Human Reliability Assessment for Physical Security: Human Responses under Extreme Threats

Md Ragib Rownak[a], Samuel Abiodun Olatubosun[a], Yunfei Zhao[a], Carol Smidts[a], and Abdollah Shafieezadeh[b]
[a] Department of Mechanical and Aerospace Engineering, The Ohio State University, Columbus, OH 43210, USA, {rownak.1, olatubosun.1, zhao.2263, smidts.1}@osu.edu
[b] Department of Civil, Environmental and Geodetic Engineering, The Ohio State University, Columbus, OH 43210, USA, shafieezadeh.1@osu.edu

Abstract: There is a paramount need today in the nuclear industry to reduce the operation and maintenance costs. This may be achieved by reducing the physical security force, which accounts for about 20% of the entire workforce in a nuclear power plant, while still meeting regulation requirements on physical security. Credible physical security risk analyses are the basis of risk-informed security force reduction. In the event of a physical attack, security guards and reactor operators take actions to respond to the attack and mitigate the consequences. Therefore, the assessment of their performance and reliability during a physical attack is an essential element of such physical security risk analyses. Although various methods for human reliability analysis have been developed in the nuclear industry, the contexts in which these methods are intended to be used are different from the context in existence during a physical attack on several important aspects, for example, the extremely high stress experienced by security guards and reactor operators during a physical attack. As a consequence, existing human reliability analysis methods need to be modified before they can be applied to the context of physical security. This research serves as a starting point to address this problem. This paper provides reviews of human behavior under extreme threats, existing human reliability analysis methods, and possible human actions during physical attacks. Based on the reviews, four typical types of human responses to extreme threats and the major factors (e.g., physical, physiological, human characteristics) that influence human behavior were identified. A preliminary list of human actions, including both security guard actions and reactor operator actions, were also identified. This paper also provides brief discussions on the modification of existing human reliability analysis methods based on these reviews, for example, by introducing additional performance shaping factors and/or considering additional levels for certain performance shaping factors.

1. INTRODUCTION

Nuclear power accounts for about 20% of the electricity generation in the U.S. and plays a critical role in reducing greenhouse gas emission. However, an accident in a nuclear power plant may cause serious consequences to both human and the environment. Therefore, nuclear power plants in the U.S. are heavily regulated. One of the regulations is related to the requirement on a large physical security workforce to protect nuclear power plants from intentional physical attacks [1]. While improving the security and safety of nuclear power plants, such requirements have significantly increased the financial burden of plant owners and operators. It is estimated that the security workforce accounts for about 20% of the entire workforce in a nuclear power plant [2].

The security workforce cost can be reduced by reducing the number of security guards in a nuclear power plant while still meeting the regulation requirements. To achieve this, one needs to develop the capability of analyzing the effectiveness of different security workforce configurations (e.g., whether security guard posts at certain locations should be kept). The results of such physical security analyses can then be used to identify the optimal security workforce configuration with the minimum number of security guards. In recent years, several modeling and simulation methods [2, 3] have been proposed for such analyses. However, considering the large uncertainty in the modeling and simulation, conservatism is typically built into the final security workforce configuration, which in turn increases the security workforce cost. This conservatism can be lifted by reducing the uncertainties in the modeling and simulation for physical security analyses.

One of the uncertainties in the modeling and simulation is related to the reliability of human operators and security guards in responding to physical attacks. In the event of a physical attack, the security guards will take actions to engage and neutralize the attackers, and the human operators may need to use flexible equipment to mitigate the damages to plant systems caused by the attackers. Successes of the human operators and security guards in taking the required actions can help reduce the consequence of a physical attack, while failures may lead to serious consequences. A thorough understanding of human reliability (i.e., the probability that the human operators and security guards successfully take a required action) during physical attacks is essential to physical security analyses used to determine the optimal security workforce configuration.

Human reliability analysis (HRA), as an integral component of probabilistic risk assessment, has received significant attention over the past 50 years and various methods for the analysis have been developed [4]. However, these methods focus on human reliability in the context of physical system failures, for example, steam generator tube rupture accidents in nuclear power plants. These contexts are different from the context of physical attacks on several important aspects. For example, during physical attacks, human operators and security guards are expected to experience an extremely high level of stress. Because of such differences, existing methods should not be directly applied to human reliability analysis for physical security.

