Available online at www.sciencedirect.com
SCIENCE DIRECT®
Reliability Engineering and System Safety 83 (2004) 229–240
RELIABILITY ENGINEERING & SYSTEM SAFETY
www.elsevier.com/locate/ress

Human error risk management for engineering systems: a methodology for design, safety assessment, accident investigation and training

P.C. Cacciabue*

Institute for the Protection and Security of the Citizen, European Commission—Joint Research Centre, Via E. Fermi, 1, I-21020 Ispra, Varese, Italy

Abstract

The objective of this paper is to tackle methodological issues associated with the inclusion of cognitive and dynamic considerations into Human Reliability methods. A methodology called Human Error Risk Management for Engineering Systems is presented that offers a 'roadmap' for selecting and consistently applying Human Factors approaches in different areas of application and contains also a 'body' of possible methods and techniques of its own. Two types of possible application are discussed to demonstrate practical applications of the methodology. Specific attention is dedicated to the issue of data collection and definition from specific field assessment.

Keywords: Design; Safety assessment of human–machine systems; Human reliability assessment; Human error management

1. Introduction

The need to include Human Factors (HF) considerations in the design and safety assessment processes of technological systems is nowadays widely recognised by almost all stakeholders of technology, from end-users to providers and regulatory bodies.

The critical role assigned to HF in design and safety assessment depends on the widespread use of automation and its impact on human errors. Automation improves performance of routine operations, aims at reducing workload, and successfully limits most human blunders at behavioural level. However, automation introduces a variety of safety critical issues due to 'errors' of cognition, which are particularly correlated and strongly affected by socio-technical contextual conditions, e.g. training and experience, physical working environment, teamwork, etc. These factors are crucial for ensuring safety and user friendliness of 'advanced' technology system [1].

Another relevant factor associated with HF in the design and safety assessment processes is that it is impossible to conceive a plant that is totally 'human-error free', as this is an intrinsic characteristic of any technological system. Therefore, the improvement of the safety level of a system can only be achieved through the implementation of appropriate measures that exploit the enormous power of both human skills and automation potential for preventing or recovering from human errors, and for mitigating the consequences of those errors that still occur and cannot be recovered. This represents a process of Human Error Management (HEM).

Over the last 25 years, the scope of Quantitative Risk Assessment (QRA) has vastly changed, progressively expanding its bearing to areas such as safety management, regulation development, and design [2]. However, the requirements and specifications of new areas and domains of application require that new methods and techniques are developed. This is particularly true for HF, which suffers from the major 'bottleneck' of providing numerical measures of the likelihood of certain events and of their associated consequences.

In order to provide a substantial contribution to QRA from the human reliability side, a variety of methods have been developed during the 1970s and 1980s. The most relevant technique and complete framework, developed in those years, for inclusion of human contribution to QRA is certainly the Technique for Human Error Rate Prediction (THERP) [3,4]. Other methods, based on the same principles, and focused on slightly different issues have been developed in the same period. All these methods are essentially focused on behavioural aspects of human performance and may be considered as 'first generation' methods of Human Risk Assessment (HRA) [5]. They are well suited for supporting basic or generic QRAs, as they

* Tel.: +39-322-78-9869; fax: +39-322-78-5813.
E-mail address: pietro.cacciabue@jrc.it (P.C. Cacciabue).

0951-8320/$ - see front matter © 2003 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2003.09.013

provide the probabilities of human errors and thus fulfil the primary requirement of reliability analysis.

However, every first generation method focuses strongly towards quantification, in terms of success/failure of action performance, with lesser attention paid to in-depth *causes* and *reasons* of observable human behaviour. In other words, first generation methods suffer from the drawback of being substantially unable to explain *why* humans did what they did and *what* triggered inappropriate behaviour. These methods neglect the cognitive processes that underlay human performances, which are instead essential for safety analyses of modern Human Machine Systems (HMS). This inability to capture root causes and understand causal paths to errors, makes it impossible for first generation methods to provide any contribution to recovery and mitigation aspects typical of HEM, which is instead a crucial component of safety assessment, as discussed above.

Another important aspect is the dependence of human errors on the dynamic evolution of incidents. From this viewpoint, the overall picture of first generation methods is even less encouraging. Only few methods make some reference to dynamic aspects of human-machine interaction, while the static dependencies are well considered in almost all approaches.

These issues are the driving subjects of most methods developed in recent years and justify their grouping into 'second-generation' techniques [5], even though not all techniques tackle these three aspects at the same time. A critical bottleneck is common to all second-generation methods: the existence and availability of adequate supporting data.

In this paper, a methodological framework of reference, called Human Error Risk Management for Engineering Systems (HERMES), will be presented. HERMES offers:

- A 'roadmap' for selecting and applying coherently and consistently the most appropriate HF approaches and methods for specific problem at hand, including the identification and definition of data; and
- A 'body' of possible methods, models and techniques of its own to deal with the essential issues of modern HRA, i.e. HEM, cognitive processes and dynamic Human Machine Interaction (HMI).

Firstly, the 'points of view' to be taken into consideration by HF analysts before starting any analysis or safety study will be reviewed. Then, models and methods for HMI will be discussed focusing on their integration in a logical and temporal sequence within the overall design and safety assessment process: This represents the methodological framework HERMES. Finally, focusing on safety assessment, it will shown how HERMES may be applied in formal HRA applications and how it was implemented for the safety audit in a real and complex environment.

2. Points of view for development of human machine systems

A number of standpoints or points of view should be onsidered when studying a HMS. These cover a wide range of issues, from considerations of specific HMI to socio-technical conditions, and must guide the designer/analyst in consistently and coherently apply HF methods. In particular five points of view are to be set at the beginning of any study:

1. Definition of the Goals of the HMS under study and/or development;
2. Concept and Scope of the HMS under study and/or development;
3. Type of analysis to be performed;
4. Area of Application; and
5. Indicators of safety level.

These five *points of view* are the essential and basic elements that the analyst must consider and resolve prior to and during the development of any design or safety assessment [6].

2.1. Goals of human machine systems

The first point of view demands that the goals of the systems under study must be clearly identified and constantly accounted for during the whole process.

