



HAL
open science

**Advanced investigation of HRA methods for
probabilistic assessment of human barriers efficiency in
complex systems for a given organisational and
environmental context**

Antonello de Galizia, Carole Duval, Emmanuel Serdet, Philippe Weber,
Christophe Simon, Benoît Iung

► **To cite this version:**

Antonello de Galizia, Carole Duval, Emmanuel Serdet, Philippe Weber, Christophe Simon, et al.. Advanced investigation of HRA methods for probabilistic assessment of human barriers efficiency in complex systems for a given organisational and environmental context. International Topical Meeting on Probabilistic Safety Assessment and Analysis, PSA 2015, Idaho National Laboratory/Idaho American Nuclear Society, Apr 2015, Sun Valley, United States. hal-01237119

HAL Id: hal-01237119

<https://hal.science/hal-01237119>

Submitted on 2 Dec 2015

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Advanced Investigation of HRA Methods for Probabilistic Assessment of Human Barriers Efficiency in Complex Systems for a given Organisational and Environmental Context

A. De Galizia¹, C. Duval¹, E. Serdet¹, P. Weber², C. Simon², B. Iung²

¹EDF R&D – Dept. of Industrial Risks Management (IRM), Clamart, France

²Research Center of Automatic Control of Nancy (CRAN) – Lorraine University, Vandoeuvre Lès Nancy, France

This paper presents the major issues concerning PSA Human Reliability Analysis (HRA) as result of the first year of a PhD in the Department of Industrial Risks Management (IRM) of Électricité de France (EDF) in the field of Integrated Risk Analysis and HRA approaches for maintenance and normal operation. In particular, we will go deep into a “state-of-the-art” of HRA methods. We will proceed to the identification of some specific analysis criteria expressly designed to compare and map methods against the criteria itself and previous works on the good practices within HRA. Such criteria deal with key issues such as the data/information/evidence required for methods to be applied, the theoretical basis underlying each method, PSF coverage (individual, operating crews or organizational) and so on. Then, we will discuss the major findings retained to provide an understanding on useful features and limitations or gaps in current HRA. Examples of these limitations are lack of an adequate interface for using qualitative analysis results for quantification of HEPs or an appropriate guidance for how to assess and use PSFs. In particular, we will focus on the problem of selecting performance influencing factors (PSFs) for the use in HRA of nominal operation and maintenance tasks in a different manner than existing methods already developed at EDF which are used for specific PSA applications (MERMOS A). Finally, we will present the conclusions and some perspectives concerning the development of a new methodology resulting from a cross-fertilization approach between a tool recently developed at EDF R&D and referred as Integrated Risk Analysis (IRA) and conventional HRA. The aim of this new methodology is to probabilistically assess human barriers efficiency in sociotechnical systems under given organizational and environmental conditions.

I. INTRODUCTION

Over the past years, technological developments have led to a decrease of accidents due to technical failures through the use of redundancy and protection. Computerized automation has been adopted in large parts of modern industrial high-risk and complex systems such as nuclear power plants, aircraft and chemical plants. However, humans still play important roles in various

parts of the design, maintenance, operation and supervision of such systems. All human activities performed in those parts are influenced by given specific working conditions or task situations, the so-called ‘context’, which is comprised of the MTO (Man, Technology and Organization) triad [1]. In human error analysis (HEA) [2] or human reliability analysis (HRA) in safety assessment, such conditions that influence human performance have been represented via several ‘context factors’. These context factors are referred to by different terms according to method: performance shaping factors (PSF), performance influencing factors (PIF), influencing factors (IF), performance affecting factors (PAF), error producing conditions (EPC), common performance conditions (CPC), and so on. The PSF or PIF are used as causes or contributors to unsafe, human actions in event analysis and also give a basis for assessing human factors in safety assessment [1], [3]. HRA has been performed as part of the probabilistic safety assessment (PSA) of large-scale systems, such as nuclear power plants. PSA is an approach that develops all the possible accident scenarios and evaluates the overall safety of a system probabilistically using the event tree (ET) and fault tree (FT) techniques. The accident scenarios are composed of two failure components, i.e. human failure events (HFEs) and hardware (system/component) failure events. HRA takes part in estimating the probability of those HFEs. There have been various approaches for evaluating human reliability. In general, those approaches can be classified into two categories, i.e. those using time–reliability correlation and those manipulating PIFs. For the methods using PIFs, some of them use a set of PIFs in adjusting the basic HEP (human error probability) such as THERP (Technique for Human Error Rate Prediction) [2], HEART (Human Error Assessment and Reduction technique) [4], CREAM (Cognitive Reliability and Error Analysis Method) [1], and others in producing HEP by rating and integrating PIFs such as SLIM (Success Likelihood Index Method) [3], etc.

In the following of this paper, the reader will be introduced to a state-of-the-art of HRA techniques through a grid of comparison, in order to highlight features of each technique and their effective applications in risk assessment. This grid of comparison is based on

the application of mapping criteria about the theoretical models methods are based on, data/information required by the method to be applied and others peculiarities characterizing each HRA technique.

II. HUMAN RELIABILITY ASSESSMENT FOR COMPLEX SYSTEMS

In response to ever changing market needs there have been a diffusion of complex industrial plants that can provide flexibility and timeliness in production. Nevertheless, because of the use of those advanced technologies specific reliability issues arose in recent years to quantify risk. For example, an important indicator is reliability itself that is defined as the probability that a system fulfill an assigned mission over time and under specific conditions. To this concept are closely related risk and workers safety that may be directly and indirectly affected by the processes in place. It is fundamental to highlight the fact that high reliable systems are not necessarily safe; in the same time, highly safe systems are not necessarily reliable. In fact, reliability and safety are different quantities and should not be confused.

In this context, it has been observed over the last decades that system failures due to human intervention are not negligible [5]; in particular, some sources report that latent human 'errors' lead to failures in the system with, in many cases, disastrous consequences for workers and the system itself.