In this research, we attempt to bridge the gap between existing methods for human reliability analysis and the application to human reliability analysis for physical security. Although the reliability of the physical attacker actions is also worthy of investigations, this research is focused on the reliability of human defenders in a nuclear plant. In the following sections, we first provide reviews of human behavior under extreme conditions, which are akin to physical attacks. We also provide reviews of existing methods for human reliability analysis. We then discuss the possible required human actions in response to physical attacks. Based on these human actions, we discuss how to adapt existing methods for human reliability analysis to the application to physical security. The paper is concluded with a summary of this research and outlook on future work.

2. A REVIEW OF HUMAN RESPONSES IN EXTREME CONDITIONS

Extreme threats, situations or conditions can be described as psychosocial and physical disruptions which occur rarely and cause major disruptions to society. Extreme conditions impose a significant cognitive, emotional, and/or physical stress on an individual, and have profound effects on the individual's performance and behavior [5]. The conditions strongly perturb the body and mind, which in turn initiate complex cognitive and affective response strategies.

Several factors influence human responses in extreme conditions. These factors include the environment where human subjects take actions, the human subject's level of fatigue, the allowable choice-reaction/response time, and the human subject's working memory/information processing. The working memory is a short-term memory through which the mind can use the information acquired from the environment to respond appropriately. In addition, there is a complex interrelationship between these factors. Human responses under extreme conditions are highly variable and understanding such variabilities requires careful studies of the interrelationships between the factors [6].

Though the often-expected behavior of an individual in an extreme situation is panic [7], studies have affirmed human response in such a situation to be rational and based on prosocial factors [8, 9].

2.1 Modes of Responses in Extreme Conditions

Reactions to a disaster vary considerably on the bases of those involved and the type of disaster [6]. However, certain general modes of responses or reactions that can be distinguished have been classified into two [6], which are:

- Agitated reactions (reactions with an increase in physical activity level, e.g., panic).
- Depressive reactions (reactions with a relative preference for mental alertness over physical processes).

Four common types of responses have been identified. The first two, flight (exit) and fight (growth), are very common under extreme situations, while the other two, freeze (hypervigilance) and fright (dormancy), are not as prevalent. These four evolutionarily established human survival responses during extreme threats and the two key underlying considerations (energy reserves and nervous system activation) are discussed extensively in [10]. Strong physical and mental coordination is required for all four responses, though with associated preferences for either physical activity or mental alertness.

Following several studies, the traditional fight-flight-freeze response [11, 12] was revised to consider freeze-flight-fight-fright response [13]. Though there are variations in practice, this sequence of human survival strategies describes the most common approach by humans when responding to extreme conditions. The sequence is based on the assessment of the situation/threat by the victim which changes with time.

In addition, socially integrative responses and affiliation have been established as common human responses in extreme conditions. These responses are due to victims' tendency to move close to familiar people and follow suit other people's steps based on trust (such as the direction they flee to or where they hide for refuge) in extreme situations.

Table 1 shows the categories, sub-types, description, and related phenomena for the responses being displayed during extreme conditions. The conditions are defined in terms of the individual's/victim's perceptions/beliefs at that instance [14].

Table 1: Classification and Examples of Human Responses in Extreme Conditions

S/N	Category of human response	Typical human survival response	Description	Related phenomena	Reference
1	Agitated reactions	Flight (Panic flight)	Flight is a relatively safe strategy and not the same as fear [15].	Trait/subjective anxiety and anxiety sensitivity.	[6, 14]
		Fight	Fight is the most aggressive response [17].		
2	Depressive reactions	Freeze (Tonic immobility)	Freeze is a state of hyper-vigilance and heightened sensitivity to environmental signals.	Trait/subjective anxiety, tonic immobility is related to significant fear.	[6, 16]
		Fright	Fright is a deliberate attempt to play dead [15].		[10, 18]
3	Pro-social (socially integrative) responses		There is usually an increase of mutual support functions among victims and others in stricken communities [19].		[14]
4	Affiliation		This seeks the proximity of familiar persons and places.		[8]

