QPI 2021, volume 3, pp. 120-131

IMPROVING OCCUPATIONAL SAFETY MANAGEMENT WITH THE USE OF HUMAN RELIABILITY ANALYSIS METHODS

doi: 10.2478/cqpi-2021-0012

Date of submission of the article to the Editor: 15/07/2021
Date of acceptance of the article by the Editor: 20/09/2021

Joanna Tabor – *orcid id: 0000-0002-7746-2970*
Częstochowa University of Technology, **Poland**

Abstract: The aim of this work is to review the basic issues and problems related to the analysis of human reliability in terms of the possibility of using the methods functioning in this area to improve occupational safety management. Methods of human reliability analysis (HRA) are related to the prediction and assessment of system failures, which are the result of incorrect actions or omissions by a human, and not the failure of a physical element in the system. The paper presents the significance of the problem of human reliability from the point of view of accidents; describes HRA as a process with three main components; characterizes human errors using various criteria as well as discusses the generations of HR analysis and evaluation methods, including factors influencing the performance. The basic guidelines for the selection of methods for reliability analysis were also presented in terms of their possible use for improving the health and safety management system.
Keywords: management system, OHS, human reliability, HRA methods

1. INTRODUCTION

The main goal of the management system improvement is to perfect the effects of its operation by making the desired changes in the right place of the system and at the adequate time. Improvement actions can be both small improvements, introduced continuously to selected elements of the system, and large changes within the entire system (Denton, 1982; Law et al., 2006, Cadiuex et al., 2006). It is important for the improvement process that the changes are based on a well-thought-out assessment of the situation using the right information and data, i.e. that they are based on a properly conducted assessment (CEN, 2009, Tabor, 2019).
The currently observed development of work processes aimed at increasing the activities related to operational service as well as the increase in the complexity of the supported technical, anthropotechnical and sociotechnical systems mean that the analysis of human reliability should be an important part of any risk or reliability analysis of each system, not only technical, but also anthropotechnical and social engineering. And, it is not only about predicting possible human failures and assessing their probability to counteract potential negative effects in the system, but also about the possibility of acting towards the improvement of such systems. The most general

definition of an occupational health and safety management system (OH & SMS) describes it as part of an overall management system that is concerned with the development, implementation, execution, review and maintenance of an organization's occupational health and safety policy (ILO, 2013; Klimecka – Tatar and Niciejewska, 2016). In practical terms, these are specific activities carried out by definite people, under certain contextual conditions, and the result of these activities may be different than intended due to many different human errors.

2. THE IMPORTANCE OF THE PROBLEM OF HUMAN RELAIBILITY

The rapid development of technology has led to the reduction of accidents caused by technical errors. As a result of the use of redundancy and technical prevention measures, the technical elements of the systems have become more reliable. At the same time, research shows that if the technical components of the systems become more safe, the system as a whole may become less safe, as people tend to have less supervision over such systems (Hollnagel, 1993).

Human is a part of every system, regardless of the stage in which this system functions (hence it is referred to as anthropotechnical or social engineering systems). Therefore, when analyzing the reliability of the system, one should take into account the risks of failure that are associated with each of the components of the system, including humans. Human element errors / failures significantly affect the other components and determine the performance of the system as a whole.

The analysis of the causes of accidents shows that human error is the main risk factor in the nuclear industry (Reason, 1990) - over 90%, in the chemical and petrochemical industry (Kariuki and Loewe, 2007) - over 80%, in marine accidents (Ren et al., 2008) - over 75% and in air accidents (Helmreich, 2000) - over 70%. The role of man in the dynamics of accident events cannot be ignored, because it is a person who can both lead to an accident and reduce the negative consequences of the events that have occurred.

Various definitions of human reliability can be found in the literature. According to the first exemplary definition, human reliability is a function of correspondence between the operator's internal factors (e.g. temperamental traits, personality, motivation, qualifications) and external factors (e.g. task features, technical and operational quality of machines, work methods, physical environmental conditions, social climate, atmosphere at work) (EN-62508, 2010).

