ResearchGate

See discussions, stats, and author profiles for this publication at: https://www.researchgate.net/publication/307946395

Testing Subtask Quantification Assumptions for Dynamic Human Reliability Analysis in the SPAR-H Method

Article *in* Proceedings of the Human Factors and Ergonomics Society Annual Meeting · September 2016

DOI: 10.1177/1541931213601345

CITATIONS
0

READS
148

2 authors:

Ronald Laurids Boring
Idaho National Laboratory

306 PUBLICATIONS **2,819** CITATIONS

SEE PROFILE

Sarah Ewing
Atos S.A.

17 PUBLICATIONS **302** CITATIONS

SEE PROFILE

All content following this page was uploaded by Ronald Laurids Boring on 03 November 2017.

The user has requested enhancement of the downloaded file.

Testing Subtask Quantification Assumptions for Dynamic Human Reliability Analysis in the SPAR-H Method

Ronald L. Boring and Sarah M. Herberger
Idaho National Laboratory
Idaho Falls, Idaho 83415, USA

Human reliability analysis (HRA) to date has relied almost entirely on static methods. To provide a more precise model of human performance, dynamic HRA has been developed. Dynamic HRA must model a range of human actions typically at a finer resolution than is accounted for by overall tasks or human failure events (HFEs) used by static HRA. Parsing HFEs into finer units of analysis requires consideration of the mathematical underpinnings of the HRA methods that will be used for quantification. This paper serves to test how conventional HRA methods scale to this level of precision. SPAR-H, a static HRA method, was evaluated as part of this research. SPAR-H, which is based on HFEs, may require further refinement before its quantification approach can be employed in dynamic HRA.

BACKGROUND ON HUMAN RELIABILITY

The legacy of human reliability analysis (HRA) is that almost all methods to date have been static (Boring et al., 2015). Static HRA supports probabilistic risk assessment (PRA) by considering the human contribution to overall system risk. HRA may be successfully integrated into PRA in a well-established process (Bell & Swain, 1983; EPRI, 1992; IEEE, 1997). The key to this integration is the human failure event (HFE), which represents a clustering of human activities related to the operation of a particular system or component. The HFE can be quantified using any of a number of HRA methods (for recent surveys, see Bell & Holroyd, 2009; Chandler et al., 2006; and Kolaczkowski et al., 2005).

The HFE is integrated into the event trees used in the PRA. Often the clustering of activities under the HFE is done using fault tree logic. In practice, the HFE is defined as the entirety of human actions related to possible and relevant human interactions with a particular system. In other words, the HFE is defined top-down, from the PRA level of interest, to encompass all human actions that can contribute to the fault of a component or system modeled in the PRA.

Static HRA mimics the predominance of static PRA. The key point in static HRA and PRA is that events are analyzed for an assumed, typical window of time. The HFE for static HRA does not change as a function of time or the event progression; the event sequences are fixed in the HRA, and the analysis represents a snapshot of time. Either the analysis represents a very generic context in which the event would occur, or the analysis is agnostic to time, meaning that time evolution is simply not factored into the calculation of the human error probability (HEP). Therefore, other performance shaping factors (PSFs) apart from time drive the quantification of the HEP.

Boring (2007), among others, explains the conceptual shift from static HRA to dynamic HRA. Key aspects of this shift are the transition from predictions based on fixed or precoded models of accident sequences into predictions based on direct simulation of an accident sequence, with explicit consideration of timing of key events. For HRA to fit into this dynamic framework, its models must follow a parallel path, shifting away from estimating the probability of a static event, and into simulating the multitude of possible human actions relevant to an event.

Traditional static HRA attempts to directly estimate or assign probabilities to pre-defined HFEs. Example HFEs are *"failure to initiate feed and bleed"* and *"failure to align electrical bus to alternative feed."* In the dynamic HRA framework, the focus shifts to simulating the human performance within a dynamic PRA framework and using the results of those simulations to assign the HEP. Dynamic HRA yields HFEs such as *"failure to initiate feed and bleed at this precise time"* or *"failure to align electrical bus to alternative feed over a window of time."*

The promise of dynamic methods is their ability to model performance more completely than the expert judgment processes required for static HRAs. The downside of dynamic methods is the increased methodological and implementational complexity.

