HUMAN ERROR QUANTIFICATION USING PERFORMANCE SHAPING FACTORS IN THE SPAR-H METHOD

Harold S. Blackman, David I. Gertman & Ronald L. Boring
Idaho National Laboratory, Idaho Falls, ID 83415, USA

This paper describes a cognitively based human reliability analysis (HRA) quantification technique for estimating the human error probabilities (HEPs) associated with operator and crew actions at nuclear power plants. The method described here, Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) method, was developed to aid in characterizing and quantifying human performance at nuclear power plants. The intent was to develop a defensible method that would consider all factors that may influence performance. In the SPAR-H approach, calculation of HEP rates is especially straightforward, starting with pre-defined nominal error rates for cognitive vs. action-oriented tasks, and incorporating performance shaping factor multipliers upon those nominal error rates.

INTRODUCTION

Human reliability analysis (HRA) assesses the safety and risk significance of pre-initiator and post-initiator human tasks performed at nuclear power plants (NPPs). The field of HRA has been called to task because of the apparent inability to tie HRA methods back to first principles in human behavior (Gertman and Blackman, 1994). Generally, HRA methods identify a set of factors believed to be related to performance, focus on classes of human error or behavior, and then manipulate those factors to arrive at a failure rate estimate for use in probabilistic risk analysis (PRA). The obvious problem with these approaches is completeness. How do we know that the set of identified factors is, in fact, complete? Moreover, are the factors flexible enough to encompass a wide range of human behaviors? The most sensible approach should begin with a comprehensive theory of human behavior, to ensure that all relevant factors are addressed and accounted for, and work to identify demonstrated, underlying mechanisms that we know influence and predict behavior.

The Standardized Plant Analysis Risk-Human Reliability Analysis (SPAR-H) method (Blackman and Byers, 1995; Byers et al., 1999; Gertman et al., 2005) was first released in 1995 as a simple-to-use approach for risk analysts to compute human error probabilities (HEPs). One way in which SPAR-H achieved simplicity was through the use of performance shaping factors (PSFs). A PSF is an aspect of the human's individual characteristics, environment, organization, or task that specifically decrements or improves human performance, thus respectively increasing or decreasing the likelihood of human error.

The cognitive and behavioral response model developed for SPAR-H evolved out of a desire to acknowledge what had been learned about human performance in early cognitive science approaches and is generally considered an information processing approach to human behavior. The factors that comprise the basic elements of this model also come from the literature surrounding the development and testing of general information processing models of human

performance. A review of operating experience in NPPs and a comparison of HRA methods revealed eight summary PSFs that contributed to human performance in NPP operation. These eight SPAR-H operational factors can be directly associated with the model of human performance and show the portion of the human information processing model with which they are associated (see Figure 1).
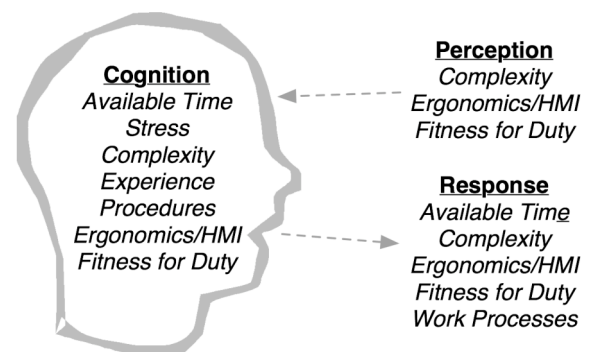

Figure 1. SPAR-H Performance Shaping Factors in the Information Processing Context.

ORIGINS OF THE SPAR-H METHOD

SPAR-H was originally called Accident Sequence Precursor (ASP) HRA (Blackman and Byers, 1995), in recognition of its use within the ASP program of the US Nuclear Regulatory Commission (NRC). The method was developed as a closely related alternative to two popular approaches at the time. A Technique for Human Error Rate Prediction (THERP; Swain and Guttman, 1983) had been formally available as a method for over ten years. THERP analyses required considerable training and topical mastery to complete. Because of the difficulty in completing a THERP analysis under strict time and resource constraints, a simplified version of THERP was commissioned in 1987 and called the Accident Sequence Evaluation Program Human Reliability Analysis Procedure (ASEP; Swain, 1987). While based on THERP, ASEP estimates diverged

from those in THERP. Moreover, the technique was often emphasized as a screening HRA method, meaning its use was primarily to provide rough estimates of error likelihood for risk determination. This approach contrasted with the nuanced results offered by THERP, offering in exchange a significant time savings and greater simplicity in terms of completing an analysis.

