Nuclear Engineering and Technology 54 (2022) 4170–4180

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

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

Original Article

Determining the complexity level of proceduralized tasks in a digitalized main control room using the TACOM measure

Inseok Jang, Jinkyun Park*

Korea Atomic Energy Research Institute, 111 Daedeok-daero, 989beon-gil, Yuseong-gu, Daejeon, 34057, Republic of Korea

ARTICLE INFO

Article history:
Received 2 March 2022
Received in revised form
12 May 2022
Accepted 18 June 2022
Available online 23 June 2022

Keywords:
Proceduralized tasks
Human reliability analysis
Level of task complexity
Unsafe act
Logistic regression
Digitalized main control room

ABSTRACT

The task complexity (TACOM) measure was previously developed to quantify the complexity of proceduralized tasks conducted by nuclear power plant operators. Following the development of the TACOM measure, its appropriateness has been validated by investigating the relationship between TACOM scores and three kinds of human performance data, namely response times, human error probabilities, and subjective workload scores. However, the information reflected in quantified TACOM scores is still insufficient to determine the levels of complexity of proceduralized tasks for human reliability analysis (HRA) applications. In this regard, the objective of this study is to suggest criteria for determining the levels of task complexity based on logistic regression between human error occurrences in digitalized main control rooms and TACOM scores. Analysis results confirmed that the likelihood of human error occurrence according to the TACOM score is secured. This result strongly implies that the TACOM measure can be used to identify the levels of task complexity, which could be applicable to various research domains including HRA.

1. Introduction

In order to ensure the safety of nuclear power plants (NPPs), the probabilistic safety assessment (PSA) technique has been applied over the past decades. One of the key results expected from PSA is the ability to perceive the risk level of NPPs through measures such as core damage frequency and large early release frequency. For obtaining reliable PSA results, therefore, it is prerequisite to apply large amounts of relevant raw information such as component failure frequencies and the human error probabilities (HEPs) of safety-critical tasks. Accordingly, the quality of the information to be used in human reliability analysis (HRA) to estimate the HEPs should be emphasized since its calculation has a significant impact on the credibility of the results of PSA.

However, the estimation of HEPs in most HRA methods largely depends on the judgments of HRA practitioners, especially when they have to decide the levels of the performance-shaping factors (PSFs), such as the levels of task complexity for predefined safety-critical tasks. Although the technical bases and detailed guidelines of HRA methods are provided, it is true that these judgments have the potential to degrade the credibility of HEP estimations. For example, in the case of the SPAR-H (Standardized Plant Analysis Risk-HRA) and K-HRA (Korean Human Reliability Analysis) methods, HRA practitioners should decide on one of three task complexity levels based on somewhat subjective evaluation criteria, as summarized in Table 1 [1,2].

The determination of task complexity levels based on the subjective evaluation criteria summarized in Table 1 may cause inconsistencies in HEP estimations in certain cases, which implies that the credibility of PSA results might not be secured due to an increase in uncertainty. In order to resolve this problem, one reasonable approach is to develop an objective, quantitative measure that can be used as a baseline for classifying the level of task complexities in a systematic way. In this regard, the application of the TACOM (TAsk COMplexity) measure is a promising candidate because it can quantify the complexity of proceduralized tasks to be conducted by human operators working in the main control room (MCR) of NPPs (for convenience, the term MCR operators will be used hereafter) [3–8]. (Note that for the application of the TACOM measure to determine the level of a task complexity involved in

* Corresponding author.
E-mail addresses: isjang@kaeri.re.kr (I. Jang), kshpjk@kaeri.re.kr (J. Park).

https://doi.org/10.1016/j.net.2022.06.018

Table 1
Determination of task complexity level in SPAR-H and K-HRA methods; adopted from Ref. [1,2].

HRA method	Level of task complexity	Evaluation criteria
SPAR-H	High complexity	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). For example, an unfamiliar equipment line-up is required that involves defeating interlocks on valves.
	Moderate complexity	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). For example, an atypical system startup is executed requiring the manual connection of backup power supplies.
	Nominal	Not difficult to perform. There is little ambiguity. An easily managed number of variables or inputs are involved. The organization of information or execution of steps is relatively straightforward with little potential for confusion.
K-HRA	Complex	The 'complex' type includes continuous control tasks or the tasks requiring comparison/integration of several sources of information.
	If-then	The 'if-then' type is applied when [two or more conditions to execute are checked or general pumps or values are operated] AND the trends of the primary parameters for operation are stable.
	Simple	The 'simple response' type is applied only if a prompt response can be possible with a simple and straightforward action.

various HRA methods, the definitions of the task complexity in the TACOM measure and in the target HRA method should be similar; this point is discussed in detail in Section 5.) Since its validity has been investigated by comparing TACOM scores with associated human performance data, it is strongly expected that a relevant technical basis can be obtained based on the TACOM measure. In this light, the objective of this study is to suggest evidence and/or criteria to determine the levels of task complexity to be faced by human operators through the use of the TACOM measure. To this end, this study performs a statistical analysis via logistic regression using TACOM scores and the occurrence of human errors observed from a full-scope simulator of a digitalized MCR installed in Korean domestic NPPs (hereafter, the term 'unsafe act (UA)' will be used instead of 'human error'; rationale is given in Section 3.2).

