Reliability Engineering and System Safety 100 (2012) 28–32

Contents lists available at SciVerse ScienceDirect

Reliability Engineering and System Safety

journal homepage: www.elsevier.com/locate/ress

Technical note

Estimating Human Error Probability using a modified CREAM

Zhiqiang Sun [a,*], Zhengyi Li [b], Erling Gong [a], Hongwei Xie [a]

[a] College of Mechatronic Engineering and Automation, National University of Defense Technology, Changsha 410073, China
[b] Department of Computer, Changsha Social Work College, Changsha 410004, China

ARTICLE INFO

Article history:
Received 17 May 2011
Received in revised form
14 December 2011
Accepted 29 December 2011
Available online 5 January 2012

Keywords:
Human Error Probability
Human reliability
CREAM
Common performance condition

ABSTRACT

Human Error Probability (HEP) point estimation is important for Probabilistic Safety Assessment (PSA) of socio-technical systems. We present a modified basic method of CREAM to provide the point estimation of HEP for PSA. Five acknowledged assumptions are introduced firstly and the HEP point estimation formula is elicited based on them. Furthermore, the reasonability of the method is discussed and the consistency with other two benchmarking HRA methods, THERP and HEART is validated. Finally, a simple example about starting up the submarine's engine is introduced and the probability of the error *forgetting the warm operation* is calculated using the modified method. The result of the method is consistent with the recorded human performance data and THERP.

1. Introduction

Human Error Probability (HEP) estimation is necessary for Probabilistic Safety Assessment (PSA) of socio-technical systems [1]. And it is one phase of many human reliability assessment (HRA) methods, such as Technique for Human Error Rate Prediction (THERP) [2], Human Error Assessment and Reduction Technique (HEART) [3], Standardized Plant Analysis Risk-Human Reliability Analysis Method (SPAR-H) [4], Cognitive Reliability and Error Analysis Method (CREAM) [5], etc. Besides, there are some methods to provide HEP estimation especially [6–10].

The well-known CREAM has been applied to the safety evaluation of many safety-critical systems, such as space-shuttles and nuclear power plants. It is presented by Erik Hollnagel in 1998, and the core idea is that human error is shaped by both context and human nature [5]. The context is described in terms of control mode which is the degree of control that an operator or one team has over the situation. Four control modes are defined as scrambled, opportunistic, tactical and strategic modes respectively. Which one is chosen depends on the context that is divided into nine Common Performance Conditions (CPCs) to be shown in Section 3. Each CPC may be in different states, and has improved, insignificant or reduced effect on human performance. The total numbers of CPCs with improved, reduced or insignificant effects are denoted as Σ_{improved}, Σ_{reduced} and $\Sigma_{\text{not significant}}$, respectively. The control mode is determined by the couple (Σ_{improved}, Σ_{reduced}) shown in Fig. 1, and Table 1 shows one-to-one correspondences between control modes and HEP intervals. The intervals are usually used as an initial screening of human failure events. Despite all that, the intervals are a bit wide. From the viewpoint of PSA, one point estimation is more desirable. Erik Hollnagel designed a new version of basic method of CREAM in Ref. [11] where the concept of balanced context was introduced. When Σ_{improved} is equal to Σ_{reduced}, the balanced context and the corresponding nominal Mean Failure Rate (MFR) are reached. Once the balanced state is broken, human MFR will increase or decrease. Moreover, a probabilistic version of the basic method of CREAM is presented in Ref. [12] to emphasize the uncertainty in the process of CPCs' states determination. Similar to Refs. [12,13] presents one modified basic method of CREAM using fuzzy logic. In Ref. [14], a simplified CREAM is presented, which can also provide HEP point estimation.

In this paper, we design a modified version of CREAM basic method to give HEP point estimation, and apply it to a simple example to demonstrate its reasonability and practicality. Finally, the new and significant characteristics of the method are summarized, and the future directions are pointed out.

2. Modified version of CREAM

In this section, the discussion is focused on the procedure of derivation of HEP estimation formula by modifying CREAM and some prerequisites for it are discussed firstly. The reasonability and consistency with other two HRA methods are explained and validated by comparison.

2.1. Assumptions

Five widely acknowledged assumptions must be presented as the prerequisites for the modified version of CREAM first of all.

(1) The control mode space is continuous [5,11].

