HUMAN RELIABILITY ASSESSMENT IN CONTEXT

ERIK HOLLNAGEL
Cognitive Systems Engineering Laboratory,
Department of Computer and Information Science,
University of Linköping, Sweden
E-mail : eriho@ida.liu.se

Received March 10, 2005

Human reliability assessment (HRA) is conducted on the unspoken premise that 'human error' is a meaningful concept and that it can be associated with individual actions. The basis for this assumption it found in the origin of HRA, as a necessary extension of PSA to account for the impact of failures emanating from human actions. Although it was natural to model HRA on PSA, a large number of studies have shown that the premises are wrong, specifically that human and technological functions cannot be decomposed in the same manner. The general experience from accident studies also indicates that action failures are a function of the context, and that it is the variability of the context rather than the 'human error probability' that is the much sought for signal. Accepting this will have significant consequences for the way in which HRA, and ultimately also PSA, should be pursued.

KEYWORDS : robust control, autonomous control, reactor control, research reactor, experimental validation

1. WHY DOES PSA NEED HRA?

Human reliability assessment (HRA) is the common name for an assortment of methods and models that are used to predict the occurrence of 'human errors'. While the origin of HRA is in Probabilistic Safety Assessment (PSA), HRA is increasingly being used on its own both as a way to assess the risks from 'human error' and as a way to reduce system vulnerability. According to [1] the three principal functions of HRA are "identifying *what* errors can occur (Human Error Identification), deciding *how likely* the errors are to occur (Human Error Quantification), and, if appropriate, enhancing human reliability by *reducing* this error likelihood (Human Error Reduction)" [1].

Practically all HRA methods and approaches share the assumption that it is meaningful to use the concept of a 'human error', hence also meaningful to develop ways of estimating 'human error' probabilities. As a consequence of this, numerous studies have been performed to produce data sets or databases that can be used as a basis for determining 'human error' probabilities. This view prevails despite serious doubts expressed by leading scientists and practitioners from HRA and related disciplines. A comprehensive criticism of HRA [2], for instance, pointed out that many HRA approaches are based on highly questionable assumptions about human behaviour. This view is supported by the experience from extensive studies of human performance in accidents, which conclude that:
... "human error" is not a well defined category of human performance. Attributing error to the actions of some person, team, or organization is fundamentally a social and psychological process and not an objective, technical one. [3]

Although the concept of 'human error' itself is the subject of much debate, it is not the intention to go into that here (but see [3, 4, 5, 6]). For the purpose of this discussion a 'human error' will simply be defined as an identifiable human action that in retrospect is seen as being the cause of an unwanted outcome[1]. (Needless to say, even the concept of a cause can be the subject of dispute, not least when it comes to the description of accidents [8].)

1.1 The Growth of HRA

In trying to understand what HRA is, and perhaps even more importantly, in trying to determine what HRA ought to be, it is necessary to take a look at how HRA has developed and how it came into use. Without attempting a complete intellectual history of HRA, it suffices to note the strong connection between the accident at Three-Mile Island (TMI) on March 28, 1979, and the growth in the number of HRA methods. As shown by

[1] The reader should compare this to the conventional definition of a 'human error' as "any member of a set of human actions or activities that exceeds some limit of acceptability, i.e., an out of tolerance action [or failure to act] where the limits of performance are defined by the system" [7].

Figure 1, most HRA methods appeared in the 1980s with the largest growth taking place in 1984. That was followed by another, but smaller, growth period around 1996, which represents the launch of the so-called second-generation methods that followed the lucid criticism of HRA [2]. (A more detailed discussion of these developments can be found in [9].)