Results of the logistic regression in this study show that the likelihood of UA occurrence can be estimated according to the quantified TACOM scores, implying that the levels of task complexity can be classified according to specific ranges of TACOM scores. Accordingly, such results make it possible to conclude that the TACOM measure could be utilized as a baseline for determining the levels of task complexity.

The remainder of this paper is organized as follows. In Section 2, the background of the evaluation approach to task complexity in various HRA methods is explained with descriptions of the TACOM measure including its validation studies. In Section 3, the extraction of UAs collected from a full-scope simulator of a digitalized MCR is described. In addition, the calculated TACOM scores with associated input data for statistical analysis are explained. In Section 4, the statistical process is briefly described, and the results of logistic regression between TACOM scores and associated UAs are given that provide the technical underpinning for determining the levels of task complexity. Finally, the conclusion of this study is provided in Section 5 with discussions including limitations and further works.

2. Background

2.1. Evaluation of task complexity in HRA methods

Various HRA methods have been developed to identify erroneous human actions with their HEPs that can impact the safety of the structures, systems, and components in NPPs [9]. In performing HRA, the conditions that influence human performance have been represented by several contextual factors. Typical factors include the characteristics of the human operators, environments, organizational characteristics, and task contents that specifically degrade or improve human performance. Therefore, HEPs can be soundly quantified by considering the level of influence from these contextual factors [10]. These contextual factors are referred to by different terms according to the particular HRA method, such as PSFs, performance influencing factors (PIFs), error producing conditions (EPCs), and common performance conditions (CPCs). Accordingly, most HRA methods suggest their own PSFs with similar and/or different definitions. Table 2 lists a catalog of PSFs and their evaluation approaches in several representative HRA methods.

As shown in Table 2, as two of the key PSFs appearing in various HRA methods, task complexity and task instructions (e.g., procedures) are clearly important for understanding the performance of human operators (see underlined text in Table 2). Numerous researchers have indeed stressed that task complexity is one of the dominant PSFs [14–17]. However, determining the levels of task complexity relies on subjective decisions of HRA practitioners and other subject matter experts. Actually, Table 1 in Section 1 illustrates the limitation of subjective decisions on the levels of task complexity.

In order to overcome this limitation, it is essential to utilize an objective measure that can quantify the complexity of the proceduralized tasks performed by human operators. This would make it possible to soundly determine the levels of task complexity based on quantitative evaluation results. In this regard, the objective measure called TACOM is briefly introduced in Section 2.2.

2.2. TACOM measure and its validation studies

The objective measure called TACOM quantifying the complexity of proceduralized tasks implemented by NPP MCR operators was previously developed [3–8], and its appropriateness has been verified by a series of studies [3,4,7].

In order to quantify task complexity with the TACOM measure, five sub-measures affecting the complexity of proceduralized tasks are defined as shown in Table 3. Based on these five sub-measures, as described in Table 4, a complexity space with three dimensions, namely task scope (TS), task structurability (TR), and task uncertainty (TU), was considered to quantify the complexity of proceduralized tasks [8].

Based on the three complexity dimensions combining the five sub-measures described in Tables 3 and 4, the TACOM measure quantifies the complexity of a proceduralized task by using the following formula. A more detailed process for quantifying the complexity of proceduralized tasks can be found in Refs. [6,8].

$$
\begin{aligned}
&TACOM = \left\{ 0.621\ (TS)^2 + 0.239\ (TR)^2 + 0.140\ (TU)^2 \right\}^{1/2} \\
&\mathrm{TS} = 0.716\ SIC + 0.284\ SSC \\
&\mathrm{TR} = 0.891\ SLC + 0.109\ AHC \\
&\mathrm{TU} = EDC
\end{aligned}
\tag{1}
$$

Table 2
Catalog of PSFs considered in representative HRA methods.