SPAR-H was born out of THERP and ASEP as a further simplification and generalization of these two approaches. Instead of mapping human activities to specific scenario templates as in THERP and ASEP, SPAR-H utilized PSFs to better generalize across the full spectrum of human performance while maintaining common quantification underpinnings. The original ASP HRA method was refined in 1999 and adopted the name of the Standardized Plant Analysis Risk (SPAR) PRA models developed in support of the US NRC (Byers et al., 1999). This latter acronym, SPAR HRA, more clearly delineated the method from ASEP. The 2005 and most recent revision (Gertman et al., 2005) adopted the acronym SPAR-H, whereby the H signified that this method was connected specifically with HRA vs. the broader PRA focus of the SPAR models.

TASK TYPES

Although it has a basis in human cognition and behavioral research, SPAR-H strives to present a simple approach to HEP estimation. Thus, in SPAR-H, quantification is based upon the assignment of tasks to one of two types, *diagnosis* or *action*. Examples of action tasks include operating equipment, conducting calibration or testing, and other activities performed during the course of following plant procedures or work orders. Tasks that constitute proceduralized actions are evaluated as action tasks. Diagnosis tasks consist of cognitively engaging planning and prioritizing activities, determining appropriate courses of action, and using knowledge and experience to understand existing conditions.

SPAR-H suggests that for cognitively engaging tasks such as diagnosis, people tend to exhibit a nominal human error rate equal to 0.01 (or 1E-2), excluding any adjustment for PSFs or dependencies within a chain of events. For tasks that are more action oriented, the nominal human error rate is equal to about 0.001 (or 1E-3), excluding adjustments.

The nominal human error rate for diagnosis activities is based on the value found in THERP Table 20-1, Item 4, corresponding to the median HEP for a control room diagnosis task within 30 minutes. This follows the so-called 30-minute rule in control room activity—a general rule for how long operators should have available before they are required to take action (IAEA, 1980). The nominal HEP for action was derived from numerous representative action tasks in THERP (see Table 20-7, Item 1; Table 20-9, Item 3; Table 20-11, Items 1 and 2; Table 20-12, Item 3; and Table 20-13, Item 1).

PERFORMANCE SHAPING FACTORS

Whether the task type is determined to be diagnosis, action, or some combination of the two, a PSF analysis is performed and adjustments are made to calculate HEPs. This section presents in order corresponding to the SPAR-H worksheets, general definitions for the PSFs that constitute part of the analysis process. Note that when an analyst encounters a PSF for which there is inadequate information, the nominal assignment is assumed. The following PSF multiplier level assignments are applied to the nominal HEP as appropriate for computing HEPs in NPPs that are at power. Due to situational factors affecting human performance, marginally different levels apply to NPPs that are at low power or in shutdown modes.   Note that the derivation of the SPAR-H PSF multipliers from THERP is documented in Boring and Blackman (2007).

Available Time

Available time refers to the amount of time that an operator or a crew has to diagnose and act upon an abnormal event. A shortage of time can affect the operator's ability to think clearly and consider alternatives. It may also affect the operator's ability to perform. Definitions differ somewhat, depending on whether the activity is a diagnosis activity or an action. The diagnosis values are provided below:

- *Inadequate time*—If the operator cannot diagnose the problem in the amount of time available, no matter what s/he does, then failure is certain. (Probability of failure = 1)
- *Barely adequate time*—2/3 the average time required to diagnose the problem is available. (Multiplier = 10)
- *Nominal time*—on average, there is sufficient time to diagnose the problem. (Multiplier = 1)
- *Extra time*—time available is between one to two times greater than the nominal time required, and is also greater than 30 minutes. (Multiplier = 0.1)
- *Expansive time*—time available is greater than two times the nominal time required and is also greater than a minimum time of 30 minutes; there is an inordinate amount of time (a day or more) to diagnose the problem. (Multiplier = 0.01)