Not subject to U.S. copyright restrictions. DOI 10.1177/1541931213601345

Static methods are based on analyzing human performance for a pre-defined set of tasks that are generally clustered as HFEs. The challenge in extrapolating from these HFE snapshots to dynamic models is that many of the basic assumptions of these methods have not been validated for dynamic applications. For example, as depicted hypothetically in Figure 1 (from Boring, 2015a), a sequence of events can be parsed in many ways. The horizontal axis divides the event along a chronological progression, in this case incremented in minutes. The dotted vertical lines demark subtasks during the sequence of events. Finally, the blue boxes denote HFEs. Each minute reveals a different outcome in terms of the dynamic human error probability (HEP) calculation. Similarly, the subtasks and HFEs track the changing HEP.

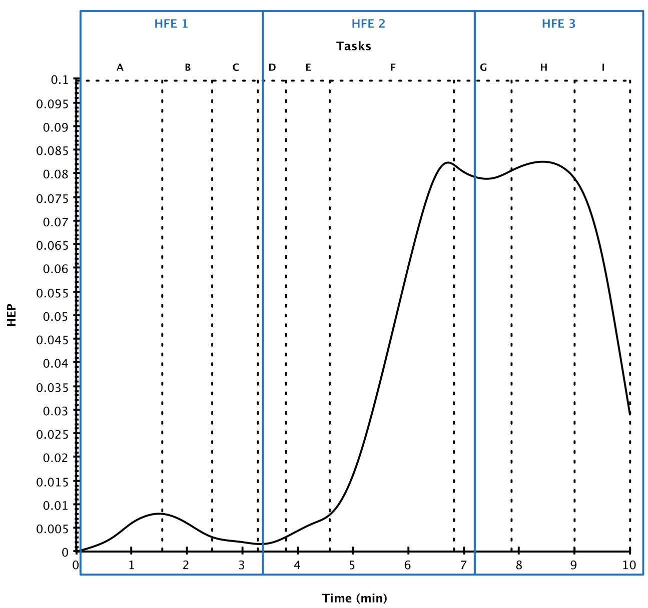

Figure 1. Human event progression according to time slices, subtasks, and HFEs.

Yet, HRA methods are not designed to track at all three levels of delineation. A HRA method that is applied successfully to three sequential HFEs as part of an event progression may not adequately cover further delimiting the HFE into 9 subtasks or 10 one-minute long time slices. To model dynamically the event progression, it is necessary to model the HFE at a finer granularity corresponding to the 9 subtasks or 10 time slices. A static HRA method may not lend itself to these different units of analysis. Moreover, the error quantification approach used may not prove accurate for the different unit of analysis (Boring, 2015b; Rasmussen and Laumann, in press).

A typical human event progression with respect to slices of time, subtasks, and HFEs is displayed in Figure 1 above. To frame such an event differently, consider the case of a major flooding incident with major damage to the plant sustained around the 4-minute mark along the timeline. HFE 1 corresponds to the pre-initiator prior to the flooding event, HFE 2 encompasses the initiating event of the flood, and HFE 3 spans the post-initiator recovery. As can be seen, the HEP remains low during the pre-initiator period, surges during the initiating event, and remains high during the recovery period. Static HRA methods, which would tend to analyze the event in terms of the three HFEs, may not fully model the changes to operator performance within each HFE. For example, the surge in error during HFE 2 (likely caused by sudden increases in stress) actually consists of several different slopes of the error plot—an initial relatively flat period, a rapidly rising period, and a plateau that shows signs of gradually declining. The flooding has differing effects on the plant and the operators, but conventional static parsing of the event may not fully map the dynamic progression of the event and the equally dynamic error curve associated with different tasks and time slices.

HRA is significantly affected when the unit of analysis is changed from an HFE to a unit of analysis suitable for dynamic modeling. Underlying this discussion is the key assumption that dynamic HRA requires a finer grain of modeling precision than the HFE. Ideally, the HFE represents a thorough human factors subtask analysis. The human reliability analyst will then quantify the event at the appropriate level of aggregation. HRA methods treat the unit of quantification differently. For example, the original HRA method, the Technique for Human Error Prediction (THERP; Swain and Guttman, 1983) quantifies at the subtask level. In contrast, the Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) method (Gertman et al., 2005) analyzes events at the HFE level, despite being derived from THERP (Boring and Blackman, 2007). Ideally, the quantification approach should accommodate different framings of the event space. The different levels of task decomposition can dramatically change the resulting error quantification, and HRA methods are generally vulnerable to spurious HEPs when not calibrated to the right level of decomposition (Rasmussen and Laumann, in press).

