Validation and Verification of HRA Methods by NPP Simulator Exercises

Gueorgui Petkov, TU-Sofia, Bulgaria

‘Expert Guesses’ vs. Simulator Exercises

- HRA methods must provide unified and quantitative description of + and - aspects of human performance.
- The expert judgement is in considerable use to overcome the *inaccuracies* in mental and situational models and simulator exercises’ conditions.
- But the *“expert guesses”* cannot be taken as the basis for **Validation and Verification of HRA methods** because “we do not have data to validate the accuracy of the guesses…” [N. Moray].

Validation and Verification of HRA

- The HRA is an integral part of the PRA.
- The V&V of the HRA methods is the key for:
  - increasing fidelity,
  - decreasing uncertainty and
  - guaranteeing acceptability of the PRA results.
- The V&V of the HRA method could be described as the “white elephant” of each PRA study.

Goal and Objectives

- Another way to overcome the difficulties of predicting operator performance is to simulate all processes in parallel and to expose the operators to variations that can occur during various scenarios.
- The simulators could be used not only for extension and coordination of training and HRA purposes but also for **V&V of advanced HRA methods**:

  More data $\Rightarrow$ More qualitative data

  More investigators $\Rightarrow$ More opinions & judgments

  Enough samplings $\Rightarrow$ Less uncertainty

HRA V & V Definitions

- The **HRA V&V** is a process of checking that a given HRA technique meets requirements, specifications and fulfills its intended purposes.
- A ***validation*** of a HRA method means to prove or confirm (to some extent) that the technique and database included are true and results are correct.
- A ***verification*** of a HRA method means checking it by careful investigation (reviewing, inspecting or testing) to establish and document the HRA procedure.

Current HRA V & V Studies

- The HRA techniques validation studies have been performed for a number of widely used HRA techniques with the aim of establishing the credibility and accuracy of the techniques.
- The preferred approach to validation is comparison of a wide range of HRA results against standard.
- The data for comparative validations are insufficient.
- Studies have therefore focused on determining the degree of correlation between techniques using limited real HEP data (*convergence validity and accuracy*).

Convergent and Comparable HRA

- The convergence is much more valuable if the validated HRA techniques are based on **different**
  - *concepts*,
  - *principles* and
  - *criteria* than having
  - similar features of techniques.
- It is very important that the comparison of HRA methods and their HEPs evaluation is implemented in the **same**
  - *scenario*,
  - *conditions*,
  - *situations* and
  - *people involved*.

How To Validate and Verify?

- The HRA methods' validation if could also be done by **comparison and confirmation** of the *basic concepts*, *capacities*, *criteria*, ***database*** and ***results*** of different HRA techniques on the NPP computer-based or full-scope simulators.
- Such studies allow comparing simulator-based data with a number of HRA techniques against a number of expert data or known HEPs.
- Validation is important because it disallows HRA technique data that can not possibly be real to be entered into a database.
- If the HRA technique based on top expert opinions or existing HRA database does not give correct evaluation of HEPs in simulated or virtual situations, it is not likely that its results for real situations would be better.

Useless Validation

- The HRA validation against an incomplete or insufficient set of criteria can lead to a state of ‘validated’, where the term ‘validated’ does not necessarily mean trustworthy.
- Thus, validation of the validation criteria is an important aspect of the HRA that is often overlooked. Establishing such validation theoretical and practical criteria can be a very difficult task when evaluating complex HMS.
- Establishing **"fitness for purpose"** is often a useful concept to support evaluation of HMS in that the approach focuses on involving operators, training personnel, HRA specialists or psychologists in establishing & reviewing the purpose.

Human Actions Types

HA1: It corresponds to those HAs which develop prior the IE and may affect system/function reliability.

HA2: It includes those HAs that may lead to an IE.

HA3: It covers HAs developed by operators during an accident, following the established procedures for mitigating the consequences of an accident.

HA4: It corresponds to HAs developed by operators during an accident, in the belief that they are the appropriate actions, as indicated in the procedures, but which in reality can worsen the conditions of an accident, complicating the mitigation process.

HA5: It corresponds to HAs that are taken by operators during an accident and that are not definitely included in the procedures; they tend to recover the operability of failed equipment or use alternative means.