At the same time, human reliability can be defined as the accurate, effective and error-free performance of tasks within the anticipated operating time, both under optimal and extreme conditions. Human reliability can be operational and biological. In this context, human operational (labor) reliability is measured by the probability of success during the performance of a job or a task being performed at a given stage of the system functioning, within a given period of time. Whereas, a measure of biological reliability is the probability of maintaining the ability to operate in a given period of time and under given conditions (Reason, 1990).

3. HUMAN RELIABILITY ANALYSIS AS A PROCESS

Human reliability analysis (HRA) is a systematic process of studying and evaluating human reliability. This analysis may be retrospective or prospective. Both qualitative and quantitative simulations are used in the analysis. Retrospective analysis is the assessment of events involving interpersonal interactions, such as accidents, incidents

or near misses. Events are analyzed and assessed in order to search in detail for the underlying cause, facts and causes that have promoted and contributed to inappropriate human behavior.
Instead, prospective analysis includes the assessment and prediction of the consequences of human-system interactions, taking into account initiating events, internal and external conditions, and the configuration of system boundaries.
Each process of analyzing and assessing human reliability consists of three main components: (a) human error identification - what errors can happen, (b) human error quantification - what is the probability of these errors, and (c) human error reduction - how the way you can improve human reliability.
The measure for quantifying human reliability is the human error probability - HEP, which is described by the formula (ANSI / ASME RA-S-2002):

$$HEP = \frac{the\ number\ of\ errors\ (mistakes)}{the\ number\ of\ opportunities\ to\ make\ a\ mistake} \quad (1)$$

The formula concerns an activity of a discrete type and refers to a specific error of one type. The probability of human error ranges from 0 to 1.

4. HUMAN ERRORS IN THE CONTEXT OF VARIOUS CLASSIFICATIONS

The guide on human aspects of dependability defines "human error" as the discrepancy between an action taken or omitted and an intended action. At the same time, this guide defines the concept of "human failure", which is understood as a deviation (variation) from the action required to achieve the goal, regardless of the reason for this deviation (variation) (EN 62508, 2010).
Since the 1980s, many different classifications of human error have been developed (see Rasmussen, 1980; Norman, 1981; Swain and Guttman, 1983; Williams, 1986, Whalley, 1988; Reason, 1990; Hollnagel, 1993; Stewart and Melchers, 1997; Gertman et al., 2004). These classifications are based on grouping errors according to specific rules. Grouping rules result from different approaches to the concept of human error. At the same time, classifications make it possible to understand the underlying causes of errors. The classifications are necessary because they not only facilitate the identification of errors, but also facilitate the selection of measures to protect against the identified errors.
The most general classification distinguishes errors that do not produce a threat and the so-called dangerous errors that have a strong relationship with the hazard and result in breakdowns, accidents and other, especially undesirable consequences. At the same time, there are errors with immediate effects, the so-called active errors (e.g. violation of health and safety regulations, mistakes, etc.) and errors with delayed effects, the so-called latent errors (e.g. defects of safety system flaws, improperly organized tasks, etc.).
Nowadays, the classification of errors according to the SRK model is very important (Rasmussen, 1980; Dougherty, 1992), which is based on the distinction of three groups of human behavior: behavior based on skills (S), behavior based on rules (R) and behavior based on knowledge (K). The term "Skill-based behavior" should be understood as meaning more or less subconscious actions resulting from remembered patterns of behavior. "Rule-based behavior" requires more mental effort to follow memorized or written instructions.

