http://dx.doi.org/10.5516/NET.04.2013.702

MEASURING THE INFLUENCE OF TASK COMPLEXITY ON HUMAN ERROR PROBABILITY: AN EMPIRICAL EVALUATION

LUCA PODOFILLINI[1*], JINKYUN PARK[2], and VINH N. DANG[1]
[1]Paul Scherrer Institute (PSI), Switzerland;
[2]Korea Atomic Energy Research Institute (KAERI), Republic of Korea
[*]Corresponding author. E-mail : luca.podofillini@psi.ch

Received March 12, 2013

A key input for the assessment of Human Error Probabilities (HEPs) with Human Reliability Analysis (HRA) methods is the evaluation of the factors influencing the human performance (often referred to as Performance Shaping Factors, PSFs). In general, the definition of these factors and the supporting guidance are such that their evaluation involves significant subjectivity. This affects the repeatability of HRA results as well as the collection of HRA data for model construction and verification. In this context, the present paper considers the TAsk COMplexity (TACOM) measure, developed by one of the authors to quantify the complexity of procedure-guided tasks (by the operating crew of nuclear power plants in emergency situations), and evaluates its use to represent (objectively and quantitatively) task complexity issues relevant to HRA methods. In particular, TACOM scores are calculated for five Human Failure Events (HFEs) for which empirical evidence on the HEPs (albeit with large uncertainty) and influencing factors are available – from the International HRA Empirical Study. The empirical evaluation has shown promising results. The TACOM score increases as the empirical HEP of the selected HFEs increases. Except for one case, TACOM scores are well distinguished if related to different difficulty categories (e.g., "easy" vs. "somewhat difficult"), while values corresponding to tasks within the same category are very close. Despite some important limitations related to the small number of HFEs investigated and the large uncertainty in their HEPs, this paper presents one of few attempts to empirically study the effect of a performance shaping factor on the human error probability. This type of study is important to enhance the empirical basis of HRA methods, to make sure that 1) the definitions of the PSFs cover the influences important for HRA (i.e., influencing the error probability), and 2) the quantitative relationships among PSFs and error probability are adequately represented.

KEYWORDS : Human Reliability Analysis, Performance Shaping Factors, Task Complexity, Empirical Evaluation

1. INTRODUCTION

Human Reliability Analysis (HRA) aims at analyzing human errors in technological systems, the influences on human performance, and ultimately assessing the so-called Human Error Probability (HEP). HRA results, often in the context of Probabilistic Safety Assessment (PSA) results, inform operating as well as regulatory decisions with important safety and economic consequences. This calls for the need to use, in HRA like in any other areas of risk analysis, methods and tools that, to the extent possible, are built from and validated on empirical evidence. Recent efforts in this direction are the International HRA Empirical Study [1, 2] and the follow-up US HRA Empirical Study [3]; these studies were aimed at evaluating HRA methods' strengths and weaknesses against evidence from operating crew performance data, collected during simulated accident scenarios. Also, recognizing the need for enhancing the empirical basis of HRA methods, programs of HRA data collection from simulated environments are active or are being activated all over the world [4].

In parallel to the collection of data, the growing interest in enhancing the empirical basis of HRA is visible in the development of HRA methods with underlying models directly built from data (where the models represent quantitative relationships between the human error probability and factors influencing it, often referred to as Performance Shaping Factors, PSFs) [5, 6]. Further, methods are being developed to validate the relationships among PSFs and the error probability, e.g., in the already mentioned HRA empirical studies [1-3] and in [7-11]. In particular, in [7], it is investigated whether the complexity of procedure-guided tasks (by the operating crew of nuclear power plants in emergency situations) can be quantitatively and objectively measured with the use of TACOM (TAsk COMplexity) measure [12]. TACOM evaluates a complexity score by combining the contributions of five different task complexity aspects (e.g., amount of information to be processed, logical complexity of tasks, and knowledge requirements).