Validation & Verification Options

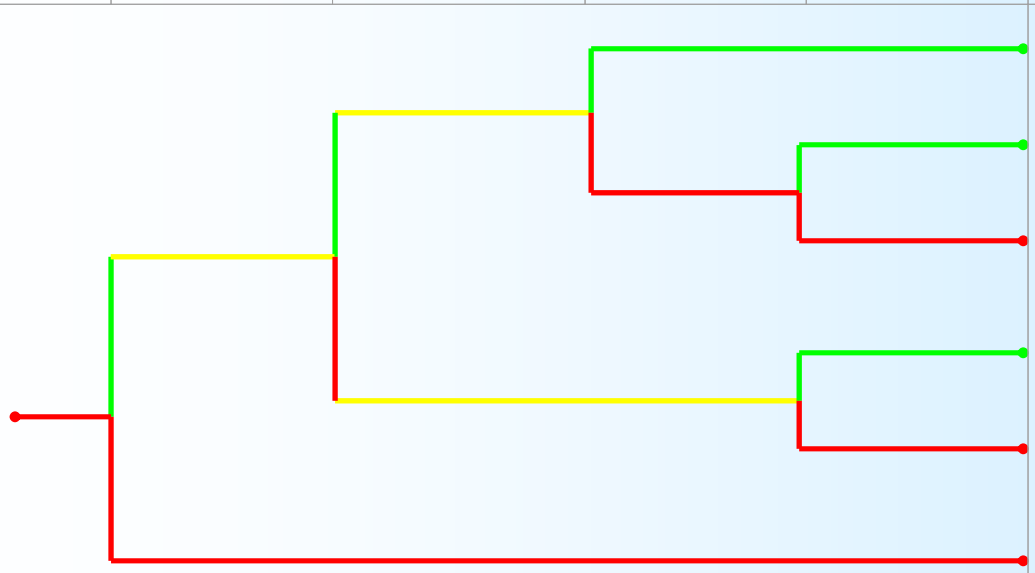
1. **Basics**: concepts, principles, criteria, databases
2. **Methods**: THERP,HCR, HEART, HDT, PET
3. **Human action types**: HA1, HA2, HA3, HA4, HA5
4. **Context**: scenario, function, conditions, situation, task, procedure, stressors, individual, crew, …
5. **Simulators**: computer-based, analogous, multifunctional and full-scope
6. **Experts:** students, operators, instructors, engineers, HRA specialists, psychologists