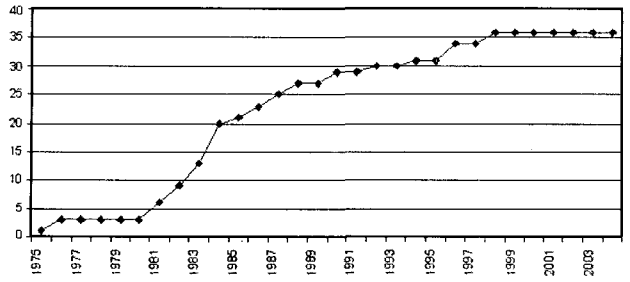

Fig. 1. Cumulated Number of HRA Methods According to Year of Publication.

While the accident at TMI was significant, it also fitted into the growing concern for system failures and helped push forward the realisation that these generally were unavoidable, hence normal occurrences rather than exceptions [10]. The anticipation of system failures was then as now guided by a scientific paradigm that relied on decomposition – in particular the decomposition of a system into its 'natural' parts, humans and machines. This paradigm has been consolidated by disciplines such as human factors (ergonomics) and human-computer interaction. Since the reliability of modern technology is quite high, the logic of the decomposition approach has forced the focus onto issue of human reliability, usually as single individuals and more rarely as groups or organisations.

1.2 Conspicuousness of the Human Factor

The TMI accident turned human factors into a central issue for both control room design and HRA in nuclear power production and elsewhere. It also raised the concept of 'human error' to a prominent level. This reinforced the already growing trend to perceive 'human error' as the main cause of accidents in complex technological systems, hence the change from viewing such systems as technical systems to seeing them as human-machine or socio-technical systems. This trend soon became so strong that the search for a human failure was the normal reaction to accidents. The detrimental consequences of such an attitude were made clear by Charles Perrow, when he wrote that:

Formal accident investigations usually start with an assumption that the operator must have failed, and if this attribution can be made, that is the end of serious inquiry. [10]

The trend nevertheless continued, so that by the end of the 20th century it was 'common wisdom' across a variety of domains to assume that the contribution of human factors to accidents was between 70% – 90%, while the contributions from other causes were correspondingly low. Yet as argued elsewhere [9], this distribution represents the *attributed* rather than the actual causes. The estimates have furthermore changed significantly over the last 40 years or so. One trend has been a decrease in the number of accidents attributed to technological failures, among other things due to an actual increase in the reliability of technological systems. A second trend has been an increase in the number of accidents attributed to human performance failures, partly due to the development of accident models sensitive to human factors and partly because of real changes in the nature of work. Most recently there have been a growing number of cases attributed to organisational factors, corresponding to the recognition of the difference between failures at the sharp end and at the blunt end [3, 11, 12]. Whereas failures at the sharp end tend to be attributed to individuals, failures at the blunt end tend to be attributed to the organisation as a separate system.

While the human factor came to the fore because of a number of tragic and severe accidents (with Tenerife in 1972, TMI in 1979, and Chernobyl in 1986 as signature cases), the trend in a perverse manner reinforced itself. Since no system has ever built itself, since very few systems operate themselves, and since furthermore no systems maintain themselves, the search for a human in the path of events leading to a failure is bound to succeed. If the fallible human is not found directly at the sharp end – as in the case of 'human errors' or unsafe acts – he or she can usually be found a few steps back in design, implementation or maintenance. The assumption that a human has failed will therefore always be vindicated. The search for a human-related cause is reinforced both by past successes and by the fact that most accident analysis methods put human failures at the very top of the hierarchy, i.e., as among the first causes to be investigated.

1.3 The Need of 'Human Error'

The attribution of a large number of accidents to 'human error' meant that there was a need to factor in 'human error' in risk assessment. Since the need came from operational practice and engineering rather than from behavioural science and human factors, the solution was to import human factors concerns into engineering practices rather than the other way around. The initial approach was to use existing PSA methods and extend them to include human actions. A typical PSA consists of the following steps [13]:

- Define the risk criterion or risk criteria. The risk criterion is used to determine which accident sequences should be included in the analysis.

