INL/CON-14-32527
PREPRINT

# Top-Down and Bottom-Up Definitions of Human Failure Events in Human Reliability Analysis

## Proceedings of the Human Factors and Ergonomics Society Annual Meeting 2014

Ronald Laurids Boring

October 2014

This is a preprint of a paper intended for publication in a journal or proceedings. Since changes may be made before publication, this preprint should not be cited or reproduced without permission of the author. This document was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency thereof, or any of their employees, makes any warranty, expressed or implied, or assumes any legal liability or responsibility for any third party's use, or the results of such use, of any information, apparatus, product or process disclosed in this report, or represents that its use by such third party would not infringe privately owned rights. The views expressed in this paper are not necessarily those of the United States Government or the sponsoring agency.

Proceedings of the Human Factors and Ergonomics Society Annual Meeting

http://pro.sagepub.com/

# Top-Down and Bottom-Up Definitions of Human Failure Events in Human Reliability Analysis

Ronald Laurids Boring

*Proceedings of the Human Factors and Ergonomics Society Annual Meeting* 2014 58: 563
DOI: 10.1177/1541931214581119

# TOP-DOWN AND BOTTOM-UP DEFINITIONS OF HUMAN FAILURE EVENTS IN HUMAN RELIABILITY ANALYSIS

Ronald Laurids Boring
Idaho National Laboratory, Idaho Falls, Idaho 83415, USA

In the probabilistic risk assessments (PRAs) 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 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 is crucial, however, as human reliability analysis (HRA) methods are generalized to new domains like oil and gas. The HFEs used in nuclear PRAs tend to be top-down—defined as a subset of the PRA—whereas the HFEs used in petroleum quantitative risk assessments (QRAs) often tend 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.

## HUMAN ERROR AND HUMAN FAILURE EVENTS

Human reliability analysis (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 human error probability (HEP).

The purpose of this paper is to review existing guidance on modeling human error in HRA and synthesize the disparate guidance 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. This paper presents initial insights derived from a literature review of applicable sources. Additional guidance will be developed and reported in the future.

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 (Dekker, 2006). 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 from 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” (2009). The HFE is therefore the basic unit of analysis used in probabilistic risk assessment (PRA) 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 range from simple to detailed, depending on the analysis needs and the level of task decomposition.

In PRAs 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 may be seen as changing depending on the approach. A top-down approach would tend to model the opportunity for HFEs and 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 PRA.

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

Not subject to U.S. copyright restrictions. DOI 10.1177/1541931214581119

nuclear PRAs tend to be top-down—defined as a subset of the PRA—whereas the HFEs used in petroleum quantitative risk assessments (QRAs) tend 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.

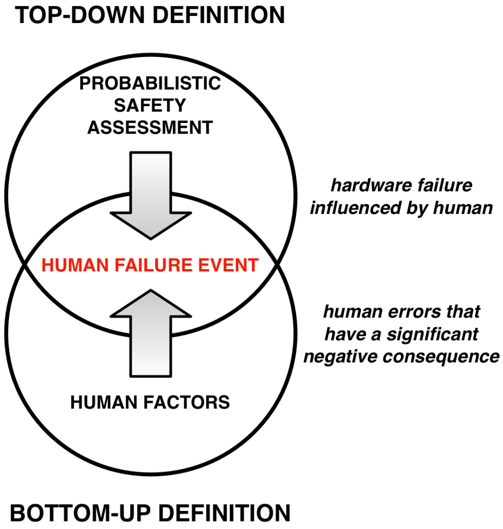
TOP-DOWN DEFINITION
PROBABILISTIC SAFETY ASSESSMENT
hardware failure influenced by human
HUMAN FAILURE EVENT
human errors that have a significant negative consequence
HUMAN FACTORS
BOTTOM-UP DEFINITION

**Figure 1. Two approaches to defining HFEs,**

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 PRAs represent a static snapshot of an HFE. The barrier analysis approach is rarely used in contemporary PRAs for the nuclear industry. Additional guidance will be necessary to link the human factors processes for identifying vulnerabilities with the PRA fault modeling in HRA (Boring and Bye, 2008).

## LIMITATIONS OF THE TOP-DOWN APPROACH

As depicted in Figure 1, there are areas covered in the bottom-up approach that aren't necessarily covered by the top-down approach (and vice versa). In this section, I discuss two noted shortcomings of the traditional top-down approach to defining HFEs—namely, errors of commission and latent errors, neither of which is adequately accounted for in traditional PRAs. I argue that the bottom-up approach provides opportunity better to incorporate these commonly omitted types of human error.

As noted, 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.

