INL/CON-14-32870
PREPRINT

One Size Does Not Fit All: Human Failure Event Decomposition and Task Analysis

Enlarged Halden Group Program Meeting

Ronald Laurids Boring

September 2014

One Size Does Not Fit All:
Human Failure Event Decomposition and Task Analysis

Ronald Laurids Boring, PhD
Idaho National Laboratory
Idaho Falls, Idaho 83415, USA
+1 208 526 0966, ronald.boring@inl.gov

Abstract

In the probabilistic safety assessments (PSAs) used in the nuclear industry, human failure events (HFEs) are determined as a subset of hardware failures, namely those hardware failures that could be triggered or exacerbated by human action or inaction. This approach is top-down, starting with hardware faults and deducing human contributions to those faults. Elsewhere, more traditionally human factors driven approaches would tend to look at opportunities for human errors first in a task analysis and then identify which of those errors is risk significant. The intersection of top-down and bottom-up approaches to defining HFEs has not been carefully studied. Ideally, both approaches should arrive at the same set of HFEs. This question remains central as human reliability analysis (HRA) methods are generalized to new domains like oil and gas. The HFEs used in nuclear PSAs tend to be top-down—defined as a subset of the PSA—whereas the HFEs used in petroleum quantitative risk assessments (QRAs) are more likely to be bottom-up—derived from a task analysis conducted by human factors experts. The marriage of these approaches is necessary in order to ensure that HRA methods developed for top-down HFEs are also sufficient for bottom-up applications. In this paper, I first review top-down and bottom-up approaches for defining HFEs and then present a seven-step guideline to ensure a task analysis completed as part of human error identification decomposes to a level suitable for use as HFEs. This guideline illustrates an effective way to bridge the bottom-up approach with top-down requirements.

1. Introduction

Human reliability analysis (HRA) methods do not have a consistent level of task decomposition at the human failure event (HFE) modeling phase. This lack of consistency can result not only in different qualitative analyses but also different human error probabilities (HEPs). The level of task decomposition affects the dependency between tasks, which may have a further effect in driving the HEP. The issue is not that different HRA methods necessarily produce different results for the same HFE; rather, different HRA methods may decompose the HFE to different levels. Thus, the quantification of the same HFE may entail different assumptions and, to some extent, different groupings of tasks across HRA methods. In other words, because of a lack of a common task decomposition framework, HRA methods may not be using the same unit of analysis when producing the HEP.

For example, the European *Human Factors Reliability Benchmark Exercise* [11], also referred to as the "Ispra Study" within the HRA community, demonstrates how central this topic is to HRA. The benchmark featured three phases of analysis to compare HRA methods. Each successive phase served to further constrain the level of decomposition that defined the HFE. The first phase included identification of HFEs and their quantification by different analysis teams. Because different HFEs were identified across methods, it was difficult to compare method results directly. The second phase involved a more explicit definition of the HFEs to ensure the analysis teams quantified the same HFE. Even with a commonly defined HFE, there was considerable variability in how analysis teams modeled the HFE. Differences in task decomposition played a significant role in the differences of the HEPs for the HFEs. Some analysis teams decomposed to a finer level, resulting in lower HEPs. However, the dependencies between HFEs were not well accounted for in the analyses with finer grained task decomposition, resulting in unrealistically low HEP values in the authors' opinion. As such, a third phase was conducted, this time with an explicit decomposition of tasks and a common HRA event tree used in quantification, which resolved much of the interanalyst variability.

The purpose of this paper is to review existing approaches to modeling human error in HRA and synthesize the disparate approaches into a simple framework that can be used in support of HRA in petroleum applications. The goal of establishing a common framework for human error modeling is to eliminate potential sources of variability in HEP quantification across methods.

2. Human Errors vs. Human Failure Events

HRA depicts a cause and effect relationship of human error. The *causes* are typically catalogued in terms of qualitative contributions to a human error, including the processes that shaped that error and the failure mechanisms. The processes—cognitive, environmental, or situational—that affect human error are typically referred to as performance shaping factors (PSFs). The resultant *effect* is the manifestation of human error—often called the failure mode. This failure mode is treated quantitatively and has an associated failure probability, the HEP.