* Corresponding author. Tel.: +86 7314573369.
E-mail address: sunzq@nudt.edu.cn (Z. Sun).

0951-8320/$ - see front matter © 2012 Elsevier Ltd. All rights reserved.
doi:10.1016/j.ress.2011.12.017

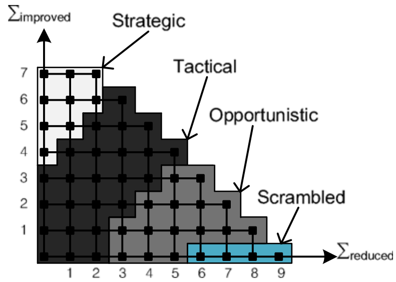

Fig. 1. Determination of control modes.

Table 1
The HEP interval corresponding to each control mode.

Control modes	HEP interval
Strategic mode	(0.00005, 0.01)
Tactical mode	(0.001, 0.1)
Opportunistic mode	(0.01, 0.5)
Scramble mode	(0.1, 1.0)

(2) There is one-to-one correspondence between each control mode variable and HEP, that means HEP is also continuous.
(3) HEP varies with the context exponentially [15].
(4) If the maximum Σ_{improved} and the minimum Σ_{improved} are reached at the same time, namely, $(\Sigma_{\text{improved}}, \Sigma_{reduced})=(7, 0)$, the context is in the most supportive state and HEP is at its minimum. On the contrary, if $(\Sigma_{\text{improved}}, \Sigma_{\text{reduced}})$ is equal to (0, 9), the highest value of HEP is reached.
(5) If the context does not affect human performance at all, the balanced state, thus the nominal HEP is reached. In this situation, the effects of the reduced CPCs are compensated by the effects of the improved CPCs. When $(\Sigma_{\text{improved}}, \Sigma_{\text{reduced}})$ is equal to (0, 0), the context is in balanced state. The HEP under this situation is denoted as HEP_0 (nominal HEP).

Assumption (5) needs some extra explanations as it is the biggest difference between our work and Ref. [11], and is a novel idea of this paper. In the opinion of Ref. [11], the balanced state of context is reached once Σ_{improved} is equal to Σ_{reduced}. It means that the reduced/improved effect caused by one CPC can be compensated by another CPC's improved/reduced effect. This idea might be problematic. The effect mechanisms and sizes are different with various CPCs, and some CPCs may bring greater effects on human performance than others. For example, if the time available is not adequate, the task will not be completed at all. However, if the man–machine interface is not supportive, the task may be completed successfully only with some delays or loss of accuracy. It means that effect sizes of the time available and the man–machine interface are different, and the difference can be reflected by the multiplier in Ref. [16]. Therefore, the effects of various CPCs must be calibrated. A straightforward way is to make the improved or reduced effects normalized. The details are shown in Section 2.2.

2.2. Procedure for estimating HEP

Let X be the control mode space, and $x \in X$ is the variable of the space X. The aim is to construct the relation between x and HEP, namely, to achieve the function $HEP=f(x)$. According to assumption (3), the function $f(x)$ can be written as follows:

$$f(x)=K\exp(\lambda x) \tag{1}$$

where K and λ are two constants to be determined. The key problem is how to define the variable x from the context. And the following two principles are introduced to guide the process of problem solving.

(1) If the context is balanced, then $x=0$. It means that x must be equal to zero when $(\Sigma_{\text{improved}}, \Sigma_{\text{reduced}})=(0, 0)$.
(2) The improved or reduced effects of various CPCs must be calibrated.

A normalized approach is designed to solve the problem, in which the variable x is equal to the result of normalized Σ_{improved} minus normalized Σ_{reduced}.

$$x=\frac{\Sigma_{improved}}{\max(\Sigma_{improved})}-\frac{\Sigma_{reduced}}{\max(\Sigma_{reduced})} \tag{2}$$

where $\max(\Sigma_{\text{improved}})$ and $\max(\Sigma_{\text{reduced}})$ are the maximum values of Σ_{improved} and Σ_{reduced}, respectively. The term $\Sigma_{\bullet}/\max(\Sigma_{\bullet})$ is denoted as normalized $\Sigma_{\bullet}$. Obviously, we have

$$\max(\Sigma_{\text{improved}})=7, \quad \max(\Sigma_{\text{reduced}})=9 \text{ and } K=f(0)=\text{HEP}_0.$$