- Create a description or representation of systems that make up the plant, including in particular a description of the interactions between the technical system and people.
- Define the associated hazards using information from the plant's operating history, if it is available, or information from similar plants. The outcome of this step is summarised in a list of the events that may initiate an accident, as well as any other event that must occur for the hazard to obtain.
- Define the accident sequences that will lead to specific hazards. Such sequences are usually described as binary-branching event trees. Each node or event may in turn be expanded and described in greater detail using e.g. fault trees or reliability block diagrams - or HRA if human actions are involved.
- Evaluate the consequences of the accident sequence, i.e., of sequences that lead to failures.

The accident sequence provides the basis for determining the frequencies and uncertainties of the consequences. The essential outcome of a PSA is a quantitative expression of the overall risks in probabilistic terms. For HRA to be useful, it therefore has to produce something that easily fits into PSA, i.e., a probability of a 'human error' being made.

In relation to HRA, the critical issue is the description of the accident sequence as an event tree. The event tree represents the accident as a sequence of events, in essence as a domino model [14] with (binary) branching. It is thus a simplified representation of what may actually happen, not least because human actions only can be described as individual events, corresponding to a node in the event tree (Figure 2).

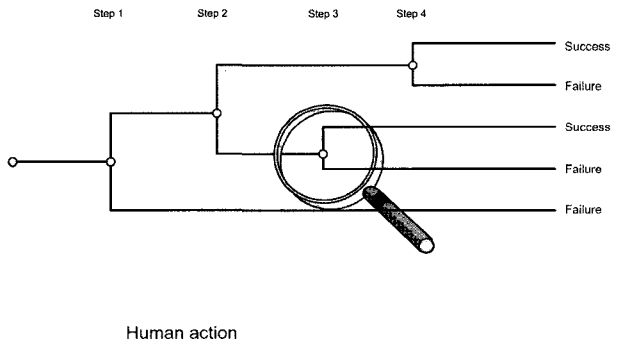

Fig. 2. Human Actions in the PSA Event Tree.

An argument in favour of this approach is that some work environments, such as nuclear power pant operation, constrain what operators can do. It is, of course, correct that the execution of a procedure must take place in a certain order, due to the nature of the physical processes and the construction of the technical systems. Yet even if operators are forced rigidly to follow procedures, it neither means that a predetermined sequence of events necessarily will match actual performance, nor that it is warranted to consider each step or event by itself.

Because of the way in which a human-machine system is decomposed, risk analysis will sooner or later reach the level of human actions. Since human actions are defined as components – or nodes – in the event tree, it becomes necessary to assess the likelihood that these nodes may fail. This is done in the same way as for other components, which are assumed to have a certain failure rate that is more or less independent of the conditions. At least the meaningfulness of referring to a failure rate is taken for granted in the case of technological components, even if the conditions must be factored in at some time.

PSA initially assumed that the human operator could be described in the same manner as a machine. Thus [15] noted in their description of the Technique for Human Error Rate Prediction (THERP) that:

The THERP approach uses conventional reliability technology modified to account for greater variability and interdependence of human performance as compared with that of equipment performance ... The procedures of THERP are similar to those employed in conventional reliability analysis, except that human task activities are substituted for equipment outputs.

Since a component has a failure probability it was natural to assume that there was a corresponding 'human error' probability. Considerable efforts were therefore dedicated to establishing tables or databases of such 'human error' probabilities, either by extensive studies in NPP training simulators [17], by analysing and refining empirical data [18], by proposing specific cognitive models [19], or by developing computer models of human operators [20]. Yet as long as HRA was carried out as a part of PSA, it was limited to consider only those human actions that could be included in the event tree. The quality of the analysis therefore critically depended on the completeness and accuracy of the PSA event tree.

1.4 The Decomposition of Cognition

The basic assumption that human failures could be described analogously with technical failures soon turned out to be invalid – with the possible exception of certain types of highly regular performance such as well-rehearsed skills. Furthermore, in cases where the human interaction comprised cognitive functions or mental acts rather than overt actions, the use of the event tree description did not make sense because the 'component' that 'failed' represented a hypothetical construct that was inferred rather than observed. It was therefore necessary to find a more realistic approach and develop descriptions or models of human actions that could provide a better basis for e.g. system design, task analysis, etc. [13].