Instead, "knowledge-based behaviors" are activities in situations that are not routine and not fully known, in which processes related to cognition and decision-making play an important role.
Another important classification of errors is External Error Modes (EEMs) (Swain and Guttmann, 1983), which divides errors into three types: skip errors, runtime errors, and so-called irrelevant activities. Skip errors are, for example, skipping an entire task or skipping a step in the task. However, execution errors include: selection errors (choosing the wrong object, choosing the wrong position of the object, issuing the wrong command, sending the wrong information), sequence errors (execution of actions in the wrong order), runtime errors (too early, too late) and errors qualitative (too little, too much).
Also popular, due to its application in the analysis of human reliability, is the SLM (Slip-Lapse-Mistake) classification (Reason, 1990). According to this classification, dangerous human activities may be intentional and unintentional. Unintentional errors include mistakes (attention errors), omissions (memory errors), errors due to lack of rules - incorrect application of good rules or application of bad rules, and errors due to lack of knowledge - unprepared solution, new situation requires thinking about the solution from scratch. By contrast, intentional errors include routine violations (habitual deviations from normal practices), exceptional violations (unusual violations caused by extreme conditions), situational violations (unusual violations caused by local conditions) and acts of sabotage.
From the point of view of improving management, the most important mistakes are made by managers at various levels of management. These errors result in unsafe working conditions and / or unsafe employee behavior, which in turn lead to accidents and failures.

5. GENERATIONS OF HUMAN RELIABILITY ANALYSIS METHODS

Various tools that can be used to analyze human reliability are presented in the literature. They are generally referred to as methods, but there are also some fairly simple techniques such as pairwise comparison technique.
The main purpose of using these tools is to assess the risk of human error / failure in the implementation of specific activities. Human reliability analysis methods were developed for Probabilistic Risk Assessment (PRA) to enable human error quantification in an accident event investigation. Therefore, these methods can be treated as PRA specialization in the area of important factors of human reliability. The rapid development of methods, especially at the end of the last century, was associated with serious accidents such as Seveso (1976), Three Mile Island (1979), Bhopal (1983) and Chernobyl (1986) in which human failure played a key role.
The methods of human reliability analysis can be divided primarily in terms of their general availability. According to this classification, generally available methods and unavailable methods have been distinguished. For generally available methods, the procedure and practical examples of the method's application can be found in scientific and / or industry literature, including on the Internet. However, non-available methods have been developed specifically for a specific industry (company) and their use is legally protected.
Incorporating a method into a specific generation is the second important classification. Although methods of analyzing human reliability have been developing since the 1980s, the chronology is not an appropriate basis for assigning them to a specific generation

(Boring and Gertman, 2005). The methodological approach is important in this classification. And in this context, the methods of the 1st generation, 2nd generation and 3rd generation have been distinguished. Instead, expert methods constitute a separate group (Tab. 1)

Table 1
Examples of human reliability analysis methods by generation

Publicly available methods	
1st generation	THERP - Technique for Human Error Rate Prediction (Swain and Guttmann, 1983; Kirwan, 1996) ASEP - Accident Sequence Evaluation Program (Swain, 1987; Everdij and Blom, 2008) HEART - Human Error Assessment and Reduction Technique (Williams, 1992, Kirwan et al., 1997) SPAR-H - Standardized Plant Analysis Risk HRA (Gertman et al., 2004, Forester et al., 2006) TESEO – Tecnica Empirica Stima Errori Operatori (Bello and Colombari, 1980)
2nd generation	HCR – Human Cognitive Reliability (Hannaman and Spurgin, 1984) ATHEANA - A Technique for Human Error Analysis (USNRC, 2000, Forester et al., 2004) CREAM - Cognitive Reliability and Error Analysis Method (Hollnagel, 1998; Kirwan, 1998)
Expert judgement	APJ - Absolute Probability Judgement (Seaver and Stillwell, 1983; Williams, 1989) PC - Paired Comparisons (Hunns, 1982, Lyons et al., 2004) SLIM - Success Likelihood Methodology (Embrey, 1983; Boring and Gertman, 2005)
Not publicly available methods	
1st generation	HRMS - Human Reliability Management System JHEDI - Justification of Human Error Data Information
2nd generation	CAHR - Connectionism Assessment of Human Reliability CESA - Commission Errors Search and Assessment CODA - Conclusion from Occurrences by Description of Actions MERMOS - Methode d'Evaluation de la Realisation des Missions Operateur pour la Surete
3rd generation	NARA - Nuclear Action Reliability Assessment

Source: own elaboration based on Bell and Holroyd, 2009

According to the human aspects of dependability guide, in the first generation methods (EN-62508, 2010) human error is handled in the same way as equipment failure, and the output from human activities corresponds to the output from the equipment. Moreover, every human action is treated in a binary way, that is, as "success" or "failure" in achieving the required result of the task.

