INL/CON-10-18620
PREPRINT

How Many Performance Shaping Factors are Necessary for Human Reliability Analysis?

PSAM 10

Ronald L. Boring

June 2010

How Many Performance Shaping Factors are Necessary for Human Reliability Analysis?

Ronald L. Boring[a*]
[a]Idaho National Laboratory, Idaho Falls, Idaho, USA

Abstract: It has been argued that human reliability analysis (HRA) has expended considerable energy on creating detailed representations of human performance through an increasingly long list of performance shaping factors (PSFs). It is not clear, however, to what extent this refinement and expansion of PSFs has enhanced the quality of HRA. Indeed, there is considerable range in the number of PSFs provided by individual HRA methods, ranging from single factor models such as time-reliability curves, up to 50 or more PSFs in some current HRA models. The US Nuclear Regulatory Commission advocates 15 PSFs in its HRA Good Practices (NUREG-1792), while its SPAR-H method (NUREG/CR-6883) espouses the use of eight PSFs and its ATHEANA method (NUREG-1624) features an open-ended number of PSFs. The apparent differences in the optimal number of PSFs can be explained in terms of the diverse functions of PSFs in HRA. The purpose of this paper is to explore the role of PSFs across different stages of HRA, including identification of potential human errors, modeling of these errors into an overall probabilistic risk assessment, quantifying errors, and preventing errors.

Keywords: Human reliability analysis, performance shaping factors, orthogonality

1. INTRODUCTION

The field of human reliability analysis (HRA) aims to identify the causes and sources of human errors and provide a numeric estimate of the likelihood of such errors. HRA typically encompasses three phases (see Figure 1), ranging from identifying error sources, to modeling these errors as part of a systemic analysis including hardware failures, to quantifying the human error probabilities (HEPs) [1]. These three phases are not found in all HRA methods. In fact, there are HRA methods that have a heavy emphasis on the qualitative phases (e.g., identification of human errors common in root cause analysis), while other HRA methods are used primarily as tools in quantification [2]. Thus, there are three broad families of HRA methods: those methods used primarily for qualitative insights into human error, those methods used primarily for quantification, and those methods that encompass the complete spectrum of HRA. Each HRA method is designed for specific purposes, and the lack of complete coverage of qualitative and quantitative phases by many methods should not necessarily be viewed as a shortcoming on behalf of those methods. In fact, HRA frameworks such as the Systematic Human Action Reliability Procedure (SHARP1) [3] provide an architecture that allows HRA methods to plug into specific phases of an analysis. This has allowed hybrid approaches like the EPRI HRA Calculator [4] successfully to mix and match several elements of different HRA methods.

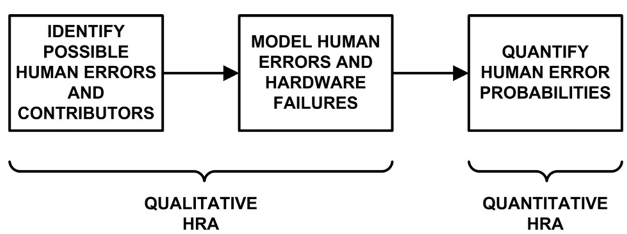

Figure 1: Three Phases of HRA

*Address correspondence to Ronald Laurids Boring, PhD, Human Factors, Controls, and Statistics Department, Idaho National Laboratory, Idaho Falls, ID 83415, USA. Email: ronald.boring@inl.gov.

Throughout the qualitative and quantitative phases of HRA, many HRA methods make use of performance shaping factors (PSFs)—aspects of behavior and context that impact human performance. Historically, PSFs were viewed primarily in terms of the deleterious effects they might exert on human performance. There has been a greater emphasis recently to catalog ways in which PSFs might also enhance performance [5]. PSFs are used in qualitative HRA to identify contributors to human performance. In quantitative HRA, PSFs are often used to derive the HEP. For example, in the Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) method, PSFs act as multipliers on a nominal HEP [6]. When the PSFs represent a positive effect, the different levels of effect for the PSFs correspond to a value less than one. Multiplying a nominal HEP by this fractional value associated with the PSF serves to decrease the overall HEP. When the PSFs represent a negative effect, the different levels of effect for the PSFs correspond to a value greater than one. Multiplying the nominal HEP by this positive integer serves to increase the overall HEP. When the PSFs are thought to have no effect, the PSF multiplier is set to 1, thereby neither increasing nor decreasing the overall HEP. The effects of PSF multipliers on the HEP are summarized in the following equation (taken from [1]):