Already in one of the key early HRA textbooks, Gertman and Blackman (1994) elaborate on how the HRA methods of that time did not account for errors of commission adequately, particularly ones that are more cognitive in nature. That is, while the earliest HRA method, THERP (Swain & Guttmann, 1983), provides failure rates for manual control actions (e.g., simple, skill-based tasks which can include errors of commission that can be quantified), there was and is still no widely accepted approach that can account for errors of commission that fall outside of slips and lapses (and the PSFs that influence these error types). To model errors of commission that are more cognitively based (i.e., not skill-based manual control actions), Gertman and Blackman state that the practice at that time was to quantify errors of commission using simplified commission models (e.g., selection errors), or to use screening values to estimate a crew's probability of successfully diagnosing an event (e.g., SHARP – Hannaman & Spurgin, 1984; ASEP – Swain, 1987). Yet, I find that these methods still do not provide enough specific and useful guidance to help identify an actionable approach to bridge this gap between the top-down and bottom up approaches.

## EXISTING METHOD GUIDANCE

In this section, I briefly review a number of available methods, guidance documents, and standards for HRA to derive potential rules for decomposing tasks to define HFEs. The methods review is centered on U.S. approaches, since these have been the sources widely used by analysts and documented in nuclear applications. Additional insights may be derived by careful study of non-US HRA methods.

## U.S. HRA Methods

***THERP.*** The task analysis model in the Technique for Human Error Rate Prediction (THERP) is described in Chapter 4 of NUREG/CR-1278 (Swain and Guttman, 1983). It uses a goal-task breakdown of human activities to answer what are the goals of the human in terms of their interface with equipment such as controls. Task analysis classifies human activities into dynamic (involving interpretation and decision-making) and step-by-step (continuous or on-going) tasks. These tasks are included in the HRA event trees as branches. Since the tasks are modeled at the level of each step in a sequence of actions, the task decomposition may be considered quite detailed. These subtasks can be combined to represent an overall human action, and THERP provides clear guidance on aggregating subtasks during quantification.

Importantly, THERP provides a dependency model—which calculates how the relationship between subtasks should be treated in mathematical terms when aggregating the HEP. In other words, related tasks should not be double-counted when computing the likelihood of error. The dependency model in THERP has been adopted by almost every subsequent HRA model. The contemporary application of dependency is, however, considerably different from the original use in THERP. In the original THERP application, dependency was used to account for subtasks that were closely related, typically in terms of using the same crew, occurring close in time, with little new contextual information. Dependency modeled intra-task relations and not inter-task relations. In fact, the point at which no relationship between tasks existed was considered the point at which the task was fully defined and constituted a complete HFE. Ironically, current use of dependency is almost exclusively for inter-task relations *between* HFEs. This is a widespread misapplication of the original THERP guidance.

***ASEP.*** The Accident Sequence Evaluation Process (ASEP) method (Swain, 1987) came about as a simplification of THERP. It does not include a unique process to model HFEs but instead defers to THERP and to PRA judgment about relevant tasks to analyzed. In contrast to THERP, there is a stronger emphasis on the need not to analyze every task, particularly for screening analyses. Thus, the clear definition for an HFE provided in THERP was loosened by the time the ASEP was released.

**SPAR-H.** The Standard Plant Analysis Risk-Human Reliability Analysis (SPAR-H) method (Gertman et al., 2005) is a simplified HRA approach based in part on THERP (Boring and Blackman, 2007). SPAR-H provides no explicit guidance on task decomposition or modeling the HFE beyond considering action and diagnosis tasks separately. SPAR-H defers to the IEEE 1082 and ASME PRA standards (discussed below) for discussion on how to model HFEs (i.e., decompose the tasks) for inclusion in the PRA. SPAR-H assumes the HFE is predefined, and the method therefore does not devote extensive time to telling the analyst how to formulate the HFE.

***ATHEANA.*** A Technique for Human Error Analysis (ATHEANA; US Nuclear Regulatory Commission, 2000; see also Forester et al., 2007) provides nine overall steps, several of which are related to identifying HFEs:

- Step 1: Define and interpret issue of concern
- Step 2: Define scope of analysis
- Step 3: Describe base case scenarios
- Step 4: Define HFEs and unsafe actions
- Step 5: Identify potential vulnerabilities
- Step 6: Search for deviations from base case
- Step 7: Evaluate recovery potential
- Step 8: Quantification
- Step 9: Incorporation into PRA

Using the ATHEANA approach, it is possible to determine if the modeled event should be considered an HFE or an unsafe action (UA). The delimiter is based on the consequence in terms of contribution to core damage—a UA is akin to a human error that is not necessarily risk significant. The ATHEANA method also provides guidance on determining errors of commission.

