Evaluation of Human Reliability Analysis (HRA) Methods Against HRA Good Practices (NUREG-1842 - Draft for Public Comment)

Erasmia Lois (USNRC)
John Forester (SNL)
Alan Kolaczkowski (SAIC)

Presentation for Public Meeting

Rockville, MD
May 23, 2006

Outline

- Background
- Evaluation of methods
  - Approach
  - Summary of results
  - Brief description of each method and some observations
  - Comparison of methods against some key characteristics
  - Implications - What methods should be used when?

Background

- The NRC has developed the "PRA Action Plan for Stabilizing PRA Expectations and Requirements," (SECY-04-0118) to address PRA quality issues
- Guidance for performing/reviewing HRAs is part of the plan
- Guidance is developed in two phases:
  – Phase 1: HRA Good Practices--NUREG-1792, completed
  – Phase 2: Evaluation of methods against the Good Practices, in progress
- Status of methods evaluation
  – Draft report submitted for internal review, including ACRS
  – Address comments from ACRS sub- and full committees and others: February 2006
  – Submit for public comment: March 2006
  – Revise/submit to publication: September 2006

Approach for HRA Method Evaluation

- Compared methods, step-by-step with Good Practices
- Independent evaluation of ATHEANA, SPAR-H, SLIM/FLIM
- Expert meeting to present initial evaluation/expert input
- Addressed recommendations
  - Look deeper to underlying technical basis (frameworks, models, data)
  - Discuss methods as intended to be used versus as practiced
  - Develop plan for next steps
- Submitted for internal NRC review and feedback
- ACRS sub - and full committees' review and feedback
- Revising draft NUREG – adding additional summary and conclusion related information

HRA Methods Reviewed

- Technique for Human Error Rate Prediction (THERP) (NUREG/CR-1278)
- Accident Sequence Evaluation Program (ASEP) HRA Procedure (NUREG/CR-4772)
- Human Cognitive Reliability (HCR)/Operator Reliability Experiments (ORE) Method (EPRI TR-100259)
- Cause-Based Decision Tree (CBDT) Method (EPRI TR-100259)
- EPRI HRA Calculator
- Standard Plant Analysis Risk HRA (SPAR-H) Method (NUREG/CR-6883)
- A Technique for Human Event Analysis (ATHEANA) (NUREG-1624, Rev. 1)
- Success Likelihood Index Methodology (SLIM) Multi-Attribute Utility Decomposition (MAUD) (e.g., NUREG/CR-3518)
- Failure Likelihood Index Methodology (FLIM)
- A Revised Systematic Human Action Reliability Procedure (SHARP1, EPRI TR-101711)

Summary of Results

- Most HRA methods are quantification tools for estimating human error probabilities (HEPs)
  - Provide guidance for obtaining HEPs
  - As such are not dealing with the HRA process per se and hence many of the good practices
- A few touch on some aspects of how to do an HRA, but how to do a good HRA is left to analysts
- An exception is ATHEANA, and to some extent THERP, that provide both HRA guidance and a quantification approach
- SHARP and SHARP1 are guidance document on how to do an HRA
- The HRA Calculator is a computerized tool that guides quantification

Summary of Results (cont.)

- All HRA quantification methods have strengths and weaknesses
- Methods reflect an evolution of how to quantify human failure
  – Early methods more simplistically address human behavior
  – Progression of methods reflects efforts to better understand/incorporate advances in behavioral and cognitive science and operational experience
  – Different approaches/capabilities for translating qualitative information into human error probabilities
- Different methods developed for different purposes (detailed versus scoping analysis)
- Some can be applied much easier than others, but at a cost (less breadth and depth of analysis)

Summary of Results (cont.)

- Strengths, e.g.,
  - Some provide clear/good technical basis of underlying model
  - Good step-by-step guidance on how to use the tool
  - Traceable analysis
- Weaknesses, e.g.,
  - Weakness in technical basis make the use of some methods questionable
  - Some address only a limited set of performance shaping factors (PSFs) and context (plant conditions)
  - Methods not always applied as intended

Summary of Results (cont.)