Besides the lack of reference HRA data, one of the

challenges for building empirically-based and validated HRA models is an objective evaluation of the performance influences (in HRA terms, the PSF ratings). Typically, these factors characterize the personnel performance in specific tasks by addressing the personnel directly (e.g., via the quality of their training, experience, and work processes), the task to be performed (e.g., the time required to complete the actions and the quality of procedural guidance), the available tools (e.g., the quality of human machine interface), and other aspects depending on the specific HRA method. Indeed, the application of current HRA methods is largely based on subjective evaluations (coming in at different stages of the analysis and to different extents, depending on the specific method and analyst knowledge/experience). On the one hand this influences the repeatability of HRA results (this issue is deeply investigated in the US HRA Empirical Study [3]). On the other hand, this challenges the empirical derivation and validation of the quantitative relationships among PSFs and HEPs, which ideally would require the availability of data points, collected for performance conditions corresponding to many different combinations of PSF ratings. The lack of objective PSF evaluations prejudices the establishment of a direct link between collected data and the associated PSF ratings, because of the need to subjectively interpret the data. For some of the PSFs typically considered by HRA methods, especially those aimed at characterizing crew behaviors (e.g., in the SPAR-H method: level of stress, fitness for duty, and work processes [13]), some (relatively high) level of subjectivity is probably unavoidable in the evaluation of these factors, due to the inherent variability of the effects of these factors on the error probability. However, there is still a substantial margin to improve on the PSF definitions and guidance to decrease subjectivity, as underscored by the results of the HRA Empirical Studies [1, 3].

In an effort to provide objective PSF measures, the present paper investigates the use of the TACOM measure to represent (objectively and quantitatively) the task complexity issues relevant to HRA methods. In previous works by one of the authors, TACOM scores have been compared with three types of crew performance indicators: the time to complete the task [12]; subjective workload scores as measured by the NASA-TLX (National Aeronautics and Space Administration – Task Load indeX) [14]; and the OPAS (Operator Performance Assessment System) scores developed by HAMMLAB (HAlden Man Machine LABoratory) of the OECD Halden Reactor Project [15]. It was shown that the TACOM scores and these human performance data have significant correlations; this supports TACOM measure as a relevant indicator of task complexity influences on crew performance. However, to evaluate the relevance of TACOM for HRA purposes, the relationship with the error probability needs to be investigated. First, qualitatively this requires that the TACOM definition should include the complexity elements that influence the human error probability. Second, quantitatively, TACOM scores should correlate with the error probabilities (the profile of the other PSFs being kept constant or their effect being properly considered, e.g., averaged out or factored out). To this end, in this paper, the TACOM measure is applied to multiple emergency tasks of different complexity, for which empirical evidence of the error probability is available (albeit with large uncertainty). In particular, Human Failure Events (HFEs) from the International HRA Empirical Study are selected [2]. In the context of the Empirical Study, pre-defined emergency tasks were simulated, obtaining evidence on the error probability and on the performance influencing factors on the corresponding HFEs. In the present paper, the TACOM scores are contrasted to the empirical evidence with the goal to evaluate both the correlation between the actual TACOM scores and the empirical HEPs of the considered HFEs. In addition, the paper investigates whether TACOM provides a difficulty characterization coherent with the empirical evidence (in other words, the ability to discriminate between, e.g., "easy" and "difficult" tasks).

The paper is structured as follows. In Section 2, a brief explanation of the TACOM measure is given. Then, the evaluation methodology applied in the present paper is described (Section 3). Section 4 presents the empirical basis of the evaluation: the data from the International HRA Empirical Study [2]. The results of the evaluation are given in Section 5. Finally, Section 6 includes concluding remarks along with summarizing the limitations of the study.

2. THE TACOM MEASURE

According to wide-spread operating experience in many industries (such as nuclear power plants, chemical plants, or aviation industries) that are based on large process control systems, it is evident that human error is one of the crucial contributors to serious accidents as well as incidents. In addition, the use of well-designed procedures is one of the practical options to reduce the possibility of human error [16].

One challenging aspect for the performance of tasks prescribed in emergency procedures is the need to cope with dynamically varying situations by using static task descriptions (i.e. the procedures). In following the emergency procedures, operators have to continuously assess the nature of the situation at hand in order to confirm the appropriateness of their response. It can be reasonably expected that the possibility of human error will increase when operators are faced with complex tasks. In this regard, one of the critical questions to be resolved is: "How complex a task is?" The TACOM measure has been developed to give a quantitative answer [12]. The TACOM measure is defined by a weighted Euclidean norm in a complexity space that consists of three dimensions, suggested by [17]. These dimensions are: task scope (TS), task structurability (TS), and task uncertainty (TU), each comprised of one

Download English Version:

https://daneshyari.com/en/article/1739980

Download Persian Version:

https://daneshyari.com/article/1739980

Daneshyari.com