$$HEP_{overall} = HEP_{nominal} \times PSF \quad \begin{cases} 0 < PSF < 1 & \Rightarrow \quad HEP_{overall} < HEP_{nominal} \quad \therefore \quad \text{reliability increases} \\ PSF = 1 & \Rightarrow \quad HEP_{overall} = HEP_{nominal} \quad \therefore \quad \text{reliability stays same} \\ PSF > 1 & \Rightarrow \quad HEP_{overall} > HEP_{nominal} \quad \therefore \quad \text{reliability decreases} \end{cases} \tag{1}$$

Note that not every HRA method follows this approach to quantification, but it is common across most HRA methods to employ formal or informal weightings on PSFs in determining the HEP. Even HRA methods such as A Technique for Human Error Analysis (ATHEANA) [7], whose quantification process uses expert elicitation, may use PSFs to guide or anchor the quantification process.

2. THE NUMBER OF PSFs

There is to date no consensus on which PSFs should be used in HRA nor the appropriate number of PSFs[1] to include in a method or analysis. Some of the earliest HRA approaches such as the time-reliability methods [8] espoused a single PSF, while recent HRA methods, e.g., [9-10], have discussed upwards of 50 PSFs. Work by Groth [11] began with a hierarchical list of 60 PSFs extracted from a combination of several HRA methods, which through a mathematical and expertise-based factor analysis yielded smaller groupings of PSFs. The resulting factor models resulted in six or nine PSF categories, depending on the level of clustering encompassed under each individual PSF.

In practice, the number of PSFs that are included in HRA methods lies between these extremes of 1 to 60 PSFs. For example, the SPAR-H method [2], which is widely used in the US nuclear industry, includes eight PSFs. The internationally widely used Cognitive Reliability and Error Analysis Method (CREAM) [12] uses nine PSFs. A recent study commissioned by the US Nuclear Regulatory Commission and entitled *Good Practices for Implementing Human Reliability Analysis* (NUREG-1792 [13]) identified 15 essential PSFs for HRA. Table 1 below presents a crosswalk of the PSFs found in the *Good Practices* [13] to the PSFs found in SPAR-H [2], CREAM [12], and Groth's Nine-Factor Model [11]. Note that the fifteenth *Good Practices* PSF—Consideration of 'Realistic' Accident Sequence Diversions and Deviations—represents a special case PSF that is an overarching principle rather than a specific measurable or quantifiable PSF. It might therefore be argued that the *Good Practices* actually only present 14 PSFs.

[1] It is beyond the scope of this paper to discuss the definitions of PSFs. The reader is referred to the source material as cited for appropriate definitions.

Table 1: Crosswalk of PSFs Between the Good Practices, SPAR-H, CREAM, and the 9-Factor Model

Good Practices [13]	SPAR-H [2]	CREAM [12]	9-Factor Model [11]
Training and Experience	Experience/Training	Adequacy of Training and Preparation	Training Knowledge
Procedures and Administrative Controls	Procedures	Availability of Procedures/Plans	Resources
Instrumentation	Ergonomics/HMI	Adequacy of HMI and Operational Support	Machine
Time Available	Available Time	Available Time	Loads/Perceptions
Complexity	Complexity	Number of Simultaneous Goals	Complexity
Workload/Time Pressure/Stress	Stress/Stressors	Number of Simultaneous Goals	Loads/Perceptions
Team/Crew dynamics	Work Processes	Crew Collaboration Quality	Team
Available Staffing	Work Processes	Adequacy of Organization	Resources
Human-System Interface	Ergonomics/HMI	Adequacy of HMI and Operational Support	Machine
Environment	Stress/Stressors	Working Conditions	Complexity
Accessibility/Operability of Equipment	Ergonomics/HMI	Adequacy of HMI and Operational Support	Machine
Need for Special Tools	Ergonomics/HMI	Adequacy of HMI and Operational Support	Resources
Communications	Work Processes	Crew Collaboration Quality	Team
Special [Equipment] Fitness Needs	Ergonomics/HMI	Adequacy of HMI and Operational Support	Resources
Consideration of 'Realistic' Accident Sequence Diversions and Deviations	--	--	--
--	Fitness for Duty	Time of Day	--
--	Work Processes	Adequacy of Organization	Organizational Culture
--	--	--	Attitude