- Overall perspective: Methods can be viewed as providing a “tool box” :
  - Some provide a tool for detailed analyses; others for screening analyses
- Using the right method for the right application is very important
- Therefore, we should use those methods that provide the best capabilities for the application
- Should use methods as they are intended to be used
- Drop any method(s) found to have unjustified technical basis

Technique for Human Error Rate Prediction (THERP) (NUREG/CR-1278)

- THERP is a method for identifying, modeling and quantifying human failure events (HFEs) in a PRA.
  - How to incorporate into PRA not covered
  - Emphasis on decomposing operator tasks into subtasks
- THERP has probably been used more than any other HRA technique
- Guidance for quantification of pre- and post-initiator HFEs
- Diagnosis contribution to error is handled with time reliability curves (TRCs) that provide no insights. Response execution HEP is added on.
- Basic HEP adjusted by PSFs
- Only a relatively small subset of PSFs actually addressed in quantifying HEPs (how to handle other PSFs left to analyst)
- Few HEPs and quantitative factors have an empirical basis

Accident Sequence Evaluation Program (ASEP) HRA Procedure (NUREG/CR-4772)

- A quantification technique for pre- and post-initiator human failure events
- Provides both screening and nominal human error probabilities for both pre- and post-initiators
- Otherwise, a simplified version of THERP meant to produce more conservative HEPs, but useable by PRA analysts with limited HRA background
- Basic HEP adjusted by PSFs
- Only a relatively small subset of PSFs actually addressed in quantifying HEPs (how to handle other PSFs left to analyst)

Human Cognitive Reliability (HCR)/Operator Reliability Experiments (ORE) Method (EPRI-TR100259)

- EPRI developed quantification technique for estimating non-response probability of post-initiator actions only
- Simulator measurement-based, time/reliability correlation (TRC) for diagnosis portion of human action
  - Does not explicitly address potential causes of human errors in diagnosis
- Needs relatively significant number of simulator exercises to produce reasonable results
- Evidence supporting use of the lognormal distribution, and thereby the standard normal distribution tables for obtaining non-response probability, is not available for public scrutiny
- Addresses both screening and nominal HEPs
- Includes Cause-Based Decision Tree (CBDT) method for longer time-frame events

Cause-Based Decision Tree (CBDT) Method
(EPRI-TR100259)

- Originally developed by EPRI to:
  - Address when HCR/ORE produces very low probability values
  - Address actions with longer time frames where "extrapolation of HCR/ORE TRC could be extremely optimistic"
- Quantification technique for estimating non-response probability of post-initiator human actions only
  - Causal approach allows consideration of 8 potential error mechanisms and factors that could contribute to those failures (diagnosis is assessed) through use of decision trees
- In more recent years, the CBDT method has frequently come to be used as a "stand alone" method
- No guidance for use under time-limited conditions
- Quantification data extrapolated from THERP, based on expert judgment

Standard Plant Analysis Risk HRA (SPAR-H) Method (NUREG/CR-6883)

- A quantification technique addressing both diagnosis and execution aspects human events
- Can be used for pre- and post-initiator events
  - SPAR-H does not use that classification nor distinguish
- Designed to provide reasonable estimates for regulatory uses
  - Accident sequence precursor program (ASP)
  - Phase 3 of the Significance Determination Process (SDP)
- Assumes basic HEP, adjusted to reflect ~8 PSFs
  - Nominal value for some PSFs usually assumed for control room actions
  - HEPs based on extrapolation of THERP and comparison with other methods
- Resolution of PSFs not appropriate for detailed HRA analysis (without expert judgment on part of the analyst)

A Technique for Human Event Analysis (ATHEANA)

(NUREG-1624 Rev. 1 & *Reliability Engineering & System Safety*, *83*: 207-220, 2004 Article on Quantification)