Therefore, the construction of $f(x)$ is transformed as the determination of the values of HEP_0 and λ. According to assumption 4), we have

$$\begin{cases} \text{HEP}_0\exp(\lambda)=\text{HEP}_{\min} \\ \text{HEP}_0\exp(-\lambda)=\text{HEP}_{\max} \end{cases} \tag{3}$$

where $\text{HEP}_{\min}$ and $\text{HEP}_{\max}$ are the minimum and maximum values, respectively. From Table 1, it can be found that $\text{HEP}_{\min}=0.00005$ and $\text{HEP}_{\max}=1$. By substituting them into Eq. (3), we have

$$\text{HEP}_0=7.07\times10^{-3}; \quad \lambda=-4.9517 \tag{4}$$

Therefore, the HEP point estimation formula is constructed as follows:

$$\text{HEP}=7.07\times10^{-3}\exp\left[-4.9517\left(\frac{\Sigma_{\text{improved}}}{7}-\frac{\Sigma_{\text{reduced}}}{9}\right)\right] \tag{5}$$

It is noticed that the accuracy of λ is with 5 digits. Considering the uncertainty and subjectivity of HEP quantification, the high accuracy seems unnecessary. However, the HEP estimation is used for the safety assessment of some safety-critical social–technical systems, whose failure rates may be less than 10^{-7} or even 10^{-9}. Therefore, small variation of the accuracy of λ may cause relatively huge changes in the final result of safety assessment. The second column in Table 1 shows that the uncertainty and subjectivity of the HEP estimation can be represented by the HEP interval.

The relationship between HEP and the couple $(\Sigma_{\text{improved}}, \Sigma_{\text{reduced}})$ is illustrated as Fig. 2.

3. Discussions

As mentioned above, some of the assumptions or principles (1)–(5) for formula (5) have been acknowledged and applied to the HRA domain, and others are inherited from CREAM directly. Therefore, it is natural to accept that the basis of formula (5) is sound. Despite all that, the reasonability and consistency with other HRA methods need to be validated furthermore.

The reasonability of the method can be justified from two aspects. On one hand, HEP should decrease/increase as $\Sigma_{\text{improved}}/\Sigma_{\text{reduced}}$ is increasing. We can validate this conclusion by inspecting directly formula (5) and Fig. 2. On the other hand, the HEP point estimation derived from formula (5) should stand within the HEP interval of basic method of CREAM. To validate it, seven

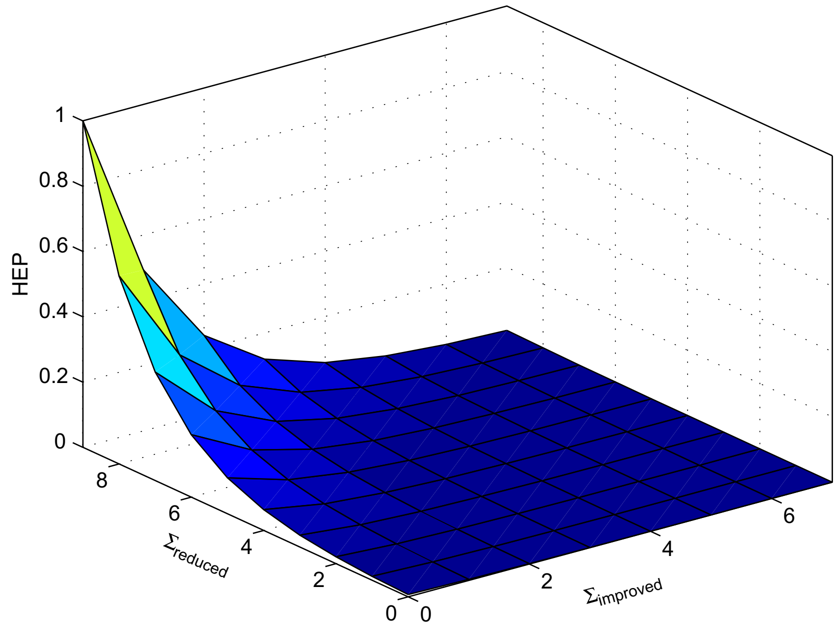

Fig. 2. Control modes and HEP intervals.

Table 2
The results of basic method of CREAM and formula (5).