The comparison in Table 1 reveals similar coverage across the methods:

- *SPAR-H:* The eight SPAR-H PSFs map to 14 of the PSFs identified in the *Good Practices*. Many of the PSFs are one-to-one matches, while some SPAR-H PSFs such as Work Processes and

Ergonomics/Human-Machine Interface (HMI) cluster multiple facets of the *Good Practices* PSFs into a single PSF. SPAR-H features a further PSF—Fitness for Duty—not found in the *Good Practices* PSFs.
- *CREAM:* The nine CREAM PSFs (called common performance conditions in the method) map similarly to SPAR-H. A few of the 14 factors in the *Good Practices* are combined under a single PSF in CREAM. Notably, many of the factors in the *Good Practices* are encompassed under a single CREAM PSF—Adequacy of HMI and Operational Support. CREAM includes a PSF—Time of Day—that is not covered under the *Good Practices* list of PSFs. Time of Day encompasses factors similar to those covered in SPAR-H under the Fitness for Duty PSF.
- *Nine-Factor Model:* The nine factors map to the *Good Practices* similarly to SPAR-H and CREAM. The Machine PSF encompasses many of the factors found in the *Good Practices*. The Nine-Factor Model features two PSFs that do not clearly map to the *Good Practices*—Organizational Culture and Attitude. These were to some extent covered in the SPAR-H and CREAM PSFs, but neither SPAR-H nor CREAM distinctly addresses these factors as unique PSFs.

As can be seen from the crosswalk in Table 1, there is considerable overlap in the PSFs, although each method brings with it a slightly different emphasis and slightly different set of PSFs. This variability in PSFs is a reflection of the vastness of factors that can influence human performance, the different approaches used to distil these into a usable set of factors (e.g., an explicit attempt to model cognitive factors in CREAM), and the different applications for which each HRA method was originally designed. Table 1 reveals that it is generally possible for the analyst to document a variety human performance conditions using the taxonomic PSF list included with each method, even when the PSFs only broadly touch on certain human performance issues.

3. A CRITICISM OF THE LARGE NUMBERS OF PSFs

Whether the HRA method provides eight or 15 PSFs, it is generally possible to model performance with that list of PSFs. Still, there is a degree of extrapolation and expert judgement required when generalizing behavioral phenomena to a limited set of PSFs. One analyst's mapping may not correspond to that by another analyst, thereby potentially introducing uncertainty and decreasing inter-analyst reliability. The parsimony of a small number of PSFs is frequently superseded by the need to create a reasonably full representation of PSFs. It is therefore no coincidence that many of the newer HRA methods, e.g., [9-10], have tended to feature a large set of PSFs. Moreover, there is a tendency to increase the number of PSFs covered by HRA methods. For example, over the course of the development of SPAR-H, the number of PSFs increased from six in the original method to the current eight [6], with the recommendation from some parties that it may be time to add a ninth PSF to the method [14].

Galyean [15] suggests that the need to drill down into the nuances of human performance by using an ever increasing list of PSFs is misguided. He suggests, instead, that it is possible to account for the entirety of human performance using only three PSFs—the individual, the organization, and the environment. The advantage of constructing a short list of PSFs, apart from simplifying the amount of effort required for an analysis, is that the PSFs could actually be orthogonal, omitting the possibility of double-counting of effects, which can have spurious effects on the HEP calculation. Initially, Galyean's criticism may seem strongly iconoclastic to the current efforts in HRA to create a nuanced and ever more complete set of PSFs. However, there is obvious validity to his arguments, particularly because the need to control for double-counting in quantification has not abated.