Approaches to Human Errors

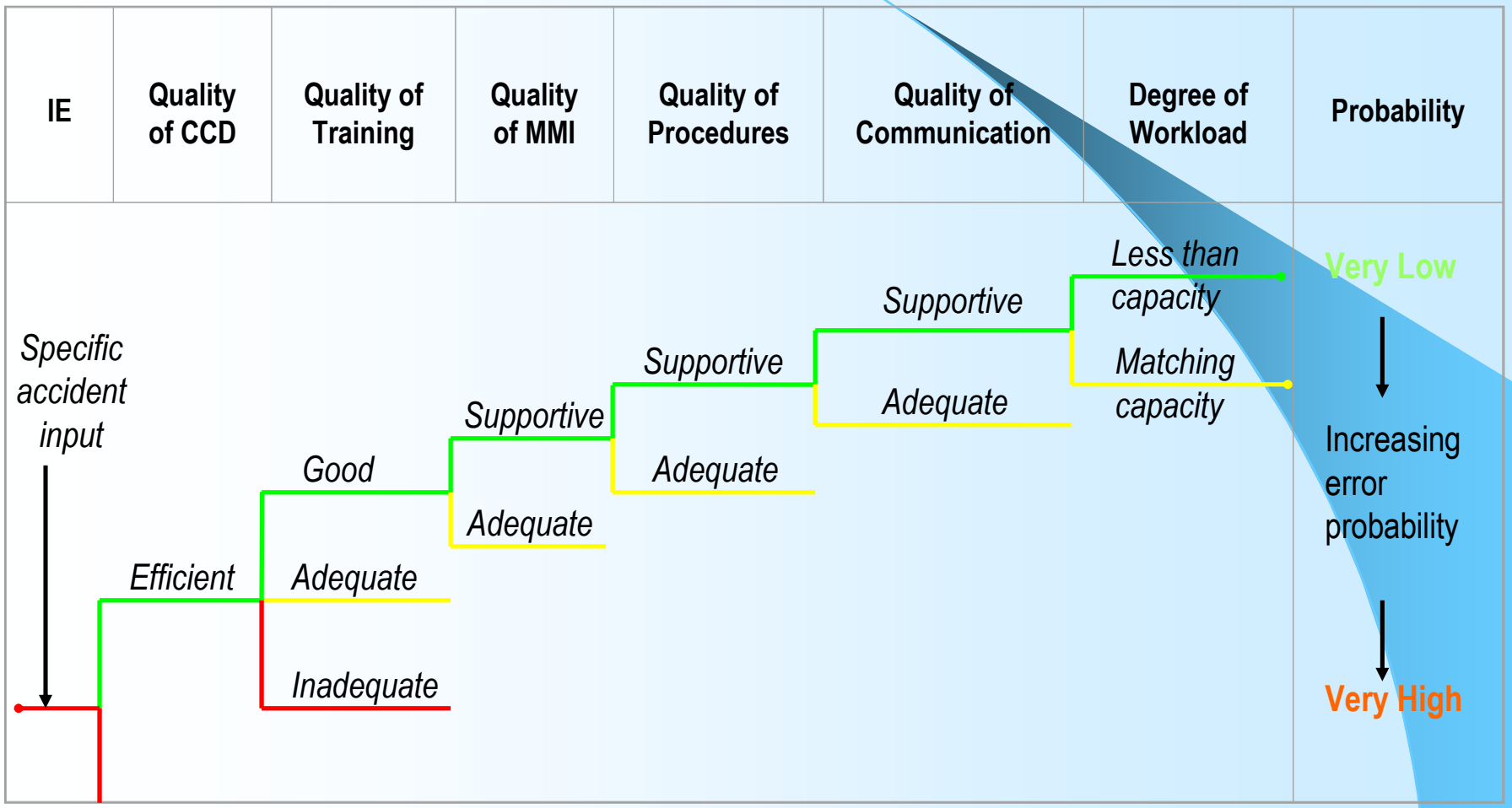
- Performance is ‘fragmented’ into separate actions
- Performance is ‘reasoned’ by trees
  - Sub-processes of cognition - **OAT**
- Performance is ‘described’ by fuzziness (influence factors - IF)
  - **THERP** (Performance Shaping Factors – PSF)
  - **HCR** (PSF)
  - **HEART** (Error Producing Conditions – EPC)
- *Holistic approaches*
  - Combination of trees and fuzziness - **HDT** (decision tree of manifestations, Influence Factors - IFs and Quality Values - QVs)