2.2 Specific Behaviors Reported in Simulated and Real Extreme Conditions

A laboratory-based exploration of the tendency to freeze during threat was reported in [15]. The study focused on determining if laboratory-based threat stressors can trigger freeze responses. Although the challenge did not involve physical attack, participants were physically confined by a breathing apparatus attached around the head, which may simulate some of the parameters relevant to freeze (considering predatory grasp). Freeze responses were predicted by hierarchical linear regression analyses. The study revealed that there is prevalence of a freeze response and relationships between freeze and panic symptoms were established (subjective anxiety was strongly associated with freeze ratings) [15]. This is in tandem with an already established fact that flight is influenced by subjective anxiety and fear.

In addition, human response to earthquake was reported in [19] with the objective of determining how people behave under earthquake threat and the factors responsible for people's behavior towards earthquake threat. Human response under a natural threat was divided into four major categories (denying behavior, acceptance of loss behavior, practical behavior, and extreme behavior). There are many theories used to explain these behaviors, although none of them can explain the behaviors perfectly. Such theories include cognitive theory [19], protection-motivation theory, the need theory, cultural/belief system and economic theory. In the study, social and cultural differences were observed to influence the decision-making process while economic differences did not lead to significant differences. Moreover, those who had recently experienced an earthquake were less likely to take any mitigating actions. Therefore, factors such as fatalism and hazard perception contribute significantly to responses to earthquakes [19]. Different models have also been used to explain and justify these responses.

2.3 Factors Influencing Human Behavior in Extreme Conditions and Their Relationships

The factors which influence human behavior in extreme conditions have been grouped into four categories of physical, physiological, psychological, and human characteristic factors. Specific factor(s) and their examples under the different categories are presented in Table 2. In addition, the relationships which exist among the factors as established in the literature are given (Table 2).

Table 2: Factors Influencing Human Behavior in Extreme Conditions and their relationships

Factors			**Relationship**
Physical factors	Temperature	Heat (too hot)	Human response and performance are affected by increase in the heart rate due to great changes in the temperature of the environment [20]. In addition, extreme fatigue and stress initiated by the physical environment have negative effects on effective functioning of cognition [21]. Furthermore, it has been established that the response time and vigilance firstly increase linearly with temperature after which they decline with temperature [22].
		Cold (too cold)	
	Radiation	Exposure to radiation affects regular activities as there is a need for better shielding when radiation levels are high, e.g., astronauts in space, nuclear accident conditions [23].	
	Smoke yield	Smoke is associated with fire disasters, earthquakes etc. It affects visibility and indirectly influences the likelihood of escape. Smoke yield is characterized by a mixture of particles and chemicals produced by incomplete combustion of carbon-containing materials [24].	
	Toxicity	Toxicity due to hazardous chemicals spillage to the environment [24].	

	Perceptual features	Perceptual features consist of visual, smell-related, audible and tangible features [24].	
Physiological	Fatigue	Physical	Extreme fatigue influences human behavior and performance negatively [25]. It also negatively impacts several functions such as vigilance and visual attention [26].
		Mental	
		Disrupted sleep-wake cycles	
Psychological factors	Psychological stress	Fear	Human responses are affected by stress and fatigue impaired memory accuracy, response time, attention, and cognitive executive functions [26, 27].
		Time pressure stress	
		Performance stress	
	Social stress	Crowding, social and cultural isolation	
Human characteristics		Individual features (thoughts, knowledge, and experience)	Individual response is influenced by the perceptions, intentions and motives of the victim trying to escape from the extreme situation [29].
		Social features (interactions between individuals)	Affiliative behavior influences human response, e.g., route choices during fire or earthquake disaster [29].
		Situational features (awareness, familiarity with the environment)	

3. REVIEW OF EXISTING HUMAN RELIABILITY ANALYSIS METHODS

In this section, several human reliability analysis methods that have been widely used in the nuclear industry are reviewed. These methods will serve as the base methods and will be adjusted as needed so that they are applicable to human reliability analysis in physical attacks. The adjustments will include, for example, consideration of additional performance shaping factors or modifications of the multipliers for certain performance shaping factors. It is worth noting that the review in this section is not exhaustive, and in the future further details of the reviewed human reliability analysis methods and additional human reliability analysis methods may need to be considered.

