



Wir schaffen Wissen – heute für morgen

Results and Insights Derived from the Intra-Method Comparisons of the US HRA Empirical Study

Julie Marble, Huafei Liao, Mary Presley, John Forester,
Andreas Bye, Vinh Dang, Erasmia Lois

presented by: Vinh N. Dang, Paul Scherrer Institute

11th Int. Conf. on Probabilistic Safety Assessment and Management (PSAM11)
PSAM11 - ESREL 2012, June 25-29, 2012, Helsinki

Results .. Intra-Method Comparisons .. US HRA Empirical Study

Many individuals and organizations made this study possible. Special thanks to the U.S. plant and its staff for providing crews, expertise, and their full-scope training simulator.

J. Marble, H. Liao, M. Presley, J. Forester, A. Bye, V.N. Dang, E. Lois



PSI's contribution supported by Swiss Federal Nuclear Safety Inspectorate (ENSI).

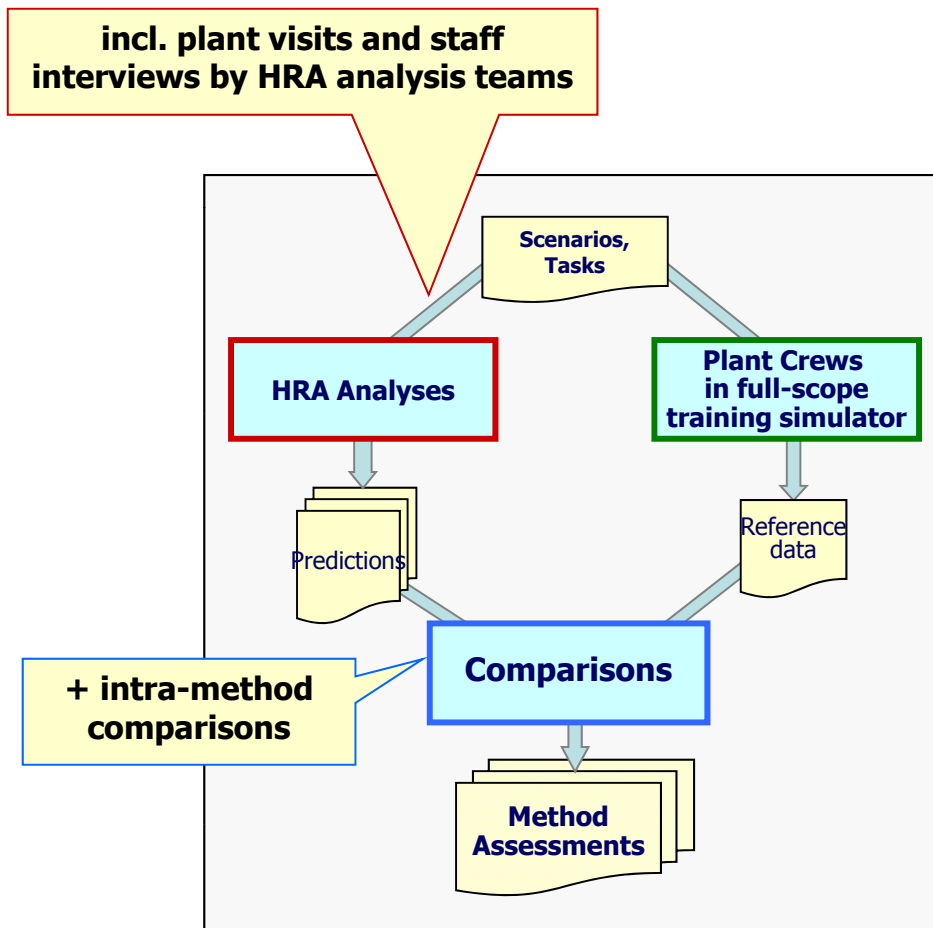
- Background to the U.S. Study
- Intra-method findings
 - general
 - few specific examples
- Conclusions

The views expressed do not necessarily represent the views of the U.S. NRC and other organizations mentioned.

The U.S. HRA Empirical Study

HRA Empirical Studies: Assessments of HRA methods based on comparisons of qualitative and quantitative predictions with reference data collected in simulator study

Among specific objectives of U.S. study: examine effect of analysts, i.e. "user effect"



Scenarios (#HFEs)

1. total LOFW followed by SGTR (2)
2. loss of CCW & RCP seal water (1)
3. textbook SGTR (1)

Licensed crews: 4

HRA analysis teams: 9

HRA Methods in Study

HRA Method	# Analysis Teams
• ASEP	2
• SPAR-H	2
• ATHEANA	2
• "HRA Calculator"	3

Multiple teams per method allowed comparisons among applications of a given method. In addition to comparisons to data, the intra-method assessments examined:

- differences in **qualitative** predictions
- differences in **quantification** and **ranking** of HFEs
- differences in the **analyses and assumptions**
- **and potential contributors to these differences**