  - Non-hierarchical fuzzy organization/classification scheme by distinction between competence & control and distinguishing causes and manifestations – **CREAM** (Common Performance Conditions – CPC)
  - Iterative contextual control loops/graphs of inventive, conscious (cognition, communication and decision-making), unconscious and physical (execution) processes by statistical measuring of HMS state randomness – **PET** (Cognition/Communication Models, Context Factors & Conditions – CFC).

Context

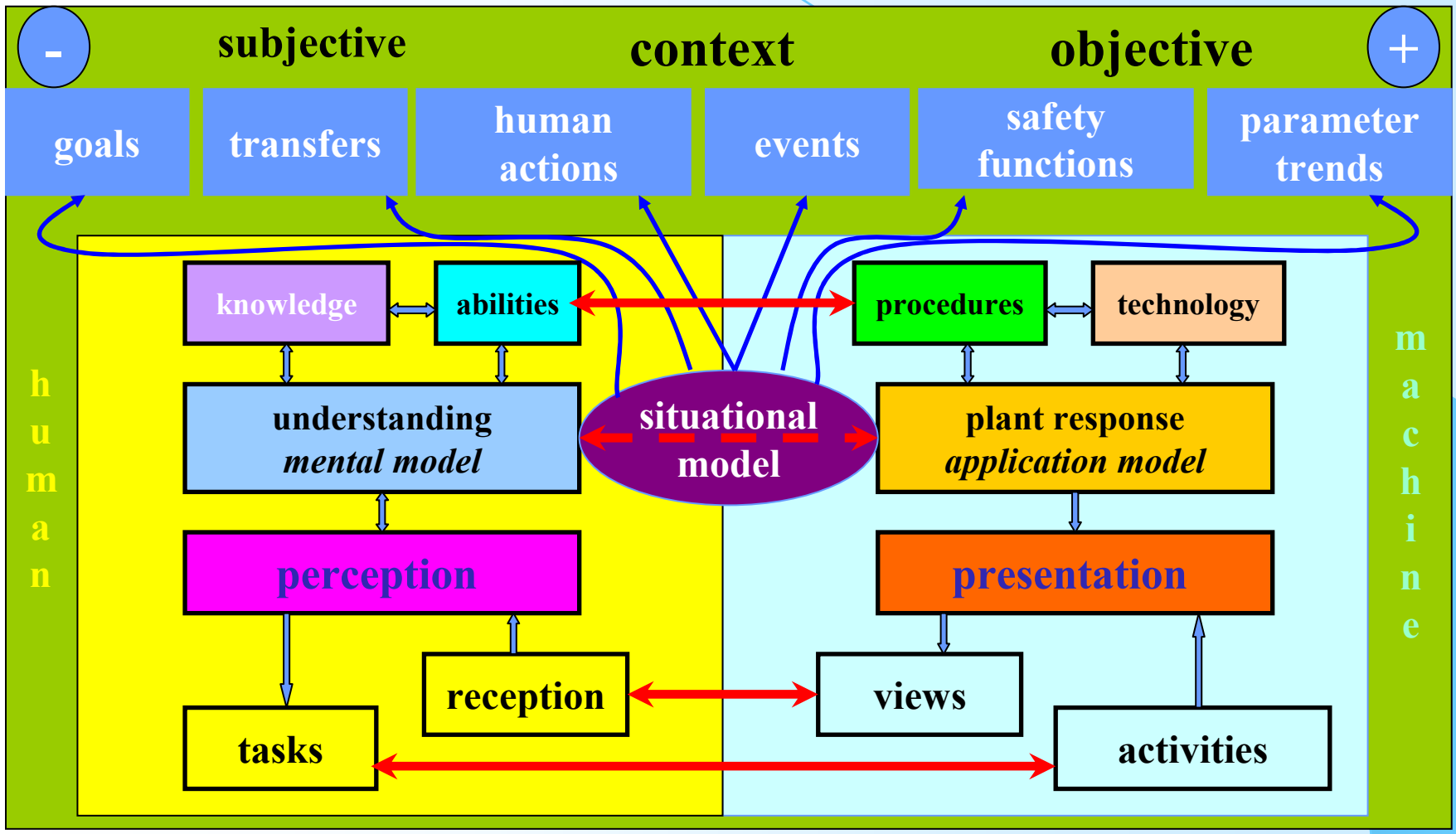
Operator Action Tree (OAT)

