# Human Reliability Analysis in Cognitive Engineering and System Design

Ronald Laurids Boring
*Risk and Reliability Analysis Department*
*Center for Nuclear Energy and Global Security Technologies*
*Sandia National Laboratories*

The field of human factors engineering (HFE) combines interdisciplinary elements of engineering, psychology, and computer science, among other fields, into a cohesive discipline (Boring, 2002). Within this single discipline, there are numerous subdisciplines, including:

- *Cognitive engineering* (CE), which addresses the cognitive facets of human-system interaction to afford systems that maximize usability (Nielsen, 1993), enjoyment (Norman, 2002), or safety (Palanque et al., 2004).
- *Human reliability analysis* (HRA), which is primarily focused on verifying the safe performance of human actions, typically as part of an overall probabilistic risk assessment (PRA) that encompasses both the hardware system and the human in the loop.

Despite similarities in focusing on human-system interactions, the main difference between CE and HRA centers on when the approaches are deployed. Whereas CE is typically implemented in the design phase of the engineering cycle, HRA is often applied only in the verification and validation phase, after systems are built.

The application of HRA primarily to as-built systems is an historical artifact. Although analysts have conducted assessments of human reliability as part of system evaluations since the 1960s (Swain, 1963), its formal foundations came in the WASH-1400 (US Nuclear Regulatory Commission, 1975) study, which purpose was to address

the safety of nuclear power plants. This method was further developed with the release of the Technique for Human Error Rate Prediction (THERP) HRA method (Swain and Guttman, 1983) and accompanying documentation (Bell and Swain, 1983), in which a systematic method for identifying, modeling, and quantifying human errors was documented. THERP and subsequent HRA methods emerged against the backdrop of the Three Mile Island incident in the US, with a corresponding call for risk-informed decision making (Kadak and Matsuo, 2007), specifically through PRA and HRA. This application required assessment of existing systems, with less of an emphasis on design than had typically been the case in HFE and CE.

**Human Reliability Process Model**

Contemporary HRA may be seen as encompassing three phases (see Figure 1):

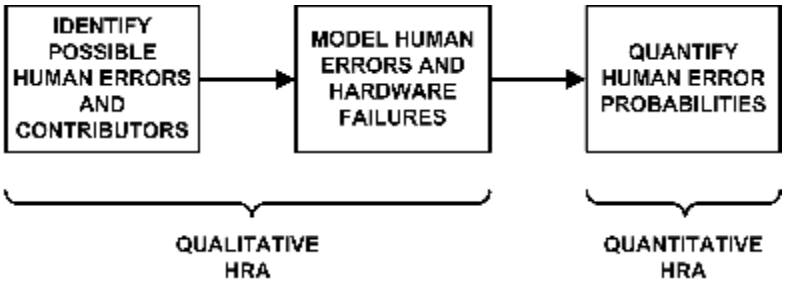


FIGURE 1 The three phases of human reliability analysis.

- *Identify the sources of errors.* Typically, this phase consists of performing a task analysis to determine human actions and then reviewing those actions for opportunities for errors—either in the form of errors committed or actions omitted. *Performance shaping factors* (PSFs) are used to determine which aspects of behavior impact upon the outcome of that action. For example, the presence or absence of clear procedures can greatly enhance or decrement human performance on a given task. The *Good Practices for HRA* document sponsored by the US Nuclear

Regulatory Commission (2005) provides a standardized list of 15 PSFs that are believed to impact human performance in the nuclear domain (see Table 1). Individual HRA methods vary from three PSFs (Galyean, 2005) to 50 PSFs (Chang and Mosleh, 2007) or more, depending on the level of detail required for capturing human activities.