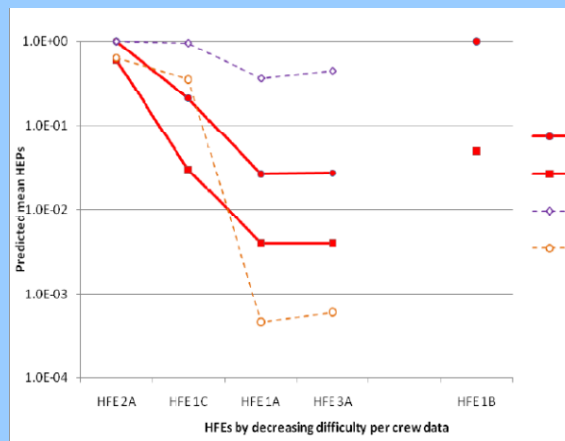
Overview of U.S. study findings

In bold: intra-method findings

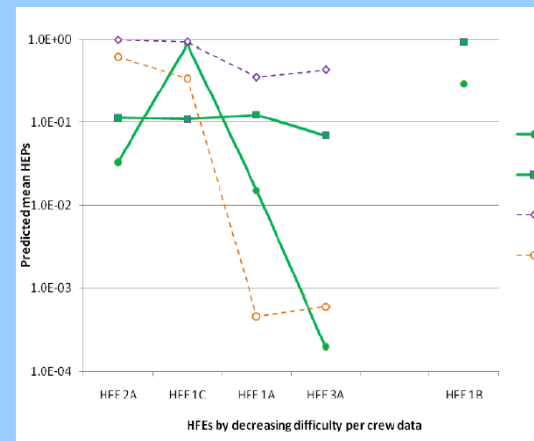
- ranking of HFEs reasonable for most methods (compared to reference data from simulator observations)
- for most HFEs, **one order of magnitude among predictions from a given method**
- some methods seem more consistent than others
- except ASEP teams, all other teams underestimated HFE 2A (Loss of CCW / RCP seal LOCA)
- **all teams ranked HFE 3A (textbook SGTR) easiest, but significant variability**

From Bye et al., "Overview and preliminary results", this conference and session.

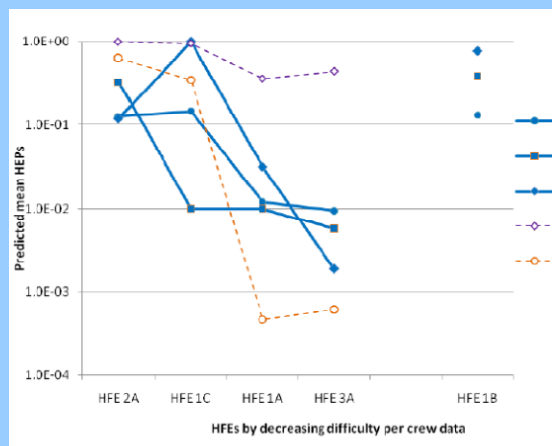
Assessments and comparisons start with **quantitative** results and comparisons but performance on qualitative predictions is weighted more strongly in assessments.



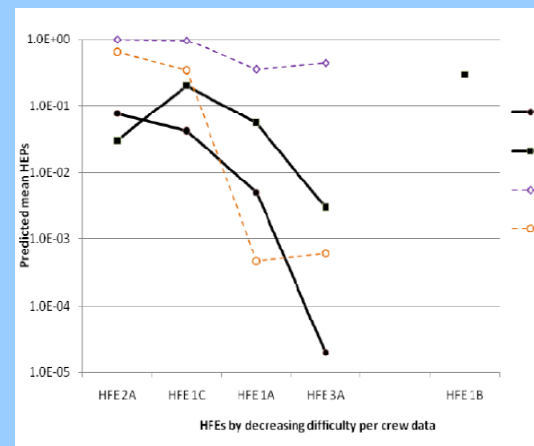
ASEP



SPAR-H



"HRA Calc" / HRAUG guide



ATHEANA

Understanding the reasons underlying differences

- in these graphs (quantitative)

as well as

- in qualitative predictions

is the goal of the intra-method comparisons.

Example intra-method findings

ASEP:

- For diagnosis, focus on evaluation of TRC led to missing consideration of diagnosis difficulties. A detailed review of procedure paths, as one team performed, can lead to better estimate of 'required time' as well as identification of potential issues for HFE success. Not much guidance within method for this.
- Method poorly equipped to quantify such difficulties, even if identified, although analysis team may compensate.
- ASEP vs. THERP execution modeling (option in ASEP): ASEP may not be conservative if number of steps is underestimated.

SPAR-H:

- One team treated HFEs as single task while other decomposed in detail. (Both approaches are "consistent" with method). Transitions (decision points) in procedures not treated by first team.
- Detailed decomposition yielded better qualitative predictions but lack of clarity in guidance for quantification of decomposed HFEs.
- Overlap of PSFs and resulting variability in modeling: lack of plant cues treated under "complexity" vs. "HMI" (indications).