4. INVESTIGATING THE CRITICISM

Part of Galyean's argument is that the PSFs used in many HRA methods are not orthogonal, i.e., they overlap to some degree. This point becomes evident when trying to operationalize PSFs in such a manner that they become observable and measurable. A review on measuring PSFs [16] demonstrates that many of the PSFs found in the *Good Practices* can only be measured in terms of their relation to one another. The definitions of PSFs—and our ability to observe and measure them—are non-orthogonal.

Evidence for non-orthogonality could be found if there was a significant co-occurrence of PSFs. In order to determine if PSFs co-occur, [14] looked at SPAR-H analyses performed for 82 incident reports publically submitted to the US Nuclear Regulatory Commission by US nuclear power plants. Tables 2 and 3 provide correlations between the eight SPAR-H PSFs across 651 subtasks documented in the reports. Since SPAR-H delineates PSFs according to Diagnosis (i.e., cognition) and Action (i.e., behavior), two separate correlation tables are required. Table 2 reflects the correlations for the Diagnosis PSFs, while Table 3 reflects the correlations for the Action PSFs. Since the incidents are self-reported and are not required to provide a detailed human performance narrative, the quality of the SPAR-H analyses depends on the availability of human performance insights captured in the reports. Because of this limitation, the data presented should be interpreted only to represent an example of SPAR-H analyses and not a definitive account of SPAR-H PSF correlations.

Table 2: Spearman Rank-Order Correlations for SPAR-H Diagnosis PSFs

For Diagnosis	Available Time	Stress/ Stress-ors	Complexity	Experience/ Training	Procedures	Ergonomics/ HMI	Fitness for Duty	Work Processes
Available Time								
Stress/Stressors	.67*							
Complexity	-.02	.15*						
Experience/Training	-.03	.06	.21*					
Procedures	-.07	.01	.25*	.28*				
Ergonomics/HMI	.01	.06	-.05	.20*	.09			
Fitness for Duty	-.03	.03	-.03	.18*	.09	.44*		
Work Processes	-.06	.00	.24*	.55*	.36*	.15*	.10	

* Marked correlations are significant at $p < 0.05$

Table 3: Spearman Rank-Order Correlations for SPAR-H Action PSFs

For Action	Available Time	Stress/ Stress-ors	Complexity	Experience/ Training	Procedures	Ergonomics/ HMI	Fitness for Duty	Work Processes
Available Time								
Stress/Stressors	.50*							
Complexity	.38*	.35*						
Experience/Training	.31*	.21*	.32*					
Procedures	.05	-.01	.12*	.08*				
Ergonomics/HMI	.10*	.04	.08*	.08*	.29*			
Fitness for Duty	.20*	.29*	.22*	.17*	.12*	.27*		
Work Processes	.00	.13*	.16*	.20*	.35*	.12*	.15*	

* Marked correlations are significant at $p < 0.05$

Because of the large sample size in terms of the 651 subtasks for which SPAR-H analyses could be completed, the significance levels may be inflated. Therefore, to err on the side of conservatism, only correlations above $|0.20|$ are considered significant in the present discussion, even if the correlation is actually reported as significant at the $p < 0.05$ level. Significant correlations make it possible to cluster heavily related PSFs. The significant correlations, particularly for the Action PSFs, suggest two groupings of PSFs:

- *Grouping 1:* Available Time, Stress/Stressors, Complexity, Experience/Training, Fitness for Duty
- *Grouping 2:* Procedures, Ergonomics/HMI, Work Processes

To frame these groupings in terms of Galyean's trio of PSFs, the first group encompasses factors related to the individual (Experience/Training, Fitness for Duty), the environment or situation (Available Time, Complexity), and a combination of the two (Stress/Stressors—a PSF that, in fact, represents the individual's stress response and the stressors present due to the environment). The second group consists of factors related to the organization and environment (Procedures, Ergonomics/HMI, Work Processes). While the clustering is not an exact match to Galyean's proposed PSFs,[2] there is clearly non-orthogonality in the SPAR-H PSFs, thereby lending credence to Galyean's concerns about the possibility of double-counting performance effects.