HRA method	PSFs	Evaluation approach
Technique for human error rate prediction (THERP) [11]	▫ Physiological stressors ▫ Psychological stressors ▫ *Task and equipment characteristics* ▫ Organismic factors ▫ Situational characteristics ▫ *Job and task instruction*	THERP relies on the *experience and judgment of human factors specialists* to assess the impact of PSFs.
K-HRA [2].	▫ Complexity of a unitary action ▫ Quality of procedure ▫ Time availability and action familiarity ▫ Time urgency ▫ Scenario severity ▫ Environmental hazard ▫ Training and education	K-HRA relies on *expert judgments* to assess the impact of PSFs.
INTENT [12].	▫ Human-Machine interface (HMI) ▫ Stress ▫ Skill, rule and knowledge based behavior ▫ Experience ▫ Safety culture ▫ Training ▫ Motivation ▫ Workload ▫ Supervision ▫ Communication ▫ *Procedures*	*Ratings of PSF importance are generated independently by each analyst*, then normalized for each analyst across the PSFs.
Human reliability management system (HRMS) [13].	▫ Time ▫ *Task complexity* ▫ Task organization ▫ *Procedures* ▫ Training/expertise/experience/competence ▫ Quality of information/interface	There are factual questions about the PSFs. *Based on these factual questions, PSF weightings are judged.*
SPAR-H [1].	▫ Available time ▫ *Task complexity* ▫ *Procedures* ▫ Fitness for duty ▫ Stress/stressors ▫ Experience/training ▫ Ergonomics/Human-System interface (HSI) ▫ Work process	PSF multipliers are *based on the authors' observation/review* of event statistics and on a comparison with data in existing HRA methods. The selection of PSF multipliers is based on *expert judgement with guidance* in SPAR-H.

Table 3
Sub-measures of TACOM; adopted from Ref. [7].

Sub-measure	Description
Step information complexity (SIC)	Complexity due to the amount of information to be processed by human operators
Step size complexity (SSC)	Complexity due to the number of actions to be conducted by human operators
Step logic complexity (SLC)	Logical complexity due to the sequences of actions to be followed by human operators
Abstraction hierarchy complexity (AHC)	Complexity due to the amount of domain knowledge required of human operators
Engineering decision complexity (EDC)	Complexity due to the amount of cognitive resources required by human operators to establish an appropriate decision criterion

Table 4
Three complexity dimensions with associated characteristics; adopted from Ref. [7].

Complexity dimension	Definition	Related characteristics for a task description
Task scope (TS)	Representing the breadth, extent, range, or general size of the task being considered	• Number of actions • Amount of information
Task structurability (TR)	Representing whether or not the sequence and the relationship between subtasks are well structured	• Logical entanglement • Amount of domain knowledge
Task uncertainty (TU)	Representing the degree of predictability or confidence in a task	• Difficulty to establish a decision criterion

Following the development of the TACOM measure, its appropriateness was validated by comparing TACOM scores with three kinds of human performance data, namely response times, subjective workload scores, and the number of UAs collected from simulated emergency situations in Korean domestic NPPs [3,4,7]. From these validation studies, it was observed that the TACOM scores are strongly correlated with the three kinds of human performance data [3,4,7]. Such results made it possible to say that the validation studies support the appropriateness of the TACOM measure in terms of quantifying the complexity of proceduralized tasks.

However, it remains unclear how to specify the different levels of task complexity to inform HRA practitioners in conducting HRA based on TACOM scores. In order to propose a baseline to determine the levels of task complexity on the basis of TACOM scores, a statistical analysis between UA data and associated TACOM scores is performed in this work. For this, the collection of UA data from a full-scope simulator of Korean domestic NPPs is explained in the next section.

3. Securing UA data

3.1. Collection of UAs from the full-scope simulator of a digitalized MCR

In order to investigate the levels of task complexity felt by MCR operators, it is promising to compare the performance data of MCR operators with the task complexities of the proceduralized tasks described in emergency operating procedures (EOPs). In other words, it is possible to grasp key insights into the determination of the levels of task complexity if the effect of task complexities on the performance of MCR operators can be properly clarified. For this reason, in this study, logistic regression based on TACOM scores and UAs observed from simulated off-normal conditions was carried out. The UA data used in this study originate from the HuREX (Human Reliability data Extraction) database developed by KAERI [18,19], which contains performance data of MCR operators collected from a full-scope simulator of Korean domestic NPPs. To analyze these performance data, all kinds of MCR operator responses were recorded using audio/video recording equipment during the simulated off-normal conditions. Extra information including component manipulation logs and the trends of key process parameters (e.g., pressure, temperature, and flow rate) were also gathered because they are helpful for confirming the detailed control behaviors conducted during the progression of the off-normal conditions. The HuREX database contains the performance data of MCR operators facing diverse off-normal conditions; in this study, data collected from four representative accident conditions were used. That is, it is necessary to specify the situations in which MCR operators strictly follow a series of proceduralized tasks because the purpose of this study is to compare the UA data of MCR operators who have to conduct a required task with its TACOM score. In this regard, since it is mandatory for MCR operators to follow EOPs when an accident occurs, UA data observed from the simulations of four representative accident conditions were used: a loss of coolant accident (LOCA), steam generator tube rupture (SGTR), loss of all feedwater (LOAF), and station blackout (SBO).