Fortunately, in recent years, technological advances have shifted human intervention from a direct commitment to the simple manual control of automatic processes. Generally, in systems reliability studies the assessment focuses on industry process and technologies that constitute it, disregarding aspects that depend on human factors and their contribution to the same system reliability. It should be noted that human error is a major contributor to the safety and reliability of many systems: over 90% in nuclear industry [3], over 80% in chemical and petro-chemical industries [5], over 75% of marine casualties [5], and over 70% of aviation accidents [6].

For this reason, starting from high-risk industrial areas, such as nuclear, aerospace and petrochemical, a need has emerged to use common techniques of risk analysis with human factor evaluation methodologies, collected under the name of Human Reliability Analysis (HRA) methods. HRA falls within the field of human factors and has been defined as the application of relevant information about human characteristics and behavior to the design of objects, facilities and environments that people use [7]. HRA techniques may be used retrospectively, in incidents analysis, or more likely prospectively to examine a system. Most methodologies are firmly grounded in a systemic approach that sees the human contribution in the context of the wider

environmental and organizational context [8]. The purpose is to examine task, process, system or organizational structure for where weakness may lie or create a vulnerability to errors, not to find fault or apportion blame. Any system in which human error can arise, can be analyzed with HRA, which in practice, means almost any process in which humans are involved [8].

In this context, the IRM Department of EDF contributed by developing methods such as MERMOS (Méthode d'Evaluation de la Réalisation des Missions Opérateur pour la Sécurité) [9], a second-generation technique capable to determine the probability of failure of a human mission in accidental conditions, in nuclear power plants. Moreover, since 2006, for nominal and incidental conditions, the same MIR Dept. has developed an innovative approach to take into account in a risk analysis not only technical but also human, crew-related and organizational factors eventually taken under specific organizational and environmental conditions. This approach is referred to as Integrated Risk Analysis (IRA) [10].

II.A. The Integrated Risk Analysis methodology

In this paragraph we will focus on IRA approach that has been formalized and developed over several years by the same two contributors of this paper, the IRM Department of EDF R&D and the Research Center in Automatic Control of Nancy (CRAN).

For some years, improving risk analysis by utilizing new and multidisciplinary approaches has been a focus of the MIR dept. In this context, a method has been developed which is called *Integrated Risk Analysis* (IRA) [10]. It considers risks associated with different 'levels' of analysis in an integrated framework which takes into account "top-down" propagation effects existing between system organization, operating crews and technical components involved in the system. In contrast with the first generation HRA methodologies and in line with most of the second generation ones, IRA does not place its ultimate boundary of analysis at the single operator dimension. In fact, it focuses more on the identification and evaluation of relevant influences existing between organization, environment and teams. According to such a view of the global system, human action efficiency is supposed to contribute to the availability of the so-called "safety barriers". This human efficiency is evaluated by characterizing system operation at three levels: technical components, management and teams related factors (i.e. training, delegation, experience etc.), and organizational factors. Bayesian Belief Networks model these influences in order to let the analyst estimate different combinations of the factors producing risks even if these factors could be of different natures – technical, human or organizational. Although the methodology has a very

solid theoretical base and it has revealed to be reliable [10], improvements in IRA are needed to go further and enhance the evaluation of human failure probabilities under given organizational and environmental constraints. Then we will propose a new approach coupling IRA and traditional HRA. The aim is to provide reliable estimates of human barriers efficiency taking into account a specific environmental and organizational context.

In order to be able to judge if conventional HRA methods might fulfill the requirements needed by IRA, in the following section we propose to the reader a traditional state of the art on HRA methodologies.

III. OVERVIEW OF HRA METHODS

Over the years several methodologies for human reliability analysis have been developed. This has led researchers to analyze accurately the information in order to understand what could be the best approach for HRA. Information refers to costs, ease of application and analysis, availability of data, reliability, and face validity [11], [12], [13]. Developed methodologies can be distinguished into two macro-categories: first and second-generation methods. First generation methods include 35-40 methods for human reliability, many of which are variations on a single method. Theoretical basis which relates most of first-generation methods are: error classification method according to the concept "omission-commission"; definition of "performance shaping factors" (PFS); SRK cognitive model (skill-based, rule based, knowledge-based). The most accredited theory to define and classify wrong action is the error classification method according to the concept "omission-commission" [14]. This concept contains the following meanings: *omission* identifies an action that is not done, is done late, or is done in advance; *commission* is the implementation of a performance by the operator that is not required by the process. Starting from these theories, prediction models of first generation have been developed and the most representative technique is THERP [2]. Since then other views have been expressed on this concept of EOC and EOO [14].

Second generation methods, term coined by Doughty [15] try to overcome limitations of traditional methods, in particular: provide guidance on possible and probable decision paths followed by operator, using mental processes models provided by cognitive psychology; extend errors description beyond usual binary classification (omission-commission), recognizing importance of so-called "cognitive errors"; consider dynamic aspects of human-machine interaction and be used as basis for simulators development of operator performance. In order to estimate and analyze cognitive reliability, is required a suitable model of human information processing. The most popular cognitive

models are based on the following theories: 'S.O.R. Paradigm' (Stimulus-Organism-Response) argues that response is a function of stimulus and organism, thus a stimulus acts on organism which in turn generates a response; 'man as a mechanism of information processing' according to this vision, mental processes are strictly specified procedures and mental states are defined by causal relations with other sensory inputs and mental states. It is a recent theory that sees man as an information processing system (IPS); Cognitive Viewpoint: in this theory, cognition is seen as active rather than reactive; in addition, cognitive activity is defined in a cyclical mode rather than sequential mode. Starting from these theories, have been developed cognitive and contextual methods of second generation; the most representative techniques are: ATHEANA (A Technique for Human Error ANALysis) [13]; CREAM [1] and MERMOS [9].