At the same time, each human task has a specific probability of failure, which is modified by the factors shaping action - PSF (Performance Shaping Factors). These factors are related to the assessment of the widely understood ergonomic environment.

The methods included in the first generation differ in the method of estimating the probability of human error (HEP) and taking into account various factors influencing the performance.

The concept of factors influencing action (PSF) is generally understood as the features of the external environment, the features of the task and the features of people that shape the individual action of a human being.

There are many different classifications of PSFs in the literature, because the models and methods of testing human reliability are various. Each of the PSF classifications can form the basis of a different checklist for assessing the environment in which a specific activity is carried out.

According to the EN-62508 (2010) standard, the results of human activity are influenced by external and internal factors influencing activity. The group of external factors includes organizational prerequisites (organizational structure and organizational dynamics) and technical prerequisites (task difficulty and situational factors). On the other hand, the group of internal factors includes performance capacity (physiological capacity and psychological capacity) and willingness (physiological fitness and psychological motivation).

The classification of L.J. Bellamy and T.A.W. Geyer (Bellamy and Geyer, 1992) distinguishes eight groups of factors influencing human performance: (1) stressors (e.g. time pressure, monotony, isolation, fatigue, etc.), (2) social engineering factors (e.g. team structure, communication, authority , etc.), (3) instructions and procedures (e.g. precision, legibility, adequacy, ease of use, etc.), 4) human-machine interface (e.g. location, identification, ease of use, readability, etc.), (5) task characteristics (frequency, repeatability, criticality, etc.), (6) individual factors (capabilities, skills, experience, personality, etc.), (7) environment (temperature, noise, lighting, etc.) and (8) task requirements (perceptual, physical, memory, etc.).

Also, eight groups of factors are included in the classification made by D. Gertman et al. (Gertman et al. 2004). The following groups were distinguished: (1) available amount of time, (2) stress and stressors, (3) experience and training, (4) complexity, (6) ergonomics, (7) procedures and (8) fitness for duty.

On the other hand, the classification made by E. Hollnagel (Hollnagel, 1993) covers nine basic groups of factors: (1) availability of procedures, (2) crew coordination quality, (3) adequacy of organization, (4) number of goals, (5) time of day, (6) adequacy of Human-Machine Interface, (7) available time, (8) working conditions, and (9) adequacy of training.

The TWIN Analysis Matrix (DOE-HDBK-1208-2012) i.e. the classification of the precursors of erroneous results of human activities should also be mentioned. The TWIN model includes four groups of fault precursors: (1) Task Demands, (2) Work Environment, (3) Individual Capabilities, and (4) Human Nature.

In the "Task requirements" group there are, for example: Time pressure (in a hurry), High workload (large memory), Simultaneous, multiple actions and Repetitive actions / Monotony. The “Work environment” group includes: Distractions / Interruptions, Changes / Departure from routine, Confusing displays or controls and Hidden system / equipment response. The group "Individual skills" includes: Unfamiliarity with task / First time, Lack of knowledge (faulty mental model), New techniques not used before and Lack of proficiency / Inexperience. In contrast, the group "Human nature" includes, for example: Stress, Habit patterns, Inaccurate risk perception and Limited short-term memory.

In contrast to the first generation methods, the second generation methods (EN-62508, 2010) take into account the influence of context on human decision-making behaviors, which behaviors can have an adverse effect on the system. The key problem in the

second generation methods is the appropriate modeling of the context, i.e. the situational conditions in which the decision-making process takes place.
Whereas, in the third generation methods, the main focus is on the possibilities of limiting the effects of incorrect / erroneous human activities or early correction of these effects already during the implementation of the activity. These methods are developing towards modeling human behavior in order to recover and avoid undesirable effects. These methods take into account the human ability to regenerate, respond to the unexpected, and think outside the box. Therefore, these methods are most often specialized and dedicated to a specific industry (aviation, nuclear industry, shipping, etc.).
Availability and generation are not the only possible guidelines for the classification of human reliability analysis methods. From a practical point of view, the following guidelines are also important: (1) key parameters appearing in the analysis, (2) the level of detail of the analysis, and (3) the data scale used. Table 2 presents examples of methods that can be distinguished within each of the above-mentioned classifications.