In practice, the final ATHEANA step—incorporation into PRA—is not as clearly articulated as the other steps, leading to some problems using ATHEANA to define HFEs. ATHEANA takes a holistic approach to HRA, and its task decomposition may be seen at the scenario or overall unsafe action level. ATHEANA considers unsafe acts, but the specific aggregation of these into the HFE remains underspecified. Importantly, ATHEANA considers deviations from nominal scenarios. These represent possibly unsafe conditions at the plant caused by operator action or inaction. The likelihood of these nominal scenarios is considered in the quantification of the overall HFE. Additional guidance (Forester et al., 2007) states that the human reliability analysts should work with the PRA team to model the HFE consistent with the PRA. This latter guidance points to a lack of clear guidance on modeling the HFE at a level consistent with the PRA. Since the SPAR-H method (Gertman et al., 2005) points analysts to the ATHEANA method specifically to identify and model human errors as needed, there is a troubling disconnect between both ATHEANA and SPAR-H and the practicable HFE in a PRA.

***CBDT.*** The Cause-Based Decision Tree (CBDT) method (EPRI, 1992), widely used in industry, uses a decision tree approach to arrive at the quantification of HFEs based on key pieces of information (decision points) about operator performance. The method uses the SHARP1 framework for task decomposition as described in the next section.

## Standards and Guidance Documents

***SHARP1.*** The Systematic Human Action Reliability Procedure Revision 1 (SHARP1; EPRI, 1992) is an extension of the original SHARP process used for integrating HRA into the PRA process. The first stage of the SHARP1 process is the identification of the HFEs that are quantified in a subsequent stage. This first stage outlines five steps to arrive at the HFE:

1. Define the human interactions with the system that are potentially of interest. These are typically those that could leave some part or function of the plant unavailable. The procedure recommends identifying both errors of omission and commission as they might impact the plant.
2. Screen these human interactions to reduce the scope of the analysis to those that are most important.
3. Break down subtasks according to procedures to identify those subtasks that may have an impact on the plant. The emphasis here is to identify any tasks that may leave specific parts or functions of the plant unavailable, even if only temporarily as part of routine plant operations.
4. Perform an impact assessment to determine what effect the human subtasks have on plant equipment and plant state.
5. Integrate the human interactions as HFEs into the plant PRA model.

The approach falls short at defining an adequate way to decompose the overall event in terms of analysis. For example, if applying the five steps, particularly Step 3, it would not be clear whether to use a method to quantify the subtasks, groupings of subtasks, or the overall human interaction with the system, which could encompass hundreds of subtasks. This lack of decomposition can result in a myriad of HEPs, as many methods are not sensitive to subtask versus task level analysis.

***IEEE-1082 (1997).*** This standard, currently under revision, advocates a “stepwise” incorporation of human actions into the PRA model. The process begins with a complete but moderately detailed inclusion of human actions. These actions are considered in terms of risk-significance, such that only human actions that truly drive core damage frequency should be considered. A screening analysis narrows the number of human actions that are considered in the PRA. Some actions may be revisited at a later time when additional detail is added to the PRA model. The HFEs that are risk significant are modeled in sufficient detail to allow quantification.

The IEEE-1082 standard is a very high level document. It better addresses screening human errors for risk significance than actually defining those errors as HFEs.

***ASME/ANS RA-Sa (2009 Revision).*** The ASME PRA standard explicates a number of important points to consider in HRA but does not provide specific recommendations on modeling HFEs beyond providing a formal definition of HFEs. The standard requires documentation of the identification, characterization, and quantification of pre-initiator, post-initiator, and recovery human actions, but it does not advocate a particular approach or recommend the appropriate level of decomposition. According to the standard, HFEs must be defined and included for each human activity that isn’t screened out and must be defined to reflect the resulting unavailability of a component, train, system, or function that is modeled in the PRA. The standard does provide guidance that several human activities may be grouped into a single HFE if the impact of the activities is similar. Three levels of HFEs are defined, differentiated by those HFEs that do not perform a task analysis, those that have a high-level task analysis (e.g., human impact at the train level), and those that have a detailed task analysis (e.g., human impact on individual components).

***Good Practices for HRA (NUREG-1792).*** As with the standards mentioned above, the *Good Practices* link the HFE to the specific hardware failure that results from the human action or inaction. The level of modeling (i.e., level of decomposition) should reflect the amount of plant hardware that is affected. Thus, the HFE may be defined at the component, train, system, or function level. Human actions may be grouped at a higher level as appropriate. For example, if multiple human actions affect multiple components in the train, the HFE should be modeled at the train level. If, however, quantification differs considerably between the component and train level of modeling, the more conservatively bounding HFE definition should be used. If grouping multiple actions masks the potential for considering subsequent dependencies, the actions should be modeled as individual HFEs.