The term *human error* is often considered pejorative, as in suggesting that the human is in him- or herself the cause of the failure mode [8]. This belies the current accepted understanding that human error is the product of the context in which the human operates. In other words, it is not the human as the ultimate cause of the error but rather the failure mechanisms that put the human in a situation in which the error is likely to occur. The colloquial term, human error, is further challenged in that a human error may manifest but have little or no risk consequence. Human errors may be recovered or may simply not have a direct effect on event outcomes. Such risk insignificant occurrences are typically screened out of the HRA model.

Thus, to denote a risk significant human error, the term *human failure event* (HFE) has been posited. According to the American Society of Mechanical Engineers (ASME), a human failure event is "a basic event that represents a failure or unavailability of a component, system, or function that is caused by human inaction, or an inappropriate action" [4]. The HFE is therefore the basic unit of analysis used in PSA to account for HRA. While an HFE may be incorporated as a simple node in a fault tree or a branch in an event tree, the documentation supporting the HFE represents an auditable holding house for qualitative insights used during the quantification process. These insights may be simple to detailed, depending on the analysis needs and the level of task decomposition.

In PSAs used in the nuclear industry, as per the ASME definition, HFEs are determined as a subset of hardware failures, namely those hardware failures that could be triggered by human action or inaction. This approach is top-down, starting with hardware faults and deducing human contributions to those faults. Elsewhere, there is a bottom-up approach. More traditionally human factors driven approaches would tend to look at opportunities for human errors first in a task analysis and then model them in terms of potential for affecting safety outcomes. The order of identifying vs. modeling HFEs as shown in Figure 1 may be seen as changing depending on the approach. A top-down approach would tend to model the opportunity for HFEs and only then identify the sources of human error. In contrast, a bottom-up approach would first identify sources of human error and then model them in the PSA.

The intersection of top-down and bottom-up approaches to defining HFEs has not been carefully studied. Ideally, both approaches should arrive at the same set of HFEs. This question is crucial, however, because the HFEs used in nuclear PSAs tend to be top-down—defined as a subset of the PSA hardware faults—whereas the HFEs used in petroleum QRAs are more likely to be bottom-up—derived from a task analysis conducted by human factors experts. The marriage of these approaches is necessary in order to ensure that HRA methods developed for top-down HFEs are also sufficient for bottom-up applications. Figure 1 depicts the top-down and bottom-up approaches to defining HFEs. As can be seen, it is possible that both approaches arrive at the same solution. However, the solution set for the top-down and bottom-up approaches should be seen in terms of two circles in a Venn diagram. The problem is not that the HFEs may indeed overlap; the problem is that these HFEs may not always be identical.

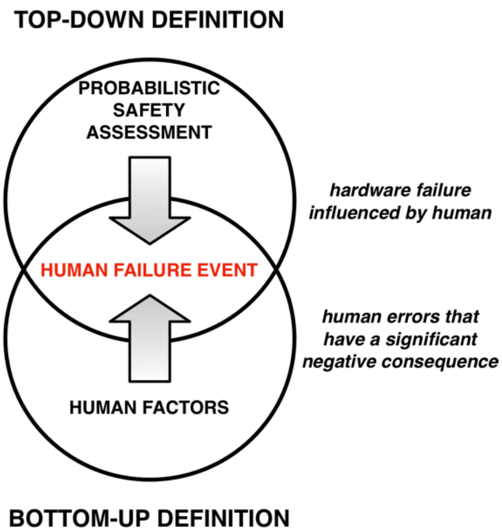

Figure 1. Two approaches to defining human failure events.

Additionally, some HFEs used in a petroleum context are derived from barrier analysis and are prospective in nature, designed to identify how the defense in depth of a system may be increased to ensure the safety of a system to be built. This approach may emphasize the evolving timescale of barrier effectiveness, whereas most conventional PSAs represent a static snapshot of an HFE. The barrier analysis approach is rarely used in contemporary PSAs for the nuclear industry. Additional guidance will be necessary to link the human factors processes for identifying vulnerabilities with the PSA fault modeling in HRA [6].