3.1 THERP

THERP is short for Technique for Human Error Rate Prediction. It is an HRA method originally developed at Sandia National Laboratories, and fully described in the "Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications - Final Report" [30]. THERP is probably the most widely used HRA method, in the nuclear industry as well as in other industries, for example, aviation.

THERP assesses human reliability based on task analysis, which is described using HRA event trees. Each branching in an HRA event tree denotes a single human operator event, for example, following a step in a procedure by an operator, that may occur in completing a task, for example, a steam generator feed and bleed operation. Each of the two limbs at a branching denotes operator success or failure, for example, successful or unsuccessful act in following the step. At the end of each path in the HRA event tree, it can be determined whether the task is successful or not.

For the limb at a branching in an HRA event tree that denotes operator failure, a human error probability is assigned. Then based on the logic embedded in the event tree, the human error probability for the task can be calculated easily. THERP provides the nominal human error probabilities for a variety of human errors in Chapter 20 in its handbook [30].

THERP considers the effects of certain performance shaping factors as well, for example, stress and experience, and uses modifiers to adjust the nominal human error probabilities. THERP also considers the effects of task dependence on human reliability. Specifically, THERP defines five levels of dependence, i.e., zero dependence, low dependence, moderate dependence, high dependence, and complete dependence. For each level of dependence, THERP provides an equation to adjust the nominal human error probability.

3.2 ASEP

As discussed in the evaluation of HRA methods against good practices report [31], HRA using THERP is resource intensive. To address this problem, a simplified version was proposed in [32]. This method is called Accident Sequence Evaluation Program, or ASEP.

Unlike THERP, which emphasizes task analysis, ASEP focuses on the quantification of human error probability for a given task and provides detailed guidance on doing so. In addition to post-initiator human failure events, ASEP also provides guidance on estimating probabilities for prior-initiator human errors. Similar to the SPAR-H method that will be introduced shortly, ASEP considers both diagnosis failure and action failure in completing a task.

When performing HRA using ASEP, one first determines the total allowable time to complete a task and the time needed to complete the action, and then determines the remaining time for diagnosis by subtracting the action time from the total task time.

Based on the time for diagnosis, one first determines the nominal human error probability, and then increases or decreases the probability considering relevant performance shaping factors, for example, whether the operator has experience with the abnormal event in question. Two performance shaping factors are considered in ASEP, which are training and stress. A similar procedure follows for estimating the probabilities for action failures. Similar to THERP, an HRA event tree can be developed based on the analyses for diagnosis failures and action failures to obtain the final human error probability.

3.3 SPAR-H

SPAR-H stands for Standardized Plant Analysis Risk - Human Reliability Analysis. It is described in detail in NUREG/CR-6883 [33].

SPAR-H segregates human failure events into diagnosis failures and actions failures and quantifies these two types of human failure events separately. For diagnosis failures, it assumes a nominal human error probability of 0.01, and for actions failures, it assumes a nominal human error probability of 0.001. SPAR-H considers the effects of specific situations using eight performance shaping factors, including "available time," "stress/stressors," "complexity," "experience/training," "procedures," "ergonomics/HMI," "fitness for duty," and "work processes." Each performance shaping factor can take one of several possible levels. For example, the performance shaping factor "stress/stressors" may be "*extreme*," "*high*," or "*nominal*." For each relevant performance shaping factor, a specific level needs to be assigned, and the corresponding multiplier is then used to adjust the nominal human error probability.

SPAR-H also considers the effect of dependency between tasks. It also considers five dependence levels, similar to the treatment in THERP. For each task, it provides a structured instruction on determining the dependence level.

Human reliability analysis using SPAR-H can be performed following the steps below [33]:

1. Determine whether the task involves diagnosis, action, or both diagnosis and action.
2. For diagnosis/action, rate the eight performance shaping factors using the corresponding worksheet provided in SPAR-H.
3. Use the performance shaping factor multipliers corresponding to the ratings to adjust the nominal human error probability for diagnosis/action.
4. Calculate the overall human error probability (if diagnosis and action both exist, add the probabilities for them to obtain the overall probability).