To consider the implications of dynamic HRA on static HRA methods, this paper reviews the topic of

HRA quantification as HFEs are translated to subtasks or time slices. This paper also serves as a modeling proof for the transferability of static HRA quantification to dynamic applications.

EXPLORATION OF SPAR-H

The SPAR-H method (Gertman et al., 2005) is a widely accepted method to determine the HEP based on expert estimation using calculation worksheets. Estimations of the effects of eight PSFs are carried out using predefined multipliers and a nominal failure probability. In many HRA methods, including SPAR-H, context-specific probabilities are generated by multiplying a nominal HEP by multipliers representing the effect of specific context elements (generally represented by PSFs), the included subset of which were deemed relevant to the problem by the method developers. This has resulted in the following general equation:

$$HEP = NHEP * \prod PSF \qquad (1)$$

where *HEP* is the overall human error probability, *NHEP* is the nominal human error probability (which is assumed to be 1E-2 or 1E-3 in SPAR-H), and $\prod PSF$ is the product of all eight PSFs in the method (Gertman et al., 2005). PSFs come in many flavors, with SPAR-H defining: *available time, stress/stressors, complexity, experience, procedures, human-machine interface/ergonomics, fitness for duty,* and *work processes*. Each PSF has different levels with a corresponding multiplier for Diagnosis (e.g., cognitive) and Action (e.g., behavioral) tasks as seen in Table 1.

Table 1. Available time PSF in SPAR-H with its respective levels and the associated action and diagnosis multipliers.

PSFs	PSF Level	Multiplier for Action	Multiplier for Diagnosis
Available Time	Inadequate Time	P(failure)=1	P(failure)=1
	Time Available ≈ Time Required	10	10
	Nominal Time	1	1
	Extra Time Available	0.1	0.1
	Expansive Time Available	0.01	0.1 to 0.01
	Insufficient Information	1	1

The application of the PSF levels and multipliers produce the following equation:

$$\begin{aligned} HEP = NHEP &* \text{available time} * \text{stress} \\ &* \text{complexity} * \text{experience} \\ &* \text{procedures} * \text{ergonomics} \\ &* \text{fitness for duty} \\ &* \text{work process} \end{aligned} \qquad (2)$$

where each PSF is substituted with the respective PSF level's multiplier.

Table 2. Available time PSF in SPAR-H with its respective levels, action multiplier, action frequency, and action probability.

PSFs	PSF Level	Multiplier for Action	Action Frequency	Action Probability
Available Time	Inadequate Time	P(failure)=1	5	0.009
	Time Available ≈ Time Required	10	36	0.065
	Nominal Time	1	500	0.898
	Extra Time Available	0.1	10	0.018
	Expansive Time Available	0.01	4	0.007
	Insufficient Information	1	2	0.004

Each level of a PSF is not equally likely. Thus, the frequency of PSF level assignments was taken from Boring et al. (2006). Additionally, for the purposes of this exploratory analysis, only the SPAR-H Action worksheet PSF multipliers are used. A small excerpt of the data used for this simulation can be seen in Table 2.

HUMAN FAILURE EVENT SIMULATION

The simulation of human failure events is based on the probabilities of a PSF level in Table 2 and equation (2). A simulation of 5,000 data points was run to represent the distribution of a single task. This is then repeated for Tasks A, B and C, so that there are a total of 15,000 simulated points (see Figure 2 left).

Without taking into consideration the frequencies provided by Boring et al. (2006) and assuming that each PSF level is equally likely, the distributions of Tasks A, B, and C tend toward the probability of 100%, as seen in Figure 2 right. Verifying the results from the simulation, a one-way analysis of variance could be used to compare means of three or more groups. However, the distributions of the HFEs are clearly not normally distributed; thus a non-parametric approach, Kruskal-Wallis, is suggested for comparison purposes. Tasks A, B and C were compared using a Kruskal-Wallis analysis and received a *p*-value of 0.6186 with 2 degrees of freedom. This is what is expected as one task is generated from the same data as the other and does not differ much from another (see Figure 2).