As depicted in Figure 1, there are areas covered in the bottom-up approach that are not necessarily covered by the top-down approach (and vice versa). Of interest, the top-down approach to defining HFEs begins by modeling those hardware systems that can fail and whose failure can be influenced by human actions or inactions. For example, if a particular electrical bus is a risk significant vulnerability to the overall system safety, the risk analyst would identify the failure of the bus as the starting point. He or she would next determine if the system is controlled by human operators. If yes, and if the human action is a significant subset of the overall risk of the bus failure, an HFE is modeled. The risk analyst must then determine what types of human errors are possible. This is often accomplished by referencing operating procedures and identifying which steps could be performed incorrectly. It is easier to identify a failure to execute particular required procedural steps than it is to postulate all the possible deviation paths the operator could follow that aren't encompassed by the procedure. In other words, the steps omitted (i.e., errors of omission) are more readily modeled than extra steps performed beyond the procedures (i.e., errors of commission). Thus, the top-down approach has exhibited far greater success in including relevant errors of omission than in anticipating possible errors of commission. We argue that the bottom-up approach, which considers all aspects surrounding task performance, provides better opportunity to incorporate these commonly omitted types of human error.

3. A Guideline for Top-Down and Bottom-Up Task Decomposition

In order to reconcile the disconnect between task analysis and HFEs in practice, this section presents a seven-step guideline for conducting a task analysis that culminates in usable input for HRA in a PSA. The guideline is appropriate for applications such as an HRA in which HFEs have not been predefined as part of the PSA. The guideline references specific methods to use, e.g., the SPAR-H HRA method [10], but the general approach is interchangeable with a variety of alternative methods. The guideline aligns the bottom-up task analysis approach with the top-down PSA approach for defining and analysing HFEs in HRA.

3.1 Step 1: Perform a Hierarchical Task Analysis (HTA)

Good explanations on how to conduct HTA can be found in any of the various summary articles by Annett published in the early 2000s [1-3]. HTA breaks down a given human-performed task according to goals and subgoals. Typical steps include:

1. Align analysis level of decomposition to the purpose of the analysis (e.g., decompose tasks according to the level that human error identification is possible, which serves as the stop rule for the analysis).
2. Determine task goals.
3. Acquire data to support and document the task decomposition.
4. Iterate the task breakdown with subject matter experts until the detail is accurate and sufficient.
5. Filter for the most significant operations (e.g., determine that the task has a reasonable probability and consequence of error—otherwise the task should not be considered further in the analysis).
6. Identify means to solve identified problems in the task analysis.

For HRA, the HTA needs to decompose to the subgoal/plan level where the analyst can look concretely at opportunities for error. For HRA specific purposes (in contrast to human factors focused on the design of new systems), an analyst would not necessarily need to look at the opportunity to remedy these potential failures, although the analyst should in his or her data collection identify opportunities for recovery from any potential failures.

3.2 Step 2: Review the Subgoals According to an Error Taxonomy

The result of the HRA is a set of goals that cluster sets of actions in the task and a set of subgoals that comprise the individual steps to achieving the goal. Each subgoal may be further subdivided into additional sub-subgoals and so on necessary to achieve that subgoal. After the completion of the HRA, the next step is to review the significant subgoals for their potential for human error. A number of taxonomies exist, e.g., the Systematic Human Error Reduction and Prediction Approach (SHERPA) [8] or the Technique for the Retrospective and Predictive Analysis of Cognitive Errors (TRACEr) [13]. The SHERPA and TRACEr error mode taxonomies, among numerous others (see [14] for a review), are functionally fairly similar and may be used interchangeably as desired. Table 1 presents the SHERPA taxonomy as an example.

Table 1. SHERPA error taxonomy.

Action Errors	**Checking Errors**
A1-Operation too long/short	C1-Check omitted
A2-Operation mistimed	C2-Check incomplete
A3-Operation in wrong direction	C3-Right check on wrong object
A4-Operation too little/much	C4-Wrong check on right object
A5-Misalign	C5-Check mistimed
A6-Right operation on wrong object	C6-Wrong check on wrong object
A7-Wrong operation on right object	**Retrieval Errors**
A8-Operation omitted	R1-Information not obtained
A9-Operation incomplete	R2-Wrong information obtained
A10-Wrong operation on wrong object	R3-Information retrieval incomplete
Information Communication Errors	**Selection Errors**
I1-Information not communicated	S1-Selection omitted
I2-Wrong information communicated	S2-Wrong selection made
I3-Information communication incomplete	

This taxonomy provides a concise account of errors of omission and commission. If the bottom-level subgoals do not present a reasonable opportunity for error according to the taxonomy, the analyst should eliminate that overarching task from the analysis. For each subgoal that presents an opportunity for error, the analyst should document the opportunities for recovery from these errors. For typical HRA, it is not necessary to complete the remedy analysis portion of SHERPA (which is useful for improving the process in a design review) but rather to identify actions the system or human might take to recover. The main purpose of the taxonomic review is simply to screen for the reasonableness that an error could occur at the subgoal level.