5. If dependency exists, determine the level of dependence and adjust the obtained overall human error probability following predefined equations. For example, for high dependence the human error probability is adjusted by $p_{w/d} = \frac{1+p_{w/od}}{2}$, where $p_{w/od}$ is the human error probability not considering task dependence and $p_{w/d}$ is the human error probability considering task dependence.

3.4 HCR/ORE

The Human Cognitive Reliability/Operator Reliability Experiments, i.e., HCR/ORE, method was developed by the Electric Power Research Institute in 1992 [34]. It focuses on the non-response of an operator after the occurrence of a disturbance and provides a model for the non-response probability as a function of the length of time following the disturbance. The HCR model was first developed, and then data collected in the simulator experiments conducted by the Electric Power Research Institute, i.e., ORE data, were used to determine the parameters in the HCR model. Plant-specific data and expert judgments can also be used to determine the model parameters.

In the original HCR model, a Weibull distribution is used to quantify the non-response probability as a function of time [35]. However, the ORE data suggests a lognormal distribution [34, 35]. The non-response probability p can be calculated as follows,

$$p = \Pr(T_r > T) = 1 - \Phi\left(\frac{\ln(T/T_{1/2})}{\sigma}\right) \tag{1}$$

In (1), T_r is the time of response following a disturbance. $\Phi(\cdot)$ is the standard normal cumulative distribution. $T_{1/2}$ and σ are two model parameters. $T_{1/2}$ is the median response time, and σ is the logarithmic standard deviation of the normalized response time, i.e., $T/T_{1/2}$. These two model parameters can be estimated using the ORE data or other data mentioned above.

The influence of human performance factors can be considered by adjusting model parameter $T_{1/2}$. The uncertainty in p can be considered through model parameter σ.

The HCR model allows one to calculate the human error probability for various types of human operator behaviors, for example, skill-based, rule-based, or knowledge-based behaviors, by determining parameters (i.e., $T_{1/2}$ and σ) in the model using data corresponding to each type of operator cognitive behavior.

3.5 ATHEANA

ATHEANA is short for A Technique for Human Event ANAlysis, an HRA method developed by the U.S. Nuclear Regulatory Commission. It is described in detail in NUREG-1642 [36]. The underlying premise of ATHEANA is that human failure events occur mainly because of the context in which human operators perform tasks. Such contexts trigger error mechanisms that eventually lead to human failures or unsafe acts. Such contexts are called Error-Forcing-Contexts (EFCs) in ATHEANA, and can be described through plant conditions, e.g., initiating events, and performance shaping factors, e.g., operator training.

Because of the above premise, ATHEANA provides a structured search process for error-forcing-contexts that may lead to human failures and unsafe acts. This search process is a major improvement over other HRA methods. The search process includes the following generic steps:
1. Define and interpret the issue
2. Define the scope of the analysis
3. Describe the base case scenario
4. Define human failure events and/or unsafe acts
5. Identify potential vulnerabilities in the operators' knowledge base

6. Search for deviations from the base case scenario
7. Identify and evaluate complicating factors and links to performance shaping factors
8. Evaluate the potential for recovery

Thanks to the structured search process for EFCs, ATHEANA is most suitable for retrospective analysis of operational events, but can also be applied to prospective analysis for probabilistic risk assessment applications.

The quantification of the probabilities of identified unsafe actions can be realized through three basic elements: 1) the probability of the error forcing context; 2) the probability of unsafe action given the error forcing context; and 3) the probability of not recovering from the initial unsafe action. The quantification of human error probabilities in ATHEANA largely relies on expert judgments.

3.6 IDHEAS

The IDHEAS HRA method is short for Integrated Human Event Analysis System. It was recently developed by the U.S. Nuclear Regulatory Commission [37]. The development of this method was aimed at providing a more structured process for human reliability analysis to reduce analyst-to-analyst variability. It also aims to provide a more solid basis on human cognitive research.