A	B	C	D	E	№	Failure / Error	Result
Event Occurs	Detection	Diagnosis	Response / Action	Recovery			
					1	A	Success
					2	AD	Recovered
					3	ADE	Lapse/Slip
					4	AC	Recovered
					5	ACE	Mistake
					6	AB	Non-response

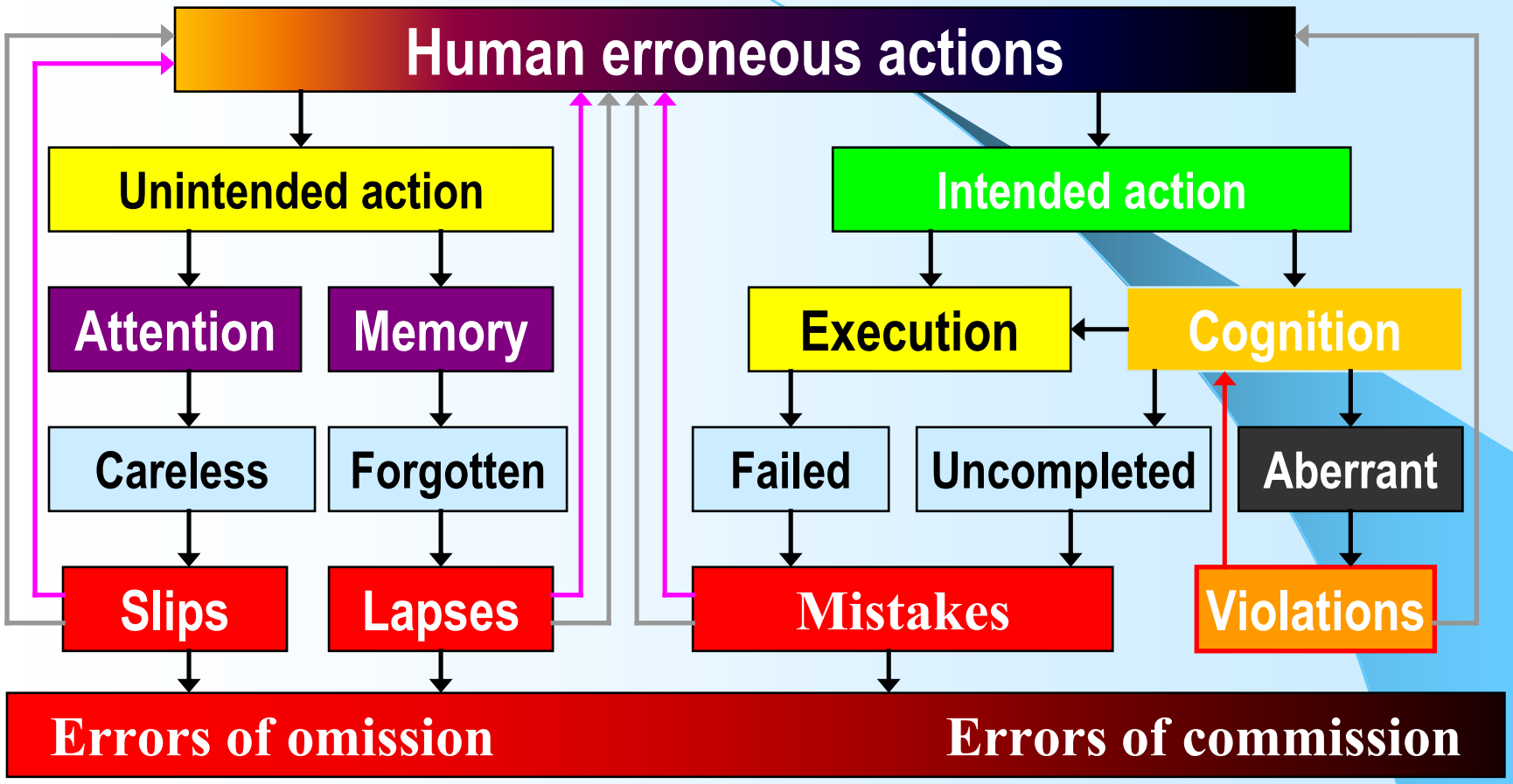
Holistic Decision Tree (HDT)