To discuss this initial standpoint, the example of safety and protection systems will be discussed. The improvement of safety of any technological system demands that appropriate measures are developed aiming at creating awareness, warning, protection, recovery, containment and escape from hazards and accidents. These are usually defined as Defences, Barriers and Safeguards (DBS) and represent all structures and components, either physical or social that are designed, programmed, and inserted in the human-machine system with the objective of making more efficient and safe the management of a plant, in normal and emergency conditions. *DBS*s should tackle one of or all three objectives typical of HEM, namely: (a) *prevention* of human errors; (b) *recovery* from errors, once they occur; and (c) *containment* of the consequences that result from their occurrence. Moreover, three fundamental principles of design of modern technological systems must be accounted for, namely: Supervisory Control, User-Centred Design (UCD) and Systems' Usability. Keeping in mind the goals of HEM and the principles of UCD constitutes therefore the initial standpoint from which to initiate a design or a safety assessment of a HMS.

2.2. Concept and scope of human machine system

Control systems are directly related to some forms of performance (either appropriate or erroneous). It is therefore

important to develop a clear understanding of human performances or behaviours and their dependence on specific dynamic context (contingencies) and socio-technical environment in which they are imbedded. This is a second point of view for the development of effective Human Machine Interfaces and HEM measures. In this perspective, the consideration for 'human factors' and 'human errors' expands the more classical definitions and embraces all behaviours that may engender dangerous configurations of a plant.

2.3. Types of analysis

A variety of models and methods are necessary for the development of HMSs. In general, they can be structured in an integrated framework that considers two types of analysis, i.e. retrospective and prospective studies, which are complementary to each other and equally contribute to the development and safety assessment of HEM measures (Fig. 1) [7].

These analyses rest on common empirical and theoretical platforms: the evaluation of socio-technical context, and the model of HMI and related taxonomies. In fact, applying consistent HMI models and theories, the data and parameters derived from evaluation of real events and working environment (retrospective studies) can be reasonably applied for predicting consequences and evaluating effectiveness of safety measures (prospective studies). The consideration of these differences and synergies between prospective and retrospective analyses is the third point of view to be assumed by designers and analysts for the development of HMSs.

In practice, retrospective analyses are oriented to the identification of 'data and parameters' associated with a specific occurrence and context. They can be carried out by combining four types of methods and models extensively formalised, namely: Root Cause Analysis (RCA); Ethnographic Studies (ES); Cognitive Task Analysis (CTA); and HF Theories and Models of HMI. Prospective analyses aim at the 'evaluation of consequences' of HMI scenarios, given a selected spectrum of Models of HMI, Data and Parameters, Initiating Events and Boundary Conditions, and Creative Thinking.

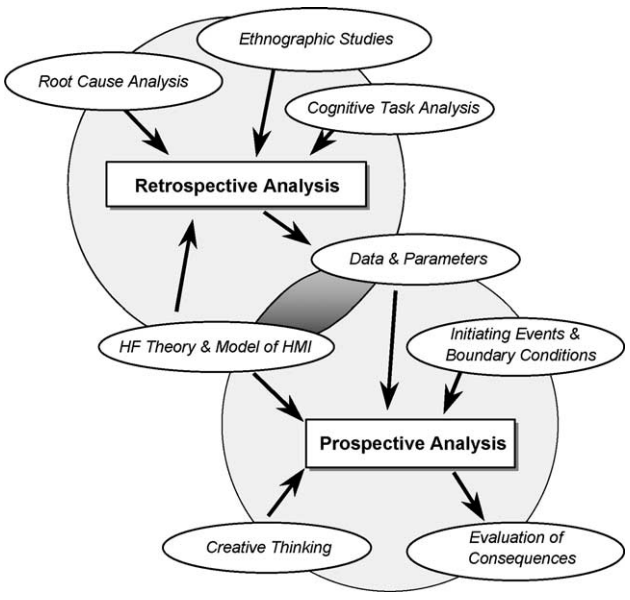

Fig. 1. Prospective and retrospective analysis.

2.4. Areas of application

When performing a process of design or assessment of a HMS, it is essential that the specific area of application of the system under study is well delineated. This differs from the point of view on identification of the goals of the system of thought there are strong links between these two standpoints. In practice, four areas of application must be considered, namely, *Design, Training, Safety Assessment*, and *Accident Investigation* (Table 1).

Each of these four areas of application encompasses specific types of assessment. The fourth point of view for the development of effective *HMS* and *HEM* measures lies in the appreciation of the fact that there are four possible areas of application, and in the links existing between different areas.

According to this fourth point of view, a variety of tools and approaches must be applied before and during the lifetime of a system for the verification that adequate safety conditions exist and are maintained.

Table 1
Areas application and type of assessment for *HMS*s

Area of application	Type of assessment
Design	*Design of control, emergency and protection systems* *Design of human–machine interfaces* *Development of normal and emergency procedures*
Training	*Classroom human factors training* *Human factors training in simulators*
Safety assessment	*Human contribution to design basis accident* *Human contribution to probabilistic risk assessment* *Evaluation of safety levels of an organisation by recurrent safety audits*
Accident investigation	*Human contribution to real accident aetiology and identification of root causes*

2.5. Indicators of safety

A final point of view is necessary for completing the process of appreciation and generation of measures for improving and safeguarding a system. This is related to the definition of appropriate safety levels of a plant. Indeed, only by applying coherent specific methods at different stages of development and management of a system, it is possible to ensure effectiveness and preservation of adequate safety margins throughout the lifetime of the plant. Consequently, in all type of analyses, and for all areas of application, it is essential that adequate *indicators* be identified that allow the estimation or measurement of the safety level of a system. As each plant and organisation bear peculiarities and characteristics specific to their context and socio-technical environment, appropriate methods and approaches must be applied for the definition of quantitative, as well as qualitative, indicators, which are unique for the plant and organisation under scrutiny.

As an example, during a Recurrent Safety Audit (RSA) process, a number of Indicators of Safety (*IoS*) are evaluated in order to assess that the plant and organisation involved are operating within acceptable safety margins, i.e. the safety measures of the system conform with current norms and standards. Moreover, the RSA must ensure that current operational procedures and emergency management systems are consistent with the originally designed and implemented safety measures.