(Σ_{improved}, Σ_{reduced})	Control mode	HEP interval in CREAM	Point estimation of HEP by formula (5)
(6, 1)	Strategic	(0.00005, 0.01)	0.00018
(4, 0)	Strategic	(0.00005, 0.01)	0.00042
(4, 1)	Tactical	(0.001, 0.1)	0.00073
(0, 2)	Tactical	(0.001, 0.1)	0.0213
(1, 2)	Opportunistic	(0.01, 0.5)	0.0105
(1, 8)	Opportunistic	(0.01, 0.5)	0.2855
(0, 8)	Scrambled	(0.1, 1)	0.5792

typical couples of (Σ_{improved}, Σ_{reduced}) are selected and the HEP intervals and point estimations are calculated, respectively, which can be seen in Table 2. It can be observed that HEP point estimations deviate from the corresponding intervals only at the couples near the boundary of two control modes, such as (4, 1). This deviation can be explained as follows. The control mode space is continuous, so it is difficult to ascertain which mode the couple near the boundary belongs to. For example, HEP point estimation for the couple (4, 1) on the boundary between the strategic and tactical modes should approach 0.001, namely, the upper limit of the HEP interval of the tactical mode. Actually, it does so. Therefore, it is reasonable to accept that the point estimation moves into the interval of other control mode.

What is more, two well-known HRA methods, THERP and HEART, are introduced as the benchmarks to validate the consistency of the aforementioned method. They have been applied to HRA domain widely, such as nuclear power plant and air traffic control.

In THERP, many generic tasks and HEPs are provided. For example, for the task *operation of valve selection*, the HEPs are given under five different situations shown as the third column of Table 3. It can be seen that the design styles and operating interfaces of the five items are different. The situation of the first item is most supportive, and the last situation is worst. From the generic descriptions in the second column, it can be found that the CPC *adequacy of MMI and operational support* is matched with them. Because other CPCs are not mentioned, it is reasonable to assume that they are all in insignificant state. For the first situation, the level of the CPC *adequacy of MMI and operational support* is supportive and has improved effect. So the couple (Σ_{improved}, Σ_{reduced}) is (1, 0), thus we can get the HEP point estimation 0.0035 using formula (5). Similarly, the couple (Σ_{improved}, Σ_{reduced}) is (0, 0) under situation 2 and 3. The point estimation of HEP is 0.0071. Under the situation 4 and 5, (Σ_{improved}, Σ_{reduced})=(0, 1) and the point estimation is 0.0123. It can be found that the orders of the results of THERP and formula (5) are identical, and the latter is a bit conservative. However, a conservative HEP point estimation is not unacceptable for safety assessment of safety-critical systems.

In HEART, task situations are summarized as nine generic scenarios, and the corresponding nominal HEP intervals are provided. For example, the generic scenario G is defined as *completely familiar, well-designed, highly practiced, routine task occurring several times per hour, performed to highest possible standards by high-motivated, highly-trained and experienced person, totally aware of implications of failure, with time to correct potential error, but without the benefit of significant jog aids*, and the nominal HEP interval is [0.00008, 0.009]. Three CPCs of *available time, adequate of training and high experience and adequacy of organization* can be derived from the scenario, and the states of the former two are both adequate, and the level of the latter one is deficient. Therefore, we have (Σ_{improved}, Σ_{reduced})=(2, 0), and the HEP is 0.0017. Obviously, it stands in the interval [0.00008, 0.009].

From the above discussions, we find that the results of the aforementioned method are consistent with any one of THERP & HEART.

4. A simple example

In this section, a simple example about diesel engine is presented to demonstrate the use of HEP estimation provided by formula (5). The diesel engine is one of the most important components of the diesel electric submarine. It provides power for the electromotor to generate electricity. The warm operation, namely, to rotate the engine at low speed for some time, must be

Table 3
The comparison with THERP.

Item	The state of the valve	Nominal HEP in THERP	HEP estimation from formula (5)
1	The valve is clearly and unambiguously labeled, set apart from valves that are similar in all of the following: size and shape, state, and presence of tags.	0.001	0.0035
2	The valve is clearly and unambiguously labeled, part of a group of two or more valves that are similar in one of the following: size and shape, state, or presence of tags.	0.003	0.0071
3	The valve is unclearly or ambiguously labeled, set apart from valves that are similar in all of the following: size and shape, state, or presence of tags.	0.005	0.0071
4	The valve is unclearly or ambiguously labeled, part of a group of two or more valves that are similar in one of the following: size and shape, state, or presence of tags.	0.008	0.0123
5	The valve is unclearly or ambiguously labeled, part of a group of two or more valves that are similar in all of the following: size and shape, state, or presence of tags.	0.01	0.0123

Table 4
The states of 9 CPCs.