PET Context Image



Human Erroneous Action Taxonomy



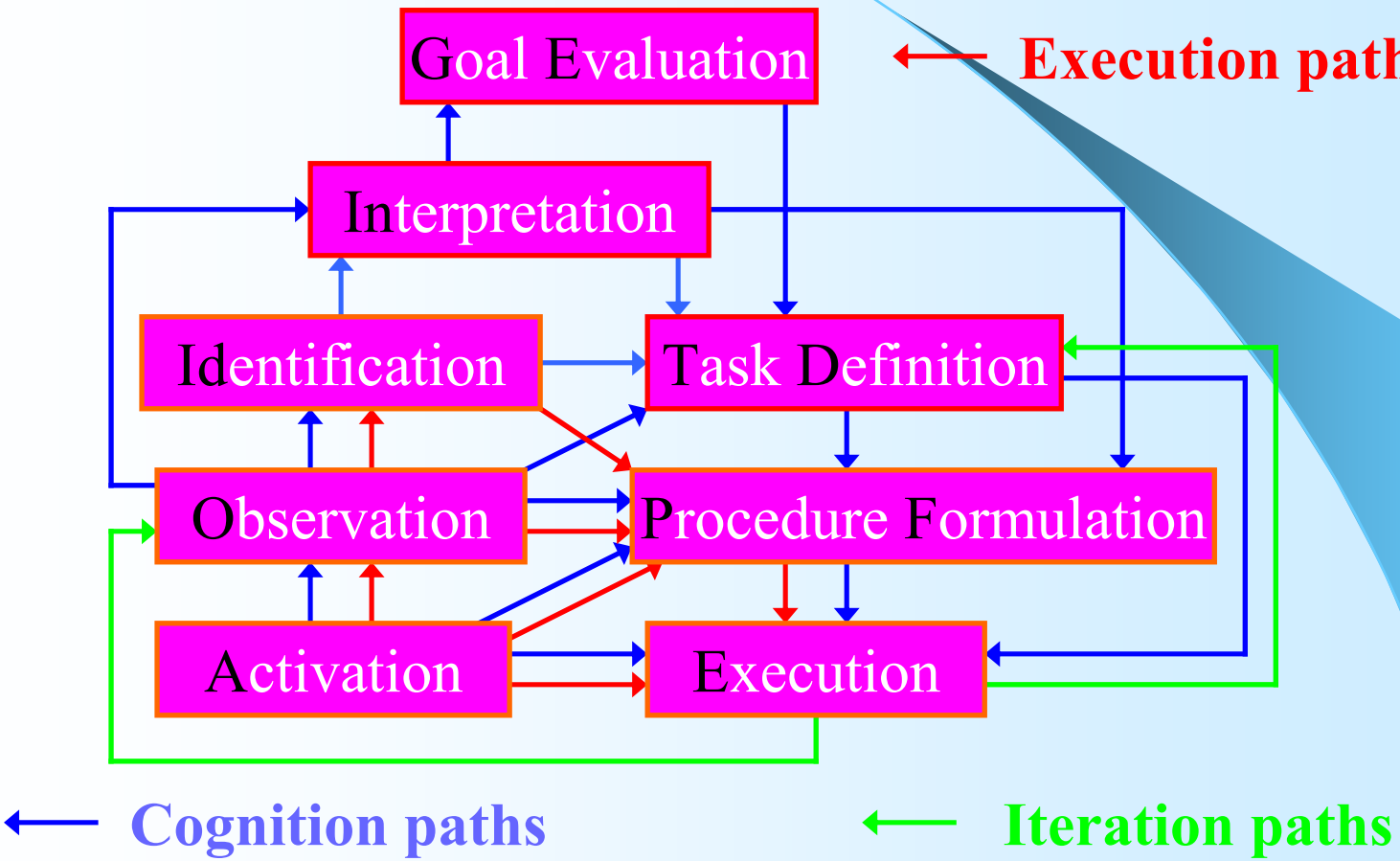
Actions / Error mode paths

Violation mode paths (short time delays)

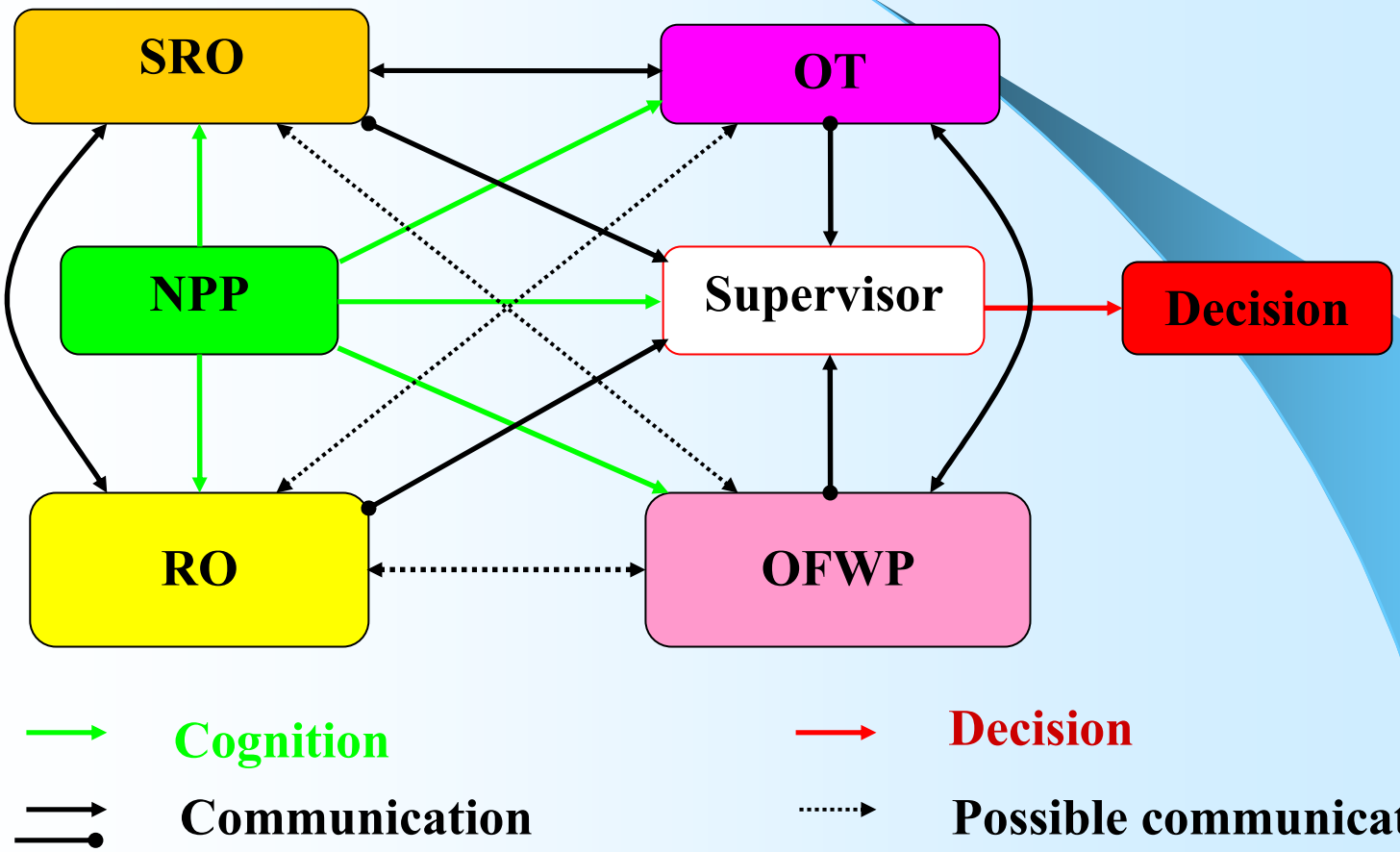
Latent error / violation mode paths (with time delays)

