An Empirical Study of HRA Methods – Overall Design and Issues

Vinh N. Dang ¹, Andreas Bye ², Erasmia Lois ³, John Forester ⁴, Alan M. Kolaczkowski ⁵, Per Øivind Braarud ²

Monterey, CA
August 26-31, 2007
Outline

• Aims of the study
• Organization
• Background
• Main study elements
  – Hammlab data collection
  – HRA analyses
  – Comparison
• Status and Outlook

Thanks to
- H. Broberg, S. Massaiu, M. Hildebrandt of HRP and R. Boring, INL
- the HRA teams!
- the many participants to the meetings and workshops
for their contributions and inputs to this study.

The views expressed are those of the authors and do not necessarily represent the views of the U.S. NRC and other organizations mentioned.
Overall aims of the study

• Assess HRA methods in light of simulator data
  – assess accuracy, strengths, weaknesses of HRA methods
  – provide the technical basis for improving HRA guidance
  – provide the technical basis for further development of HRA methods, if needed

• Explore ("pilot test") the study methodology
  – establish and test study methodology
  – what HRA “predictions” are testable in a simulator?
Organization and Participants
HRP Joint Programme and Own Funding

- **Steering Group**
  - E. Lois, US NRC
  - A. Bye, HRP
  - V.N. Dang, PSI
  - J. Julius, Scientech / EPRI
  - P. LeBot, EDF
  - P. Pyy, NEA (observer)

- **Organizing / Assessment Group**
- **HRP Experimental Staff**
- **Nuclear power plant crews**
- **HRA Teams**

- **Utilities**
- **Regulators**
- **Research**

US – NRC + Consultants, EPRI (Scientech), UMaryland
Korea – KAERI
France – IRSN, EDF
Czech Rep. – NRI
Denmark – Riso
Finland – VTT
Germany – GRS
Italy – Politecnico di Milano
Sweden – Vattenfall, Ringhals
Switzerland – PSI
HRA Methods represented in study
Classical, recent, 2nd gen, as well as simulation approaches

**THERP**: NRC staff + Consultants

**THERP w Bayesian Enhancement**: VTT

**ATHEANA**: NRC staff + Consultants

**SPAR-H**: NRC staff + Consultants, INL

**CBDT**: EPRI (Scientech)

**Decision Trees + ASEP**: NRI

**MERMOS**: EDF

**PANAME**: IRSN

**HEART**: Vattenfall & Ringhals

**KHRA**: KAERI

**CREAM**: NRI

**CESA**: PSI

**Simulations**

**Microsaint**: Alion

**IDAC**: UMaryland

**QUEST-HP**: Riso
Benchmarking / validation experience in HRA field

Limited experience with respect to benchmarking to data and particularly to simulator data

Related experience

• **Ispra benchmark**: multiple teams applying multiple methods. No reference data.
  - The results of the methods from each team were compared against each other.
  - **Outcome**: large spread of estimates across teams, across methods

• **EPRI HCR validation**: single team evaluating HCR correlation against large set of data
  - HCR: normalized time-to-response related to skill-, rule-, knowledge-based “tasks”
  - BWR and PWR, several utilities each, two or more countries, multiple scenarios
  - Range of data collected. Mainly, the time-to-response data was fitted/compared to the HCR correlations. Practically all crews responded successfully to scenarios.
  - **Outcome**: the HCR correlation was rejected (tasks could not be classified as S-, R-, K-). The HCR/ORE correlations and method were developed from the data (normalized time-to-response related to “cue structures”). These have not been validated.
Experience in HRA field

Related experience (cont.)

- Various U.K. benchmarks (Kirwan et al.): validation of few methods against HEPs
  - Most of the available HEPs relate to execution (considered less problematic for HRA methods).
  - Practically no empirically-based HEPs for decision-making performance.
  - Outcome: several methods found to be relatively accurate (factor of 3 and factor of 10 measures) and the different analysts for a given method were reasonably consistent across the different events.
  - The experience of the teams with the methods and the information about the scenarios/events being evaluated was controlled carefully.
# Types of comparisons and the Empirical Study

## Convergent validity
- Method-to-method

### Task / Scenario
- Method 1
- Method 2

### Result
- Result 1 ↔ Result 2
- reference HEP for Task

## Predictive validity
- Method-to-data

### Task
- Method 1
- Method 2

### Result
- Result 1 ↔ Result 2

## This study
- Method-to-data

### Scenario / Tasks
- Method 1
- Method 2

### Simulator sessions
- Result 1 ↔
- Measured performance (aggregated over observ.)
Three Main Elements

1. HRA Analyses
   - range of methods
   - multiple analysis teams

2. Empirical (Reference) Data
   - reference is performance in Hammlab
   - PRA scenarios

3. Comparison Predicted vs. Empirical Outcomes
   - HRA predictions vs. reference
   - (Additional comparison among teams and among methods may be performed, not focus of this initial study)

Essential study design issues

- Differences in the scope of the methods
  - e.g. older methods often assumed an external identification process while recent methods tend to include identification
  - errors of commission, decision errors

- selection of tasks for methods
  - what do we expect of the methods? probabilities, key factors, likely errors, information useful for error reduction?