3.3 Step 3: Identify Opportunities for Cognitive Errors

The SHERPA taxonomy considers mainly action, perceptual, and communication errors. It does not explicitly consider decision errors, although they are implied in some taxonomic items like *A10–Wrong operation on wrong object*. It is recommended that additional Decision items be added to the SHERPA taxonomy to ensure that decision errors are considered. Note that these Decision errors may overlap with other items in the taxonomy. This overlap is inconsequential to the analysis. Suggested decision errors to augment SHERPA's taxonomy are found in Table 2. Note that some taxonomies may be more complete with regard to cognitive types of errors and may not require this separate guidance in Step 3.

Table 2. Additional cognitive items to augment the SHERPA taxonomy.

Decision Errors
D1-Correct decision based on wrong/ missing information
D2-Incorrect decision based on right information
D3-Incorrect decision based on wrong/ missing information
D4-Failure to make a decision (impasse)

3.4 Step 4: Identify Risk Significant Events

Not all errors are risk significant. In order to align the analysis with the level of task decomposition appropriate in PSA, each identified task should be considered in terms of its opportunity to have an impact on safety, namely to cause or exacerbate a fault in a hardware system. If the task is not connected to a hardware system related to the safety of the facility or process, it can usually be discarded from the analysis. An exception occurs where safety of individuals (e.g., loss of life) is a modeled risk and is caused by human action or inaction.

3.5 Step 5: Synthesize Human Failure Events

Although there remains no authoritative definition of what essential elements comprise the HFE used in PSA, for the present purposes consider the HFE to be all actions tied to a particular hardware failure outcome. This means that typically all tasks related to a single hardware system or a single safety outcome should be grouped together as a single HFE. Failure to cluster tasks at the right level of HFE decomposition can result in spurious results as HRA methods are used for human error quantification. HFEs are typically modeled in the PSA as nodes in fault trees. Depending on the logic used to connect HFEs together, the overall logic model may represent multiplicative (AND gate) or summative (OR gate) functions of the HEPs. If a series of unaggregated tasks from the HTA is treated as individual HFEs, this can have the effect of lowering the overall HEP for AND gates or raising the overall HEP for OR gates.

3.6 Step 6: Consider Drivers on Performance

Many contemporary HRA methods use PSFs to account for the mechanisms that increase or decrease performance. A commonly used PSF-based HRA method is the Standardized Plant Analysis Risk-Human (SPAR-H) method [10]. The SPAR-H method is strictly a quantification approach that provides no guidance on defining the HFEs. As such, the method assumes pre-defined HFEs from the PSA, and it is particularly susceptible to generating spurious HEPs when the HFEs are not defined at the appropriate level of decomposition. SPAR-H was originally developed in support of the U.S. Nuclear Regulatory Commission's Accident Sequence Precursor program [5] and is therefore optimized to set the PSF levels retrospectively according to information available in an incident that has already occurred. As such, selecting the PSF

Table 3. SPAR-H performance driver table.

Available Time Would you have plenty of time (POSITIVE) or not enough time (NEGATIVE) to complete the tasks?	POSITIVE	NEGATIVE	N/A
Stress Would you feel stressed during the scenario? Would it improve (POSITIVE) or hinder (NEGATIVE) your performance?	POSITIVE	NEGATIVE	N/A
Scenario Complexity Would it be easy (POSITIVE) or difficult (NEGATIVE) to diagnose the situation properly?	POSITIVE	NEGATIVE	N/A
Execution Complexity Would it be easy (POSITIVE) or difficult (NEGATIVE) to carry out the activities in the scenario?	POSITIVE	NEGATIVE	N/A
Experience/Training Would your previous practice and training on this type of scenario improve (POSITIVE) or hinder (NEGATIVE) your performance?	POSITIVE	NEGATIVE	N/A
Procedures Would procedures adequately aid (POSITIVE) or fail to aid (NEGATIVE) the completion of the tasks?	POSITIVE	NEGATIVE	N/A
Human-Machine Interface Is the labeling and organization of information clear and helpful (POSITIVE) or unclear and misleading (NEGATIVE)?	POSITIVE	NEGATIVE	N/A
Fitness for Duty Are there physical factors like being wide awake (POSITIVE) or tired (NEGATIVE) that would affect your performance?	POSITIVE	NEGATIVE	N/A
Work Processes Would organizational factors help (POSITIVE) or get in the way (NEGATIVE) of resolving this scenario?	POSITIVE	NEGATIVE	N/A