3. A methodological framework

A methodology that aims at supporting designers and analysts in developing and evaluating safety measures must consider the basic viewpoints discussed above. In particular, a specific stepwise procedure may be considered according to each area of application and to the key objective of the safety measure under analysis, namely prevention, recovery or protection. In addition, a set of methods and models can be considered as the functional content supporting the application of procedures.

3.1. Procedural content

Firstly it is necessary to ensure consistency and integration between prospective and retrospective approaches (Fig. 2). As already discussed, both analyses rest on a common empirical and theoretical platform: the evaluation of the socio-technical context, and the theoretical models and related taxonomies of HMI. Moreover, the evaluation of socio-technical context represents an essential condition that leads to the definition of data and parameters for prospective studies, and supports the analyst in identifying conditions that favour certain behaviours and may foster accidents.

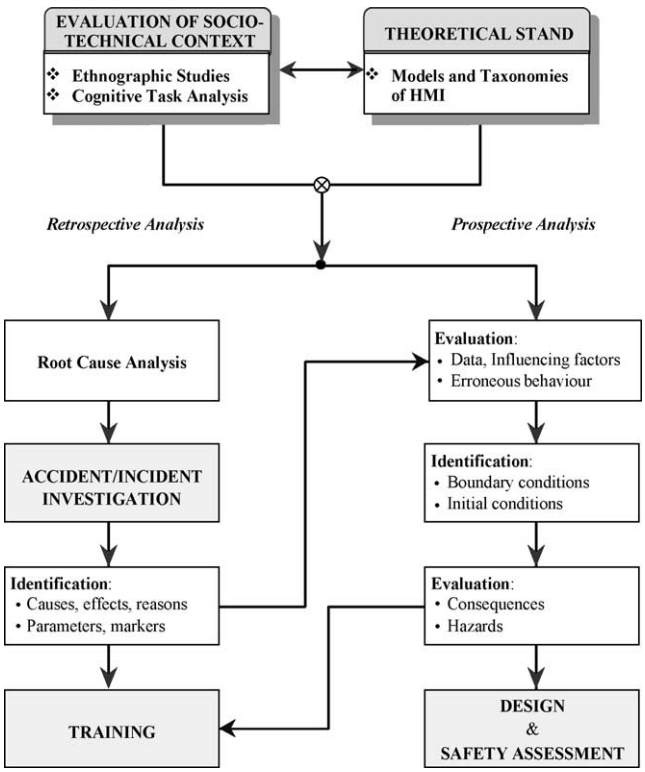

Fig. 2. Framework for human error risk management for engineering systems (HERMES).

The study of socio-technical contexts is performed by field studies, such as observations of real work processes and simulators, interviews, questionnaires. These represent a set of empirical methods that can be classified as 'Ethnographic Studies'. Further investigation is conducted by theoretical evaluation of work processes, i.e. 'Cognitive Task Analysis'. The selection of HMI model and related taxonomy is equally important, as they outline the correlation between humans and machines that are considered in performing prospective studies. At the same time, in retrospective studies taxonomies help in identifying items of HMI that may influence incidental conditions.

These two forms of preliminary assessment of HMS are correlated by the fact that all empirical studies and ethnographic evaluations can be implemented in prospective and retrospective applications only if they can feed their observations in a formal structure offered by the HMI model and taxonomy selected by the analyst.

The Methodology described in Fig. 2, called HERMES [7,8], can be applied for all safety studies that require HF analysis and must be supported by existing models and specific methods for performing each step of the procedure.

The steps to be carried out in the application of HERMES are the following:

Firstly, it is necessary to select a common theoretical platform for both retrospective and prospective types of

analysis. This is done by defining a set of:

- Models of human behaviour;
- Models of the systems; and
- Models for the HMI,

that are adequate for the working environment and technological complexity under study. At the same, data and parameters typical of the system, i.e. the common empirical platform for both types of analysis, are derived from the evaluation of the socio-technical environment by:

- Ethnographic studies; and
- Cognitive task analysis.

With the retrospective analysis, a set of data, influencing factors and erroneous behaviours is evaluated by:

- Investigation on past events that is governed by appropriate RCA and leads to the identification of causes of accidents; and
- Identification of parameters and markers of cognitive and factual behaviour.

Resulting from these analyses, additional insights can be utilized for the prospective study in the form of causes, effects and reasons of human erroneous behaviour with respect to cognitive functions and mechanisms during the dynamic interaction with the working environment.

Then, for a complete prospective study the analyst and designer needs to apply his/hers experience and creativity for:

- Identifying boundary and initial conditions for performing predictive safety studies; and for
- Evaluating unwanted consequences and hazards, by applying an adequate QRA technique.

In this way, HMI methods can be consistently and coherently applied for design, safety assessment and accident investigation as well as for tailored training.

3.2. Functional content

The HERMES methodology offers the analyst a functional content based on methods, models and techniques ready for application. Some of them will briefly discussed with the objective to give the reader the flavour of what is available.

3.2.1. Model and taxonomy RMC/PIPE

The most important characteristic of HERMES consists of the model of human behaviour (Fig. 3). The proposed model is called Reference Model of Cognition, RMC/PIPE

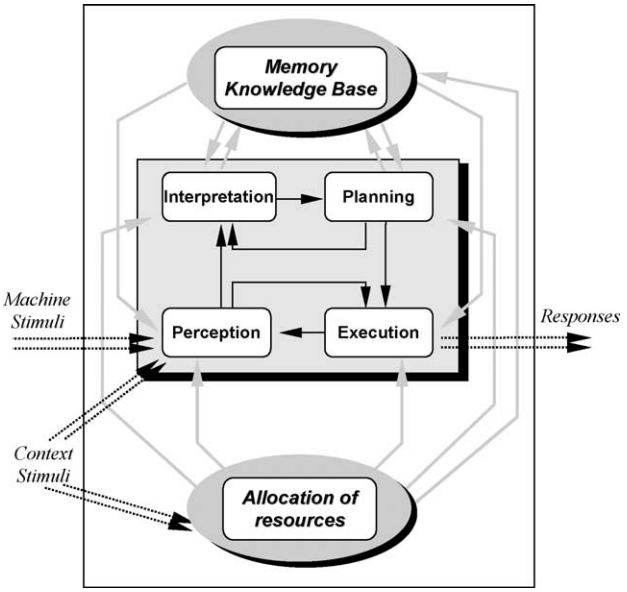

Fig. 3. Reference model of cognition, RMC/PIPE [9].

