ARTICLE IN PRESS

Safety Science xxx (xxxx) xxx–xxx

Contents lists available at ScienceDirect

Safety Science

journal homepage: www.elsevier.com/locate/safety

Incorporating individual differences in human reliability analysis: An extension to the virtual experimental technique

Mashrura Musharraf[a], Jennifer Smith[a], Faisal Khan[a,*], Brian Veitch[a], Scott MacKinnon[b]

[a] Centre for Risk, Integrity and Safety Engineering (C-RISE), Faculty of Engineering & Applied Science, Memorial University of Newfoundland, St John's, Newfoundland and Labrador A1B 3X5, Canada

[b] Department of Mechanics and Maritime Sciences, Chalmers University, Gothenburg, Sweden

ARTICLE INFO

Keywords:
Human reliability
Human factors
Adaptive training
Virtual environment
Safety training

ABSTRACT

The data required to perform Human Reliability Analysis (HRA) for emergency conditions are not readily available and are difficult to retrieve from accident investigations. In the absence of emergency conditions data, the conventional approach of gathering data for HRA is using expert judgment. Expert judgment often suffers from uncertainty, subjectivity, and incompleteness, which makes the reliability of this data collection technique questionable. A more recent approach is to collect data by conducting experiment in virtual environments with human subjects. Though virtual experimental technique addresses the issues of uncertainty, subjectivity, and incompleteness, it still does not consider individual differences while assigning the influence of different factors on human performance. This paper proposes to advance the virtual experimental technique by enabling the consideration of individual differences. An experiment using virtual environment was done to observe performances of 36 individuals during offshore emergency evacuation. By integrating the data collected from the virtual environment into an HRA model, the reliability of each individual was assessed. Sensitivity analysis was then performed to identify the most influential factors that contributed to failure in emergency conditions. This analysis can help identify specific weaknesses that a participant might have. For example, if a participant is found to be more sensitive to a particular factor, then training scenarios with different variations of the factor can be provided to the participant until an accepted level of competency is reached. Identification of a weakness can be combined with adaptive human factor training so that each individual can obtain competence more quickly.

1. Introduction

Human reliability is defined as the probability that a person correctly performs system-required activities in a designated time period (Swain and Guttmann, 1983). There are many human reliability quantification techniques available today to assess how reliable humans are in different contexts. Examples include: Success Likelihood Index Methodology (SLIM), Technique for Human Error Rate Prediction (THERP), and A Technique for Human Error Analysis (ATHENA) (Kirwan, 1994; Cooper et al., 1996). The Bayesian network (BN) approach has also been applied to human reliability analysis (HRA) (Baraldi et al., 2009). Most of the human reliability quantification techniques involve the calculation of human error probability (HEP), which is the probability that a person will fail to carry out a task as required (Kirwan, 1994). Performance shaping factors (PSFs) are often used to calculate HEP (Blackman et al., 2008). Human performance, and hence error, is influenced by PSFs, and therefore the relationship between PSFs and human errors must be defined to calculate HEP. Due to lack of real or ecologically-valid data, the majority of the human error prediction techniques (i.e. SLIM, THERP, BN) often use expert judgment to define this relationship. Though expert judgement is a valuable technique, it can suffer from uncertainty, subjectivity, and incompleteness. Significant conflict among judgements may also arise when collected from multiple experts. Recent works (Musharraf et al., 2014) have proposed the use of virtual experimental technique as an alternative to expert judgement. This technique collects empirical evidences required to perform a human reliablity assessment by conducting experiments in virtual environments with human subjects. However, this work does not account for individual differences when it comes to the influence or importance of PSFs on human errors. Humans are inherently different and therefore the role that different PSFs play on performance may vary from individual to individual. For example, consider a case where complexity and visibility are two different PSFs that can influence one's performance during an evacuation. While

* Corresponding author.
E-mail address: fikhan@mun.ca (F. Khan).

http://dx.doi.org/10.1016/j.ssci.2017.07.010
Received 2 November 2016; Received in revised form 6 July 2017; Accepted 23 July 2017

complexity can play a more important role than visibility for one individual, it can be the other way around for another individual. This paper proposes an expansion of the virtual experimental technique to account for individual differences during the HRA process. In this paper, the term individual difference refers to the difference between the sensitivity of two individuals to external PSFs. It does not cover the more general aspects that might differ between individuals such as gender, education, and physical characteristics.

The HRA technique used in this paper is the BN approach. BNs have proven to be a powerful tool for HRA for the following reasons: (1) this approach can consider the dependencies among PSFs and the associated actions, (2) it can incorporate new evidence and update the HEP, and (3) it can support the root-cause analysis of human error (Podofillini and Dang, 2013; Sundaramurthi and Smidts, 2013). BNs have been widely used to model the impact of different PSFs on human performance or human error (Baraldi et al., 2009; Dang and Stempfel, 2012). Kim and Seong (2006), Cai et al. (2013) and Martins and Maturana (2013) show examples of using the evidential reasoning aspect of BN to find the underlying causes of human error. Also, the BN model allows the incorporation of multiple sources of data into a single predictive HRA model (Groth and Mosleh, 2012b). A more comprehensive list of the demonstrated benefits of BN for HRA in different domains can be found in Groth and Swiler (2013) and Mkrtchyan et al. (2015).