Similar to other HRA methods, the first two steps in IDHEAS are to understand the scenario under study and to define the human failure events (HFEs). Following these two steps, IDHEAS consists of three important steps. The first of the three steps is to develop the crew response diagram (CRD). A CRD describes a sequence of human activities that are essential to the success of the human response to a disturbance. Failure of each human activity will lead to a HFE that has been defined in the earlier step. Each human activity may be further defined as one or more critical tasks. The second of the three steps is to determine one or more crew failure modes (CFMs) for each critical task defined earlier. An example of CFMs is failure to attend to a cue. More details can be found in the report describing IDHEAS [37]. The third of the three steps is, for each CFM, to develop a decision tree (DT) and use the DT to calculate the human error probability (HEP) for the specific situation under study. The results in the analyses for all CFMs can be integrated to obtain the HEP for the HFE and the result can then be integrated into the overall probabilistic risk assessment. To improve transparency and traceability, IDHEAS also provides guidance on documentation.

4. HUMAN OPERATOR AND SECURITY GUARD ACTIONS IN PHYSICAL ATTACKS

The actions taken by human operators and security guards in the event of a physical attack on a nuclear power plant are generally different from the actions observed in the event of a physical system failure (e.g., a steam generator tube rupture accident). To analyze human reliability in the context of physical attacks, it is necessary to first identify the possible human actions during physical attacks. In this section, a preliminary review of such possible human actions is performed.

Human defender actions in the event of a physical attack can be divided into two groups, the first group for actions taken by security guards, and the second group for actions taken by reactor operators.

When a physical attack occurs, the security guards' role is to detect and respond to physical attacks. According to [3, 38], the actions that may be taken by security guards include the acknowledgement and assessment of alarms indicating a physical attack, communications to notify the responsible security force, initiation of the response to the attack, reaching critical points to interrupt the attackers, and engaging and neutralizing the attackers. Depending on the applications, additional and more detailed actions may need to be considered. Such human actions are generally not considered in conventional human reliability analysis and therefore should be paid particular attention in the analysis for physical security.

Security guards may fail to or only partially successfully respond to a physical attack. This means that one or more nuclear systems or components may be sabotaged by the attackers. In this case, reactor operators can take mitigative actions to minimize the consequence of the attack. It is expected that besides following emergency operating procedures or abnormal operating procedures used for responding to physical system failures, reactor operators may take actions specific to physical attacks. In [2], the use of Diverse and Flexible Mitigation Capability (FLEX) equipment in responding to a physical attack is considered and its contribution to physical security improvement is assessed. The FLEX equipment is implemented by the operator following a predefined procedure. The example procedure in [2] used for responding to attacks on the offsite power, the emergency diesel generators, and the turbine driven pumps includes operator actions such as get keys and open doors, assess condition of plant system & equipment, connect FLEX steam generator makeup pumps' hose, establish configuration to support FLEX 480V ac installation. Other operator actions may be found in the procedures for the use of other FLEX equipment and other mechanisms to defend against physical attacks.

5. THOUGHTS ON THE ADAPTATION OF EXISTING HUMAN RELIABILITY ANALYSIS METHODS TO PHYSICAL SECURITY

To be applicable to physical security, existing methods for human reliability analysis may be modified in the following ways.

The first way is to expand human behaviors considered in existing human reliability analysis methods to accommodate for the context of physical attacks. For example, additional cognitive activities beyond diagnosis and action may be considered. One example is that the human subject's mind during a physical attack may be in a freeze state, which means that the human subject will not do anything.

The second way is to consider additional performance shaping factors or to redefine existing performance shaping factors to describe the operator action-taking context in the event of a physical attack. Such performance shaping factors may be derived based on the factors identified from the review of human behavior under extreme threats, for example, the one described in Table 2. These new factors may describe physical factors specific to physical attacks, for example, the blast effect.

The third way in adapting existing human reliability analysis methods is to extend the levels for existing performance shaping factors. For example, for stress in the SPAR-H method introduced in Section 3.3, besides the existing levels (i.e., *nominal*, *high*, *extreme*), a higher level than *extreme* may need to be considered. For each new level, the multiplier used in calculating the human error probability needs to be determined, based on either review of relevant studies in the literature or data collected from experiments involving human subjects under the situation of interest (e.g., a specific level of stress).

6. CONCLUSION

Physical security force reduction can help effectively decrease the operation and maintenance cost in a nuclear power plant, which in turn helps improve the economic competitiveness of nuclear energy. Such reductions should be done based on the thorough analysis of the physical security risk for a given security force. In this paper, we focus on a specific aspect of physical security risk analysis, that is the assessment of security guard and reactor operator performance and reliability in responding to a physical attack. Considering the differences between physical system failures, which are the context for which existing human reliability analysis methods are intended to be used, and physical attacks, there is a need to adapt existing human reliability analysis methods so that they are applicable to physical security.