[9], and is based on very neutral principles of cognition with building-up of four cognitive functions, namely Perception, Interpretation, Planning and Execution (PIPE), and two cognitive processes, namely, Memory/Knowledge Base and Allocation of Resources. The underlying paradigm of human behaviour reflects a general consensus on the representative characteristics of human cognition, as emanating from early studies in cybernetics [10] and, more recently, in models based on information processing system analogy [11], and on psychological theories of cognition [12,13].

A taxonomy of human erroneous behaviour is correlated to the RMC/PIPE model, that allows the distinction and correlation of possible human errors at the level of different cognitive function. On the basis of this taxonomy it is possible to:

- Define causes and effects/manifestations of human erroneous behaviour in relation to the four cognitive functions;
- Link causes/general effects of erroneous behaviour with neighbouring cognitive functions and/or sudden causes, affecting at a cognitive function; and
- Identify general categories of external (system related) and internal (person related) factors affecting human behaviour.

Further on, the causes of erroneous behaviour are represented by a very useful concept of ‘not looking for a scapegoat’, that means person related causes are the root of inappropriate behaviour, but external causes do generate, trigger and enhance them. Person related causes are generically and objectively defined avoiding subjective ‘personal terms’ like inner motivation, inclination or mood.

In this framework, it is also possible to focus on organisational issues.

3.2.2. DYLAM method

For systematic studies integrating Human–Machine simulation, the Dynamic Logical Analytical Method (DYLAM) method enables to combine the stochastic nature of system failures and human (erroneous) behaviour [14]. DYLAM permits the evaluation of time dependent behaviour of human-machine systems when the boundary conditions and failure modes of its components are defined.

DYLAM generates and follows a series of possible incidental sequences, which can arise during the running time, due to failures or inappropriate behaviour of some components of the plant. Failures occur at time instants, which are unpredictable at the beginning of the analysis, as they are calculated either by probabilistic algorithms or by logical correlations between events and occurrences. Failure conditions are chosen or defined by the user at the beginning of the analysis.

Possible *states* related to human behaviour can be identified as performances adhering perfectly to procedures, or inappropriate performances of manual and cognitive activities. *Causes of transitions* between states of human behaviour can be related to environment and contextual conditions as well as to random events. The taxonomy associated to the RMC/PIPE model may be directly applied to this purpose.

In a typical case, DYLAM, as dynamic simulation manager, triggers time dependent failures in the plant components and human inappropriate behaviours according to probabilistic/logical criteria. An example of DYLAM sequences can be seen in Fig. 4. These sequences are generated by time dependent evolutions of physical system variables and human characteristics that describe the actual transient response of the human-machine system. In Fig. 4(a) it is possible to identify 17 sequences (S_1–S_{17}), generated in addition to the 'nominal sequence' (S_0), which is calculated following the initiating event and all physical components and human behaviours performing as expected, i.e. in *nominal state*. Fig. 4 (b) and (c) show the time evolution of some physical variables (Q_{QVCT}, Q_{MKUT}, and L_t) during Sequence S_1 and the corresponding time variation of the function *Stress*. The function Stress, correlated to the physical variables, generates possible human errors and therefore sequences, at times t_1, t_3, and t_5 (Fig. 4(a)). All 17 sequences are generated at various times and are distinguished from each other because of different failure modes of components and human errors.

4. Application of HERMES for human reliability assessment

In this section, following the procedural framework and the functional content provided within HERMES the performance of an extended HRA is outlined [15–17]. It will be shown how the prospective HF analysis can be sustained by a conservative method or technique instead of the more advanced methods being offered by HERMES (e.g. DYLAM for dynamic event tree analyses), when the objective is the performance of a more conventional approach.

4.1. Theoretical stand

The Model RMC/PIPE and the correlated taxonomy of human erroneous behaviour can be used for the extended HRA performance. Thereby, the generic external causes provided have to be specified into a set of context specific expressions: at first, a structured exploitation of the extensive plant specific knowledge is conducted in a general way, to be specified later on. The theoretically defined manifestations of human erroneous behaviour have to be analysed and reflected in the particular working and socio-technical environment, initially on a general level with further specifications in the course of the analysis.

4.2. Evaluation of the socio-technical context

Firstly, on a macro level, the working and socio-technical environment and human–machine interaction have to be analysed in order to become acquainted with the procedures and practices that are implemented within an organisation, as well as with the philosophy and policy that affect them. In the course of this process, issues like design, system performance, safety and control system, human tasks, working conditions, organisational and management aspects are observed, reviewed and evaluated. In particular, the existing operating procedures with the most relevant HF related aspects are analysed in relation to environmental and contextual (external) conditions and performance influencing factors (PIF), that may affect human behaviour.

In a subsequent phase, at a micro level, a CTA is performed, referring to actual manifestation of behaviour. Such CTA is mainly subject of the QRA to be carried out, but also of the integrated RCA and Accident Investigation. The novelty of conducting a CTA consists in the structured identification and analysis of behavioural and cognitive performance. That means, error modes and causes, PIF and error mechanisms referring to cognitive functions can additionally be identified and taken into account.

For the practical conduct of a QRA, the possible visible forms of erroneous behaviour have to be identified and analysed, in relation to external contextual factors and to cognitive functions.