- Identification, modeling, and quantification of post-initiator human actions, including treatment of errors of commission
  - Concepts applicable to pre-initiators, but little specific guidance provided
- Addresses potential cognitive and implementation failures for a human action and the situations that could cause them
  - Strives to address a wide range of scenario and performance conditions (context) and unsafe actions
  - Intent is to address both nominal and deviation scenarios (i.e., not just "near-average" PRA context)
- Formal, facilitator-led expert elicitation process for quantification
- Guidance for addressing broad range of factors relevant to the <u>nominal case</u> needs to be strengthened (emphasis is on error forcing context)
- Detailed context development to determine the most appropriate influencing factors can be complicated and time and resource intensive
- If deviation scenarios need to be identified, analysis will take additional time

Success Likelihood Index Methodology (SLIM) Multi-Attribute Utility Decomposition (MAUD) and Failure Likelihood Index Methodology (FLIM)

- Quantification methods with a primary focus on post-initiator diagnosis failures
- Assumes that relative importance weights and ratings of PSFs, obtained from expert judges and related to a task, can be multiplied together and then summed across PSFs to arrive at the Success Likelihood Index (SLI).
  - FLIM (developed by PLG) is similar but provides scaling guidance for a suggested 7 PSFs (in some applications more)
- Requires events with known HEPs as calibration events (anchor values), and an assumption of a logarithmic-linear relation between the desired HEP and the SLI
- Identifying appropriate calibration data can be problematic
- Questions exist regarding the appropriateness of the linear model to reflect the experts' judgments
- Software tool for SLIM/MAUD not available

EPRI HRA Calculator

- Software tool – not a method
- Automates HCR/ORE, CBDT, THERP annunciator response model to address diagnosis of post-initiator HFEs
  - No guidance for which method to use
  - Includes aspects of SPAR-H for comparison purposes
- Uses THERP for response execution portion
- Uses THERP and ASEP to quantify pre-initiator HFEs
- Relies on SHARP1 as the HRA framework
- Not all PSFs discussed/addressed appear to be handled within the software quantification (this is being improved)
- Limited flexibility to address other PSFs (focus on standardization)

A Revised Systematic Human Action Reliability Procedure (SHARP1, EPRI TR-101711)

- SHARP1 is a guidance document for performing many aspects of an HRA in the context of a PRA (including identification and modeling issues)
- Covers both pre- and post-initiator human actions
- While it does not provide a quantification method for either, it does provide a summary of quantification methods available at the time.
- Generally consistent with the ASME standard for performing an HRA and with the NRC's HRA good practices guidance
- Insufficient guidance on identification of PSFs and context and on the consideration of errors of commission

Comparison of Methods Against Some Key Characteristics

- Discuss important (selected) HRA characteristics
- Address how the different methods cover those characteristics
  – Characteristics of quantification process

Overall Quantification Approach

- Uses concept of a basic/initial HEP that is subsequently adjusted and/or set tables, curves (generally limited and fixed set of PSFs) – THERP, ASEP, CBDT, SPAR-H
- Estimates HEP directly based on context & experience/judgment – SLIM/FLIM, ATHEANA
- Based on empirical or judged measures of timing for actions – HCR/ORE

Addresses Dependencies

- Has a model to address dependencies – THERP (among subtasks), ASEP uses simplified version of THERP, SPAR-H and sometimes FLIM uses THERP
  - Generic model requiring expert judgment to assess the level of dependence
- Discussed and to be considered as part of the context and included in the estimated HEP for given HFE – ATHEANA and to some extent SLIM/FLIM
- Discussed, but specific quantitative estimates not proposed. Effect on quantification left to the analysts – HCR/ORE, CBDT.
- SHARP1 provides overall good discussion, but does not address quantification of dependencies

Uncertainty Estimates

- Not context specific, stated to cover aleatory and epistemic but cannot be separated-THERP, ASEP, SPAR-H
- Limited guidance – HCR/ORE, CBDT
- Elicitation results from each team of judges are used to obtain a distribution of the HEP for each action; this distribution represents the team’s estimate, including uncertainties. Intent is primarily on capturing the epistemic uncertainty - SLIM, FLIM
- More context specific- largely aleatory since judges are asked to directly consider aleatory influences in obtaining distributions for HEPs. Aspects of epistemic uncertainty may be captured when judges consider HEP estimates for different quantiles.- ATHEANA