III.A. First generation HRA methods

These tools were the first to be developed to help risk evaluators to predict and quantify the likelihood of human error. First generation approaches tend to be atomistic in nature; they encourage the evaluator to break a task into component parts and then consider the potential impact of modifying factors such as time pressure, equipment design and stress. First generation methods focus on the skill and rule base level of human actions and are often criticized for failing to consider such things as the impact of context, organizational factors and errors of commission. Despite these criticisms they are useful and many are in regular use for quantitative risk assessments.

III.B. Second generation HRA methods

The development of 'second generation' tools began in the 1990s and is on-going. Benefits of second generation over first generation approaches are yet to be established. They are in process to be validated by nuclear authorities. Kirwan [16] reports that the most notable tools of the second generation are ATHEANA, CREAM and MERMOS. Literature shows that second generation methods are generally considered to be still under improvement. Nevertheless, even in their current form, these methods can provide useful insights for human reliability issues. It is convenient to consider HRA methods on the basis of a taxonomy that has been recently produced by the IRM department for classifying human reliability techniques. We will resume this taxonomy to present methods listed in Table 2. It is a first classification step that should allow the identification of three main groups of HRA methods. Thus, we distinguish between methods that are:

- *Factorial*
- *Contextual*
- *Based on expert judgment.*

III.C. Factorial methods

These methods mainly use performing shaping factors (PSF) mostly related to the work environment and which may result in specific types of errors. Factorial methods consider that PSF have a direct impact on the task performance. Examples are time available to perform the action, man-machine interface, training, procedures, organization and complexity of the task. The effect of these factors is taken into account in the HEP quantification. Here below we describe the most representative of this class of methods, i.e. THERP.

III.C.1. THERP

THERP (Technique for Human Error Rate Prediction) began already in 1961, but the main work was done during 1970s and resulting in the so-called *THERP handbook* [2].

This is a first generation methodology which means that its procedures follow the way conventional reliability analysis models a machine or mechanical components.

The aim of this methodology is to calculate the probability of successful performance of activities needed for the realization of a task. THERP involves performing a task analysis to provide a description of performance characteristics of human tasks being analyzed. Results are represented graphically in an HRA event tree, which is a formal representation of required actions sequence.

THERP relies on a large human reliability database containing HEPs (Human Error Probabilities), which is based upon both plant data and expert judgments. The technique was the first approach in HRA to come into broad use and is still widely used in a range of applications even beyond its original nuclear setting.

THERP is probably the best known of first-generation HRA methods. it describes both how events should be modeled and how they should be quantified. Dominance of HRA event tree (ET), however, means that classification scheme and model necessarily remain limited, since ET can only account for binary choices (success-failure). A final feature of THERP is the use of performance shaping factors to complement task analysis. The use of this technique to account for non-specific influences is found in most first-generation HRA methods [12]. The separate use of performance shaping factor is relevant for a full evaluation of operator operation model. However, it suggests that the model by itself is context independent.

III.D. Contextual methods

Contextual techniques model human activity primarily using the concept of ‘EPC’ (Error Producing Conditions). These are context properties related to the history of the facility, the organization of the system, the characteristics of the interface, and they influence the nature and content of the performance of the task entrusted to the operator. They are used to identify main types of errors and propose measures to reduce them. These methods consider that the error is mainly due to the context of the activity.

III.D.1. ATHEANA

ATHEANA (A Technique for Human Error ANALysis) is both a retrospective and prospective HRA methodology developed by the US NRC in 2000. It was developed in the hope that certain types of human behavior in nuclear plants and industries, which use similar processes, could be represented in a way in which they could be more easily understood. It seeks to provide a robust psychological framework to evaluate and identify PSFs - including organizational/environmental factors - which have driven incidents involving human factors, primarily with intention of suggesting process improvement. Essentially it is a method of representing complex accident reports within a standardized structure, which may be easier to understand and communicate.

The probability of a HFE in ATHEANA, given a particular initiator, is determined by summing over different error forcing conditions associated to HEF, taking account of likelihood of unsafe actions given the EFC, and likelihood of no recovery action given the EFC and the UA. The most significant advantage of ATHEANA is that it provides a much richer and more holistic understanding of the context concerning the Human Factors known to be the cause of the incident, as compared with most first generation methods. Compared to many other HRA quantification methods, ATHEANA allows for the consideration of a much wider range of performance shaping factors and also does not require that these be treated as independent. This is important as the method seeks to identify all the interactions, which affect the weighting of the factors of their influence on a situation. In contrast some criticisms are: the method is cumbersome and requires a large team; the method is not described in sufficient detail that one could be sure that different teams would produce the same results; in the quantification method no HEP are produced [17].

III.D.2. CREAM

CREAM (Cognitive Reliability and Error Analysis Method) was developed by Eric Hollnagel in 1998 [1] following an analysis of the methods for existing HRA. It is the most widely utilized second generation HRA

technique and is based on three primary areas of work; task analysis, opportunities for reducing errors and possibility to consider human performance with regards to overall safety of a system.

This methodology is a technique used in HRA for the purposes of evaluating the probability of a human error occurring throughout the completion of a specific task. From such analyses measures can then be taken to reduce likelihood of errors occurring within a system and therefore lead to an improvement in the overall levels of safety. HRA techniques have been utilized in a range of industries including healthcare, engineering, and nuclear, transportation and business; each technique has varying uses within different disciplines.

Compared to many other methods, it takes a very different approach to modeling human reliability. There are two versions of this technique, the basic and the extended version, both of which have in common two primary features; ability to identify the importance of human performance in a given context and a helpful cognitive model and associated framework, usable for both prospective and retrospective analysis. Prospective analysis allows likely human errors to be identified while retrospective analysis quantifies errors that have already occurred. Cognition theory is included in the model through use of four basic ‘control modes’ which identify differing levels of control that an operator has in a given context and characteristics which highlight occurrence of distinct conditions.

The particular control mode determines the level of reliability that can be expected in a particular situation and this is in turn determined by collective characteristics of relevant Common Performance Conditions (CPCs).