- measures of performance of the methods

- data analysis and representation of aggregate performance
Study participants – the 4 roles

- **HRA Teams (analysts)**
  - apply HRA methods to predict performance in the PRA scenarios => **predicted outcomes**.

- **Assessment group: Halden and “neutral” participants**
  - compile **information package** for HRA teams
  - answer team questions (clarify info)
  - assess testability and fairness, “normalize” HRA predictions
  - compare **predicted** vs. **empirical** outcomes

- **Experimental staff of the Halden Reactor Project (HRP)**
  - design scenarios, together with assessment group
  - collect performance and other data on-line
  - post-session debriefing data
  - analyze experimental data => **empirical outcomes** = reference data

- **Nuclear power plant crews**
  - their performances in PRA scenarios in the Hammlab simulator = (raw) reference data
Empirical Study of HRA Methods

Method 1

Application

docum. HRA analysis
- model
- HEPs

Express as predictions in common form

predicted outcomes method 1

Method 2

Application
docum. HRA analysis
- model
- HEPs

Express as predictions in common form

predicted outcomes method 2

Method 3

crews

Simulator Runs, On-line and Debrief data collection

“measurements” & data from sessions

Analysis raw data

Data analysis for HRA study

reduced and refined data

experimental outcomes

- scenario descriptions, procedures
- walk-throughs, …
- way of working, training, other info

crews - walk-throughs, …
- way of working, training, other info

predicted outcomes method 2, method 3, …

scenario descriptions, procedures

methods predictions in common form

predicted outcomes
HRA Team Inputs (Information Package)

1. Overview, instructions, schedule
2. Administrative and agreement forms
3. Organization of the study
4. Information on Hammlab
5. Scenario description and HFES (actions of interest)
6. Characterization of the crews, their work practices and training
7. Procedures used in Hammlab
8. Forms for the response of the HRA teams
Reporting of Predicted Outcomes by HRA Teams

- free-form, open-ended questions on important factors and expected behaviors, for each HFE
  1) HEP
  2) Summarize most influencing factors on the crews’ behavior with respect to this HFE and why they are important.
  3) Discuss the difficulty or ease associated with these HFES in operational and scenario-specific terms

- structured responses using a pre-defined set of influencing factors, based on HERA classification
  - uniform terminology
  - reduction of effort required for data analysis, for deriving the aggregate performance or performance tendencies

- Documentation of HRA analyses according to PSA good practice
Comparison of predicted and experimental outcomes

- factors that most influence the performance
- level of difficulty associated with actions of interest
  - reason for difficulties
  - how difficulties are expressed in operational and scenario-specific terms

To be performed by assessment group
- all teams to review and comment on comparison results
- teams will have access to the experimental outcomes (processed to address privacy issues, etc) and be able to verify comparisons or perform their own.
Status and outlook

- **Nov-Dec 2006** 14 crews in **Hammlab** in 4 scenarios
  
  (2 types, base and complex variants)
  
  - dual use of Hammlab data for PSF/Masking and Empirical Study of HRA
  
  - data analysis is currently on-going

- **Info packages to HRA teams**
  
  Jan ’07
  
  - no information on performances in Hammlab!
  
  - clarification period (Q&A)

- **HRA responses April ’07**, feedback and revisions

- **June**: first results of experimental data analysis (experimental outcomes), start review and comparison

- **Draft comparison results** from assessment group

- **Oct 2007**: Workshop on study results, Washington (hosted by NRC)

- Finalization of comparison results

  *Data from second scenario type reserved for follow-up study.*
Expected results

On performance of the HRA methods (although preliminary due to pilot character of study)

- does a method identify the main influences on performance observed in this scenario?
- does the method’s model of how these factors interact correspond to observed interactions?
- does the empir. performance provide evidence for the performance levels predicted by HRA?
- what are areas where guidance could be developed to improve HRA predictions?

On study methodology

- What do HRA teams need? Strengths and drawbacks of info package.
- Empirical data analysis: Adequacy of the study’s approach to aggregating the observed performances at a level compatible with HRA results?