CPC	Situations in the submarine	Level (bold)	Effect
Adequacy of organization	The persons in the submarine are organized as one framework of leaders–operators. The operator cannot do anything until he accepts the instruction from the leader. The leaders can correct some but not all errors made by the operators.	Very efficient, **efficient**, inefficient, deficient	Not significant
Working conditions	The room of the cabin is very small, and noising. The temperature is high.	Advantageous, compatible, **incompatible**	Reduced
Adequacy of MMI and operational support	The area of operating panel is little and filled with many similar buttons. And the warm operation button is not isolated with other similar buttons.	Supportive, adequate, tolerable, **inappropriate**	Reduced
Availability of procedures/plans	The operation procedure is well-defined, and but importance of the warm operation is not emphasized	Appropriate, **acceptable**, inappropriate	Not significant
Number of simultaneous goals	There is only one goal to be implemented, which is to start up engine.	**Fewer than capacity**, matching current capacity, more than capacity	Not significant
Available time	Under normal situation, there is adequate time to start up engine. The operator is not under high time pressure.	**Adequate**, temporary inadequate, continuously inadequate	Improved
Time of day (circadian rhythm)	In the submarine, the circadian rhythms of the operators are scrambled because of the difficulty to know the day or night under the sea.	Adjusted, **unadjusted**	Reduced
Adequacy of training and experience	The operator must be trained at the specific school to learn the knowledge about the submarine and gain the operating ability. The operator is certified to join the submarine only if he passed the test. And the operator is not allowed to operate the engine until he has served at least two years in the submarine.	**Adequate, high**; adequate, limited; inadequate	Improved
Crew collaboration quality	This is a one-person task. So the CPC is not used.	Not used	Not used

executed before the engine is started up. Otherwise, the engine might be injured after startup. Unfortunately, the operator may sometimes start the engine up without the warm operation for some reasons. Thus the error *forgetting the warm operation* will occur. But what is the probability for this error? That's what we want to know when doing the safety assessment of the submarine.

First of all, the working context must be evaluated. That is to say, the states of 9 CPCs should be determined. After a thorough analysis of the submarine, the states of 9 CPCs are determined as Table 4.

When we substitute the couple $(\Sigma_{\text{improved}}, \Sigma_{\text{reduced}})=(2, 3)$ acquired from Table 4 into formula (5), we have

$$HEP_{es} = 0.0090 \tag{6}$$

where HEP_{es} is the point estimation of HEP.

Ever since 1/1/2009 until the time of this paper being written (10/12/2011), we have recorded the operation data continuously on the simulation platform of submarine. From the recorded data we find that there are 3611 opportunities for the error *forgetting warm operation*, and only 22 errors occurred actually. The error frequency is 0.0061 and is less than 0.0090. The main cause of it may be that the data are recorded from the laboratory, and the context is relatively supportive.

We also find that the potential error *forgetting the warm operation* is matched with the generic error mode *Omitting a step or important instruction from a formal or ad hoc procedure*, with the nominal HEP 0.003 and error factor (EF) 3 in THERP. Obviously, HEP_{es} is exactly the upper limit of THERP's result, which is the product of the nominal HEP and EF, 0.009.

5. Conclusions and outlook

A HEP quantification method is presented by modifying the basic method of CREAM. Similar with the ancestor, it can be used as an initial screening of human failure events. It is more valuable for safety engineers because of the ability to provide HEP point estimation. In Section 3, we have demonstrated the reasonability, and validated the consistency of the method by comparing it with two benchmarking methods. Compared with the two benchmarks, the method shows a bit conservative. However, a conservative HEP is not a bad thing to those safety-critical systems. From Section 3, it can be found that the result of the method is consistent with recorded human performance data. Therefore, it can be concluded that the method is practical and has the potential for engineering applications.