III.D.3. MERMOS

MERMOS means ‘Assessment method for the performance of safety operation’ (once again, in French it is ‘*Méthode d’Evaluation de la Réalisation des Missions Opérateur pour la Sécurité*’) [9], [18]. It is an improvement of the EDF’s previous HRA method, and was designed initially to guide EDF’s analysts in taking human factor into account in the “level 1” PRA for the N4 French series nuclear power plants. In the methodology a “human factor mission” constitutes the interface between PRA and HRA. For each initiator, a functional analysis will determine the “missions” that have to be performed to recover or mitigate the accident. Failures of one mission or several consecutive missions lead to unacceptable consequences. Among the missions, Human Factor missions (HF mission) refer to safety critical actions that the operating system, comprising the crew interacting with the procedures, the systems, the organization and the layout, has to initiate and carry out to handle the situation. One of the purpose of HRA is to assess the failure of post-initiator HF mission. MERMOS considers that the

performance of HF mission is the responsibility of a system called “emergency operations system” (EOS) [19], [20]. Strategy, Action, Diagnostic (SAD) are the three functions involved in the performance of HF missions assumed by the EOS [21]. The actual functioning of the system is modeled with the help of a new concept, named CICA (“Important Characteristics of Emergency Operation”). CICA refer to particular ways of operating the plant adopted by the EOS in the course of the emergency situation [20]. The aim of the MERMOS qualitative analysis is to identify plausible scenarios referring to SAD model leading to the HF mission failure.

III.E. Methods based on experts’ judgment

Methods that belong to this group focus on determining error probabilities from estimates of expert judgment. The reliability of estimates is highly dependent on the complexity of the situations analyzed, on the experts selected and on how judgments are aggregated to produce estimations.

The most representative for this category are APJ, SLIM, HORAAM and PC. We do not go deep into details in describing these last ones because of the less usage that have been done in HRA.

More relevant conclusions on HRA methodologies require a deeper understanding defining some key criteria to compare methods. We will present these specific criteria in the next paragraph.

IV. MAPPING CRITERIA TO COMPARE HRA METHODS

In the preceding paragraphs we have provided the elements to understand the way first and second generation methods operate and how they are used: theoretical approaches they are based on, the temporal evolution of HRA techniques and finally some details such as the development context, the estimation procedure of HEPS, validation status and application areas. Although this approach may provide an overview on how human reliability has been assessed in the last decades, we wanted to continue our analysis by a further analysis. Thus, analytical criteria were developed to evaluate and compare the methods in analysis and conclude brings their specific needs.

In this paper, we address the so-called socio-technical systems where the consideration of human actions and the environment is an essential part of the process of risk management. In the nuclear industry, as in other areas, the assessment of human reliability methods have been implemented and integrated with risk analysis but it is still difficult to identify exactly why the application of some rather than others. The identified criteria seek to meet this need. They briefly presented in Table 1 below.

Table 1
Criteria identified for selecting HRA methods for use.

<i>Name</i>	<i>Taxonomy</i>	<i>Description</i>
Main theoretical framework	▪ Behaviorist	Theoretical basis and assumptions to model human behavior for both quantitative and qualitative phase of analysis.
	▪ Cognitive	
	▪ Situated Cognition ¹	
Origins of Data	▪ Inputs	It distinguish between data already existing in the model and input data coming from other sources: feedback, test results, simulator investigations, , nominal error probabilities etc.
	▪ Predefined	
Nature of Data	▪ Quantitative	We refer to the method being used utilize data having a nature mostly qualitative (nuns, etc.) or a quantitative one (numbers, failure rates, etc.).
	▪ Qualitative – Ordinal ²	
	▪ Qualitative – Nominal ³	
Processing Approach	▪ Frequentist	It refers to the treatment for the quantitative assessment in processing data. It depends on the type of data processed.
	▪ Bayesian	
Analysis Target	▪ Operator	This is the domain of operability of the method.
	▪ Operating Crew	
	▪ System	
Coverage of PSFs	▪ Human	Human : personal characteristics and working capabilities of humans operator Task : procedures and task features required for the operator System: MMI, hardware systems, and physical features of plant process Environment: team and organizational factors, and physical working environment.
	▪ Task	
	▪ System	
	▪ Environment	

¹ Situated cognition theorists suggest that knowing/learning cannot be separated from social, cultural and physical contexts.

² Data have an ordinal (or ordered) qualitative nature if they naturally have an order, or can be arranged along a scale (for example, the ‘poor’, ‘good’ and ‘excellent’ attributes or days in a week).

³ Data have a nominal qualitative nature if they do not have a natural order (e.g., diseases or eyes color).

These criteria have been applied to compare preselected HRA methods that resulted particularly appropriate to nuclear power PRA. They are listed in Table 2 below.

Table 2
HRA methods selected for use in nuclear safety studies.

<i>Method</i>	<i>References</i>	<i>Date</i>
THERP*	Swain & Guttman – Rapport WASH1400/ NUREG 75/014	1983
SLIM*	Embrey & Kirwan – NRC NUREG/CR-3518	1984
HCR*	Hannaman & Spurgin – EPRI RP 2170-3	1984
TRIPOD-DELTA*	SHELL	1985
HEART*	Williams – NRC	1988
CREAM**	E. Hollnagel – Halden (Norway)	1994
ATHEANA**	Cooper et al. – NRC NUREG/CR-6350	1996
MERMOS**	Le Bot et al. – EDF	1998
SPAR-H**	Gertman et al. – NRC NUREG/CR-6883	1999
NARA***	British Energy (UK)	2005

* First generation
** Second generation
*** Third generation [16]

Considering a consistent review of existing literature on the argument and basing on expertise in MIR dept. on the field of human reliability and human factors, we proceed to the construction of a benchmark grid aimed to compare method in Table 2. This grid of analysis is shown in Table 3, which summarizes the results of applying criteria to the HRA methods. Accordingly to the benchmark grid, it should be noted that all methods use models and other knowledge as the underlying basis for how they approximate the realities of human performance. In addition, all use assumptions and other judgments that, given the current state of the art in HRA, still need to be supported with appropriate data. Some bases for some methods are weaker than others and, with the continued advances and expected evolution in HRA methodology, it is expected that some methods will become less used while others, or even new methods become more prevalent.