levels prospectively can be challenging. It is important as part of data collection to consider *what-if* scenarios such as “Is time a factor in overall performance on this task?” or “Could the quality of procedures affect the outcome of this task?”

Per recent step-by-step guidance on SPAR-H [15], it is appropriate to consider only those PSFs that have a dominant effect on the outcome of the HFE. These PSFs should be considered in terms of their positive and negative effect. It is assumed that only the strongest or dominant drivers on performance should be considered in a prospective SPAR-H analysis. Table 3 may help to classify relevant PSFs for each HFE. Each HFE should be walked through with subject matter or process experts according to the SPAR-H PSFs. Any strongly positive or negative drivers should be discussed and used to select an appropriate PSF level. Note that recent work in the petroleum sector has redefined the basic list of SPAR-H PSFs to be more suitable to petroleum HRA [12]. Where appropriate, Table 3 may be modified to reflect a different list of PSFs, either through customization of the SPAR-H PSFs or use of another HRA method’s PSFs.

3.7 Step 7: Perform the Human Error Quantification

Once reasonable errors have been identified, the HFEs have been defined, and the drivers on performance have been predicted, it is finally possible to use this information to complete the human error quantification. Again, using SPAR-H as an example, in the final stage, the analyst would complete the quantification worksheets. The SPAR-H process entails:

1. Determination if the HFE is either Diagnosis (cognitively engaging), Action (behaviorally based), or both. This determines the starting or nominal HEP (NHEP).
2. Assign the PSF levels for the HFE.
3. Mathematically compute the basic HEP by taking the product of the nominal HEP and the PSF multipliers that correspond to the PSF assignment levels. If there are both Diagnosis and Action worksheets, the basic HEP is computed separately for Diagnosis and Action and then added:

$$HEP_{BASIC} = NHEP_{DIAGNOSIS} \times \prod PSF + NHEP_{ACTION} \times \prod PSF \qquad (1)$$

4. Apply a correction factor to ensure no HEP is greater than 1.0.
5. Adjust the HEP for dependency between HFEs where appropriate.

It should be remembered that most HFEs include both Diagnosis and Action tasks [15]. Only dominant drivers on performance should be considered as PSFs, and any assumptions that inform PSF level assignment should be carefully considered and documented. Recovery actions identified as part of the HRA and SHERPA analysis (see Steps 1 and 2 above) should be documented. If recovery actions are modeled separately in the PSA event or fault trees, they should simply be documented as part of the analysis. As noted in Section 2.8 of [10], recovery may also be modeled within an HFE as positive effects on PSFs. This latter approach must be clearly called out in the analysis documentation in order to avoid double-counting recovery actions in the overall PSA. Where a sequence of successive HFEs is present, dependency should be considered in the analysis per the SPAR-H worksheets. Note that recent guidance [15] has suggested that dependency corrections to the HEP should in many cases be omitted, as it can result in overly conservative values. Again, all such assumptions should be carefully documented to aid in the potential verification and reuse of HRAs. Similar processes can be followed with other HRA methods as preferred.

4. Discussion

This paper has reviewed the difficulty in defining HFEs for human factors analysts who are completing HRAs where no PSA is readily available or the PSA does not include HFEs. As HRA is applied to novel domains beyond its traditional use in nuclear power, the completeness of PSAs, particularly with regard to treatment of human actions, will vary considerably. In the absence of predefined HFEs in the PSA, it is important to have a standard, scrutable approach to facilitate HRAs that can serve as standalone analyses and can be incorporated into PSAs with the right level of information and granularity. The challenge for the analyst in completing the bottom-up, task analysis driven approach is that there has been no clear process to bridge the task analysis to the completed HRA. This paper has provided initial guidance to help the analyst in such cases. In particular, it is important that the bottom-up defined tasks properly match the granularity required of the HFEs to fit the PSA. It is hoped the guidance in this paper will aid analysts in making this match and in ensuring interanalyst consistency in future HRAs. This guidance should, of course, serve only as a starting point, and analysts should refine and revise it according to their needs.