The immediate solution was to decompose human

actions into their constituent cognitive functions, further to decompose such cognitive functions into their assumed components, and finally to describe their relations by means of smaller event trees. Assume, for instance, that an operator action can be decomposed into the following four segments:

- *Problem identification*, where the operator must detect that something has happened, and that the situation deviates from what it should be. The operator must further identify or diagnose the situation.
- *Decision making*, where the operator must select an appropriate action to carry out, based on the preceding identification of the problem.
- *Execution*, where the operator must execute or perform the chosen action. The execution must be correct, i.e., according to the prescribed procedure or established practice.
- *Recovery*, which offers the operator the possibility of determining whether the action had the expected effect. If not, the action may have been incorrect, and the operator may have a chance to recover, i.e., correct it, provided the nature of the process and the characteristics of the system allow that.

This decomposition can be described graphically as shown in Figure 3. (The additional distinction between *response + recovery* is dictated by the needs of a PSA.) The approach illustrated by Figure 3 limits the decomposition to three segments, being *identification, decision*, and *execution*. This is pragmatically speaking the smallest number of segments that makes sense, although several information processing models offer a far greater variety. Given the uncertainty about the nature of human cognition, not least in the light of the current demise of the information processing paradigm, there may actually be little reason to go into further detail. The apparent gain from a larger number of details will soon be lost in the increased uncertainty.

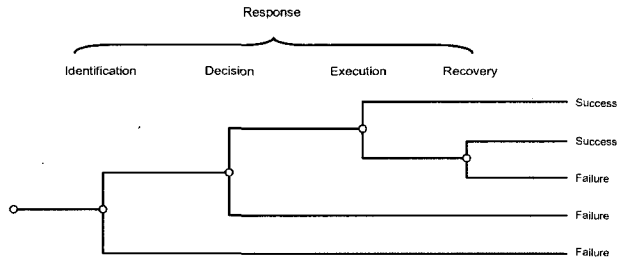

Fig. 3. Decomposition of Manual Interventions.

Although the hope for a while existed that the development of cognitive models of human performance would meet the needs of HRA and PSA [16], the experience from practice soon damped the optimism. On balance, the convenient use of the binary event tree to describe the details of responses and cognitive functions turned out to go against the common understanding of the nature of human action. It resulted in a description that was computationally simple but psychologically unrealistic. Quite apart from that, it was inherently difficult to find basic error probabilities for the functions that were a result of the decomposition, e.g., identification, decision, and execution (or even recovery). The need for this kind of data was in fact determined by the way in which the decomposition was made, hence reflected the structure of the underlying model rather than the requisite variety.

A further setback came from an insightful but critical review of what later became known as first-generation HRA approaches [2]. This criticism pointed out a number of significant shortcomings of HRA (see also [21]) and additionally threw doubt on the role of HRA as a part of PSA. This led a number of people to question whether HRA should try to go beyond the PSA-*cum*-HRA construct. One reason was the need to investigate the larger perspective where humans are involved in the design and construction of a system, in the operation and maintenance, and in the management [22]. Human reliability can obviously play a role in every phase of a system's life-cycle, although the outcome of action failures in many cases may not be immediately visible. It was therefore necessary to develop a comprehensive understanding of human action in context, and that in turn created a need to revise existing HRA approaches.