4.3. Retrospective analysis

Additional insights on causes and effects on human erroneous behaviour of socio-technical context can

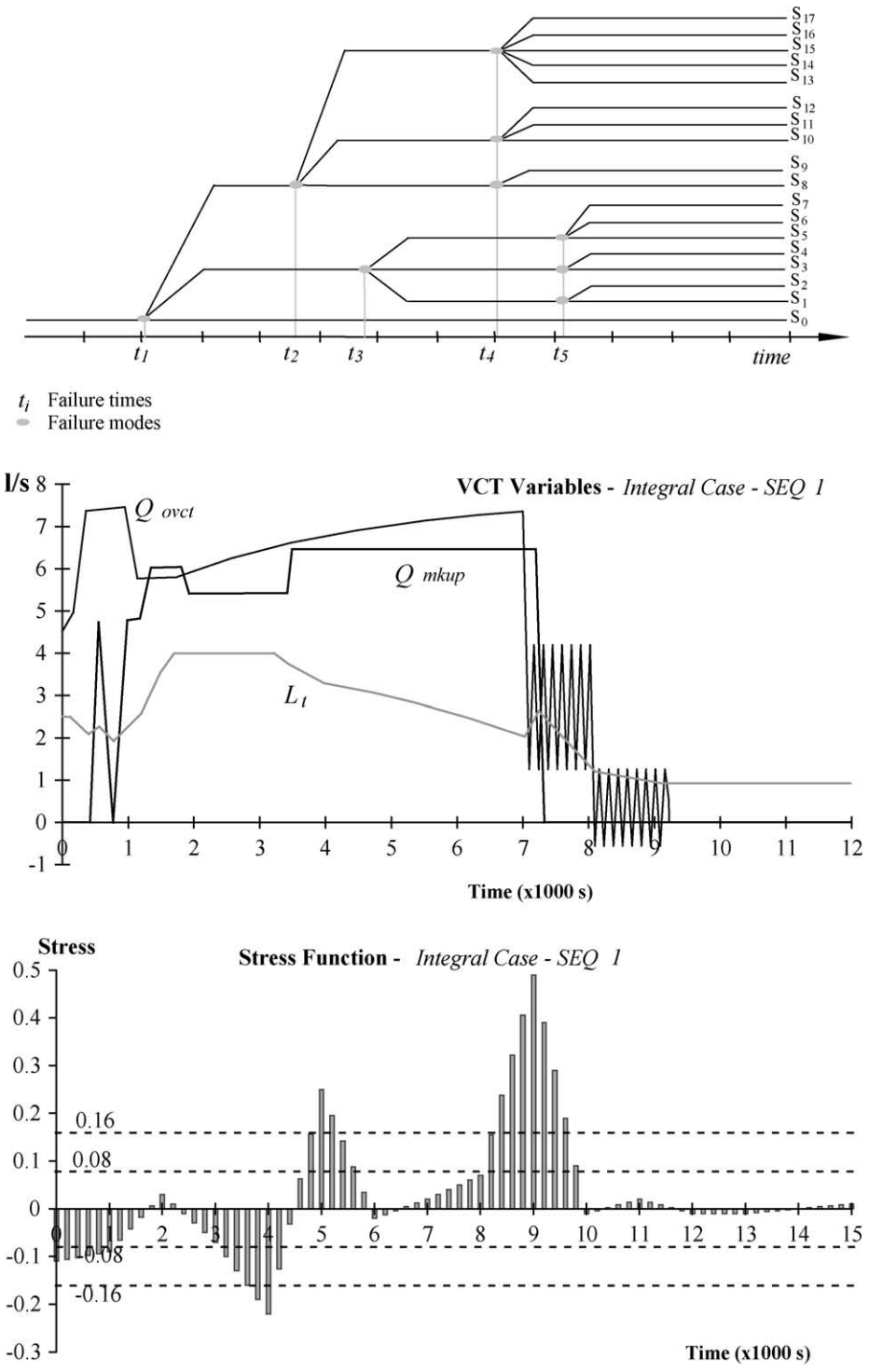

Fig. 4. The DYLAM method for dynamic management of HMI simulation.

be gained as result of integrated RCA and Accident Investigations.

This part of the HRA performance corresponds clearly to the known and accepted approach in the system technique, where plant specific data and information are gathered and evaluated from the operational experiences, in order to more adequately reflect, represent and predict expected behaviour of engineering systems in the context of a prospective QRA.

4.4. Prospective analysis

Main objective of a prospective QRA lies in the qualitative and quantitative prediction of the unwanted consequences resulting from an incidental evolution of an engineering system. Thereby, the human behaviour in the dynamic interaction with the socio-technical context has systematically to be considered and assessed. With

the objective of demonstrating the performance of a static QRA, the utilisation of the conventional reliability technique THERP [3] will be 'embedded' in the methodological framework of HERMES.

4.4.1. Information and data

The input for the prospective analysis are a thorough understanding of the working environment, the socio-technical context and the actual manifestation of behaviour by performing tasks and control actions. From this, environmental and contextual (external) conditions and PIF are analysed and determined, that may affect human behaviour. In addition to a pure conventional approach, causes and effects of human erroneous behaviour with respect to cognitive functions and mechanisms in the interaction with the working and socio-technical environment are identified and can thus be considered. This holds also for the additional insights gained from the retrospective analysis of the real operational experiences.

4.4.2. Identification

All information and data gathered up to now has eventually to be evaluated for the actual performance of the prospective analysis, i.e. the identification of data, PIF, cognitive functions and error mechanisms as well as the evaluation of possible forms or erroneous behaviour identified (error modes). These are based on causes and effects-triggered by external, context related factors—related to cognitive functions according to the RMC/PIPE model and correlated taxonomy.

Additionally, boundary and initial conditions, initiating events, systemic processes and failure modes have to be evaluated and a set of data and parameters for modes, types and frequencies of occurrence of human erroneous behaviour and systemic failures has to be developed. The latter aspect is also tackled in the following section.

4.4.3. Evaluation

At this stage of the HRA performance an appropriate reliability method has to be developed or applied, that combines causes and effects—triggered by external, context related performance factors—of human erroneous behaviour. Herewith, the investigation and analysis of unwanted consequences and/or hazards and their associated frequencies of occurrence and uncertainty distributions can be carried out.

The practicability of this stage of the HRA performance, i.e. the utilisation of THERP embedded in the methodological framework of HERMES is described hereafter explicitly for an exemplary case.

The following task shall be investigated: In the course of the accidental evolution of a Loss of Coolant Accident (LOCA) in a Boling Water nuclear Reactor (BWR), the Low Pressure (LP) injection is automatically initiated. In case of a systemic failure of the main steam isolation valves (MSIV), the interruption of the LP-injection is anticipated—to prevent from flooding—and to intermittent feeding afterwards. The objective of this measure is to avoid the flooding of the main steam piping—due to erroneously open MSIV, e.g. due to a systemic common cause failure—with a possible consequent failure of this piping in the reactor building and flooding.