TABLE 1 Performance shaping factors found in the *Good Practices for HRA*.

| | | |
|---|---|---|
| Applicability and suitability of training and experience | Workload, time pressure, and stress | Accessibility on operability of equipment |
| Suitability of relevant procedures and administrative controls | Team and crew dynamics | Need for special tools |
| Availability and clarity of instrumentation | Available staffing and resources | Communications strategy and coordination |
| Time available and time required | Ergonomic quality of human-system interface | Special fitness needs |
| Complexity of required diagnosis and response | Environment | Off-normal operations and situation |

- *Model the errors in an overall risk model such as a PRA*. Human activities of interest to HRA do not generally occur in isolation but rather in interaction with hardware systems. Hardware systems modeled in PRA feature reliability curves for systems and components to address the mean time before failure. A failed hardware system can cause humans to fail at their prescribed task, or a human error can cause a hardware system to fail. Likewise, a hardware system may be designed as a failsafe backup for human actions that fail, e.g., an automatic pressure venting valve can mitigate system damage should the human fail to regulate a pressurized system properly. Perhaps often overlooked, humans are often the key to saving a failed hardware system: positive human intervention can prevent the escalation of a

hardware failure. In HRA, human activities are modeled as part of a fault or event tree (see Figure 2) to show the interaction of human activities with the hardware system functioning.