The outcome of this revision was not just the development of second-generation HRA methods such as ATHEANA [23], CREAM [9], or MERMOS [24], but also a questioning of the assumptions about the nature of human performance that were the foundation of both accident and risk analysis. Although this revision has not yet been completed, it is now commonly acknowledged that humans should not be considered as mere components of systems. Indeed, most systems of interest are themselves too complex for a linear description. In cases where performance fails, when the outcome is worse than expected, we therefore worry and try to look for the cause. Yet in cases where the opposite is the case, i.e., when systems work better than expected, we gratefully accept the good fortune but rarely try to understand why. (A recent example is the success of the two Mars rovers Opportunity and Spirit that were designed to work for 90 days but so far has worked for more than a year.) From a psychological perspective, the underlying phenomenon is nevertheless the same, namely the inherent variability of human performance. It is therefore a serious mistake to try to model the negative outcomes only, not least if it is done in terms of simple 'mechanistic' models of human information processing. The focus should instead be on how to understand the nature of human performance

variability and eventually how to describe and analyse it.

2. SIGNAL AND NOISE

Due to the influences described above, the common approaches to performance predictions have focused human performance failures. Performance prediction, as practised by HRA, has confined itself to an investigation of the ways in which individual actions can possibly fail. In doing so, the likelihood of failure has been seen as an attribute of human actions *per se*, usually expressed in terms of a 'human error probability' (HEP). This is quite consistent with the information processing view, where specific internal error mechanisms are assumed to exist – and where such constructs furthermore are rather easy to invent on an *ad hoc* basis [25]. The logic seems to go along these lines: if a function can be seen as an attribute of a component, it follows that the possibility of function failure can be considered for the component by itself, although it is usually acknowledged that the circumstances or context may have some influence. In HRA the circumstances have from the very beginning been encapsulated by a set of performance shaping factors [26], which were assumed to exert their influence in a simple, additive fashion. Yet the likelihood of a component function failure, i.e., a 'human error', was calculated or assessed prior to, hence independent of, the effects of the performance shaping factors.

2.1 'Human Error' As a Signal

HRA has in common with many accident analysis methods the assumption that it is reasonable to consider the inherent variability of human performance by itself, hence that a performance failure is an attribute of the human component rather than of the circumstances during which actions take place. In this sense the 'human error' is – metaphorically, at least – the sought for signal rather than the noise. This assumption is strangely inconsistent with one of the main tenets of the information processing approach, which states that:

A man, viewed as a behaving system, is quite simple. The apparent complexity of his behavior over time is largely a reflection of the complexity of the environment in which he finds himself. [27]

If this tenet was accepted as the basis for risk analysis and the anticipation of human performance failures, then the focus would be on the variability of the environment or circumstances and not on the possibility of a failure of the human component. In other words, the possibility of failure would be an attribute of the context and not of the human. More recently, a similar notion has been expressed specifically addressing the issue of error management:

The evidence from a large number of accident inquiries indicates that bad events are more often the result of error-prone situations and error-prone activities, than they are of error-prone people. [12]

It is, indeed, the general experience from the analysis of accidents in a wider sense is that they usually are due to the combination of a number of factors, rather than to single causes. Maintaining the notion of 'human error' as a central concept in HRA furthermore disregards the fact that performance usually is the outcome of the activities of a team rather than of an individual. This alone is a compelling reason to look for ways of addressing the PSA need to describe human performance reliability without making individual 'human errors' the pivotal concept. The validity of this argument has been recognised by the second-generation HRA approaches, where some of the better known methods emphasise that the likelihood of something being done incorrectly is determined by the performance conditions rather than by inherent 'human error' probabilities. Despite this, many practitioners of HRA blissfully continue to treat 'human error' as a meaningful concept and to suggest new ways to bestow the elusive 'human error' probability with a solid empirical basis.

2.2 'Human Error' As Noise

Since it does not make much sense to think of an action or of an action failure without a context, and since the context often may be the 'error forcing condition' that leads to the failure, it seems reasonable to consider whether the coveted 'error probability' can be determined directly from a characterisation of the context. This would first of all render irrelevant the question of whether the failed action was of an individual or of a team. It would furthermore put into focus that it is performance as a whole that fails or is unsuccessful, and that we should seek the likelihood of this rather than the probability that a specific type of action goes wrong.