For the first part of the task two error mechanisms are investigated. The following task should be carried out: after instruction to carry out the task, the level has to be observed by the operator; if the level is ≥ 14.92 m, then the LP-injection has to be switched off. In the framework of the CTA, causes and effects of human erroneous behaviour—triggered by external, context related factors—related to the cognitive functions (observation) PERCEPTION, INTERPRETATION, PLANNING (decision making) and EXECUTION are described.

The first error mechanism identified (**case A**) is described hereafter: during the observation of the level, the limiting value is erroneously perceived (Incorrect estimation of the measure due to an *attention failure* triggered by an *external distraction*; operator erroneously recognises 14.92 m, in reality the level is 9.42 m). This leads to an incorrect recognition of the value in the cognitive function PERCEPTION. Due to the linked cause from the PERCEPTION function (Incorrect recognition of the value), the state is incorrectly recognised in the INTERPRETATION function ('Level ≥ 14.92 m'). In the PLANNING function, an incorrect choice of alternatives occurs due to the linked cause from the INTERPRETATION function ('Decision to switch off, because level increases limiting value'). Eventually, in the EXECUTION level the premature (level = 9.42 m) switching off the LP-injection takes place.

In this error mechanism, an error at level PERCEPTION results in the erroneous behaviour of execution of an action before the established time; it can be modelled in a static binary HRA event tree, following the reliability technique THERP (Fig. 5).

In order to identify an Human Error Probability (HEP) for this error in the cognitive function PERCEPTION, chapter 20 of THERP, can be consulted. So, for the 'sub task' RECOGNITION OF VALUE according to THERP, table 20-10 (2), an estimated HEP for an error in reading and recording quantitative information from an unannunciated display, here a digital indicator, can be identified, i.e. an $HEP_{(median)}$ of 0.001 (EF = 3).

The PIF identified, i.e. an attention failure triggered by an external distraction, can be assessed following the suggestions in THERP, table 20-16, how to modify estimated nominal HEPs; in this case a modifier of $\times 2$ could be appropriate.

That means, the utilisation of the methodological framework HERMES and its functional content leads to the additional classification of an erroneous action, i.e. a typical *Error of Commission*, that in the presently used pure conventional approaches would not have been identified.

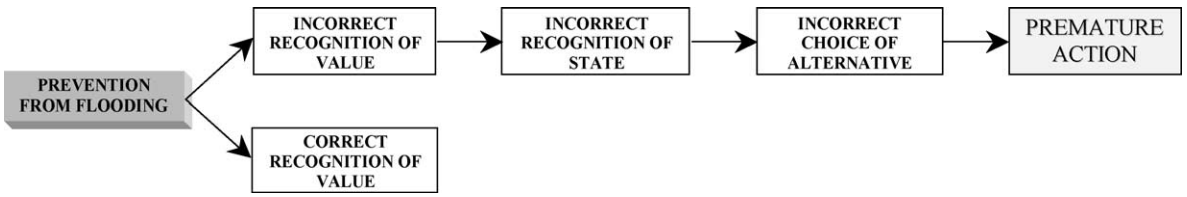

Fig. 5. Error of commission (premature execution)—error of PERCEPTION.

In **case B** a second error mechanism is identified: during the observation of the level, the real measure of 14.92 m is correctly read; this results in a correct recognition of the value in the cognitive function PERCEPTION. In the INTERPRETATION function the correct recognition of the state takes place ('Level $\geq$ 14.92 m'). At PLANNING level, a wrong decision rule is used: the limiting value is erroneously recalled at 16.34 m in the memory (Use of wrong decision rule due to *recent failures* triggered by an *external distraction*). This leads to an incorrect choice of alternatives ('Decision to **not** switching off, because level has not yet reached limiting value'). Eventually, on the EXECUTION level the delayed (level = 16.34 m) switching off the LP-injection takes place.

In this second error mechanism identified, an error on the cognitive function level PLANNING results in the manifestation of a human erroneous behaviour, i.e. the execution of an action after the established time, what—in a conventional understanding—would correspond to an error of omission referring to the objective to avoid the flooding of the main steam piping; this error mechanism can also be modelled in a static binary HRA event tree (Fig. 6).

In order to identify an HEP for this error on the cognitive function PLANNING, chapter 20 of THERP, can be consulted. Though from a behavioural point of view the manifestation of this error results in the known conventional *error of omission* (referring to the action goal 'preventing a flooding because of a certain systemic failure'), the identification of an HEP consulting THERP is slightly more demanding.

Let us assume, the instruction for the carrying out of the action is given orally: "Switch off LP-injection at 14,92 m [to avoid possible failure of the main steam piping, because MSIV are not closed]; perform intermittent feeding afterwards". Then, an estimated HEP for an error in recalling oral instruction items according to THERP, table 20-8 (6c), i.e. an $HEP_{(median)}$ of 0.001 (EF = 3), can be estimated for the error of using a wrong decision rule, that results in an incorrect choice of alternatives.

The PIF identified, i.e. an error due to recent failures triggered by an external distraction, can be assessed following the suggestions in THERP, table 20-18, how to modify estimated nominal HEPs in order to achieve conditional probabilities of failure given the failure on a preceding task performance; in this exemplary case the assumption of low level of dependence could be appropriate.

The erroneous action identified of not preventing a flooding can then-in a quite standard manner-be modelled, taken into account and followed up in a usual, static approached event tree analysis.

4.5. HERMES versus THERP

What is then the difference for this exemplary case in comparison to a 'pure' THERP analysis? The latter comes from a systemic perspective, i.e. asking for the HEP of a certain action needed for the success of a systemic function. That means, the actual action performance is investigated with respect to possible errors from a behavioural perspective, i.e. what *observable* human errors can occur, without explaining, *why* humans act in a certain way when performing tasks and control actions and *what* may trigger inappropriate human behaviour.

In the very end, only the *visible* forms of erroneous actions, i.e. the error modes or manifestations of human erroneous behaviour are explicitly taken into account in a QRA. Nevertheless, the consideration of the methodological framework HERMES and its functional content leads to the identification of error mechanisms and modes which otherwise would not have been reflected or predicted.