Table 3
Benchmarking grid comparing HRA methods of Table 2.

Method	Main theoretical framework	Predefined Data	Nature of Data	Processing Approach	Analysis Target	Coverage of PSF
THERP	Behavioral	TRC curves and tables	Quantitative	Frequentist	Operator	Human Task
SLIM-MAUD	Behavioral	PSF coefficients	Quantitative	Bayesian	Operator	Human Task
HCR/ORE	Cognitive	Curves	Quantitative	Frequentist	Operator	Human Task
HEART	Cognitive	Tables	Quantitative	Frequentist	Operator & Crew	Human Task
CREAM	Cognitive	Nominal HEP tables	Quantitative	Frequentist	Operator & Crew	Human Task System
ATHEANA	Behavioral & Cognitive	Nominal HEP	Quantitative Qualitative	Frequentist	System	Human Task System Environment
MERMOS	Situated Cognition	Inputs (simulator investigations, scale for expert judgment elicitation)	Qualitative ordinal	Bayesian	System	/
SPAR-H	Cognitive	Nominal HEP	Quantitative	Frequentist	System	Task System
NARA	Cognitive	Nominal HEP (Tables) NARA database	Quantitative	Frequentist	System	Task System Environment
TRIPOD-Delta	Behavioral	Expert judgments	Qualitative nominal	Bayesian	Operator & Crew	Human Task System

This does not suggest that current methods cannot be used successfully in the sense that for many applications, reasonable estimates of HEPs can be obtained and potential problem areas can be identified. In fact, for the risk informed decisions that need to be made, there have been successful uses of PRA and HRA for general risk-assessments of operating plants and for applications such as ranking components for the maintenance rules, changing technical specifications, and performing evaluations in the risk significance of activities concerned by industrial processes.

V. MAJORS FINDINGS – CONBINING CONCLUSIONS FROM THE BENCHMARK GRID ANALYSIS WITH THE LESSONS LEARNED FROM HRA LITERATURE REVIEW

The comparison of the methods against the analysis criteria and previous work on the good practices within HRA provides an understanding on useful features and limitations or gaps in current HRA methods. This knowledge was expanded upon with the lessons learned from more and less recent international reports on advancements on HRA [11], [13], [22], [23], [31]. This paragraph aim to verify and extend the insights resulting from analyzing bibliographic works and reviews on HRA in combination with the described above t benchmarking in Table 3. This study permitted to gain further information about strengths and weaknesses of HRA methods and practices and identified holes that require to be filled in HRA. The major findings are discussed here below.

V.A. Theoretical basis about human information processing

First of all, we identified for the larger part of the methods above limitations in modeling and quantifying human performance under ‘various’ conditions [31]. This effect can be partially attributed to the lack of an adequate underlying theoretical basis to guide the analysis, particularly with respect to cognitive activities [11] associated with more challenging situations. The assumptions about how operators can fail and why when performing a specific task are made on the basis of analysts’ understanding of plant and human behavior. HRA methods provide a technical basis for determining human performance issues and developing assumptions about how and why teams may not accomplish a safety action [24]. This could be the reason why they are mostly expressed in terms of performing shaping factors (PSFs), which ultimately are used in the estimation of HEPs. Empirical studies [11] show that deficiencies in these theoretical models impact analyst capability to

appropriately characterize the tasks analyzed and the associated PSFs, limit the development of a good operational understanding of the human information processing, and finally can have a large effect on the HEP.

V.B. Relation between qualitative analysis and quantification

Many newer methods focus on identifying failure mechanisms, including contextual factors that drive or cause them (e.g. ATHEANA [25], or MERMOS [18]) and these methods generally produce a superior qualitative analysis (richer in content and better operational stories). However, superior qualitative analysis itself does not necessarily produce more reasonable HEPs [25]. Therefore, a good relation between the qualitative analysis and the quantitative analysis is needed. Most methods have inadequate guidance on how to use the information from qualitative analysis to determine HEPs (i.e., translating the information into the inputs to the quantification of HEPs). That is, even when the analyst went beyond the guidance provided by a given method for performing the qualitative analysis, it was often difficult to make an efficient and consistent use of information and qualitative results.

V.C. Range of PSFs covered

Most methods do not seem to cover an adequate range of PSFs or influencing factors in attempting to predict operating crew performance for all circumstances [26]. That is, important aspects of accident scenarios were not always captured by the factors – especially for those that considered only individual-based PSF – considered by given first generation methods.

V.D. Judging PSFs influences and choosing the right PSFs

Looking across methods (similar and different), there are inconsistent judgments about which PSFs (e.g., high vs. low workload, adequacy of indications) are important and how strongly PSFs affect HEPs in a given situation. The methods do not provide adequate guidance for these judgments [26].

V.E. Team variability

Dealing with team variability in HRA is still a difficult issue [23]. The objective of HRA is to model/assess average performance and many methods (e.g., SPAR-H, HEART) are designed to evaluate “average” team performance. While detailed contextual methods like

ATHEANA [25] and MERMOS [9] can in principle address this team variability issue, it is difficult to observe enough crews in enough situations to be able to make reasonable inferences about systematic effects for a prospective analysis for use in a PRA. How to address crew variability remains an outstanding issue in HRA. In light of these limitations, it is possible to conclude that none of the HRA methods discussed may support and feed the IRA technique in order to accomplish its objectives. A cross-fertilization approach is then proposed to overcome such limitations and improve both HRA and IRA.

VI. CROSS-FERTILIZATION APPROACH BETWEEN HRA AND INTEGRATED RISK ANALYSIS

Limitations of the various methods for HRA were identified and the major findings have been discussed above. Therefore, it became apparent that newer methodologies should go much further to overcome these gaps and improve conventional HRA practice.