Interestingly enough, a number of HRA methods indirectly support this view. The classical principle of time-reliability correlation (TRC [28]) is an expression of the idea that the likelihood of failing in performing an activity is a function of time – although in this case it is time after the onset of an accident rather than time available as such. A more sophisticated version of the same principle is found in the notion of 'error forcing conditions', where a determining factor is time available rather than elapsed time [23]. The sophistication is due both to the set of conditions that may 'force' an error and the more detailed description of possible error modes. The common feature is that the possibility of performance failure is an attribute of the conditions rather than of the humans.

A closer inspection of a well-known HRA method (HEART, [29]) also reveals the dominance of the circumstances over the individual. Firstly, HEART refers to the possible failure of an action, but not to specific failure types. Secondly, the characterisation is related to different tasks, which actually means different task conditions. This can be substantiated by a gentle

Table 1. Description of Failure Types and Causes in HEART.

Generic tasks	Context or set of circumstances
Totally unfamiliar, performed at speed with no idea of likely consequence.	High time pressure, unfamiliar situation
Shift or restore system to a new or original state on a single attempt without supervision or procedures	Lack of supervision and procedures
Complex tasks requiring high level of comprehension and skill	High task complexity
Fairly simple task performed rapidly or given scant attention.	Simple tasks of limited significance
Routine, highly-practised, rapid task involving relatively low level of skill.	Routine or highly familiar tasks
Restore or shift system to original or new state following procedures, with some checking .	Following a procedure
Completely familiar, well-designed, highly practised routine task, oft-repeated and performed by well-motivated, highly trained individual with time to correct failures but without significant job aids.	High-routine task with no time pressure
Respond correctly to system event when there is an augmented or automated supervisory system providing accurate interpretation of system state.	Task with monitoring and highly supportive MMI
Miscellaneous tasks for which no description can be found.	No specific characteristics

reinterpretation of the basic HEART table, as shown in Table 1.

Even if the objectivity of the reinterpretation may be disputable, it is a demonstrable fact that the major source of variability, which determines the likelihood of a failure, is ascribed to the context or circumstances. In other words, the specific working conditions are the signal while the individual HEP is the noise. The possibility of performance failure is thus an attribute of the conditions rather than of the humans.

2.3 Failures without Errors

The consequence of this line of argument is that the variability of human performance constitutes the noise rather than the signal. Conversely, the main determinant of performance quality – and specifically of performance failure – comes from the context or the circumstances. The possibility of failure thereby becomes an attribute of the joint human-machine system rather than of any of its components [30]. The anticipation of system failures should consequently concentrate on developing effective ways of describing how joint system performance depends on the conditions rather than on the potential for human failures.

Specifically, predictions should be about how the joint system can lose control of the situation, rather than about whether the human will make an isolated failure. This would also acknowledge the fact that a human failure is just a single event that requires other conditions to result in an accident rather than a sufficient cause by itself. A practical implementation of this principle can be found in the basic method for performance prediction that is part of CREAM [9]. Here an assessment of the common performance conditions leads to an overall prediction of how likely the operator, hence the joint system, is to lose control. This prediction is made without considering the failure probability for specific actions, or even describing the tasks at the level of component actions. In an application of this approach, [31] developed a systematic process of calculating mean failure rates as a function of Common Performance Conditions, but without making any assumptions about individual human actions. The method provided an efficient way of screening various scenarios, thereby limiting the efforts needed to carry out the more detailed analyses. The authors concluded that:

In terms of data needs, the acknowledgement of the importance of the performance conditions means that there is little reason to conduct massive data collection exercises on the level of individual performance. Instead of using human performance characteristics as the starting point for speculations about internal 'failure mechanisms', models should be developed of how working conditions may influence the way in which people adjust their actions to make ends meet. By taking the performance of a joint human-machine system as the unit of analysis, discussions about the influence of organisational factors are also given a new meaning, since the organisation obviously is but one of several constituents of the context. [31]

Yet another reason for ditching 'human error' is that there must be a symmetry between analysis and prediction. In the field of accident analysis, the development has gone from sequential models based on simple cause-

effect chains, over epidemiological models that can account for the effects of latent factors, to systemic models that explain accidents as emergent phenomena [8]. Yet in risk analysis and HRA, there has been no comparable development. The prediction of how humans and technological systems can fail is still mainly based on the sequential models that accident analysis has generally abandoned.