Recovery mode paths (with short/long time delays)

PET SLM of Individual Cognition



PET Crew Decision-Making



HRA Data Collection System

- The HRA and operators' training needs are different
- Their coordination leads to the definition of DCS processes and formation of a database for the performance assessment of operators and crews during operation, maintenance & decision-making.
- In the design of DCS there are 3 basic arguments to consider what is logical, theoretical and practical:
  - What is logically required to be collected?
  - What is the theoretical underpinning for the data needs?
  - What can be practically collected?

Extensive Data Mining Needs

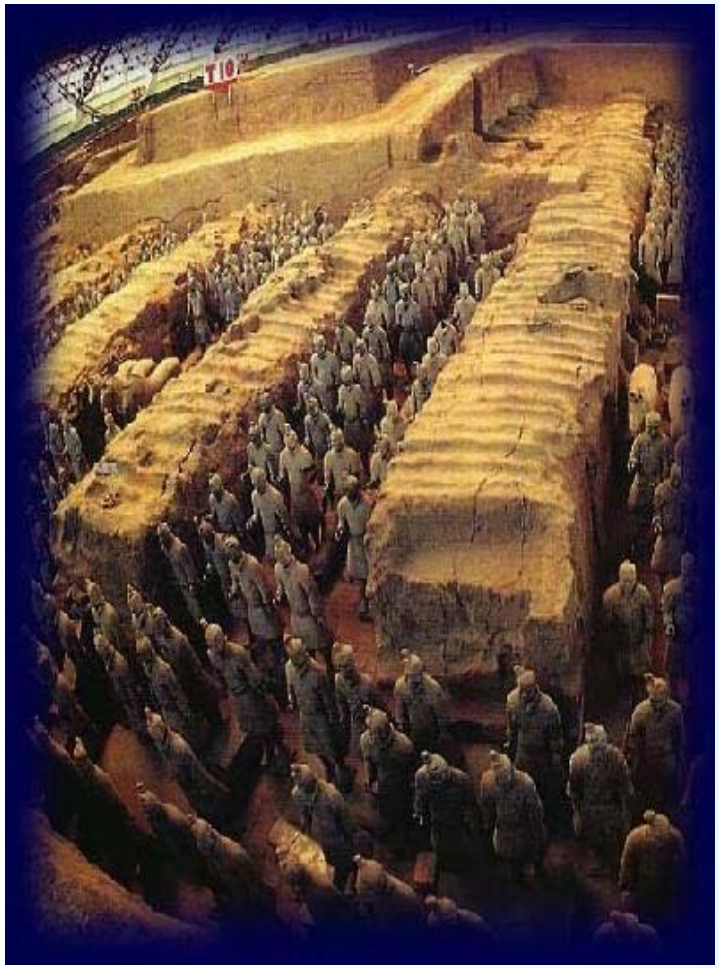
The second-generation HRA methods shift the problem from quantification of the operator behaviour and HEP to *determination of the error-forcing context* (ATHEANA) as a function of time “on a second-by-second basis” (CREAM):

- ⇒ the need for detailed and dynamic determination of operator’s performance context
- ⇒ the need for continuous monitoring and diagnostics
- ⇒ the need for automatic data mining on the control rooms or simulators
- ⇒ the need to reduce the use of experts to make judgements on the effect of context.

Measurement for HRA Methods V&V

- *Emperor Qin Shihuangdi (Cheen She hwang dee). (221-209BCE) knew that having common system for measurement would help.*
- Most DCSs record simulator response data, operator actions, alarms, etc. A few DCS systems also include the observational data made by instructors or similar individuals of the impact of communications, leadership aspects, stress during transients, etc.
- The current Full-Scope Simulators (FSSs) have progressed to the stage of laboratory tests of simulator data mining where the exercises could be extensive, specific and as representative of reality as possible (**from 'field tests' to 'full experiments'**).