This guidance is helpful in establishing the boundaries between HFEs in an event evolution, although it fundamentally reflects the top-down definition of the HFE.

## Other Considerations

**Dynamic modeling.** In developing human performance simulation models, the issue of task decomposition resurfaces. Simulation models like ADS-IDAC (Chang and Mosleh, 2007) feature the ability to model human performance at the very detailed subtask level, such as decision points and simple manual actions. While quantification is possible at the subtask level, HRA methods do not provide guidance for combining subtask HEPs into the HFE level appropriate for a PRA. Such combinatorial quantification must consider dependencies between subtasks, but there is the possibility to inflate HEP values if the aggregation algorithm does not properly consider the small subtasks that the simulation models use (Boring, 2007). Additionally, dynamic modeling reveals the need for PSF latency—namely, that PSFs must not be considered discretely without the lingering effect from one time point to another. For example, stress cannot simply be turned off because the underlying cause of that stress has disappeared. This insight suggests that PSFs may need to play a role in defining the HFEs, or at least the boundaries between HFEs.

## CONCLUSIONS

Defining an HFE for use in novel HRA applications still remains somewhat elusive. Although general guidance exists for the top-down approach, there remains a large PRA expertise requirement for actually decomposing groups of subtasks into an HFE suitable for inclusion in the PRA. While approaches exist for bottom-up definitions, these still do not adequately address topics such as latent errors or errors of commission. Nonetheless, several candidate principles of HFE modeling have emerged from my review in this paper:

- Until clear guidance is available to identify commonalities and differences between the top-down and bottom-up approaches, it is desirable to employ a combination of both approaches to define the HFE.
- When adopting the top-down approach, the definition of the HFE should start broad, identifying those human actions and inactions that may trigger the unavailability of components, systems, or functions.
- These broad HFEs should be screened to determine the risk significant activities. The risk significant activities are the primary HFEs that are modeled in greater detail in the HRA.
- Task analysis of these risk significant activities may reveal additional sources of failures that may not be anticipated in the initial definition of the HFE. This represents the bottom-up approach. The definition of the HFE and screening should be an iterative process to arrive at a complete and relevant model of the human contribution to the overall system risk.
- Bottom-up approaches should consider errors of commission and latent errors in crafting the HFEs.
- Subtasks may reasonably be grouped into a single HFE provided that they are logically related; they do not represent different tasks, personnel, or equipment; and they do not mask dependencies that need to be accounted for.
- The earliest HRA methods used a simple equipment-level task decomposition. This is the level of flipping a switch. As interfaces have progressed in complexity, the interaction of the human with the equipment may represent a much higher level of decomposition that includes more cognitive or diagnostic activities. It is insufficient to define HFEs in terms of simple tasks—it must include a significant cognitive component as well.

These principles will be refined and developed into comprehensive guidance for defining HFEs. Ultimately, one key goal of is to bridge the gap in existing HRA guidance and application to the petroleum domain. Current practice follows a somewhat vague top-down approach of using predefined HFEs from the PRA. As HRA is refined for oil and gas applications, it will need to include a clear bottom-up approach compatible with QRA.

## DISCLAIMER

This work of authorship was prepared as an account of work sponsored by 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.

## REFERENCES

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.

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, Volume 2: Thematic Topics. Proceedings of the European Safety and Reliability Conference* (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.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.

Chang, Y.H.J., and Mosleh, A. (2007). Cognitive modeling and dynamic probabilistic simulation of operating crew response to complex system accidents, part 1: Overview of the IDAC model. *Reliability Engineering and System Safety, 92*, 997-1013.

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

EPRI. (1992). *An Approach to the Analysis of Operator Actions in Probabilistic Risk Assessment, TR-100259.* Palo Alto: Electric Power Research Institute.

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

Forester, J., Kolaczkowski, A., Cooper, S., Bley, D., and Lois, E. (2007). *ATHEANA User's Guide, NUREG-1880.* Washington, DC: US Nuclear Regulatory Commission.

Gertman, D. I., & Blackman, H. S. (1994). Human reliability and safety analysis data handbook. John Wiley & Sons, New York, NY.

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.

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

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

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

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

US Nuclear Regulatory Commission. (2000) *Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA)*, NUREG-1624, Rev. 1. Washington, DC: US Nuclear Regulatory Commission.