In addition to correlations, Groth [11] performed a factor analysis on PSF data to develop her nine-factor model. Groth found four groupings to be the best fit for the data:

- *Grouping 1:* Training, Team, Loads/Perceptions, Complexity
- *Grouping 2:* Organizational Culture, Attitude, Knowledge
- *Grouping 3:* Organizational Culture, Attitude, Loads/Perceptions, Complexity
- *Grouping 4:* Resources, Complexity

Note that the groupings are not entirely orthogonal, suggesting complex interactions between the PSFs. These groupings also partially validate Galyean's notion of three PSFs. The first grouping relates closely to the individual (perhaps with the exception of Complexity, which is definitionally related to the environment or situation but manifests itself individually). The second and third groupings relate to the intersection of the individual and the organization. The fourth grouping might be seen as a link between the environment (in terms of Resources) and the individual. Note that a final PSF—Machine—was not part of any grouping. This PSF forms an important part of the environment, although it was not thus clustered in the data.

Certainly, any attempt to put a label to a cluster introduces a large degree of subjectivity. It is therefore inappropriate to suggest that the clusters conclusively validate Galyean's approach. Nonetheless, there is conclusive evidence that the PSFs are non-orthogonal. Moreover, there is at least preliminary evidence to hint that categories like Galyean's three PSFs may encompass a much larger family of PSFs.

5. COUNTERARGUMENT TO GALYEAN'S THREE-PSF MODEL

Given that most PSFs seem to cluster into a very small number of PSFs, is there value in continuing to use, develop, and research a larger number of PSFs? To answer this question, it is important to reconsider the uses of PSFs. Recall that the main analysis implication in Galyean is the possibility of double-counting performance effects when quantifying the HEP. This consideration applies to the quantification phase of HRA, but what about the other phases of HRA? The following observations may be made across the three phases of HRA identified in Figure 1:

- *Identification Phase:* During identification, PSFs assist the analyst in classifying and categorizing errors. The availability of a rich selection of PSFs can aid the analyst in identifying and pinpointing error types that might otherwise be overlooked.
- *Modeling Phase:* During modeling, PSF information is minimally used, and there is little advantage afforded by having a large vs. small number of PSFs.
- *Quantification Phase:* During quantification, as discussed, PSFs may serve as multipliers in calculating the HEP. A small number of PSFs is generally sufficient to arrive at a screening value, although a large number of PSFs can help the analyst arrive at a nuanced error estimate. It must be noted, however, that a large number of PSFs does not guarantee valid quantification. PSFs, when properly derived from empirical data sources of human performance, serve to narrow the uncertainty range in the HEP calculation. If these PSFs do not have a solid foundation in empirical data, the nuanced estimation is no more accurate than a screening value.

Beyond these three phases of HRA, it is also valuable to consider the applications of HRA. There is new research on HRA for design applications [17-18], considering not only HRA for its typical use in

[2] A factor analysis was not completed using the data set, making it difficult to determine the best grouping of PSFs from the correlations.

as-built systems but also for use in ensuring the safety of novel designs. In nuclear power, this takes the form of HRA as a design tool for next-generation operating principles such as advanced control rooms. HRA may be used not only to certify that an advanced control room design is safe to operate (in the sense of having minimal negative PSFs to degrade performance), it may also be used to influence the design of that advanced control room. HRA can be used to pit the risks of one design against another prior to actual build out or human-in-the-loop testing. It can also serve as a screening tool for new designs to ensure that all relevant factors that might affect operator performance are considered. In this context, having a concise list of PSFs offers little that can improve the design process. For example, offering that the new control room design may degrade the operators' "environment" does little to pinpoint how it degrades performance and, more importantly, how that control room interface might be improved. Having a larger, more nuanced list of PSFs allows the analyst to make more precise predictions. For such an application, the "environment" PSF should be teased apart into more detail.