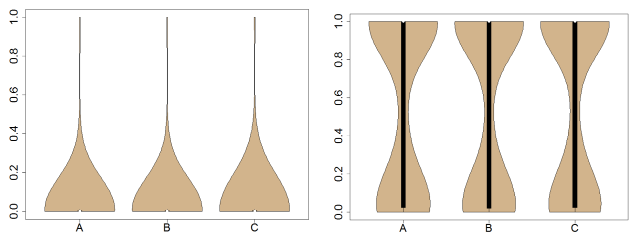

Figure 2. Violin plots of Tasks A, B and C taking into consideration PSF frequencies from Boring et al. (2006) (left). Tasks A, B and C assuming each PSF level is equally likely (right). Because A, B, and C are generated in the same manner, for 5,000 iterations A, B, and C are expected to have the similar distributions.

Multiple tasks are often grouped as single HFEs. SPAR-H assumes the unit of analysis is the HFE. If HFE 1 is comprised of Tasks A, B, and C (see Figure 1), there are then several ways to calculate the HFE based on a PSF multiplier or group of PSF multipliers (see Figure 3). The Maximum HFE calculation selects the largest values across Tasks A, B and C. The assumption is that the analysis should capture the strongest or most conservative manifestation of the PSF, even if the PSF changes across the evolution of the HFE. Median HFE selects the median value of the three tasks, and Average HFE calculates the average of the three tasks. The respective distributions for the different HFEs can be seen in Figure 3.

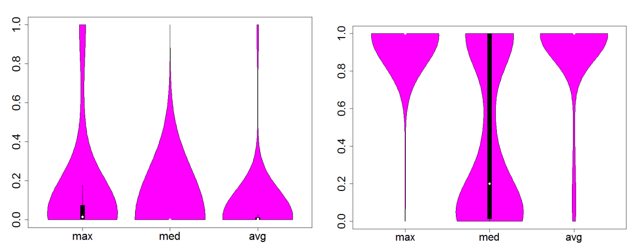

Figure 3. Violin plot of HFEs calculated three different ways from Tasks A, B, and C. The Maximum (max) calculation selects the largest of the three tasks. Median (med) selects the median value of the three tasks. Average (avg) calculates the average of the three tasks. The left plot is calculated using frequencies from Boring et al. (2006), while the right is calculated assuming a uniform frequency for all PSF levels.

Tasks A, B, and C and Maximum HFE were compared using a Kruskal-Wallis analysis and received a p-value < 0.001 with 3 degrees of freedom. Tasks A, B, and C and Average HFE were compared using a Kruskal-Wallis analysis and received a p-value < 0.001 with 3 degrees of freedom. Both of these p-values indicate that Maximum HFE and Average HFE are significantly different from Tasks A, B, and C (Figure 4).

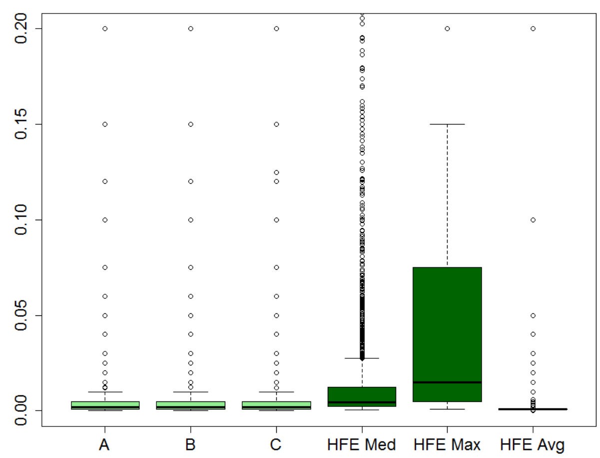

Figure 4. Task A, B, C, HFE Median, HFE Maximum, and HFE Average with frequencies from Boring et al. (2006). Each task was sampled 5,000 times from each PSF with frequencies.

Additionally, Tasks A, B, and C and Median HFE were compared using a Kruskal-Wallis analysis and received p-value < 0.001 with 3 degrees of freedom. While still significant, visually Median HFE is the closest in distribution to the three tasks, and the graphical representation can be seen in Figure 4. Generally, Maximum HFE overestimates Tasks A, B, and C, and Average HFE underestimates Tasks A, B, and C.