Uniform Measurement Helps for V&V



Terra-Cotta Army

- *A result of measurement can be seen today in an "army" that Qin had built from terra cotta, a brownish-red clay.*
- *Thousands of life-sized statues of horses with carriages and soldiers with weapons were all made to look like a different person or animal, but built to the same size even though they were built by hundreds of different workers.*
- *Having a standard idea of how to measure sizes, the workers were able to produce the entire "army" on the same scale or matching sizes.*

DCS for the HA Context Measurement

- HRA data comparison collected by different DCS on the consecutive simulator generations could give us valuable tips on how to match high quality HRA data.
- An automated context measurement system should be designed and operated.
- If there is no idea how to determine the effect of context as a quantitative measure, then an expert opinion may have to be used to compensate for imperfections of the automatic context fixation system by qualitative descriptions.

Practical Relationships between HRA & Training

- To ‘measure’/evaluate individual & crew performance
- To improve the efficiency of education & simulator use
- To establish explicit standards for simulator training
- To establish an operators’ reliability database to predict the success or failure of the operators’ actions
- To be focused on the accident indications, displays, procedures, individual actions and group process for the particular scenario, i.e. the operators’ performance context.

Difference between HRA & Training

- The **training** tries to improve individual and crew performances based on current monitored *macro-context* (**possible manifestations**).
- The **training** relies on an analysis of engineering factors and conditions evaluation in terms understood by operators.
- The **HRA** considers systematic (applied across all crews) and individual (applied to the activities of a crew or individual personnel) effects on the *micro-context* (**probable causes**).
- The **HRA** uses probabilistic concept and more fuzzy psychological factors.

Consistency between HRA & Training

- This difference requires maintaining reasonable consistency between training & HRA terms, concepts for context determination and communication.
- It brings up the basic problems that must be solved:

1. What type of simulator data mining should be used?
2. How to use the simulator data of operators' performance assessment for training needs and for HRA purposes?

Comparison of Five HRA Techniques for Data Mining

- The data mining practice shows that the quality and applicability of the data obtained from research and plant simulators strongly depends on:
  - designed DCS,
  - planned experimental effects,
  - execution and
  - methods for data accumulating and processing.
- The complaints of a lack of suitable HRA data with psychological realism, scenario's validity, plant-specific relevance and PRA applicability are continuing.
- The question arises as to whether there is an overlapping of the basic theoretical concepts of five HRA techniques (HCR, THERP, HEART, HDT and PET) that will be used for data mining, V&V based on proposed NPP simulator studies.

Non-overlapping of the Five HRA method's concepts and capacities

Concept	THERP	HCR	HEART	HDT	PET
Human and machine as a common system	+	-	+	+	+
Influential vs. contextual approach	-	-	+/-	+/-	+
Decomposition of a task into sub-tasks and the HEP of the overall task is given by the sum of sub-tasks' HEPs	+	+	+	+	-
Accounting for differences in the environment, under which a task is performed by multipliers – PSFs.	+	+	-/+	-/+	-
Association different circumstances with modifiers (PSFs) to correct the basic TRC	-	+	-	-	-
Nominally the PSFs are the same for all crews, yet this does not explain the crew variations	+	+	+	-	-
Rasmussen's SRK concept is applicable	+	+	+	-	-
Rasmussen's SLM is applicable for DM process.	+	-	-	-	+
Usually, two iterative steps are used.	-				+

Spectrum Approach for HRA Data Mining

- Three projects for HRA data collection by the THERP, HCR, HEART, HDT and PET techniques on the computer-based, multifunctional and full-scope NPP simulators have been proposed.
- The face fidelity refers to the subjective reality as experienced by participants in simulator exercises. It is a key issue for operators' performance and context determination.
- The face fidelity is very dependent on the V&V options: basics, methods, HA types, simulators, scenarios, experts and contexts (PSFs) selected.
- For example, in training conditions the participants of course realize that they are on a simulator and it is unappropriate to assume a stress level (a V & V option) as high as in real plant operation.
- In similar case, a given V & V option can be specified only by comparison of spectrum of identical other options on different simulators where an option is an independent variable.
- Then some extrapolation is possible. It means that experimental adequacy and face fidelity can be manipulated and the results depend on the HRA technique used.

Conclusions

- A methodology based on the common use of five HRA methods for quantitative assessment of student and MCR operators and crew performances has been elaborated for three types of simulators:
- simple students' computer-based simulator of TUS,
- Kozloduy NPP multifunctional simulator (WWER-440)
- full-scope simulator (WWER-1000) of Kozloduy NPP.
- Nine normal, abnormal and emergency scenario models have been developed for simulators with different extent of details.
- They were adapted for the different type of simulators.

Thank you for your attention!