The goal of a cross-fertilization approach between the IRA technique and traditional HRA has been considered in order to develop a tool that could be applied with minimal adaptation to address a wide range of HRA domains and situations. In particular, the aim is to improve the probabilistic assessment for human actions efficiency in risk-informed decision-making. Especially, the major challenge, i.e. the calculation of the human efficiency probability for a specific task, has been tackled starting by incorporating organizational and team related PSFs upon the existing structure defining a human barrier model in IRA as well as by a deep study of the state of the art on PSFs from conventional HRA [26], [27].

As suggested from many recent studies on HRA good practices, we are convinced that team and organizational factors are the roots of human efficiency in performing tasks in nominal conditions (maintenance activities) as well as in pre-incident situations. For this reason we searched for an orthogonal and proper PSFs set, which are assumed to reduce overlaps between IRA predefined factors while still permitting previously established dependencies to exist. As described in [29], [30], a Bayesian Network based approach has been used till today in IRA methodology for the estimation of the causal dependencies among factors and leading to the target of analyses, i.e. the human efficiency probability for tasks performed in nominal and pre-initiator situations.

There is substantial evidence that methods focusing on the identification of failure mechanisms (ways the crews could fail a particular task) and the contextual factors that enable them (e.g., ATHEANA and MERMOS), tend to produce richer content in the qualitative analysis than the PSF-based methods (e.g., SPAR-H, THERP, CREAM

and HEART) and the resulting operational stories reflected a more detailed prediction of what could or would occur in responding to the scenario [11]. However, richer operational stories did not necessarily lead to more accurate HEPs, so other factors are involved, e.g., reliable processes and associated guidance for translating the richer information into HEPs [27]. Nevertheless, it seemed clear that across the variety of possible conditions that can occur in an accident scenario, a thorough assessment of failure mechanisms and organizational context will be needed for more reliable results (but see discussion of PSFs below).

This is the reason why we consider that a cross-fertilization approach could fill many gaps existing on both sides, i.e. IRA technique and existing HRA methods at the IRM dept. [30]. Considering requirements to be fit for increasing IRA completeness, candidate PSFs that resulted from covering these requirements have been extracted by doing a full PSF benchmark related to the HRA methods presented in Table 2 and those situational characteristics and human factors relevant to IRA domain of application. Through an iterative process, a set of PSF composed of 12 representative PSFs and their sub-items have been defined and we presented it in Table 4. The proposed taxonomy is a straightforward consequence of what we said above, regarding the importance of organizational and management related factors. PSFs set will need a verification process in parallel with further developments.

As IRA domain of application consists in nominal situations (including maintenance activities) and specific incidental ones in nuclear power plants (support systems) and under given specific boundary conditions, i.e. organizational and environmental contexts, PSF were properly identified to tackle these kinds of situations in a more efficient and reliable way.

A schematic representation of a *human barrier* resulting from first steps in combining IRA model and conventional HRA for nominal and incidental conditions after incorporating PSF is shown in Figure 1. As IRA bases its human efficiency analysis on the idea that major hazardous situations mostly come from organizational factors [10], dependencies among factors are intended in a “top-down” direction, [27], which means that the causal relationships are established starting from root factors in the model, i.e. the OFs, and weighing this influence by evaluating the impact of OFs PSFs.

Further developments are needed to extend the model: definition of probability distributions on the root variables; degree of discretization on values assumed by each PSF, depending on the specific action considered in the analysis; and so on. It should be noted that OFs are intended as *neutral factors* that could have a negative as well positive impact on the PSFs. By this way, OFs can be described as discrete variables having n values associated to different consequences at the PSFs level.

Table 4
Candidate PSF selected for the IRA model.

Group	PSF	Description
TEAM-based Factors	<i>Cohesiveness</i>	Frequency with which the crew is trained or performs duty as a team. Mutual assistant behavior. Willingness to sacrifice the right judgment for maintaining team cohesiveness
	<i>Coordination</i>	The level at which different individual roles and responsibilities (including backup responsibilities) are clarified to each team member.
	<i>Composition</i>	Team size, homogeneity/heterogeneity, and compatibility. The team stability or turnover rate (For example, if too many members are replaced, particularly when replaced by less skilled members, team performance is likely to degrade
	<i>Communication Availability</i>	The sufficiency and availability of communication means for use. The physical proximity of operators and communication equipment (if any). Failure or functional unavailability of communication system (e.g., signal jam due to a high volume of communication signals). The likelihood that communication equipment is unmanned.
	<i>Communication Quality</i>	Distorted signal or degradation of the equipment due to equipment faults (e.g., old or poorly maintained equipment).Human fault (e.g., heavy accent, linguistic ambiguity, and unclear instruction).
	<i>Leadership</i>	Ascendancy, rank, experience, and reliability of the decision maker. Leadership quality, comprising of three elements: direction-setting, gaining followers' commitment, and overcoming obstacles to change situation. The level of commitment could be measured in three dimensions: identification with work, identification with co-workers, and identification with the organization
MANAGEMENT Factors	<i>Work Process Design, Tasking & Direction</i>	Work sequence/schedule. Completeness and correctness of task description Timeliness of work assignment
	<i>Human-System Interface</i>	Quality of system design for ease and accuracy of visual, audio, and cognitive information perception. Appropriateness of workload distribution between automation and operator manual controls.
	<i>Safety & Quality Culture</i>	Policy (i.e., clear emphases on safety/quality policy). Senior management commitment to safety/quality. Response and commitment of individuals. Violations and errors recorded in operation log Investigation of accidents or near-miss events
	<i>Work Environment (physical)</i>	Workplace habitability (e.g., illumination, temperature, humidity, vibration, noise) Sufficient work space
	<i>Special Tool Availability</i>	Needed tools are available, well organized, and accessible
	<i>Special Tool Adequacy and Quality</i>	Availability of specially designed tools for certain tasks
	<i>Procedure Availability</i>	Existence and accessibility of procedures Content accessibility (e.g., document indexing)
	<i>Procedures quality and adequacy</i>	Document fidelity (e.g., adequacy of the level of detail, completeness, and correspondence of procedures to actual tasks) legibility and readability (e.g., page layout). Usability (e.g., provision for check-listing) Easiness for distinguishing different procedures