Example intra-method findings (cont.)

HRA Calculator (per Users Group guidance):

- In complex scenarios, method's basic assumption -- that operators can/will follow procedures and can/will trust their cues -- breaks down. Difficulties treating HFE-specific aspects of procedural guidance, diagnosis complexity, and teamwork. Differences in how analysis teams compensated led to variability.
- Timing estimates for HCR/OR for diagnosis and execution varied widely. In many cases, failed to account for extra time due to distractions, parallel actions, resulting in optimistic estimates.
- Some analyses decomposed the HFE to consider multiple failure mechanisms in HFE subtasks while others did not.

ATHEANA:

- One team quantified MC simulation of a "scenario map" (set of potential plant-crew responses) with expert-elicited probabilities and delay distributions. Led to clear and traceable link between qualitative analysis and resultant HEPs.
- Second team more holistic in identifying operational challenges (less emphasis on time and more on potential difficulties) and quite effective even if less detailed. Used point estimates for durations and did not quantify Unsafe Actions (UAs) separately.
- Substantial differences in how teams performed qualitative analysis as well as quantification. First team's HEPs consistently lower but too many differences to identify main causes.

Conclusions (focusing on intra-method results)

- **Variability is examined at two levels**
 - **in qualitative findings:** issues that are identified in qualitative analysis
 - **in translation of qualitative findings into quantification** inputs

Assessment of method takes each into consideration.
- **Of the 4 methods, only ATHEANA guides the qualitative analysis** so differences in qualitative analysis approach of the teams
- **Teams varied in quality of information obtained in plant visit :** differences in interview skills.
 - So teams using the same method not making the same assumptions.
- PSF issues
 - **Lack or shortcomings of scales for PSF ratings.** Given similar qualitative assessment of HFE, different PSF ratings. *SLIM-like scales.*
 - **Overlaps among PSFs,** allowing same qualitative to be modeled using different PSFs (having different impacts).
 - Analysis teams interpret PSFs "broadly" to address **factors poorly covered by set of PSFs** provided by the method.

Outlook for reducing intra-method variability

- **variability due to inadequate method guidance** (lack of clarity, specificity, examples on aspects of the method)
 - **extend** guidance
 - **test this guidance** for analyst-to-analyst interpretation
- **variability due to differences in how analysts compensate for method shortcomings** (e.g. lack of coverage of a given factor or type of performance issue)
 - **method needs modification**
 - not a guidance issue per se

Note: poor method quantitative calibration, e.g. biased low or high, is also not a guidance issue.

- **variability due to scope and depth of the qualitative analysis** (PSF characterization and driver identification)
 - **suggest need for guidance for structured, consistent qualitative analysis**
 - then, coupling to quantification method needs to be assured
- **variability in information from plant visits / staff interviews**
 - Existing **interview, walk-through, talk-through guidance could be adapted for HRA analysis**, e.g. to avoid biasing responses or incorrect framing of HFE.

Back-up slides

Methodology and Pilot

- Halden Work Report HWR-844
(available from nrc.gov as NUREG/IA-0216)

Assessments of each method and detailed comparison results

- for SGTR, HWR-915
(NUREG/IA-0216 Vol. 2)
- for LOFW, HWR-951 (forthcoming NUREG/IA-0216 Vol. 3)

Overall Empirical Study assessment results

- Halden Project Report and NUREG (later in 2012)

U.S. Study assessment results

- Halden Project Report and NUREG (later in 2012)

Scenarios, HFEs, HEP table

- 1. LOFW with induced SGTR**
 - 1A - B&F given manual trip
 - 1B - B&F given automatic trip (*no observations*)
 - 1C - isolate SGTR and reduce pressure
- 2. Loss of CCW and cooling of RCP seals**
 - 2A – trip RCP pumps and start PDPs
- 3. Basic SGTR**
 - 3A – isolate SGTR and reduce pressure

	# crews	# failures	5th	mean HEP	95th	qualitative rank
1A	4	0	4.6 E-4	0.1	3.6 E-2	3
1B	-			n/a		
1C	4	3	3.5 E-1	0.7	9.5 E-1	2
2A	4	4	6.4 E-1	0.9	1.0	1 (most diff.)
3A	3	0	6.0 E-4	0.125	4.4 E-1	4 (easiest)

Table 7.1, draft report

Assessment Criteria: performance of method+team

Qualitative predictive power – driving factors

- Prediction of observed driving factors and performance issues
- Predictions not supported by the data

Qualitative predictive power – operational expressions

- Prediction of observed failure mechanisms

Quantitative predictive power

- Optimism wrto most difficult HFEs
- Consistency of the ranking of the HFEs (by predicted HEP) with reference difficulty ranking
- Predicted HEPs relative to the confidence/uncertainty bounds of the reference data
- Quantitative differentiation of the HFEs by HEP

Guidance and traceability

Insights for error reduction

HWR-915