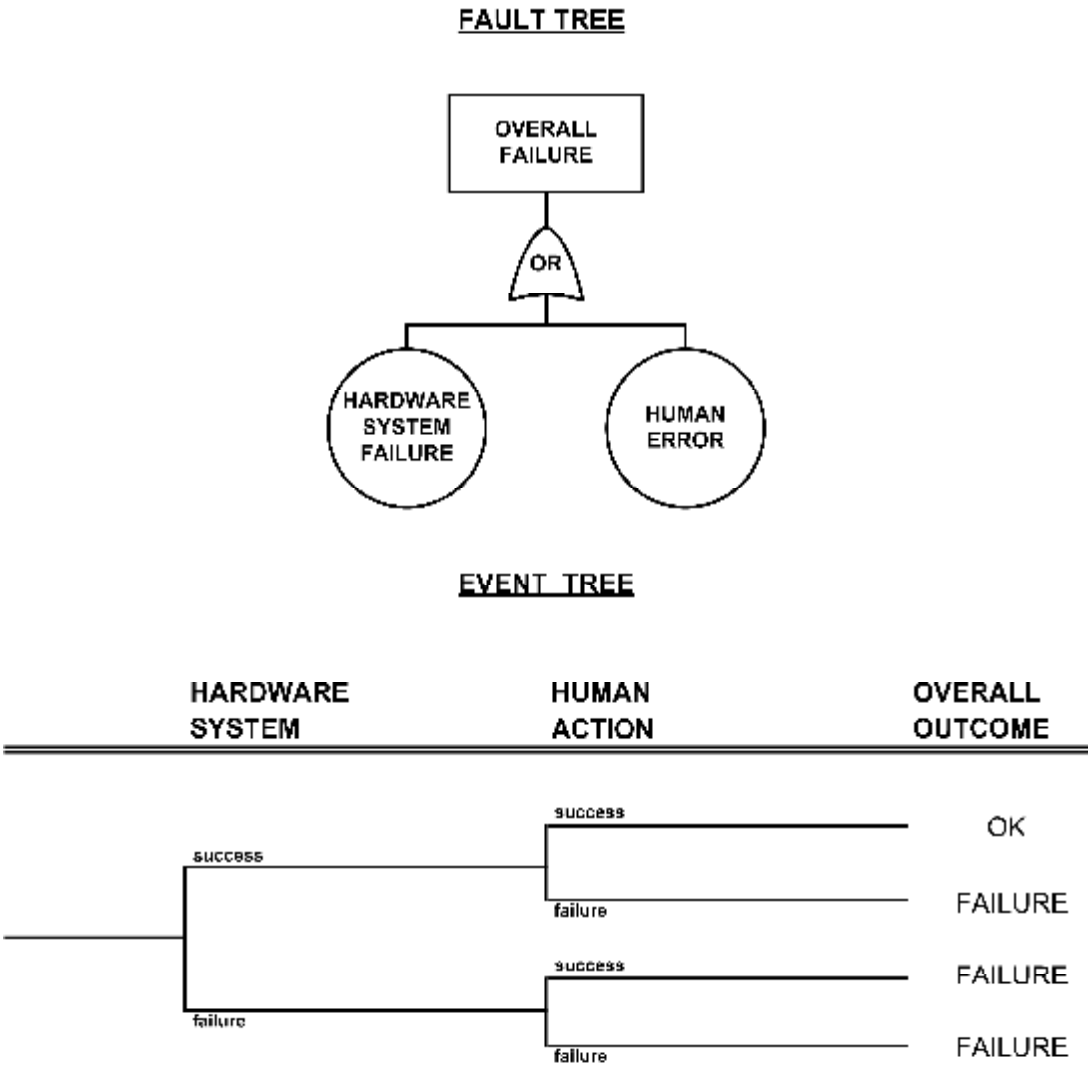


FIGURE 2 A logical "OR" gate connecting hardware system failure and human error in the form of a fault tree (top) and event tree (bottom).

- *Quantify the errors.* Much of HRA has a goal to provide a probabilistic expression of the likelihood of a failed human action, called the *human error probability* (HEP). The various approaches to error quantification are the primary differentiators among

dozens of HRA methods. Quantification approaches tend to follow a common approach of beginning with *a nominal HEP*—a generic or default error rate for types of human activities—and then modifying the nominal HEP according to the method's specific PSFs. Often, these PSFs are treated as multipliers. For example, the positive effect of good procedures might consist of a value less than one; hence, the product of the nominal HEP and the PSF multiplier is less than the nominal HEP, resulting in an overall decrease in the HEP and corresponding increase in human reliability. Conversely, the negative effect of poor procedures might consist of a value greater than one; hence, the product of the nominal HEP and the PSF multiplier is greater than the nominal HEP, resulting in an overall increase in the HEP and corresponding decrease in human reliability (see Figure 3).

**Factor Increasing Human Reliability ($HEP_{overall} < HEP_{nominal}$)**

$HEP_{overall} = HEP_{nominal} \times PSF$, *where* $0 < PSF < 1$

**Factor Decreasing Human Reliability ($HEP_{overall} > HEP_{nominal}$)**

$HEP_{overall} = HEP_{nominal} \times PSF$, *where* $PSF > 1$

FIGURE 3 Increasing and decreasing human error probability through PSF multipliers.

As depicted in Figure 1, HRA is sometimes delineated into qualitative and quantitative HRA. Qualitative HRA encompasses the identification and modeling phases described above. Qualitative HRA converges on approaches such as root cause analysis, where the goal is not to determine the likelihood of error but rather the cause of error.

**Application of Human Reliability Analysis to System Design**

HRA has been applied in retrospective and prospective analyses. *Retrospective HRA* focuses on assessing the risk of something that has already happened, such as an incident or accident. The purpose of such an analysis is to determine the likelihood that something should or could have happened the way it actually did—was it an anomalous activity, or would it be expected that such an activity could occur again given the same situation? *Prospective HRA* attempts to assess the risk of something that hasn't actually happened, such as determining the characteristics of an extremely rare event like human performance in a nuclear power plant control room during a seismic event or fire. Note that while prospective HRA holds tremendous opportunity to anticipate breakdowns in the human-system interface, prospective applications of HRA have not commonly centered on incorporation of such information into the early-stage design of a system. However, as noted in Hirschberg (2004), HRA is actively used for improvement of existing processes and systems. HRA pinpoints weaknesses and allows prioritization of fixes. Thus, HRA is typically used not in the initial system design phase but rather in the assessment and iterative improvement of existing technologies. This after-the-fact use of prospective HRA is artificially limiting. There are tremendous opportunities to apply prospective HRA more broadly—not just on as-built systems but also on systems that are still being designed. This emerging movement toward HRA as a design tool arguable aligns HRA with CE and HFE.

The HRA for design approach is evident in three recent developments:

- *The need for new human certified safety-critical systems.* Recent regulatory design guidance such as the *Human Factors Engineering Program Review Model* (O'Hara et al., 2004) for nuclear power or the *Human-Rating Requirements* (NASA, 2005) for

aerospace suggest using HRA as part of the design process to complement existing human factors design best practices (Boring, 2007a). As new nuclear power and aerospace systems are built, HRA need no longer become a tool for as-built systems. The opportunity exists for HRA to be used before systems are built. Qualitative HRA may be used in a complementary fashion to other HFE and CE techniques to anticipate sources of human errors and, ultimately, to design the system to prevent those errors from occurring. Further, quantitative HRA may be used to help determine the likelihood and consequence of specific errors and to prioritize those error-likely design issues that have the greatest impact to the safety of the users or the integrity of the system.