The objective of a QRA, i.e. the qualitative and quantitative prediction of the unwanted consequences

Fig. 6. Error of omission—error of PLANNING.

resulting from an incidental evolution of an engineering system can thus be achieved in a conventional, static approach, but with an extended HRA performance. Thereby, the cognitive aspect of human behaviour in the interaction with the socio-technical context can explicitly be tackled and considered appropriately.

5. Application of HERMES for safety audit

5.1. Problem statement

The methodology HERMES was applied for auditing a railway organisation and offering a set of safety indicators that would provide a structured format for regularly checking safety state and identifying areas of concern. This is a typical approach for preventive safety assessment based on RSA. The application of HERMES was limited to the identification of safety critical factors, or *IoS*, and their distribution into *RSA-Matrices* that serve the purpose of defining the existing level of safety within the organisation and defining the reference measures for future audits.

The dimension of the railway organisation, called RC for convenience, consisted of more than 100,000 employees, engaged every day in the management of the railway traffic, with a population of more than 70% of train drivers. The technology and means available at RC consist of a high number of 'trains' presenting a wide variety of technological solution on board, from the most modern fully automatic controlled, high-speed machines, to very conventional locomotives, based practically on full manual control. At the time of the study, the railway network covered almost 20,000 km over a vast territory and a variety of terrains. The RC company managed also the underground systems of the some metropolitan areas. Approximately 10,000 trains per day ensured the movement of several hundreds of thousand passengers.

As a consequence of the complexity and dimension of the RC organisation, the application of HERMES for the Safety Audit alone covered an exercise of several months of work by a team of HF experts (HF Team).

5.2. Procedural and functional application of HERMES

The work was performed in three correlated and timely distributed phases in order to gain the maximum possible experience and insight by the HF analysts. At the same, a goal of the study was to transfer to the personnel of the company the methodological know-how and information that would allow RC to carry out future audit exercises and safety evaluation from inside of the organisation, which is the most efficient approach. The three phases were structured as follows:

Phase 1

This phase included the setting up of the HF Team and working groups supporting the study, and covered the initial steps of HERMES, namely:

- Acquisition of information and knowledge in the field about the working environments and practices of train driving (Ethnographic Studies);
- Study of work requirements by task and job analysis (Cognitive Task Analysis); and
- Identification of theoretical models and techniques to apply (Selection of Models and Taxonomies of HMI).

This activity involved some staff members and managers at different levels within the organisation. The focus of this phase was to select a consistent number of train drivers, or the most representative depots, that would represent the core of further detailed interviews and data collection.

Training procedures and books with norms, regulations and standards were also made available as part of theoretical documentation.

Phase 2

The second phase of work was dedicated to the extensive field assessment and, thus, to the collection of the most important source of information and data through questionnaires and interviews. The analysis of all collected data aimed at identifying possible areas of concern.

The questionnaires were distributed to a population of 2500 train drivers and more than 700 answers were colleted. More than 300 TDs and RC managers were involved in the interviews.

The work of data analysis was performed with the support of a limited selected number of experts of the company, including train drivers and managers.

Phase 3

The third phase was totally dedicated to the development of the safety audit and to the preparation of the recommendations for the organisation. According to the HERMES framework, this phase was performed exploiting the data and results obtained in phase 2 in combination with adequate boundary and initial conditions, selected by the analysts on the basis of their experience and creativity. The development of IoS and RSA-Matrices concluded the work of Safety Audit.

5.3. Results of application

The final step of this study focused on the generation of a set of *IoS*s and *RSA-Matrices* that would allow RC to perform successive evaluations of its own safety level.

The generic format for applying HERMES in the case of Safety Audit requires that *IoS* are identified and grouped in four 'families' of Indicators according to whether they can be considered Organisational Processes (OP), Personal and External Factors (PEF), Local Working Conditions (LWC), and DBS [6]. In the *RSA-Matrices*, the contribution of each *IoS* is associated with its prevention, recovery, or containment safety characteristics.

Table 2
Identification of *IoS* with reference to PIFs

Performance influencing factors		Family of *IoS*	
Communication within RS	Serious problems encountered in the contacting managers for discussing rules and standards	*OP*:	Unwritten rules; reporting systems
	Uncertainty about future—low morale	*PEF*:	Mental conditions
	Unions as unique channel for communicating with top management level	*OP*:	Role of unions vs. management
Communication means	Obsolete technology for communication	*LWC*:	Quality of tools/actuators
	Inadequate maintenance on communication means	*LWC*:	Maintenance of tools
	Unclear rules for communication	*DBS*:	Policies, standards
Technological interfaces	Poor ergonomics of interfaces of train cabins	*LWC*:	Workplace design
	Inconsistency between signals track-cabin	*LWC*:	Signals; automation
Maintenance of trains/railway	Inadequate and insufficient maintenance of trains and tracks	*DBS*:	Safety devices
		PEF:	Stress, system conditions
Comfort of working contexts	Obsolete technology for communication	*LWC*:	Quality of tools/actuators
	Poor comfort of cabin, rest areas, etc.	*LWC*:	Workplaces
	Lack of development plans	*PEF*:	Morale, mental condition
Roster and shifts planning	Heavy and too stiff shifts	*PEF*:	Physical conditions
	Inadequacy of 'fixed' couple for max safety	*DBS*:	Training; supervision
Regulations/rules	Excess of rules and regulations	*DBS*:	Procedures, standards
Training methods and simulators	Inadequate training	*DBS*:	Training standards
	Insufficient experience/expertise of instructors	*OP*:	Human relationship

The complete application of the approach foresees that the *RSA-Matrices* defining the 'current safety state' of an organisation are evaluated and confronted with the corresponding *RSA-Matrices* associated to: (a) the safety state of the organisation at a reference time, e.g. the time of initial operations for a plant; and (b) the *IoS* required by safety regulations and standards for ascertaining acceptable safety levels and thus operability of a plant and an organisation. However, in this study only the *RSA-Matrices* and the *IoS* associated with the current state of the organisation were performed.

