaping Factors and Response Time Variation

The quantification of SPAR-H is centered on eight PSFs, which are the only factors that can differentiate any situation from one that has eight nominal values for the PSFs. In order to prevent calculating HEPs greater than unity, the PSFs are corrected by a factor if there are more than three PSFs. This method recognizes that there are problems with multiplying scalars with probabilities.[3] Beyond the eight discussed PSFs, there is no guidance on how to include a factor that may not fall into one of the predefined PSFs. Dependency is modeled using THERP. Both the terminology used and modification factors taken from THERP tables are used by SPAR-H.[18,4]

When using HCR/ORE as part of the EPRI HRA Calculator, the analysts can choose between the Sigma Decision Tree or the cue-response structure. PSFs are only considered by their impact on the TRC.[13] Using CBDT with the HRA Calculator allows incorporation of any relevant PSFs into the analysis in theory, but direction is only provided for the PSFs included in the decision trees.

ATHEANA needs PSFs to account for variation in the response time. These PSFs are not incorporated the same way as most HRA methods. Instead of using the PSFs to determine the HEP, ATHEANA uses the PSFs as factors dependent on the context, so for different conditions varying sets of PSFs are "triggered." [5] The method allows for any PSFs to be considered that are important to the EFC; however, the experts are left to determine how the combined PSFs for a scenario will affect the HEP quantification. Unless there are multipliers determined a priori, there will be no other way to quantify the PSF.[2]

5.5 Recovery

SPAR-H does not model recovery as a PSF or other influence that directly changes the quantification of an HEP with a multiplicative factor. This is a weakness of the simple eight PSF system that SPAR-H uses to quantify HEPs. The two suggested methods that can force the HEP to reflect recovery are as follows (NUREG/CR-6883): first, the analyst can perform more detailed analysis and update the logic structures to incorporate recovery, and the second option is to adjust the appropriate subset of PSFs. There is no guidance on which PSFs are appropriate for a situation.

The goal of the ATHEANA search process is to identify EFCs that are least likely to be recovered by the operators. Recovery is used as a screening tool to eliminate EFCs because recovery makes the scenario less important than other EFCs that are not likely to be recovered.[4] There is an option for the HRA team to perform a THERP based analysis for recovery if it finds this desirable; however, recovery will be less likely than usual due to expectations of short time frames that make recovery unlikely.[5]

The EPRI HRA Calculator addresses recovery with the CBDT method, which uses THERP tables. HCR/ORE does not model recovery because it assumes that the operators will perform the correct diagnosis and action.[12] To compare, ATHEANA does not have a fundamental need for recovery because the search process is based on finding EFCs that operators are not likely to recover from. This provides a more thorough analysis despite not handling the issue of recovery explicitly. The other two models should represent recovery, but do not provide defensible methods for incorporating the recovery into quantification of HEPs. This is because SPAR-H does not have an official mechanism for quantifying recovery, so there are two paths for adjusting the either the logic structure or PSFs to include recovery in the HEP.[4]

5.6 Documentation

Documentation is important for verifying and reproducing the results of an HRA. None of the methods discussed provide substantial discussion of documentation, but they do all address the issue implicitly. For ATHEANA, detailed documentation is more important than for the other methods because the analysts have more freedom at every stage of the HRA process from the identification and screening to the quantification of HEPs. The reasoning behind the actions of the HRA team is not transparent or intuitively obvious as it may be for SPAR-H. For example, using SPAR-H, the levels of PSFs are all clearly defined, so by filling out the work sheets completely, a range of scenarios can be inferred based on the values of the PSFs selected for the quantifications of the HEP. This is not the case for ATHEANA, because the experts may have agreed on a particular HEP curve for reasons that are not obvious and the reasoning behind their decision would be lost without proper documentation.[5] ATHEANA implicitly addresses the issue of documentation by listing the results that each step needs to pass on to the next step of the process. Assuming all of the results of each step are recorded, the HRA should be reproducible.

The EPRI HRA Calculator essentially documents itself by saving the inputs that the experts use to quantify the HEPs. This electronic documentation is the digital version of the worksheets of SPAR-H. Similarly to ATHEANA the search process, SHARP1, would be well documented if the analysts recorded how the products from each step of SHARP1 were determined to justify the work.[15]

6 Conclusion

The HRA field has made great progress in the past decade with the development of ATHEANA, multiple revisions of the EPRI HRA Calculator, and continuing work on second generation HRA models. After comparing the capabilities and limitations of ATHEANA, SPAR-H, and the EPRI HRA Calculator, some conclusions and future needs have been identified.

ATHEANA is very resource intensive when used for the entire HRA portion of a PRA, but it can be utilized to identify and quantify specific EFCs that would benefit from more rigorous modeling. By quantifying a select number of critical HEPs with ATHEANA while using another quantification method for HEPs screened as less critical, the much of the benefit of ATHEANA may be realized without expending the resources required to use ATHEANA for the entire HRA. Providing an alternate quantification method that does not rely so heavily on expert elicitation may reduce the resources required to use the method and allow it to become more widely employed.