3. CONCLUSIONS

Anticipating failures of joint human-machine systems requires an underlying model. This should not be a model of human information processing in disguise, but a model of how human performance is determined by – hence reflects – the context or circumstances, i.e., a model of joint system performance rather than of individual human actions. This type of model corresponds to the notions of distributed or embedded cognition [32], although neither of these have been used to consider performance prediction specifically. A concrete expression of these ideas is found in the contextual control models [9], which emphasises human-machine co-operation (coagency) rather than human-computer interaction.

Traditional PSA aims to calculate the probability that a component or subsystem will fail. When HRA is carried out in this context, the corresponding question becomes what the probability is that a human operator will make an 'error'. Yet if we realise that 'human error' is a consequence of performance variability, we must also realise that performance always is variable due to the underspecification of the work environment. The question therefore does not have a meaningful answer, hence should not be posed. Indeed, it may be argued that the concept of 'human error', philosophically speaking, is a category mistake.

An alternative is to adopt a systemic perspective and ask what the probability is that an event will get out of control. When HRA is carried out in this context, the corresponding question becomes when concurrencies of human and system performance variability will occur[2]. This question can be given a meaningful answer by looking to the descriptions of concurrence that have been developed by the epidemiological and systemic types of accident models [8].

The conclusion is therefore that HRA is of limited value as an input generator for PSA, mainly because it harbours an oversimplified conception of human performance. There are better and more realistic ways of analysing risks, both qualitatively and quantitatively. Indeed, in the light of the development of systemic accident models it may well be asked whether PSA itself has not become inadequate for its purpose. But that is another debate for another time.

[2] Concurrency: the temporal property of two or more things happening at the same time.

REFERENCES

[1] Kirwan, B. A guide to practical human reliability assessment. London: Taylor & Francis, 1994.

[2] Dougherty, E. M. Jr. Human reliability analysis - Where shouldst thou turn? *Reliability Engineering and System Safety*, 29, 3 (1990), pp. 283-299.

[3] Woods, D. D., Johannesen, L. J., Cook, R. I. & Sarter, N. B. Behind human error: *Cognitive systems, computers and hindsight. Columbus*, Ohio: CSERIAC, 1994

[4] Hollnagel, E. & Amalberti, R. The Emperor's New Clothes, or whatever happened to "human error"? 4th International Workshop on *Human Error, Safety and System Development*, June 11-12, 2001, Linköping, Sweden.

[5] Rochlin, G. I. Safe operation as a social construct. *Ergonomics*, 42, 11 (1999), pp. 1549-1560.

[6] Woods, D. D. & Cook, R. Nine Steps to Move Forward from Error. *Cognition, Technology, and Work*, 4, 2 (2002), pp. 137-144.

[7] Swain, A. D. *Comparative evaluation methods for human reliability analysis* (GRS-71). Köln, Germany: Gesellschaft für Reaktorsicherheit (1989).

[8] Hollnagel, E. *Barriers and accident prevention*. Aldershot, UK: Ashgate Publishing (2004).

[9] Hollnagel, E. *Cognitive reliability and error analysis method*. Oxford, UK: Elsevier Science Ltd. (1998).

[10] Perrow, C. Normal accidents: *Living with high risk technologies*. New York: Basic Books, Inc. (1984).

[11] Reason, J. T. The contribution of latent human failures to the break down of complex systems. *Philosophical Transactions of the Royal Society (London), Series B*. 327 (1990), pp. 475-484.