Stress and Stressors

Stress (and level of arousal) have been broadly defined and used to describe negative as well as positive motivating forces of human performance. Stress as used in SPAR-H refers to the level of undesirable conditions and circumstances that impede the operator from easily completing a task. Stress can include mental stress, excessive workload, or physical stress (such as that imposed by difficult environmental factors). It includes aspects of narrowed attentional field or muscular tension, and can include general apprehension or nervousness associated with the importance of an event. Environmental factors often referred to as stressors, such as excessive heat, noise, poor

ventilation, or radiation, can induce stress in a person and affect the operator's mental or physical performance. It is important to note that the effect of stress on performance is curvilinear—some small amount of stress can enhance performance, and should be considered nominal, while high and extreme levels of stress will negatively affect human performance.

The assignment and definition of levels of stress or stressors is identical across action and diagnosis and action tasks:

- *Extreme*—a level of disruptive stress in which the performance of most people will deteriorate drastically. This is likely to occur when the onset of the stressor is sudden and the stressing situation persists for long periods. This level is also associated with the feeling of threat to one's physical well-being or to one's self-esteem or professional status, and is considered to be qualitatively different from lesser degrees of high stress (e.g., catastrophic failures can result in extreme stress for operating personnel because of the potential for radioactive release). (Multiplier = 5)
- *High*—a level of stress higher than the nominal level (e.g., multiple instruments and annunciators alarm unexpectedly and at the same time; loud, continuous noise impacts ability to focus attention on the task; the consequences of the task represent a threat to plant safety). (Multiplier = 2)
- *Nominal*—the level of stress that is conducive to good performance. (Multiplier = 1)

Complexity

Complexity refers to how difficult the task is to perform in the given context. Complexity considers both the task and the environment in which it is to be performed. The more difficult the task is to perform, the greater the chance for human error. Similarly, the more ambiguous the task is, the greater the chance for human error. Complexity also considers the mental effort required, such as performing mental calculations, memory requirements, understanding the underlying model of how the system works, and relying on knowledge instead of training or practice. Complexity can also refer to physical efforts required, such as physical actions that are difficult because of complicated patterns of movements.

In general, a task with greater complexity requires greater skill and comprehension to complete successfully. Multiple variables are usually involved in complex tasks. Concurrent diagnosis of multiple events and execution of multiple actions at the same time is more complex than diagnosing and responding to single events. The multipliers are:

- *Highly complex*—very difficult to perform. There is much ambiguity in what needs to be diagnosed or executed. Many variables are involved, with concurrent diagnoses or actions (i.e., unfamiliar maintenance task requiring high skill). (Multiplier = 5)

- *Moderately complex*—somewhat difficult to perform. There is some ambiguity in what needs to be diagnosed or executed. Several variables are involved, perhaps with some concurrent diagnoses or actions (i.e., evolution performed periodically with many steps). (Multiplier = 2)
- *Nominal*—not difficult to perform. There is little ambiguity in the diagnoses or actions. Single or few variables are involved. (Multiplier = 1)
- *Obvious diagnosis*—diagnosis becomes greatly simplified. There are times when a problem becomes so obvious that it would be difficult for an operator to misdiagnose it. (There is no *obvious action* PSF level counterpart. Easy to perform actions are encompassed in the nominal complexity rate.) (Multiplier = 0.1)

Experience and Training

This PSF refers to the experience and training of the operator(s) involved in the task. Included in this consideration are years of experience of the individual or crew, and whether or not the operator/crew has been trained on the type of accident, the amount of time passed since training, and the systems involved in the task and scenario. Another consideration is whether or not the scenario is novel or unique (i.e., whether or not the crew or individual has been involved in a similar scenario, in either a training or an operational setting). Specific examples where training might be deficient are guidance for bypassing engineered safety functions, guidance for monitoring reactor conditions during reactivity changes, and guidance for monitoring plant operation during apparently normal, stable conditions for the purpose of promoting the early detection of abnormalities. The multipliers are as follows:

- *Low*—less than six months experience and/or training. This level of experience/training does not provide the level of knowledge and deep understanding required to adequately perform the required tasks; does not provide adequate practice in those tasks; or does not expose individuals to various abnormal conditions. (Multiplier = 10)
- *Nominal*—more than six months experience and/or training. This level of experience/training provides an adequate amount of formal schooling and instruction to ensure that individuals are proficient in day-to-day operations and have been exposed to abnormal conditions. (Multiplier = 1)
- *High*—extensive experience; a demonstrated master. This level of experience/training provides operators with extensive knowledge and practice in a wide range of potential scenarios. Good training makes operators well prepared for possible situations. (Multiplier = 0.1)

Procedures

This PSF refers to the existence and use of formal operating procedures for the tasks under consideration. Common problems seen in event investigations for

procedures include situations where procedures give wrong or inadequate information regarding a particular control sequence. Another common problem is the ambiguity of steps. PSF levels differ somewhat, depending on whether the activity is a diagnosis activity or an action. In situations where multiple transitions between procedures are required to support a task or group of tasks, SPAR-H suggests that the analyst adjust the PSF for complexity accordingly. If the procedures themselves are problematic, i.e., inadequate, then, the HRA analyst should assess the procedures and determine whether they should be assigned an "inadequate" or "poor" rating. The multipliers are as follows:

- *Not available*—the procedure needed for a particular task or tasks in the event is not available. (Multiplier = 50)
- *Incomplete*—information is needed that is not contained in the procedure or procedure sections; sections or task instructions (or other needed information) are absent. (Multiplier = 20)
- *Available, but poor*—a procedure is available but it is difficult to use because of factors such as formatting problems, ambiguity, or such a lack in consistency that it impedes performance. (Multiplier = 5)
- *Nominal*—procedures are available and enhance performance. (Multiplier = 1)
- *Diagnostic/symptom oriented*—diagnostic procedures assist the operator/crew in correctly diagnosing the event. Symptom-oriented procedures (sometimes called function-oriented procedures) provide the means to maintain critical safety functions. These procedures allow operators to maintain the plant in a safe condition, without the need to diagnose exactly what the event is, and what needs to be done to mitigate the event. There will be no catastrophic result (i.e., fuel damage) if critical safety functions are maintained. Therefore, if either diagnostic procedures (which assist in determining probable cause) or symptom-oriented procedures (which maintain critical safety functions) are used, there is less probability that human error will lead to a negative consequence. This being said, if the symptom-based procedure is found to be inaccurate or awkwardly constructed, then the procedures PSF should be negatively rated. (Multiplier = 0.5; diagnosis only)

Ergonomics and Human Machine Interaction

Ergonomics refers to the equipment, displays and controls, layout, quality, and quantity of information available from instrumentation, and the interaction of the operator/crew with the equipment to carry out tasks. Aspects of human machine interaction are included in this category. The adequacy or inadequacy of computer software is also included in this PSF. Examples of poor ergonomics may be found in panel design layout, annunciator designs, and labeling. The PSF levels are:

- *Missing/Misleading*—the required instrumentation fails to support diagnosis or post diagnosis behavior, or the instrumentation is inaccurate (i.e., misleading). Required information is not available from any source (e.g., instrumentation is so unreliable that operators ignore the instrument, even if it is registering correctly at the time). (Multiplier = 50)
- *Poor*—the design of the plant negatively impacts task performance (e.g., poor labeling, needed instrumentation cannot be seen from a work station where control inputs are made, or poor computer interfaces). (Multiplier = 20)
- *Nominal*—the design of the plant supports correct performance, but does not enhance performance or make tasks easier to carry out than typically expected (e.g., operators are provided useful labels; the computer interface is adequate and learnable, although not easy to use). (Multiplier = 1)
- *Good*—the design of the plant positively impacts task performance, providing needed information and the ability to carry out tasks in such a way that lessens the opportunities for error (e.g., easy to see, use, and understand computer interfaces; instrumentation is readable from workstation location, with measurements provided in the appropriate units of measure). (Multiplier = 0.5)

Fitness for Duty

Fitness for duty refers to whether or not the individual performing the task is physically and mentally fit to perform the task at the required time. Things that may affect fitness include fatigue, sickness, drug use (legal or illegal), overconfidence, personal problems, and distractions. Fitness for duty includes factors associated with individuals, but not related to training, experience, or stress. The levels are:

- *Unfit*—the individual is unable to carry out the required tasks, due to illness or other physical or mental incapacitation (e.g., having an incapacitating stroke). (Probability of failure = 1)
- *Degraded fitness*—the individual is able to carry out the tasks, although performance is negatively affected. Mental and physical performance can be affected if an individual is ill, such as having a fever. Individuals can also exhibit degraded performance if they are inappropriately overconfident in their abilities to perform. Other examples of degraded fitness include experiencing fatigue from long duty hours; taking cold medicine that leaves the individual drowsy and nonalert; or being distracted by personal bad news (such as news of a terminal illness diagnosis of a loved one). (Multiplier = 5)
- *Nominal*—the individual is able to carry out tasks; there is no known performance degradation. (Multiplier = 1)

Work Processes

Work processes refer to aspects of doing work, including

inter-organizational, safety culture, work planning, communication, and management support and policies. How work is planned, communicated, and executed can affect individual and crew performance. If planning and communication are poor, then individuals may not fully understand the work requirements. Work processes include consideration of coordination, command, and control. Work processes also include any management, organizational, or supervisory factors that may affect performance. Examples seen in event investigations are problems due to information not being communicated during shift turnover, as well as communication with maintenance crews and auxiliary operators. Measures could include amount of rework, risk worth of items in utility corrective action program backlog, enforcement actions, turnover, performance efficiencies, etc. The multipliers used in Work Processes include:

- *Poor*—performance is negatively affected by the work processes at the plant (e.g., shift turnover does not include adequate communication about ongoing maintenance activities; poor command and control by supervisor(s); performance expectations are not made clear). (Multiplier = 2)
- *Nominal*—performance is not significantly affected by work processes at the plant, or work processes do not appear to play an important role (e.g., crew performance is adequate; information is available, but not necessarily proactively communicated). (Multiplier = 1)
- *Good*—work processes employed at the plant enhance performance and lead to a more successful outcome than would be the case if work processes were not well implemented and supportive (e.g., good communication; well-understood and supportive policies; cohesive crew). (Multiplier = 0.8)

DISCUSSION

In an effort to characterize more completely those factors influencing human performance, HRA methods have proposed various numbers of PSFs. SPAR-H attempts to provide reasonable coverage of the spectrum of human performance influences in nuclear power plant operations within the framework of the minimum reasonable number of PSFs. The decision to use eight PSFs was based on a review of cognitive activities suggested by a basic information processing model and refined through field use. These PSFs harmonize with more recent human performance characterizations, including the *Good Practices for Implementing HRA* (Kolaczkowski et al., 2005). Further, the SPAR-H quantification values used for the PSFs were based on available data within HRA, especially data provided in the THERP method, thus bringing generalizing THERP to a broader range of human activities.

The SPAR-H method provides a potent extension of THERP that allows the analyst flexibility and generalizability beyond narrowly defined tasks and scenarios. This approach does not guarantee valid HEP estimates. It does nonetheless provide a useful tool for categorizing and quantifying human contributions to risk and for facilitating risk-informed decision making.

DISCLAIMER

This paper 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, 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 paper, or represents that its use by such third party would not infringe privately owned rights.

REFERENCES

Blackman, H.S., & Byers, J.C. (1995). *ASP Human Reliability Methodology Development, INEL-95/0139*. Idaho Falls: Idaho National Engineering Laboratory.

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.

Byers, J.C., Gertman, D.I., Hill, S.G., Blackman, H.S., Gentillon, C.D., Hallbert, B.P., & Haney, L.N. (1999). *Revision of the 1994 ASP HRA Methodology, INEEL/EXT-99-00041*. Idaho Falls: Idaho National Engineering and Environmental Laboratory.

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

Gertman, D., Blackman, H., Byers, J., Haney, L., Smith, C., & Marble, J. (2005). *The SPAR-H Method, NUREG/CR-6883*. Washington, DC: US Nuclear Regulatory Commission.

International Atomic Energy Agency (IAEA). (1980). *Protection System and Related Features in Nuclear Power Plants, Safety Series No. 50-SG-D3*. Vienna: International Atomic Energy Agency.

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

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

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