Table 2
Classifications of methods of human reliability analysis

Classification	Examples of methods
Due to the key parameters used in the analysis	- error related methods: e.g. THERP, ASEP - time reliability related methods: e.g. HCR - PSF related methods: e.g. SLIM, HEART
Due to the level of detail of the analysis	- decomposition methods: e.g. THERP, ASEP - holistic methods: e.g. HCR, SLIM, HEART
Due to the type of data scale in the analysis	- methods with a relative (nominal) scale of the data: e.g. THERP, ASEP - methods with an ordinal scale of data: e.g. HCR, SLIM, HEART - methods without scale: e.g. SHARP

Source: own elaboration based on NEA/CSNI/(98)1

It is impossible to indicate which method is the best, each method has specific advantages and limitations. A given method may be more or less appropriate depending on (1) the context in which it is used; (2) the available resources (financial, human, time, etc.); and (3) the skills of the analysts.

6. GUIDELINES FOR THE SELECTION OF METHODS FOR HRA

In the literature, there are comparative studies of the most commonly used methods of human reliability analysis (e.g. Kirwan, 1996, Boring et al., 2005), but there are too few of these studies to prepare a complete and consistent comparison of the methods. However, guidelines can be presented that should be followed when selecting a method for testing and assessing human reliability. These are: (1) Accuracy and repeatability of results; (2) Modeling fidelity; (3) Utility; (4) Use of resources; (5) Acceptability of the method; (6) Maturity of the method; (7) Simplicity of the method; (8) Labor intensity of the method, as well as (9) Access to or collection of nominal HEP data (Humphreys, 1988). Table 3 contains the characteristics of individual criteria.

Table 3
Criteria for the selection of methods for the human reliability analysis

Criterion	**Description**
Accuracy and repeatability of results	Accuracy of the final HEP prediction (agreement with empirical values); repeatability of the results in subsequent analyzes of the same situation by different people.
Modeling fidelity	Accuracy to reflect the situation (task) and factors influencing human reliability; compliance with theoretical knowledge about human behavior in controlling technical systems; the extent to which the method appears convincing to the potential user; compliance of the results with those obtained by other methods.
Utility	Possibility to formulate recommendations as to measures to improve human reliability in the analyzed situation; the ability to analyze the sensitivity of the result to changes in input data; the scope of the method's applicability in various industries; the scope of the method's applicability in various types of tasks, behaviors and mental activities of the operator.
Use of resources	Requirements in terms of the number of people needed for the analysis, their availability, time and service needed to complete the analysis; requirements for the quantity and form of input data (qualitative and quantitative); the ability of the method to provide reliable results despite incomplete data; the degree of decomposition needed to perform the analysis; requirements for the qualifications, training and skills of the people involved in the analysis.
Acceptability of the method	Acceptance of the method by safety regulators, by the scientific community as well as by the evaluating experts; the ability to audit and control the correctness of analyzes carried out; the possibility of supervision of analyzes carried out by external experts.
Maturity of the method	Current technical maturity of the method, proven in its practical application; the ability of the method to develop further in terms of meeting one or more of the criteria in question to a higher degree.
Simplicity of the method	Necessary skills of the assessor, his professional preparation, type and level of education; professional practice in the field from which the analyzed problem originates; experience in using methods of human reliability analysis; computer skills and computing software.
Labor intensity of the method	The amount of data that needs to be prepared to perform the analysis; the degree of complexity of the work needed to perform analyzes; the amount of "manual" work required for the analysis; the ratio of workload to the achieved effects.
Access to or collection of nominal HEP data	Availability of input data (HEP baseline values) for the considered types of tasks; the possibility of obtaining this data from expert assessments or external sources (literature, computer databases); the possibility of experimental measurements in conditions similar to real ones.