[12] Reason, J. T. *Managing the risks of organizational accidents*. Aldershot: Ashgate Publishing Limited (1997).

[13] Dougherty, E. M. Jr., & Fragola, J. R. Human reliability analysis. A systems engineering approach with nuclear power plant applications. New York: John Wiley & Sons (1988).

[14] Heinrich, H. W. Industrial accident prevention. McGraw-Hill (1931).

[15] Miller, D. P. & Swain, A. D. *Human Error and Human Reliability*. In G. Salvendy (Ed.) Handbook of Human factors. New York: Wiley (1987).

[16] Rasmussen, J. Trends in human reliability analysis. *Ergonomics*, 28, 8 (1985), pp. 1185-1195.

[17] Parry, G. W. & Mosleh, A. *Control room crew operations research project* (EPRI TR-105280). Palo Alto, CA: Electrical Power research Institute (1995).

[18] Gertman, D. I., Gilmore, W. E. & Ryan, T. G. *NUCLARR and human reliability: Data sources and data profile*. Conference Record for 1988 IEEE Fourth Conference on Human Factors and Power Plants, June 5-9, 1988, Monterey, CA. (pp. 311-314).

[19] Hannaman, G. W., Spurgin, A. J. & Lukic, Y. D. *Human cognitive reliability model for* PRA *analysis* (NUS-4531). Palo Alto, CA: Electric Power Research Institute (1984).

[20] Yoshikawa, H. & Wu, W. An experimental study on estimating human error probability (HEP) parameters for PSA/HRA by using human model simulation. *Ergonomics*,

42, 11 (1999), pp. 1588-1595.
[21] Swain, A. D. Human reliability analysis: Need, status, trends and limitations. *Reliability Engineering and System Safety*, 29 (1990), pp. 301-313.
[22] Hollnagel, E. & Wreathall, J. HRA at the turning point? In P. C. Cacciabue & I. Papazoglou (Eds.), *Probabilistic safety assessment and management* '96. Berlin: Springer Verlag (1996).
[23] Cooper, S. E., Ramey-Smith, A. M., Wreathall, J., Parry, G. W., Bley, D. C., Luckas, W. J., Taylor, J. H. & Barriere, M. T. A *technique for human error analysis* (ATHEANA) (NUREG/CR-6350). Washington, DC: US Nuclear Regulatory Commission (1996).
[24] Bieder, C., Le Bot, P., Desmares, E., Bonnet, J.-L. & Cara, F. MERMOS: EDF's new advanced HRA method. *PSAM 4*, New York, pp. 129-134, Springer-Verlag, London (1998).
[25] Dekker, S. W. A. & Hollnagel, E. Human factors and folk models. *Cognition, Technology & Work*, 6 (2004), pp. 79-86.
[26] Swain, A. D. & Guttman, H. E. Handbook of human reliability analysis with emphasis on nuclear power plant applications (NUREG CR-1278). Washington, DC: NRC (1983).
[27] Simon, H. A. *The sciences of the artificial.* Cambridge, MA.: The M. I. T. Press (1972).
[28] Hall, R. E., Fragola, J. & Wreathall, J. Post event human decision errors: *Operator action tree / time reliability correlation* (NUREG/CR-3010). Washington, DC: U. S. Nuclear Regulatory Commission (1982).
[29] Williams, J. C. A *data-based method for assessing and reducing human error to improve operational performance.* Proceedings of IEEE 4th Conference on Human factors in Power Plants, Monterey, CA, 6-9 June 1988.
[30] Hollnagel, E. & Woods, D. D. Joint cognitive systems: Foundations of cognitive systems engineering. Boca Raton, FL: CRC Press (2005).
[31] Fujita, Y. & Hollnagel, E. Failures without errors: Quantification of context in HRA. *Reliability Engineering and System Safety*, 83, 2 (2004), pp. 145-151.
[32] Hutchins, E. *Cognition in the wild.* Cambridge, MA: MIT Press (1995).