- *The emergence of resilience engineering.* A recent trend in engineering is an understanding that the negative consequences of an incident may be greatly mitigated by the quality of the underlying human interactions with the system. Called *resilience engineering* (Hollnagel, 2006), this approach attempts to identify what qualities make humans, processes, and systems robust or resilient in the face of adverse events. Resilience engineering shares many conceptual underpinnings with HRA but has been treated as a distinct approach. The key to reconciling resilience engineering with HRA is in considering HRA applied to system design. HRA provides a standardized way to assess vulnerabilities in human actions—those things that make actions less robust. HRA can even be used to define the characteristics of resilience—e.g., those PSFs that mark resilient actions vs. less resilient or brittle actions. In applying HRA to design, the goals of resilience engineering and HRA are

unified, whereby HRA can be used to aid in the design of resilience processes and systems.

- *Development of HRA for human performance modeling*. Cacciabue (1998) and others (e.g., Lüdke, 2004; Boring, 2007b) have outlined the importance of simulation and modeling of human performance for the field of HRA. In human performance modeling, a virtual human (in the form of a cognitive simulation) interacts with virtual systems to reveal areas where human performance is degraded or enhanced in human-system interactions. Such simulations address the dynamic nature of human performance in a way that has not been found in classic static HRA methods. The chief advantage of incorporating HRA into a human performance modeling system is the ability to estimate the safety of novel equipment and configurations. It is anticipated that in many cases, there is a significant cost advantage in utilizing modeling to screen new equipment virtually vs. the cost of configuring a simulator with new equipment and enlisting appropriate personnel (e.g., control room staff) to perform representative tasks. Human performance modeling is already established as a powerful system design tool in HFE (Foyle and Hooey, 2007), one that must utilize insights from CE in order to render a reasonable fidelity in the simulation. When elements of HRA—such as dynamically assigned PSFs—are included in human performance modeling, this opens the door not only to simulating if humans will interact successfully with a system but also for understanding the basis of the performance decrements or enhancements offered by a particular system configuration.

**Conclusion**

This chapter has briefly outlined the three process phases typically associated with HRA: identification, modeling, and quantification. These three process phases represent an historic evolution that can and should now include a forth phase: error prevention (see Figure 4). The insights learned from 25 years of formal HRA are now available in a process more closely aligned with HFE and CE. By incorporating HRA in the design phase of a system, insights on the types and causes of human errors, as well as the likelihood and consequences of those errors, can facilitate the design of safe systems.

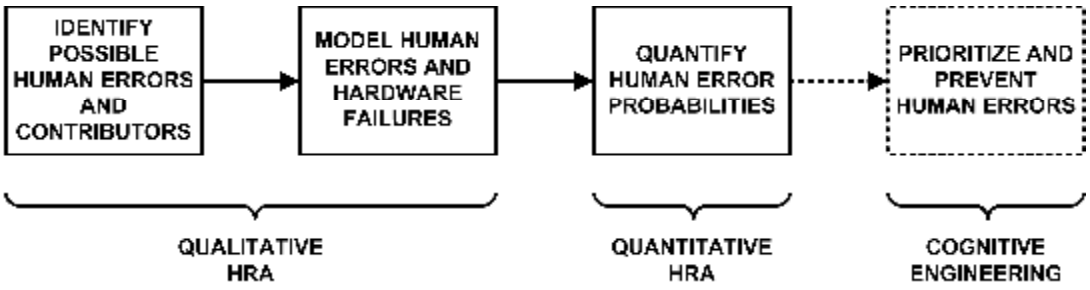


FIGURE 4 The four phases of human reliability analysis integrated with cognitive engineering.

**Disclaimer**

Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy's National Nuclear Security Administration under contract DE-AC04-94AL85000. 

necessarily state or reflect those of Sandia Corporation, the United States Government, or any agency thereof.