Source: Own study based on (Humphreys, 1988)

Based on the method descriptions available in the literature (Kirwan, 1998, Stanton et al., 2005; Forester et al., 2006, Everdij and Blom, 2008), taking into account in particular their advantages and limitations, a subjective assessment of the basic, most often relevant methods was performed in terms of the analyzed criteria:
(1) Accuracy and repeatability of results: THERP > SLIM > HEART > HCR > TESEO
(2) Modeling fidelity: THERP > SLIM > HEART > HCR > TESEO

(3) Utility: SLIM > HEART > TESEO > THERP > HCR
(4) Use of resources: THERP > SLIM > HEART > HCR > TESEO
(5) Acceptability of the method: THERP > SLIM > HEART > HCR > TESEO
(6) Maturity of the method: THERP > SLIM > HEART > HCR > TESEO
(7) Simplicity of the method: HEART > SLIM > TESEO > HCR > THERP
(8) Labor intensity of the method: THERP > HEART > SLIM > HCR > TESEO
(9) Access to or collection of nominal HEP data: THERP > HCR > HEART > SLIM > TESEO

This means that from the point of view of 7 out of 9 evaluation criteria, the THERP method is the most appropriate. However, there may be times when a single methodology may not be appropriate for all contexts and elements in the analysis being performed. Therefore, you may need to use two or three different methods.

7. SUMMARY AND CONCLUSIONS

Human reliability analysis is a natural and required supplement to technical risk analysis in anthropotechnical systems. In addition, effective human reliability analysis helps organizations develop a safety culture and improve risk management. Human action is complex and very often counterintuitive, while recognizing the possibility of employees making a mistake can help improve the system design, organizational structure or processes.

Although the key goal of human reliability testing and evaluation is to develop a system that is resistant to human errors, it is also very important to prepare people to work in the complex and changing environment that is characteristic of contemporary social engineering systems.

Nowadays, the literature shows a departure from the classical understanding of the concept of "error", which is associated with assigning blame to someone. It seems more reasonable to use the concept of "behavior" when referring to an individual, group or organization. Talking about inappropriate behavior does not carry such negative connotations.

On the other hand, unfavorable effects can be the result of quite rational and reasonable human activities. And that is why it is necessary to examine and evaluate not only human errors as such, but all behaviors that are crucial for the effective functioning of the system.

From the review of the basic issues related to the analysis of human reliability made in the work, several key problems appear, related to the possibility of using methods of human reliability analysis to improve the occupational safety management system.

As there are no ready-made solutions in this area, first of all, the appropriate reference model of the OSH management system should be selected. Another problem is the choice of methods (two or three) and adapting them to the specificity of management activities. It is also related to the selection of the appropriate taxonomy of human errors and the method of their evaluation. Building a logical set of factors shaping human performance at various levels of management or in various areas of OSH management will also be crucial. These factors will be subject to specific expert assessment or their impact may be defined top-down. Finally, it will be important to propose possible actions to reduce the identified errors to improve human reliability.

The considerations presented in the work are original and may constitute a significant contribution to the improvement of management systems, with particular emphasis on dedicated OSH systems, because human errors in this system most often lead to accidents and serious failures.

REFERENCE

ANSI/ASME RA-S-2002. Standard for Probabilistic Risk Assessment for Nuclear Power Plant Applications.

Bell, J, Holroyd, J., 2009. *Review of human reliability assessment method*, Research Report RR679, HSE, Health and Safety Laboratory, Buxton: Derbyshire.

Bellamy, L.J., Geyer, T.A.W, 1992. *Organizational, Management and Human Factors in Quantified Risk Assessment – Report 1*, HSE Contract Research Report 33/92.

Bello, G.C., Colombari, V., 1980. *Empirical Technique to estimate operator's error (TESEO)*, Reliability Engineering, 1(3), 3-24.