5. Disclaimer

A portion of this paper previously appeared as “Task Decomposition in Human Reliabilty Analysis” published in the Proceedings of the 12th Probabilistic Safety Assessment and Management Conference (PSAM) [7]. Idaho National Laboratory (INL) is a multi-program laboratory operated by Battelle Energy Alliance LLC, for the United States Department of Energy under Contract DE-AC07-05ID14517. This work has been carried out as part of The Research Council of Norway project number 220824/E30 “Analysis of human actions as barriers in major accidents in the petroleum industry, applicability of human reliability analysis methods (Petro-HRA)”. Financial and other support from The Research Council of Norway, Statoil ASA and DNV-GL are gratefully acknowledged. This paper represents the opinion of the authors, and does not necessarily reflect any position or policy of the above mentioned organizations. 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.

6. References

[1] Annett, J. (2000). Theoretical and practical influences on task analysis methods. In J.M. Schraagen, S.F. Chipman, and V.L. Shalin (eds.), *Cognitive Task Analysis.* Mahwah: Lawrence Erlbaum Associates.

[2] Annett, J. (2003). Hierarchical task analysis. In D. Diaper and N.A. Stanton (eds.), *Handbook of Task Analysis in Human-Computer Interaction.* Mahwah: Lawrence Erlbaum Associates.

[3] Annett, J. (2003). Hierarchical task analysis. In E. Hollnagel (ed.), *Handbook of Cognitive Task Design.* Boca Raton: CRC Press.

[4] American Society of Mechanical Engineers. (2009). Addenda to ASME/ANS RA-S–2008, Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications, ASME/ANS RA-Sa-2009. New York: American Society of Mechanical Engineers.

[5] 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.

[6] Boring, R.L., & Bye, A. (2008). Bridging human factors and human reliability analysis. *Proceedings of the 52nd Annual Meeting of the Human Factors and Ergonomics Society*, 733-737.

[7] Boring, R.L., & Joe, J.C. (2014). Task decomposition in human reliability analysis. *Proceedings of the 12th Probabilistic Safety Assessment and Management Conference.*

[8] Dekker, S. W. A. (2006). *The Field Guide to Understanding Human Error.* Aldershot, UK: Ashgate Publishing Co.

[9] Embrey, D.E. (1986). SHERPA: A systematic human error reduction and prediction approach. *International Meeting on Advances in Nuclear Power Systems.*

[10] Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C., and O'Reilly, P. (2005). *The SPAR-H Human Reliability Analysis Method, NUREG/CR-6883.* Washington, DC: US Nuclear Regulatory Commission.

[11] Poucet, A. (1989). *Human Factors Reliability Benchmark Exercise, Synthesis Report, EUR 12222 EN.* Luxembourg: Office for Official Publications of the European Communities.

[12] Rasmussen, M., and Laumman, K. (2014). The suitability of the SPAR-H "Ergonomics/HMI" PSF in a computerized control room in the petroleum industry. *Proceedings of the 12th Probabilistic Safety Assessment and Management Conference.*

[13] Shorrock, S.T., and Kirwan, B. (1999). The development of TRACEr: A technique for the retrospective analysis of cognitive errors in ATC. In D. Harris (ed.), *Engineering Psychology and Cognitive Ergonomics, 3*. Aldershot, UK: Ashgate Publishing Co.

[14] Stanton, N.A., Salmon, P.M., Rafferty, L.A., Walker, G.H., and Baber, C. (2013). *Human Factors Methods: A Practical Guide for Engineering and Design, Second Edition.* Aldershot, UK: Ashgate Publishing Co.

[15] Whaley, A.M., Kelly, D.L., and Boring, R.L. (2012). Guidance on dependence assessment in SPAR-H. *Joint Probabilistic Safety Assessment and Management and European Safety and Reliability Conference*, 27-Mo4-4.