3.2. UAs extracted from simulation records and corresponding TACOM scores

In order to extract UAs from the responses of MCR operators in dealing with the simulated accident conditions, the records of their responses such as audio/video recording, component manipulation logs, and process parameters logs were analyzed with

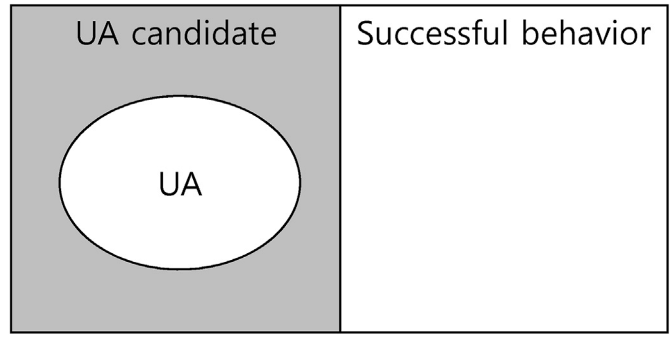

Fig. 1. Classification scheme of UA and UA candidate; adopted from Ref. [19].

consideration of the definition of UA that is the basis for constructing the HuREX database. In the HuREX database, MCR operators' responses are divided into three categories as shown in Fig. 1: "(1) *successful behavior* satisfying all kinds of requirements and performance standards pertaining to the operation of NPPs, (2) *UA candidate* implying all kinds of human behaviors that deviate from procedures such as EOPs or operation practices (e.g., technical specifications and the conduct of operation), and (3) *UA* indicating human behaviors that have the potential to give a negative impact on the operational safety of NPPs in a direct or indirect way" [19, p. 37].

Of them, the third category corresponds to the definition of UA. According to this UA definition, a catalog of UAs was distinguished from the responses of MCR operators who have to carry out individual tasks to cope with the four accident conditions. Table 5 summarizes the number of UAs with associated TACOM scores that are quantified with respect to each task. For example, in the case of 'LOCA-04' in Table 5, its TACOM score is 2.823, and no UA was observed even though this task was conducted 38 times. In contrast, the 'LOCA-10' task, with a TACOM score of 4.275, indicates that MCR operators showed 10 UAs in conducting this task 31 times. Here, since the TACOM score of the latter is relatively higher than that of the former, it is possible to assume that the effect of task complexity on the occurrence of UAs can be scrutinized by using a relevant statistical analysis such as logistic regression.

In addition to the data for individual tasks shown in Table 5, in order to clarify the threshold TACOM score at which the performance of MCR operators is rapidly reduced using the logistic regression analysis, it is much more advantageous to see if UAs occur over a wide range of TACOM scores. In this light, arbitrary group tasks consisting of 10 successive individual tasks were extracted from the LOCA and SGTR procedures. For example, in the case of the group task denoted 'LOCA-04-13' in Table 6, it was counted from the HuREX database that MCR operators conducted 10 individual tasks in a row (i.e., from LOCA-04 to LOCA-13) 29 times. According to Table 5, the total number of UA occurrences corresponding to these 10 individual tasks is 23. As the TACOM score with respect to the hypothetical group task containing these 10 individual tasks can be easily calculated [8], it is promising to construct the contents of Table 6. Similar to Table 5, Table 6 provides the group task opportunities, the TACOM score of the group tasks, and the number of UAs with respect to the group tasks.

4. Statistical analysis

4.1. Logistic regression

Generally, a logistic regression model is adopted to investigate the relation between a binary/categorical (1 or 0) variable and

Table 5
Task opportunities, UA occurrences, and associated TACOM scores for emergency tasks in the four accident conditions.