Boring, R. L, Gertman, D.I., 2005. *Atomistic and holistic approaches to human Reliability analysis in the US nuclear power industry*, Safety and reliability, 25(2), 21-37, DOI: 10.1080/09617353.2005.11690802

Cadieux, J., Roy, M., Desmarais, L., 2006. *A preliminary validation of a new measure of occupational health and safety,* Journal of Safety Research, 37(4), 413-419, DOI: 10.1016/j.jsr.2006.04.008

CEN, 2009, EN-ISO 9004: 2009. *Managing for the sustained success of an organization – A quality management approach*, CEN.

Denton, D., 1982. *Safety management: Improving Performance,* McGraw-Hill, New York, USA.

DOE-HDBK-1208-2012, *Accident and Operational Safety Analysis, Volume I: Accident Analysis Techniques,* US Department of Energy.

Dougherty, E.M., 1992. *SRK- It just keeps a rolling*, Reliability Engineering and System Safety, 38, 253-255.

Embrey, D.E., 1983. *The Use of Performance Shaping Factors and Quantified Expert Judgment in the Evaluation of Human reliability: An Initial Appraisal*, NUREG/CR-2986, Brookhaven National Laboratory.

EN-62508: 2010. *Guidance on human aspects of dependability,* (IEC 62508), CENELEC, Brussels.

Everdij, M.H.C, Blom, H.A.P, 2008. *Safety Method Database*, http://www.nlr.nl/documents/flyers/SATdb.pdf

Forester, J., Bley, D., Cooper, S. Lois, E., Siu, N., Kolaczkowski, A, Wreathall, J., 2004. *Expert elicitation approach for performing ATHEANA quantification*, Reliability Engineering and System Safety, 83, 207-220, DOI: 10.1016/j.ress.2003.09.011

Forester, J. Kolaczkowski, A., Lois, E., Kelly, D., 2006. *Evaluation Analysis Methods against Good Practices*, Final Report. NUREG-1842, US Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington DC.

Gertman, D., Blackam, H., Marble, J., Byers, C , Smith, C, 2004. *The SPAR-H human reliability analysis method*, NUREG/CR-6883. Idaho National Laboratory, US Nuclear Regulatory Commission Office of Nuclear Regulatory Research Washington, DC.

Granerud, R.L., Rocha, R.S., 2011. *Organizational learning and continuous improvement of health and safety in certified manufacturers*, Safety Science, 49, 1030-1039, DOI: 10.1016/j.ssci.2011.01.009

Hannaman, G.W., Spurgin, A.J., 1984. *Systematic Human Action Reliability Procedures*, EPRI NF 3583.

Helmreich, R.L., 2000. *On error management: lessons from aviation*, British Medical Journal, 320(7237) 781–785, DOI: 10.1136/bmj.320.7237.781

Hollnagel, E., 1993. *Human reliability analysis: Context and control*, Academic Press.

Hollnagel, E., 1998. *Cognitive Reliability and Error Analysis Method (CREAM),* Elsevier Science, Amsterdam.

Humphreys, P. (ed.) 1988. *Human Reliability Assessor Guide,* Safety and Reliability Directorate, UK, RTS 88/95Q.

Hunns, D.M., 1982. *The method of paired comparisons*, [In:] A.E. Green (ed.) High risk safety technology, Wiley, Chichister.

ILO, 2013. *Audit Matrix for the ILO Guidelines on Occupational Safety and Health Management Systems* (ILO-OSH 2021), ILO, Switzerland.

Kariuki, S.G., Loewe, K., 2007. *Integrating human factors into process analysis,* Reliability Engineering and System Safety, 92, 1764-1773, DOI: 10.1016/j.ress. 2007.01.002

Kirwan, B., 1996. *The validation of three human reliability quantification techniques, THERP, HEART and JHESI: Part 1 - technique descriptions and validation issues*, Applied Ergonomics, 27(6), 359-337, DOI: 10.1016/s0003-6870(96)00044-0

Kirwan, B., Kennedy, R., Taylor-Adams, S., Lambert ,B., 1997. *The validation of three human reliability quantification techniques THERP, HEART and JHEDI: Part II – results of validation exercise*, Applied Ergonomics, 28(1), 17-25, DOI: 10.1016/ S0003-6870(96)00045-2

Kirwan, B., 1998. *Human error identification techniques for risk assessment of high risk systems – Part 1: Review and evaluation of techniques*, Applied Ergonomics, 29(3), 157-177, DOI: 10.1016/S0003-6870(98)00010-6

Klimecka – Tatar, D., Niciejewska, M., 2016. *The OHS management system in the small-sized production company*, Production Engineering Archives, 13(4), 49-52.