There are also some other modified versions of CREAM able to provide HEP point estimation [11,14]. Compared with them, the aforementioned method has some unique and significant characteristics as follows:

(1) Nominal human MFR is required in the modified CREAM in Ref. [11], and varies with different contexts or tasks. It must be elicited from historical data or by expert judgments. In the aforementioned method, the nominal HEP under the balanced context is fixed. It makes the method more flexible and practical.
(2) The effect sizes and mechanisms of various CPCs are different. Therefore, the opposite effects caused by two CPCs cannot compensate for each other. In the aforementioned method, the effects of various CPCs are calibrated by normalizing Σ_{improved} and Σ_{reduced}. It means that the effects caused by the CPCs are dealt with as a whole and not as separate parts.

A minor problem at the PSA project of the submarine is how to determine the CPCs' states reasonably. Until now, the determination of the states is completely dependent on the domain knowledge and experiences of the HRA analyzers and engineers. Consequently, the HEP estimation result is a bit subjective inevitably. Actually, each CPC can be divided into various sub-level performance-related factors. The CPC's state is determined by these factors, thus it is reasonable to design one approach to determine the CPC's state from its sub-level factors. This topic has been discussed in Ref. [14]. It is believed to be a good beginning for the future work. Furthermore, the extended method of CREAM is more valuable but complex than the basic method. In the future, we will make a further step to do some modifications on the extended method and try to simplify it.

Acknowledgments

The authors express their sincere appreciation to the two anonymous reviewers for the useful comments. And the work is co-supported by the Specialized Research Fund for the Doctoral Program of Higher Education (No. 20114307120032), the National Natural Science Foundation of China (Grant No. 61005084) and the Research Fund of Civil Aviation Administration of China (CAAC).

References

[1] Boring RL, Hendrickson SML, Forester JA, et al. Issues in benchmarking human reliability analysis methods: a literature review. Reliability Engineering & System Safety 2010;95(6):591–605.
[2] Swain AD, Guttmann HE. Handbook of human reliability analysis with emphasis on nuclear power plant applications. Washington, DC: Nuclear Regulatory Commission; 1983.
[3] Williams JC. A data-based method for assessing and reducing human error to improve operational performance. In: Proceedings of IEEE 4th conference on human factor and power plants. 1988. Monterey, California: p. 436–53.
[4] Gertman DI, Blackman HS, Marble J, et al. The SPAR-H human reliability analysis method. Washington DC: US Nuclear Regulatory Commission; 2005.
[5] Hollnagel E. Cognitive error and error analysis method. London: Elsevier; 1998.
[6] Cepin M. DEPEND-HRA: a method for consideration of dependency in human reliability analysis. Reliability Engineering & System Safety 2008;93(10): 1452–60.
[7] Park KS, Lee J. A new method for estimating human error probabilities: AHP-SLIM. Reliability Engineering & System Safety 2008;93(4):578–87.
[8] Sun ZQ, Xie HW, Shi XJ, et al. Engineering approach for human error probability quantification. Journal of Systems Engineering and Electronics 2009;20(5):1144–52.
[9] Cepin M. Importance of human contribution within the human reliability analysis (IJS-HRA). Journal of Loss Prevention in the Process Industries 2008;21(3):268–76.
[10] Chang YHJ, Mosleh A. Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents part 5: dynamic probabilistic simulation of the IDAC model. Reliability Engineering & System Safety 2007;92(8):1076–101.
[11] Fujita Y, Hollnagel E. Failures without errors: quantification of context in HRA. Reliability Engineering & System Safety 2004;83(2):145–51.
[12] Kim MC, Seong PH, Hollnagel E. A probabilistic approach for determining the control mode in CREAM. Reliability Engineering & System Safety 2006;91(2):191–9.
[13] Sun ZQ, Xie HW, Li XX, et al. Determination of control modes in CREAM method based on fuzzy logic [in Chinese]. China Safety Science Journal 2007;17(5):21–6.
[14] He XH, Wang Y, Shen ZP, et al. A simplified CREAM prospective quantification process and its application. Reliability Engineering and System Safety 2008;93(3):298–306.
[15] Apostolakis GE, Bier VM, Mosleh A. A critique of recent model for human error rate assessment. Reliability Engineering & System Safety 1988;22(1–4):201–17.
[16] Blackman HS, Gertman DI, Boring RL. Human error quantification using performance shaping factors in the SPAR-H method. In: Proceedings of the 52nd Annual Meeting of the Human Factors and Ergonomics Society; 2008.