SPAR-H is not a method that can reliably incorporate all of the information from a scenario into the eight PSFs and provide defensible results. The handling of multiple PSFs with an adjustment factor to prevent HEPs greater than one does not take into consideration which PSFs are acting together. Interaction between PSFs is not well managed when there are more than two PSFs identified for a scenario. There are suggestions in the SPAR-H documentation that guide interactions between some PSFs that are likely to interact, such as the available time and stress/stressors PSFs; however, the treatment of multiple PSFs does not account for the dependency of the PSFs. The dependency ratings are based on THERP and are recognized as not "exhaustive," but do "bring a degree of standardization." (SPAR-H REF) Inclusion of information into the SPAR-H method provides generic HEPs without rigorous consideration of how the specifics of the scenario can interact and affect the HEP.

Many HRA models are reliant on THERP data, which is based on old non-nuclear data, so a new database, formed specifically for the nuclear industry, would provide more applicable values for NPP PRAs. Such a database could be composed from the history of nuclear power plant operation and also from simulator studies. There are some groups working on similar tasks as this, but even after the completion of such a database, the HRA community will need to adopt the new data for use as the underlying data that HRA methods are based on.

Documentation within the models should be more heavily emphasized because the validation of the calculated HEPs and selected HEPs from the search process is dependent on understanding the process that the experts used to come to their conclusions. There are two distinct goals of documenting HRA: to provide traceability and reproducibility. The former refers to the ability of analysts to reference the documentation and completely understand how the HEPs were calculated for the situation. This includes an understanding of the situation based on the documentation that allows the work of the analysts to be traced from the context to the quantification.

The reproducibility requires that the actual values calculated for the HEPs can be duplicated based on the documentation. Many current HRA methods allow the analyst the freedom to determine how detailed the documentation needs to be, so the recommendation for documentation is that expectations and standards need to be explicitly defined to ensure traceability.

(This page intentionally left blank)

7 References

[1] Blackman, H. S. and Gertman, D. I., 1994. *Human Reliability & Safety Analysis Data Handbook.* New York: John Wiley & Sons, Inc.

[2] "Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA)," NUREG-1624, Rev. 1, Report, US Nuclear Regulatory Commission, Washington, D.C., May 2000.

[3] Gertman, D.I., H.S. Blackman, J. Byers, L. Haney, C. Smith and J. Marble, "The SPAR-H Method." NUREG/CR-6883, US Nuclear Regulatory Commission, Washington, D.C., 2005

[4] J. Forester, A. Kolaczkowski, E. Lois, and D. Kelly, "Evaluation of Human Reliability Analysis Methods Against Good Practices," NUREG-1842, Sandia National Laboratories, September 2006

[5] Presley, M. P., 2006. *On the Evaluation of Human Error Probabilities for Post-Initiating Events*. Thesis, (SM). Massachusetts Institute of Technology.

[6] "ATHEANA User's Guide," NUREG-1842, Draft Copy, US Nuclear Regulatory Commission, Washington, D.C., 2006.

[7] "An Empirical Investigation of Operator Performance in Cognitively Demanding Simulated Emergencies," NUREG/CR-6208, US Nuclear Regulatory Commission, Washington, D.C., July 1994.

[8] Forester, J., D. Bley, S. Cooper, E. Lois, N. Siu, A. Kolaczkowski, and J. Wreathall, "Expert elicitation approach for performing ATHEANA quantification." *Reliability Engineering and System Safety* 83 (2004), pp. 207-220.

[9] Kolaczkowski, A., et. al. "Example Application of ATHEANA: Pressurized Thermal Shock (PTS) Analyses," a Presentation to the Advisory Committee on Reactor Safeguard, PRA and Human Factors Subcommittees, Rockville, MD, June 28, 2006.

[10] Dawson, P.D., 2007. *Evaluation of the Economic Simplified Boiling Water Reactor Human Reliability Analysis Using the SHARP Framework.* Thesis, (SB). Massachusetts Institute of Technology.

[11] Orvis, D.D., A.J. Spurgin, C.T. Laio, J.P. Spurgin, and P. Moieni. Operator Reliability Experiments at Maanshan: Final Report. San Diego, CA: Accident Prevention Group, APG Report #8910, 1989.

[12] Hannaman, G.W., A.J. Spurgin, and Y.D. Lukic, Human Cognitive Reliability Model for PRA Analysis, NUS-4531, Electric Power Research Institute, 1984.

[13] "An Approach to the Analysis of Operator Actions in Probabilistic Risk Assessment;" EPRI TR100259, Final Report, June 1992

[14] EPRI: Benchmark of Systematic Human Action Reliability Procedure (SHARP), NP-5546, December 1987.

[15] "SHARP1 – A Revised Systematic Human Action Reliability Procedure;" EPRI NP-7183-SL, Interim Report, December 1990.

[16] Swain, A.D. and H.E. Guttmann, "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications," NUREG/CR-1278, SAND80-0200, Sandia National Laboratories, August 1983.

[17] "SHARP1 – A Revised Systematic Human Action Reliability Procedure;" EPRI TR-101711, Final Report, December 1992.

[18] Swain, A.D. and H.E. Guttmann, "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications," NUREG/CR-1278, SAND80-0200, Sandia National Laboratories, August 1983.

[19] Daneshkhah, A.R. Uncertainty in Probabilistic Risk Assessment: A Review. 2004. http://www.shef.ac.uk/content/1/c6/03/09/33/risk.pdf.