Law, W.K., Chan, A.H.S., Pun, K.F., 2006. *Prioritizing the safety management elements*, Industrial Management and Data Systems, 106(6), 778-792, DOI: 10.1108/02635570610671470

Lyons, M., Adams, S., Woloshynowych, M., Vincent, C., 2004. *Human reliability analysis in healthcare: A review of techniques,* International Journal of Risk & Safety in Medicine, 16, 223-237.

NEA/CSNI/R(98)1, 1998. *Critical operator actions: Human reliability modeling and data issues*, Principal Working Group, NE Agency, Committee on the Safety of Nuclear Installations.

Norman, D.A., 1981. *Categorization of action slips*, Psychological Review, 88, 1-15.

Rasmussen, J., 1980. *What Can be Learned from Human Error Reports*? [In:] Duncan, K.D., Gruenberg, M.M., Wallis, D. (Eds.). John Wiley and Sons Ltd., London

Rasmussen, J., 1982. *Human errors: a taxonomy for describing human malfunction in industrial installations,* Journal of Occupational Accidents, 4, 311-335.

Reason, J., 1990. Human Error. Cambridge University Press, New York.

Ren, J., Jenkinson, I., Wang, J., Xu, D.L., Yang, J.B., 2008. *A methodology to model causal relationships in offshore safety assessment focusing on human and organizational factors*, Journal of Safety Research, 39, 87-100.

Stewart, M.G., Melchersand, R.E., 1997. P*robabilistic Risk Assessment of Engineering Systems*, Chapman and Hall, London, Berlin.

Seaver, D.A., Stillwell, W.G., 1983. *Procedures for using expert judgement to estimate human error probabilities in nuclear power plant operations*, NUREG/CR-2743, US Nuclear Regulatory Commission, Washington DC.

Stanton, N.A., Salmon, P.M., Walker, G.H., Baber, C., Jenkins, D.P., 2005. *Human factors methods: a practical guide for engineering and design*, Ashgate Pub. Co., Aldershot, England, 2005.

Swain, A.D., Guttmann, H.E., 1983. *Handbook of human reliability analysis with emphasis on nuclear power plant applications*, NUREG/CR-1278, US Nuclear Regulatory Commission, Washington, DC.

Swain, A.D., 1987. *Accident Sequence Evaluation Program Human Reliability Analysis Procedure,* NUREG/CR-4772, US Nuclear Regulatory Commission, Washington, DC.

Tabor, J., 2019. *Fuzzy TOPSIS in the Assessment of OHS Management System*, System Safety: Human - Technical Facility - Environment, 1(1), 18-25, DOI: 10.2478/czoto-2019-0003

USNRC, 2000. *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.

Whalley, S. P., 1988. *Minimizing the cause of human error,* [In:] 10th Advances in Reliability Technology Symposium, G.P. Libberton (ed.), London: Elsevier.

Williams, J., 1986. *HEART – A proposed method for assessing and reducing human error,* [In:] Ninth Advances in Reliability Technology Symposium, University of Bradford, NCRS, UK, pp. B3/R/1-B3/R/13.

Williams, J.C., 1989. *Validation of human reliability assessment techniques*, Reliability Engineering, 11, 149-162.

Williams, J.C., 1992. *Toward and improved Evaluation Analysis Tool for Users of HEART*, [In:] Proceedings of the International Conference on Hazard Identification and Risk Analysis, Human Factors and Human Reliability in Process Safety, 15-17 January 1992, Orlando, Florida.