CONCLUSIONS

SPAR-H in a dynamic simulation can be seen in Figure 1. Viewing the situation from a more simplified position, how to define a Task or HFE into subtasks becomes burdensome if calculations change depending on the resolution defined, and whether the data is point estimates or units of time.

This paper reviewed basic assumptions of quantification and extrapolated from static modeling to dynamic modeling. The support of dynamic HRA requires the modeling of a range of human actions, typically at a finer resolution than is accounted for by the HFE used in static HRA. As the HFE is parsed into finer units of analysis, it is critical that the mathematical underpinnings of the HRA methods be used for quantification. This research evaluated quantification assumptions of the SPAR-H HRA method. With SPAR-H comes the use of the HFE as a unit of analysis, which is problematic when modeling the dynamic evolution of the event. This lack of fine resolution can result in spurious HEPs when using SPAR-H. Future efforts to develop a dynamic HRA approach will seek to refine these static methods for better application in dynamic contexts. Quantification in additional candidate HRA methods will also be explored.

DISCLAIMER

This work of authorship was prepared as an account of work sponsored by Idaho National Laboratory, an agency of the United States Government. Neither the United States Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately-owned rights. 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.

REFERERNCES

Bell, B.J., & Holroyd, J. (2009). *Review of Human Reliability Assessment Methods, RR679*. Buxton, UK: Health and Safety Executive.

Bell, B.J., & Swain, A.D. (1983). *A Procedure for Conducting Human Reliability Analysis for Nuclear Power Plants, Final Report, NUREG/CR-2254.* Washington, DC: U.S. Nuclear Regulatory Commission.

Boring, R.L. (2015a). A dynamic approach to modeling dependence between human failure events. In L. Podofillini (ed.), *Safety and Reliability of Complex Engineered Systems* (pp. 2845-2851). London: Taylor & Francis Group.

Boring, R.L. (2015b). Defining human failur events for petroleum applications of human reliabilty analysis. *Procedia Manufacturing, 3,* 1335-1342.

Boring, R.L. (2007). Dynamic human reliability analysis: Benefits and challenges of simulating human performance. In T. Aven & J.E. Vinnem (Eds.), *Risk, Reliability and Societal Safety, Vol. 2: Thematic Topics. Proceedings of the European Safety and Reliability Conference (ESREL 2007)* (pp. 1043-1049). London: Taylor & Francis.

Boring, R. L., & Blackman, H. S. (2007). *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.

Boring, R., Mandelli, D., Joe, J., & Groth, K. (2015). *A Research Roadmap for Computation-Based Human Reliability Analysis, INL/EXT-15-36051.* Idaho Falls: Idaho National Laboratory.

Boring, R.L., Whaley, A.M., Tran, T.Q., McCabe, P.H., Blackwood, L.G., & Buell, R.F. (2006). *Guidance on Performance Shaping Factor Assignments in SPAR-H, INL/EXT-06-11959.* Idaho Falls: Idaho National Laboratory.

Chandler, F.T., Chang, Y.H.J., Mosleh, A., Marble, J.L., Boring, R.L., & Gertman, D.I. (2006). *Human Reliability Analysis Methods: Selection Guidance for NASA*. Washington, DC: NASA Office of Safety and Mission Assurance Technical Report.

Electric Power Research Institute (EPRI). (1992). *SHARP1—A Revised Systematic Human Action Reliability Procedure, EPRI-101711.* Palo Alto: Electric Power Research Institute.

Gertman, D., Blackman, H., Marble, J., Byers, J., & Smith, C. (2005). *The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883.* Washington, DC: U.S. Nuclear Regulatory Commission.

Institute of Electrical and Electronics Engineers (IEEE). (1997). *uide for Incorporating Human Action Reliability Analysis for Nuclear Power Generating Stations, IEEE 1082.* New York: Institute of Electrical and Electronics Engineers.

Kolaczkowsk, A., Forester, J., Lois, E., & Cooper, S. (2005). *Good Practices for Implementing Human Reliability Analysis (HRA), Final Report, NUREG-1792.* Washington, DC: U.S. Nuclear Regulatory Commission.

Rasmussen, M., & Laumann, K. (in press). Decomposition level of quantification in human reliability analysis. *Proceedings of the 2016 European Safety and Reliabilty (ESREL) Conference.*

Swain, A. D., & Guttman, H. E. (1983). *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR-1278.* Washington, DC: U.S. Nuclear Regulatory Commission.

View publication stats