In this paper, a BN model is developed to observe the impact of two PSFs (complexity and visibility) on human error during an offshore emergency evacuation. In this model, PSFs and errors are all random variables, and the probability of an error occurring is conditionally dependent on the PSFs. To define conditional dependencies in the BN, necessary data were collected from a study conducted in a virtual environment with 36 participants. At the beginning of the study each participant was assigned to one of two training groups: (1) G1: high level training and (2) G2: low level training. The training level assigned to each participant remained unchanged for the rest of the experiment. Virtual emergency scenarios were created with different levels of visibility (clearly visible vs. blackout conditions) and complexity (low complexity, such as a muster drill vs. high complexity, such as a dynamic emergency situation). Participants' performance in the series of virtual emergency scenarios were observed. By integrating the performance data into the BN, the reliability of each subject was assessed. Next, sensitivity analysis was performed to find the relative contribution of the PSFs to failure.

Section 2 gives an overview of the BN approach to HRA and the virtual environment used in the experiment. Section 3 describes the methodology, data collection and integration using a case study of offshore emergency evacuation. Section 4 presents and explains the results. The limitation of the study and future works are discussed in Section 5. Section 6 summarizes and concludes the paper.

2. Background

2.1. Bayesian network (BN) approach to HRA

A BN approach was used to calculate the HEP. According to Pearl (1988), BNs are acyclic directed graphical models that represent conditional dependencies among a set of random variables. While performing a task or exercise, errors can occur at different steps of the process. Each error is regarded as the outcome of the joint influence of different PSFs (as depicted in Fig. 1). In the BN approach to HRA, error is the critical node which depends on several PSFs that can influence the occurrence of the error. For example, in an offshore emergency evacuation situation, interacting with hazards (e.g. smoke or fire) is an error that may occur because the visibility is compromised (PSF_1), or the operator is not familiar with the complexity of the situation (PSF_2), or both. Fig. 1 shows the relationship between human error and PSFs. This paper investigates the impact of only two PSFs (visibility and complexity) on human error. A comprehensive list of PSFs can be found in Groth and Mosleh (2012a) and Mearns et al. (2001).

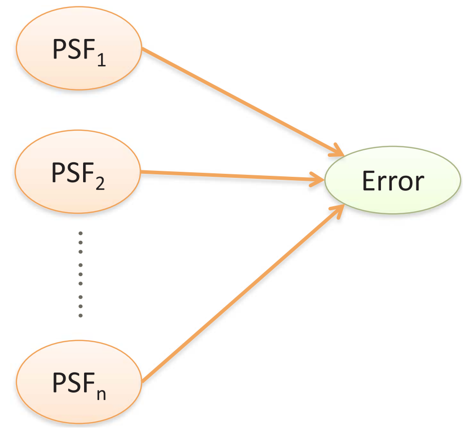

Fig. 1. Relationship between PSFs and human error. Error is the outcome of joint influence of PSF_1 to PSF_n.

To define the relationship between a human error and PSFs, two parameters are needed: (1) the prior belief (in terms of probabilities) of the PSFs and (2) the conditional belief (in terms of probability distribution) of the human error. In this case, prior probabilities of all possible states of a PSF are assumed equal (50% if the PSF is binary). The difficult part is to define the conditional probabilities, which represent the conditional dependency of human error on PSFs. This paper uses data collected in a virtual environment to define these conditional dependencies. Conditional dependencies are defined separately for each individual to reflect the fact that influence of PSFs on error may vary from individual to individual. Section 3 illustrates the approach in detail.

Once the probabilities of different errors during a task are calculated, they can be combined using the definitional/synthesis idiom, rather than a causal relationship, to achieve an overall failure probability for the task (Fenton and Neil, 2012). For example, in an offshore emergency evacuation situation, if an operator is interacting with a smoke hazard ($Error_1$) while keeping all fire doors open throughout the evacuation process ($Error_2$), then these errors can be combined to get an overall failure probability of the operator for the task evacuation. To reduce the computational complexity, errors ($Error_{1-n}$) are first classified into categories (CT_{1-n}) and then combined to get an overall failure (F) probability. The different categories of error considered in this paper are as follows: perception error, recognition error, procedural error, and lack of situational awareness. Each error can be classified into one or more categories. For example, interaction with a hazard can be categorized as a failure to perceive the severity of the hazard (perception error) and keeping fire doors open can be categorized as a procedural error. Fig. 2 shows how error probabilities in different categories can be combined to quantify the overall failure probability.

As shown in Fig. 2, there are two relationships that need to be defined: (1) the relationship between the errors ($Error_{1-n}$) and different categories (CT_{1-n}) and (2) the relationship between different categories (CT_{1-n}) and overall failure (F). Two parameters are needed to define these relationships: (1) the conditional belief (in terms of probability distribution) of the categories (CT), and (2) the conditional belief (in terms of probability distribution) of the overall failure (F).

To demonstrate how conditional probability distribution of *CTs* can be defined, a simple case is considered where the category variable CT_1 is binary and can have two possible states: *acceptable* and *not acceptable*. CT_1 is assumed to be dependent on $Error_1$ and $Error_2$. Table 1 shows the conditional probability table for CT_1. As shown in the table, $P(CT_1 = Acceptable)$ becomes zero if either $Error_1$ or $Error_2$ occurs. The only case when $P(CT_1 = Acceptable)$ becomes one is when none of the errors have occurred.

The conditional probability table for the failure node F can be