As a starting point to fill the gap between existing human reliability analysis methods and the physical security context, in this research, human behavior under extreme threats, that share certain similarities with physical attacks, are first reviewed. Four types of human responses to extreme threats and the major factors that influence human behavior are identified. Then reviews of existing human reliability analysis methods and possible human actions in the event of physical attacks are performed. Based on the reviews, suggestions on how to modify existing human reliability analysis methods are also given.

In future research, efforts on drawing insights from human behavior under extreme threats to develop human reliability analysis methods applicable to physical security will be continued. Experiments involving human subjects will be designed and implemented as needed to estimate the parameters in such models.

Acknowledgements

This research is being performed using funding received from the DOE Office of Nuclear Energy's Nuclear Energy University Programs.

References

1. US NRC. Title 10, Code of Federal Regulations, Part 73. “Physical Protection of Plants and Materials”. https://www.nrc.gov/reading-rm/doccollections/cfr/part073/
2. Yadav, V., Prescott, S.R., Christian, R., St Germain, S.W. and Weathersby, J.H., 2020. Integration of FLEX Equipment and Operator Actions in Plant Force-On-Force Models with Dynamic Risk Assessment (No. INL/EXT-20-59510-Rev000). Idaho National Lab.(INL), Idaho Falls, ID (United States).
3. Cohn, B., Aldemir, T., Haskin, T., Noel, T., Cardoni, J., and Osborn, D., 2021. Integrated Safety and Security Analysis of Nuclear Power Plants Using Dynamic Event Trees. 2021 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2021), November 7–12, 2021.
4. Bell, J. and Holroyd, J., 2009. Review of human reliability assessment methods (Tech. Rep. No. RR679). Derbyshire: Health and Safety Executive.
5. Paulus, M. P., Potterat, E. G., Taylor, M. K., Van Orden, K. F., Bauman, J., Momen, N., ... Swain, J. L. (2009). A neuroscience approach to optimizing brain resources for human performance in extreme environments. Neuroscience & Biobehavioral Reviews, 33(7), 1080-1088.
6. Fritz, C. E., Marks, E. S. (1954). The NORC studies of human behavior in disaster. Journal of Social Issues.
7. Quarantelli, E. L. (1954). The nature and conditions of panic. American Journal of Sociology, 60(3), 267-275.
8. Mawson, A. R. (2005). Understanding mass panic and other collective responses to threat and disaster. Psychiatry: Interpersonal and biological processes, 68(2), 95-113.
9. Perry, R. W., Green, M. R. (1982). The role of ethnicity in the emergency decision-making process. Sociological inquiry, 52(4), 306-334.
10. Mithani, M. A., Kocoglu, I. (2020). Human and organizational responses to extreme threats: a comparative developmental approach. Management Decision.
11. Cannon, W.B. (1932), The Wisdom of the Body, W.W. Norton & Company, New York.
12. Roelofs, K. (2017). Freeze for action: neurobiological mechanisms in animal and human freezing. Philosophical Transactions of the Royal Society B: Biological Sciences, 372(1718), 20160206.
13. Gray, J.A., McNaughton, N. (2003), The Neuropsychology of Anxiety: An Enquiry into the Function of the Septo-Hippocampal System, Oxford University Press, New York.
14. Perry, R. W., Lindell, M. K. (2003). Understanding citizen response to disasters with implications for terrorism. Journal of Contingencies and Crisis Management, 11(2), 49-60.
15. Lang, P.J., Davis, M. (2006), Emotion, motivation, and the brain: reflex foundations in animal and human research, Progress in Brain Research, Vol. 156, pp. 3-29.
16. Schmidt, N. B., Richey, J. A., Zvolensky, M. J., Maner, J. K. (2008). Exploring human freeze responses to a threat stressor. Journal of behavior therapy and experimental psychiatry, 39(3), 292-304.

17. Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., Updegraff, J. A. (2000). Biobehavioral responses to stress in females: tend-and-befriend, not fight-or-flight. Psychological review, 107(3), 411.
18. Gray, J.A., McNaughton, N. (2003), The Neuropsychology of Anxiety: An Enquiry into the Function of the Septo-Hippocampal System, Oxford University Press, New York.
19. Asgary, A., Willis, K. G. (1997). Household behaviour in response to earthquake risk: an assessment of alternative theories. Disasters, 21(4), 354-365.
20. Rodahl, K. (2002). Occupational health conditions in extreme environments. Annals of Occupational Hygiene, 47(3).
21. Enander, A. E. (1989). Effects of thermal stress on human performance. Scandinavian journal of work, environment & health, 27-33.
22. Grether, W. F. (1973). Human performance at elevated environmental temperatures. Aerospace Medicine, 44(7), 747-755.
23. Williams, S., Fiedler, E. R., Harrison, A. A. (2008). Human Performance in Extreme Environments.
24. Kobes, M., Helsloot, I., De Vries, B., Post, J. G. (2010). Building safety and human behaviour in fire: A literature review. Fire Safety Journal, 45(1), 1-11.
25. Poussel, M., Laroppe, J., Hurdiel, R., Girard, J., Poletti, L., Thil, C., ... Chenuel, B. (2015). Sleep management strategy and performance in an extreme mountain ultra-marathon. Research in Sports Medicine, 23(3), 330-336.
26. Bourne, L., Yaroush, R. (2003). Stress and cognition: A cognitive psychological perspective. Final Report. National Aeronautics and Space Administration. https://humanfactors. arc. nasa. gov/flightcognition/download/misc/Stress% 20and% 20Cognition. pdf.
27. Karatsoreos, I. N., McEwen, B. S. (2010). Stress and allostasis. In Handbook of behavioral medicine (pp. 649-658). Springer, New York, NY.
28. Taylor, M. K., Gould, D. R., Hardy, L., Woodman, T. (2006). Factors influencing physical risk taking in rock climbing. Journal of Human Performance in Extreme Environments, 9(1), 2.
29. Kobes, M., Helsloot, I., De Vries, B., Post, J. G. (2010). Building safety and human behaviour in fire: A literature review. Fire Safety Journal, 45(1), 1-11.
30. Swain, A.D. and Guttmann, H.E., 1983. Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report (No. NUREG/CR-1278; SAND-80-0200). Sandia National Labs., Albuquerque, NM (USA).
31. Forester, J., Kolaczkowski, A., Lois, E., and Kelly, D. (2006). Evaluation of human reliability analysis methods against good practices. Technical Report NUREG-1842, U.S. Nuclear Regulatory Commission, Washington, DC.
32. Swain, A. D. (1987). Accident sequence evaluation program human reliability analysis procedure. Technical Report NUREG/CR-4772, U.S. Nuclear Regulatory Commission, Washington, DC.
33. Gertman, D., Blackman, H., and Marble, Julie Smith, C. (2005). The spar-h human reliability analysis method. Technical report, U.S. Nuclear Regulatory Commission.
34. Parry, G., Lydell, B., Spurgin, A., Moieni, P., and Beare, A. (1992). An approach to the analysis of operator actions in probabilistic risk assessment. Technical Report TR-100259, Electric Power Research Institute.
35. Moieni, P., Spurgin, A., and Singh, A. (1994). Advances in human reliability analysis methodology. part i: Frameworks, models and data. Reliability Engineering & System Safety, 44(1):27–55.
36. U.S. Nuclear Regulatory Commission (2000). Technical basis and implementation guidelines for a technique for human event analysis (atheana). Technical Report NUREG-1624, Revision 1, U.S. Nuclear Regulatory Commission, Washington, DC.
37. Xing, J., Parry, G., Presley, M., Forester, J., Hendrickson, S., and Dang, V. (2017). An integrated human event analysis system (idheas) for nuclear power plant internal events at-power application. Technical Report NUREG-2199 (Vol. 1), U.S. Nuclear Regulatory Commission, Washington, DC.
38. Zamanali, J. and Chwasz, C., 2013. Nuclear Power Plant Security Assessment Guide. United States Nuclear Regulatory Commission, Office of Nuclear Security and Incident Response.