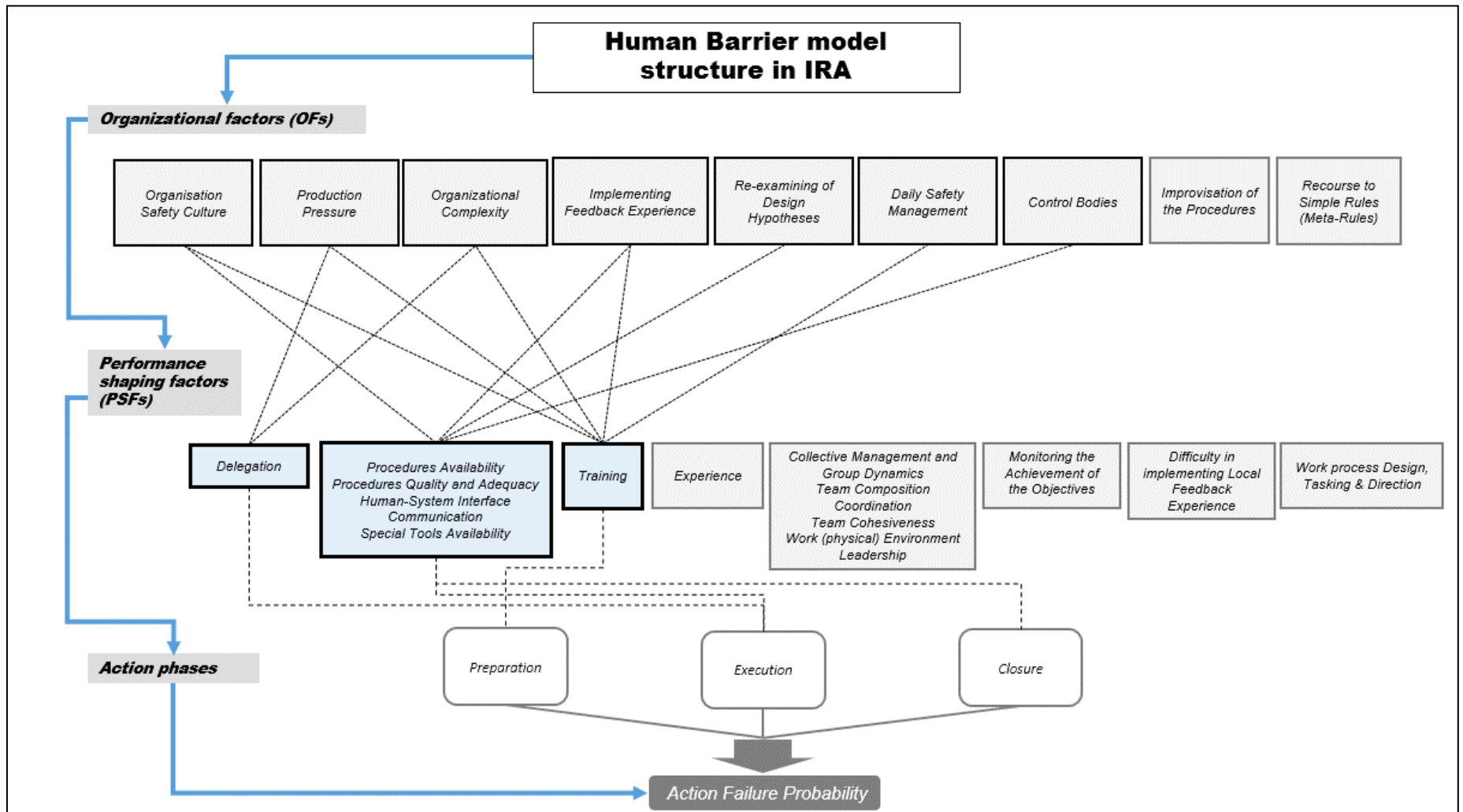


Figure 1. The human barrier model based on OF, PSF and phases and an example of causal links between factors resulting in a specific human action failure probability.

VII. CONCLUSIONS

In this paper we wanted to propose a state-of-the-art on current HRA, with some focus on the most representative methods being used in nuclear PRA. Then, we have discussed selected HRA methods and compared them by using specific criteria of analysis in order to catch and highlight limitations and gaps to be overcome. The benchmark grid built applying these criteria allowed us drawing some conclusions and major findings for further developments.

In the last part, we presented a cross-fertilization approach between the Integrated Risk Analysis (IRA) methodology which have been developed at EDF R&D and conventional HRA techniques. In particular, we proposed a complete PSF set to be integrated in the existing IRA framework. Finally, we presented the schematic model of a human barrier based on this new method which combines different approaches coming from conventional HRA and the IRA methodology. In the future some results will be presented associated to a simplified case study from nuclear industry.