The value of a more nuanced list of PSF is exemplified in research by Boring and colleagues [19-20], which created a usability technique based on HRA in order to evaluate safety critical domains such as control rooms in nuclear power plants. Whereas much of usability has focused on factors such as efficiency, ease of use, and appeal of interfaces, safety critical domains have a first priority of ensuring the safety of the application. Any HMI issues that hinder safe operation of the system need to be identified. The approach adopted in [19-20] was to adapt usability heuristics [21] as PSFs, where usability heuristics are simply factors known to influence the quality of the HMI. The list of heuristics as PSFs and their associated weights is shown in Table 4. Note that these HMI PSFs were taken verbatim from a common list of usability heuristics [21] and have not been optimized to meet safety critical HMI issues. The weights have also not been validated and are simply taken from the weights used for the Ergonomics/HMI PSF in SPAR-H. Nonetheless, this proof-of-concept list demonstrates how a single PSF—the HMI PSF used in many HRA methods—can actually be further delineated in considerable detail. When this detail is present, it can provide considerable qualitative insights into what aspects of the HMI need to be addressed. As the weights on such multipliers are fine-tuned, they also allow the analyst to prioritize the PSFs that have the greatest impact on the safety of the HMI.

Table 4: Heuristic PSFs for HMI

Heuristic	Multipliers				
Simple and natural dialog	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Speak the users' language	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Minimize users' memory load	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Consistency	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Clearly marked exits	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Shortcuts	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Good error messages	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Prevent errors	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent
Help and documentation	10 Poor	5 Available	1 Nominal	0.2 Good	0.1 Excellent

6. CONCLUSIONS

Facets of HRA requiring greater detail—notably error identification and quantification—as well as applications like HRA for design clearly benefit from the richer representation of the error space that is typically provided by a larger number of PSFs. The greatest advantage of a nuanced list of PSFs comes for qualitative analysis, in which the PSFs can help to pinpoint exact causes of errors. A higher level grouping of PSFs such as found in Galyean [15] may not provide enough detail to elucidate the qualitative causes for degraded or enhanced human performance. In those cases, a larger list of PSFs is essential to a useful and complete analysis. However, there are limitations to a conclusion that favors a large number of PSFs. For example, a quantitative analysis aimed to screen human failure events typically will not require an extensive list of PSFs in order to derive the HEP. Moreover, an extensive list of PSFs may not provide any greater precision in quantification than a screening analysis if the PSFs are not calibrated to validated data points.

Ultimately, the crucial issue is not the number of PSFs but rather the use for which those PSFs are engaged. Not all phases or applications of HRA require an extensive list of PSFs, and HRA would be well served to pay heed to Galyean's arguments for a smaller number of PSFs. But, Galyean's argument does not hold universally, because there are HRA applications that directly benefit from the increased precision afforded by a larger number of PSFs. Future HRA development efforts should be sensitive to the applications for which the HRA is intended and adjust the number of PSFs accordingly. Such efforts should continue to expand the list of PSFs where it is desirable to achieve greater nuance and detail in the analysis, and they should seek to consolidate PSFs where the application requires orthogonal definitions and values.

Acknowledgements and Disclaimers

The author wishes to thank William Galyean and Dana Kelly (Idaho National Laboratory), Tuan Tran (Sprint Corporation), and Johanna Oxstrand (Vattenfall Ringhals Nuclear Power Station) for the engaging and sometimes spirited discussions that led to the writing of this paper.

Idaho National Laboratory is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517.

References

[1] R.L. Boring, *"Human reliability analysis in cognitive engineering and system design,"* in Frontiers of Engineering: Reports on Leading-Edge Engineering from the 2008 Symposium, National Academy of Engineering, Washington, DC, 103-110, 2009.

[2] D. Gertman, H. Blackman, J. Marble, J. Byers, L. Haney, and C. Smith, *"The SPAR-H human reliability analysis method, NUREG/CR-6883,"* US Nuclear Regulatory Commission, 2005, Washington, DC.

[3] Electrical Power Research Institute, *"SHARP1—A Revised Systematic Human Action Reliability Procedure, EPRI TR-101711,"* EPRI, Palo Alto, 1992.