**References**


Bell, B.J., and Swain, A.D. 1983. A Procedure for Conducting a Human Reliability Analysis for Nuclear Power Plants, NUEG/CR-2254. Washington, DC: US Nuclear Regulatory Commission.

Boring, R.L. 2002. Human-computer interaction as cognitive science. Pp. 1767-1771 in Proceedings of the 46th Annual Meeting of the Human Factors and Ergonomics Society. Santa Monica: Human Factors and Ergonomics Society.

Boring, R.L. 2007a. Meeting human reliability requirement through human factors design, testing, and modeling. Pp. 3-8 in Risk, Reliability and Societal Safety, Volume 1: Specialisation Topics. Proceedings of the European Safety and Reliability Conference (ESREL 2007), T. Aven and J.E. Vinnem, eds. London: Taylor & Francis.

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

Cacciabue, P.C. 1998. Modelling and simulation of human behaviour for safety analysis and control of complex systems. Safety Science 28: 97-110.

Chang, Y.H.J., & Mosleh, A. 2007. Cognitive modeling and dynamic probabilisitic simulation of operating crew response to complex system accidents. Part 2: IDAC performance influencing factors model. Reliability Engineering and System Safet, 29: 1014-1040.

Foyle, D.C., and Hooey, B.L. 2007. Human performance modeling in aviation. Boca Raton, FL: CRC Press.

Galyean, W.J. 2006. Orthogonal PSF taxonomy for human reliability analysis. Paper PSAM-0281, pp. 1 – 5 in Proceedings of the 8th International Conference on Probabilistic Safety Assessment and Management, May 14-18, 2006, New Orleans, Louisiana, USA.

Kadak, A.C., and Matsuo, T. (2007). The nuclear industry's transition to risk-informed regulation and operation in the United States. Reliability Engineering and System Safet, 92: 609-618.

Hirschberg, S. (2004). CSNI Technical Opinion Papers No. 4, Human Reliability Analysis in Probabilistic Safety Assessment for Nuclear Power Plants. OECD NEA No. 5068.

Hollnagel, E. 2006. Resilience—the challenge of the unstable. PP. 9-17 in Resilience Engineering: Concepts and Precepts, E. Hollnagel, D.D. Woods, and N. Leveson, eds. Burlington, VT: Ashgate Publishing Company.

Lüdke, A. 2005. Kognitive Analyse formaler sicherheitskritischer Steuerungssysteme auf Basis eines integrierten Mensch-Maschine-Models. Berlin: Akademische Verlagsgesellschaft Aka GmbH.

National Aeronautical and Space Administration (NASA). 2005. Human-Rating Requirements for Space Systems, NPR 8705.2A. Washington, DC: NASA Office of Safety and Mission Assurance.

Nielsen, J. 1994. Usability Engineering. San Francisco: Morgan Kaufman.

Norman, D.A. 2002. Emotion and design: Attractive things work better. Interactions Magazine, IX (4): 36-42.

O'Hara, J.M., Higgins, J.C., Persensky, J.J., Lewis, P.M., and Bongarra, J.P. 2004. Human Factors Engineering Program Review Model, NUREG-0711, Rev. 2. Washington, DC: US Nuclear Regulatory Commission.

Palanque, P., Basnyat, S., Blandford, A., Bernhaupt, R., Boring, R., Johnson, C., & Johnson, P. 2007. Beyond usability for safety critical systems: How to be SURE (safe, usable, reliable, and evolvable)? Pp. 2133-2136 in CHI 2007 Conference Proceedings, Extended Abstracts. New York City: Association for Computing Machinery.

Swain, A.D. 1963. A Method for Performing a Human Factors Reliability Analysis, Monograph SCR-686. Albuquerque: Sandia National Laboratories.

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

U.S. Nuclear Regulatory Commission. 1975. Reactor Safety Study: An Assessment of Accident Risks in US Commercial Nuclear Power Plants, WASH-1400, NUREG-75/014. Washington, DC: US Nuclear Regulatory Commission.

US Nuclear Regulatory Commission. 2005. Good Practices for Implementing Human Reliability Analysis (HRA), NUREG-1792. Washington, DC: US Nuclear Regulatory Commission.