REFERENCES

- [1] E. Hollnagel, *Cognitive Reliability and Error Analysis Method (CREAM)*, Elsevier, Ed., 1998.
- [2] Swain and H. Guttman, "Handbook of human reliability analysis with emphasis on nuclear power plant applications", NUAREG-NRC, Washington, DC, NRC 83, 1983.
- [3] D. Embrey, "Application of human reliability assessment techniques to process plant design," *Ergonomics Problems in Process Operations*, 1984, Pages 65-80, pp. 65-80, 1984
- [4] J. Williams, "A data-based method for assessing and reducing human error to improve operational performance," in *Proceedings of the IEEE Fourth Conference on Human Factors and Power Plants*. Monterey, California : IEEE, New York, 5-9 June 1998, pp. 436-450.
- [5] J. Reason, "The contribution of latent human failures to the breakdown of complex systems," *Philosophical Transactions of the Royal Society of London*, vol. B327, no. 1241, pp. 475-484, 1990.
- [6] R. Helmreich, "On error management: lessons from aviation," *British Medical Journal*, vol. 320, no. 7237, pp. 781-785, 2000.
- [7] P. Humphreys, "Human reliability assessor's guide," *Human Factors in Reliability Group*, SRD Association., Tech. Rep., 1995.
- [8] S. Kariuki and K. Lowe, "Integrating human factors into process analyses," *Reliability Engineering & System Safety*, vol. 92, pp. 1764- 1773, 2007.
- [9] P. Le Bot, and H. Pesme, "MERMOS: an extended second generation HRA method" *IEEE / HPRCT*, Monterey CA, USA, 26-31 Aug. 2007.
- [10] C. Duval, G. Fallet, B. Iung, P. Weber, and E. Levrat, "A Bayesian network-based integrated risk analysis (IRA) approach for industrial systems: Application to heat sink system and prospects development," *Proceedings of the Institution of Mechanical Engineers, Part O: Journal of Risk and Reliability*, vol. 226, no. 5, 2012, pp. 488-507.
- [11] E., Dang V.N., Forester J., Broberg H., Massaiu S., Hildebrandt M., Braarud P.O, Parry G., Julius J., Boring R., Mannisto I., Bye A. "International HRA Empirical Study-Phase 1 Report", NUREG/IA-0216, November 2009.
- [12] Swain, "Comparative evaluation of methods for human reliability analysis," Tech. Rep. GRS-71, 1989.
- [13] J. Bell and J. Holroyd, "Review of human reliability assessment methods". Health and Safety Executive, *Health and Safety Laboratory*. Harpur hill. buxton, derbyshire. sk179jn," 2009.
- [14] Le Bot, "How to deal with commission and omission errors in HRA", ESREL, Prague, 2009.
- [15] E. Doughty, "Human reliability analysis - where shouldst thou turn?" *Reliability Engineering & System Safety - RELIAB ENG SYST SAFETY* , vol. 29, no. 3, pp. 283-299, 1990.
- [16] B. Kirwan, H. Gibson, R. Kennedy, J. Edmunds, G. Cooksley, and I. Umbers, "Nuclear action reliability assessment (NARA): A data-based HRA tool," *Safety & Reliability Journal, Safety and Reliability Society*, Manchester., 2005.
- [17] U. N. R. Commission, "HRA method review - ASEP in expert workshop on the evaluation of HRA methods with respect to HRA good practices," Rockville, Maryland, NUREG 1792, June 27-28 2005.
- [18] P. Le Bot and E. Serdet, "MERMOS-A: A new method to analyze pre-initiator for specific HRA", *PSAM11 & ESREL 2012 congress*, Helsinki, Finland, 25-29 June 2012.
- [19] P. Meyer, P. Le Bot, and H. Pesme, "MERMOS: an extended second generation HRA method," *IEEE / HPRCT*, Monterey CA, USA, 26-31 Aug. 2007.
- [20] P. Le Bot, "The Safe Regulation Model: A Grid to describe EOS (Emergency Operating System)", *Workshop ISV (Integrated System Validation)*, Halden, 2009.
- [21] S. Massaiu, P. Braarud, and P. Le Bot, "Including organizational and teamwork factors in HRA: the EOS approach", *EHPG 2013 Enlarged Halden Program Group. Man-Technology Sessions*. IFE, 2013.
- [22] M. Lyons, A. Sally, M. Woloshynowych, and C. Vincent, "Human reliability analysis in healthcare: A review of technique," *International Journal of Risk & Safety on Medicine*, 2004.
- [23] Mosleh, R. Boring, D. Gertman, and Y. Chang, "Human reliability analysis methods. Selection guidance for NASA," NASA, Tech. Rep., July 2006.
- [24] Kirwan, "The validation of three human reliability quantification techniques - THERP, HEART, JHEDI: Part i - technique descriptions and validation issues," *Applied Ergonomics*, vol. 27, no. 6, pp. 359-373, 1996.
- [25] U.S. Nuclear Regulatory Commission (USNRC). "Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA)". NUREG- 1624. Division of Risk Analysis and Applications. Office of Nuclear Regulatory Research, Washington DC. May 2000.
- [26] R. Boring, "How many performance shaping factors are necessary for human reliability analysis?" *PSAM 10*. Idaho Falls, Idaho, USA. : Idaho National Laboratory, June 2010.
- [27] Z. Mohaghegh, A. Mosleh, "Incorporating organizational factors into probabilistic risk assessment of complex socio-technical systems: Principles and theoretical foundations", *Safety Science*, Elsevier - Volume 47, Issue 8, October 2009, Pages 1139-1158.

- [28] C. Simon, P. Weber, and A. Evsukoff, "Bayesian network inference algorithm to implement Dempster Shafer theory in reliability analysis," *Reliability Engineering & System Safety*, vol. 93, no. 7, pp. 950–963, 2008.
- [29] P. Weber and L. Jouffe, "Complex system reliability modelling with dynamic object oriented bayesian networks (doobn)," *Reliability Engineering & System Safety*, vol. 91, no. 2, pp. 149–62, 2006
- [30] P. Le Bot, H. Pesme, and F. Ruiz, "Methodological validation of MERMOS by 160 analyses," *PSA'02 International Topical meeting on Probabilistic Safety Assessment*. Detroit, Michigan, October 2002.
- [31] J. Forester, A. Kolaczowski, E. Lois and D. Kelly, "Evaluation of Human Reliability Analysis methods against Good Practices, NUREG-1842, Washington, 2006.
- [32] E. Hollnagel, J. Pariès, D. D. Woods, and J. Wreathall, "Resilience Engineering Perspectives Volume 3: Resilience Engineering in Practice". Farnham, UK: Ashgate.E, 2011.