The PIF and possible types and modes of errors identified during the two initial phases of work have been considered in detail for identifying adequate *IoS*s. Table 2 shows the generic *IoS*s that have been identified.

6. Conclusions

In this paper a methodology called HERMES is presented with its procedural framework aiming at a consistent and coherent application of the most appropriate HF approaches and methods for a specific problem at hand, including the identification and definition of data. Further, some of the functional content, i.e. the body of methods, models and techniques offered by HERMES to deal with the essential issues of modern HRA is briefly discussed.

Further on, the consideration of five points of view for the development of safety studies is recommended, prior to the analysis, in order to appropriately put into perspecive the issues to be resolved.

The utilization/application of HERMES is presented for two areas of application: the performance of an extended HRA in the framework of a QRA, and the implementation of a safety audit in a railway organisation. In the latter case, the methodology was applied to a very large organisation, and supported the analysts in a two ways:

- Firstly, it offered a consistent ground for selecting the most adequate models and methods for analysing working contexts and socio-technical conditions; and
- Secondly, it provided the procedure for determining the crucial *IoS*, by applying in a consistent and logical manner the selected methods and the available information and past experience existing within the organisation. The *IoS*s are the base for determining the safety level of an organisation and for identifying weak points and areas of intervention.

These types of approaches are becoming more and more important nowadays for ensuring safe operation of complex systems. The HERMES methodology has shown efficiency and effectiveness in a real and complex application.

For a QRA, the steps to be followed according to the procedural framework and the functional content provided within HERMES are outlined. In a sample case, it was shown how a prospective HF analysis can be sustained by a conservative technique, instead of a more advanced method offered by HERMES, when the objective is the performance of a more conventional approach that requires probabilities of errors. To this purpose,

the utilisation of THERP, embedded in the methodological framework of HERMES, was explicitly described. However, a specific improvement could be devised. Indeed, it resulted that error mechanisms and modes could be identified, which otherwise would not have been detected or predicted: root causes of Errors of Commission and causal paths to conventional Errors of Omission.

Last but not least: what about the major bottleneck of innovative HF approaches, i.e. the lack of readily available data? The data retrieval while applying such a methodology certainly demands a not negligible, accurate and extended analysis of the socio-technical context under study. However, also the application of a conventional, rigid and static method or technique requires a laborious and eventually considerable effort for the HRA performance [18]. Therefore, it is worthwhile to start utilising advanced HF methodologies, as they are certainly able to furnish different, in-depth and particularly safety relevant insights.

Acknowledgements

The author wishes to express many thanks and deep appreciation to C. Spitzer, for her detailed and accurate work of revision of the methodology during its theoretical formalisation and while performing safety assessment of real working situations. The sample case discussed in the section 'Application of HERMES for HRA' has been drawn from her work of implementation.

References

[1] Cacciabue PC. Evaluation of human factors and man–machine interaction problems in the safety of nuclear power plants. Nucl Engng Des, NED 1988;109:417–31.
[2] Apostolakis G, editor. Risk perception and risk management. Special issue of reliability engineering and system safety, RE&SS, 59, (1). Amsterdam: Elsevier; 1998.
[3] Swain AD, Guttmann HE. Handbook on human reliability analysis with emphasis on nuclear power plant application. NUREG/CR-1278. SAND 80-0200 RX, AN. Final Report; 1983.
[4] Hannaman GW, Spurgin AJ. Systematic human action reliability procedure (SHARP). EPRI NP-3583, Project 2170-3, Interim Report. San Diego, CA, US: NUS Corporation; 1984.
[5] Hollnagel E, Cacciabue PC. Reliability of cognition, context, and data for a second generation HRA. Proceedings of International Conference on Probabilistic Safety Assessment and Management, San Diego, California, 20–25 March 1994.
[6] Cacciabue PC. Guide to applying human factors methods. London, UK: Springer; 2003.
[7] Cacciabue PC. Human factors impact on risk analysis of complex systems. J Hazardous Mater 2000;71:101–16.
[8] Cacciabue PC. Human error management impact on design and assessment of safety measures. In: Bonano EJ, Majors MJ, Camp AL, Thompson R, editors. Proceedings of PSAM 6—International Conference on Probabilistic Safety Assessment and Management. Puerto Rico, 23–28 June. Amsterdam: Elsevier; 2002. p. 473–80.
[9] Cacciabue PC. Modelling and simulation of human behaviour in system control. London, UK: Springer; 1998. ISBN 3-540-76233-7.
[10] Ashby WR. An introduction to cybernetics. London: Chapman and Hall; 1956.
[11] Neisser U. Cognitive psychology. New York: Appleton–Century–Crofts; 1967.
[12] Reason J. Human error. Cambridge UK: Cambridge University Press; 1990.
[13] Rasmussen J. Information processes and human–machine interaction. An approach to cognitive engineering. Amsterdam, The Netherlands: Elsevier–North Holland; 1986.
[14] Cacciabue PC, Cojazzi G. A human factor methodology for safety assessment based on the DYLAM approach. Reliab Engng Syst Saf, RE&SS 1994;45:127–38.
[15] Spitzer C. Data collection and Evaluation as well as Methods for incorporation into PSAs: recommendations due to the experiences gathered during the course of the assessment of psas performed by utilities. Final Report—IAEA Coordinated Research Programme on Collection and Classification of Human Reliability Data for Use in PSA, May 1998.
[16] Spitzer C. Improving operational safety in nuclear power plants: extended consideration of the human factor issue in PSAs. In: Kondo S, Furuta K, editors. Proceedings of the Fifth International Conference on Probabilistic Safety Assessment and Management (PSAM 5). November 27–December 1, 2000, Osaka, Japan. Tokyo, Japan: Universal Academy Press, Inc.; 2000. p. 627–32.
[17] Carpignano A, Piccini M, Cacciabue PC. Human reliability for the safety assessment of a thermoelectric power plant. ESREL Conf 1998; 17–19.
[18] Spitzer C. PSA Reviews: experiences and insights from methodological points of view. In: Mosleh A, Bari RA, editors. Proceedings of the Fourth International Conference on Probabilistic Safety Assessment and Management (PSAM 4)New York City, USA, 13–18 September 1998. London, UK: Springer; 1998. p. 833–9.