Range of Contexts Considered

Mainly the plant- related characteristics (plant conditions) that might vary for a given PRA scenario

- Largely an expected context based on PRA definition of scenario - THERP, ASEP, CBDT, SPAR-H, SLIM/FLIM – Depends on analyst to some extent
- Investigates nominal and related but different contexts (including so-called deviation scenarios) that all fit within the PRA definition of scenario ATHEANA
- Context not explicitly addressed other than as represented in the simulator runs (when used) Range of contexts requires many simulator runs - HCR/ORE.

Range of Specific PSFs Considered

- Most methods cover a relatively small range of PSFs and commonality is less than might be expected.
- THERP – For diagnosis, discusses a wide range of PSFs, but model addresses only a few - time available, event specific training, task load, redundant signals, stress, experience.
- SLIM and ATHEANA do not specify a fixed set - ATHEANA provides range of examples
- Only ATHEANA (and SLIM if modified) considers potential interactions between PSFs

Implications for Use of HRA Methods
(What methods should be used when?)

It all depends on the issue and decision being made

- HRA process
  - When issue/decision clearly affects just one or very few already identified HFEs with no need to worry about dependencies nor interactions with the rest of the PRA, then detailed identification and modeling processes etc. are not important
  - When issue/decision affects multiple HFEs or requires interactions with the rest of the PRA to be accurate (e.g., need to account for dependencies and the correct component rankings), then following the HRA process correctly becomes more important

Implications for Use of HRA Methods
(What methods should be used when?) (cont.)

It all depends on the issue and decision being made (cont.)

- HRA quantification and qualitative analysis of HFE
  - When the risk-related decision being made is not very sensitive to the specific qualitative and/or quantitative results from the method because, for instance, screening analysis or sensitivity studies show that the conclusions do not change, or
  - When level of PRA analysis (extent of PRA conditions being considered) is not intended to include detailed HRA considerations (e.g., ASP analyses), or
  - When, based on prior experience, seems likely that the most important influencing factors affecting the human action of interest are easily and directly handled using a less detailed, easier to use method,

  then simpler quantification methods may be used.

Implications for Use of HRA Methods
(What methods should be used when?) (cont.)

- Simpler quantification methods can provide helpful answers to the decision process as long as:
  - the primary weaknesses of a method are avoided
  - the method is not asked to give answers it cannot provide, for example,
    - determine causal influences to a diagnosis error using a simple TRC
    - assess the potential effects of communications when "communications" is not addressed directly by the method or easily interpreted as part of another factor that is covered by the method.

Implications for Use of HRA Methods
(What methods should be used when?) (cont.)

It all depends on the issue and decision being made (cont.)

- HRA quantification and qualitative analysis of HFE
  - The more the decision requires the "best" answer we can provide because the decision is very sensitive to the probabilistic inputs and the associated results from the HRA, the more important it is that the HRA process be rigorously followed and that a more detailed, broader scope quantification method needs to be used.
    - Whatever quantification method is used, it needs to be justified as to why it is appropriate for the decision being made
  - If one needs, for example,
    - A reasonably accurate estimate of the HEP - whether the probability of failure is high or low
    - To understand what the drivers for success/failure are and what conditions could create problems for the crew (so as to identify fixes),

    Then a detailed analysis that considers a reasonably broad range of conditions is needed

Implications for Use of HRA Methods
(What methods should be used when?) (cont.)

- Analysts/reviewers/users should avoid selecting a method first and then making the decision/issue *fit* the method
- The HRA process should be the other way around
  - Determine what is needed from the HRA to address the decision/issue
  - Select the appropriate method(s) accordingly AND justify the selection as well as the assumptions and judgments made in implementing the method(s)
  - Perform sensitivities to make results even more robust