[4] J.A. Julius and J.F. Grobbelaar, *"Integrating human reliability analysis approaches in the EPRI HRA Calculator,®"* Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, American Society of Mechanical Engineers, New Orleans, 2006.

[5] B. Hallbert, R. Boring, D. Gertman, D. Dudenhoeffer, A. Whaley, J. Marble, and J. Joe, *"Human event repository and analysis (HERA) system overview, NUREG/CR-6903, Volume 1,"* US Nuclear Regulatory Commission,Washington, DC, 2006.

[6] R.L. Boring and H.S. Blackman, *"The origins of the SPAR-H method's performance shaping factor multipliers,"* Official Proceedings of the Joint 8th IEEE Conference on Human Factors and Power Plants and the 13th Annual Workshop on Human Performance/Root Cause/Trending/Operating Experience/Self Assessment, 177-184, 2007.

[7] J. Forester, A. Kolaczkowski, S. Cooper, D. Bley, and E. Lois, *"ATHEANA user's guide, NUREG-1880,"* US Nuclear Regulatory Commission, Washington, DC, 2007.

[8] A.J. Spurgin, *"Human reliability assessment theory and practice,"* CRC Press, Boca Raton, 2009.

[9] US Nuclear Regulatory Commission, *"Technical basis and implementation guidelines for a technique for human event analysis (ATHEANA), NUREG-1624,"* US Nuclear Regulatory Commission, Washington, DC, 2000.

[10] Y.H.J. Chang and A. Mosleh, *"Cognitive modeling and dynamic probabilisitic simulation of operating crew response to complex system accidents, Part 2: IDAC performance influencing factors model,"* Reliability Engineering and System Safety, 29, 1014–1040, (2007).

[11] K. Groth and A. Mosleh, *"A data-informed model of performance shaping factors for use in human reliability analysis,"* Center for Risk and Reliability, University of Maryland, College Park, 2009.

[12] E. Hollnagel, *"Cognitive reliability and error analysis method,"* Elsevier, Oxford, 1998.

[13] A. Kolaczkowski, J. Forester, E. Lois, and S. Cooper, *"Good practices for implementing human reliability analysis, Final report, NUREG-1792,"* US Nuclear Regulatory Commission, Washington, DC, 2005.

[14] R.L. Boring, A.M. Whaley, T.Q. Tran, P.H. McCabe, L.G. Blackwood, and R.F. Buell, *"Guidance on performance shaping factor assignments in SPAR-H, INL/EXT-06-11959,"* Idaho National Laboratory, Idaho Falls, 2006.

[15] W.J. Galyean, *"Orthogonal PSF taxonomy for human reliability analysis,"* Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, American Society of Mechanical Engineers, New Orleans, 2006.

[16] R.L. Boring, C.D. Griffith, and J.C. Joe, "The measure of human error: Direct and indirect performance shaping factors," Official Proceedings of the Joint 8th IEEE Conference on Human Factors and Power Plants and the 13th Annual Workshop on Human Performance/Root Cause/Trending/Operating Experience/Self Assessment, 170-176, 2007.

[17] R.L. Boring, J. Oxstrand, and M. Hildebrandt, *"Human reliability analysis for control room upgrades,"* Proceedings of the 53rd Annual Meeting of the Human Factors and Ergonomics Society, 1584-1588, 2009.

[18] N.A. Stanton and C. Baber, *"Error by design: Methods for predicting device usability,"* Design Studies, 23, 364-384, (2002).

[19] R.L. Boring, D.I. Gertman, J.C. Joe, and J.L. Marble, *"Proof of concept for a human reliability analysis method for heuristic evaluation of software,"* Proceedings of the 49th Annual Meeting of the Human Factors and Ergonomics Society, 676-680, 2006.

[20] R.L. Boring, T.Q. Tran, D.I. Gertman, and A.S. Ragsdale, *"A human reliability based usability evaluation method for safety-critical software,"* Fifth International Topical Meeting on Nuclear Plant Instrumentation, Controls, and Human Machine Interface Technology, 1275-1279, 2006.

[21] R. Mohlich and J. Nielsen, *"Improving a human-computer dialogue,"* Communications of the ACM, 33, 338-348, (1990).