Individual task ID	Task opportunity	TACOM score	UA occurrence
LOCA-04	38	2.823	0
LOCA-05	38	4.021	2
LOCA-06	37	3.025	1
LOCA-07	37	2.647	0
LOCA-08	37	4.569	0
LOCA-09	36	4.055	5
LOCA-10	31	4.275	10
LOCA-11	31	3.624	2
LOCA-12	31	4.387	3
LOCA-13	29	3.588	0
LOCA-14	28	2.007	0
LOCA-15	22	3.789	6
LOCA-16	10	3.756	1
LOCA-17	6	3.4708	1
LOCA-18	5	3.617	0
LOCA-19	5	2.898	0
LOCA-20	5	3.0713	0
LOCA-21	4	3.072	0
LOCA-22	2	2.941	0
LOCA-23	1	5.116	0
LOCA-38	7	3.233	0
LOCA-46	5	2.007	0
LOCA-47	5	3.617	0
LOCA-48	4	2.898	0
LOCA-49	4	3.071	0
LOCA-50	4	3.072	1
LOCA-51	3	3.328	0
LOCA-52	3	4.175	0
LOCA-53	3	2.941	0
LOCA-54	1	5.127	0
LOCA-55	1	1.914	0
LOCA-56	1	2.730	0
LOCA-57	1	4.353	0
SGTR-04	27	2.823	0
SGTR-05	26	4.021	0
SGTR-06	26	3.025	2
SGTR-07	25	2.648	1
SGTR-08	25	3.027	0
SGTR-09	25	3.470	6
SGTR-10	24	3.532	1
SGTR-11	24	4.389	13
SGTR-12	23	4.084	0
SGTR-13	23	3.436	0
SGTR-14	23	3.205	3
SGTR-15	20	4.013	5
SGTR-16	15	3.617	0
SGTR-17	13	2.580	0
SGTR-18	10	5.048	0
SGTR-19	10	4.171	0
SGTR-20	9	1.914	0
SGTR-21	9	2.730	0
SGTR-22	8	2.999	0
SGTR-23	6	3.829	0
SGTR-24	6	2.391	0
SGTR-25	5	3.372	0
SGTR-26	4	3.071	0
SGTR-27	4	3.072	0
SGTR-28	3	2.381	0
SGTR-29	1	3.328	0
SGTR-30	1	4.053	0
SGTR-31	1	4.0882	0
SGTR-32	1	3.599	0
SGTR-33	1	3.608	0
SGTR-34	1	4.249	0
SGTR-35	1	4.106	0
SGTR-36	1	2.648	0
SGTR-37	1	1.390	0
SGTR-38	1	2.375	0
LOAF-04	10	1.706	0
LOAF-05	10	3.418	2
LOAF-06	10	4.319	2
LOAF-07	10	3.172	1
SBO-04	5	3.124	0
SBO-05	5	3.809	0
SBO-06	5	3.315	0
SBO-07	5	1.914	0
SBO-08	5	3.300	1
SBO-09	5	3.741	2
SBO-10	3	2.867	0
SBO-11	3	2.518	0
SBO-12	3	2.007	0
SBO-13	3	4.146	0
SBO-14	3	2.730	0
SBO-15	3	3.588	0
SBO-16	2	2.855	0
SBO-17	2	3.470	0
SBO-18	1	2.826	0
SBO-19	1	4.038	0
SBO-20	1	3.073	0

diverse types of variables including categorical or continuous variables. With those variables, if the binary dependent variable is defined as Y, the predicted conditional probability under condition x can be obtained as $\Pr(Y=1|X=x)=p(x)$. To obtain the probability $p(x)$ in a logistic regression model, the concept of odds is defined as $p(x)/(1-p(x))$, which indicates the ratio of the probability that the event occurs to the probability that the event does not occur. Then, the logit of odds, $\ln\left(\frac{p(x)}{1-p(x)}\right)$, is used as the dependent variable in an ordinary linear regression [20].

Formally, the logistic regression model is that:

$$\ln\frac{p(x)}{1-p(x)}=\beta_0+\beta_1 x \tag{2}$$

Solving for $p(x)$,

$$p(x)=\frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}}=\frac{1}{1+e^{-(\beta_0+\beta_1 x)}} \tag{3}$$

where β_0, β_1 are regression coefficients and x is the independent variable in the logistic regression model.

Using this logistic regression model, it is possible to not only predict the effect of the independent variables on a binary response variable but also classify observations by estimating the probability that an observation is in a particular category (such as whether a UA occurred or not as in this study).

4.2. Dependent variable and independent variable

According to the purpose of this study, the dependent variable can be defined as UA occurrence, of which the variable type is dichotomous. If an UA occurs, the dependent variable (UA occurrence) is recorded as a value of 1. If the operator performs a given task successfully, the dependent variable is recorded as a value of 0. In order to scrutinize the relation between the estimated probability of UA occurrence and the associated TACOM score, the independent variable is the TACOM score, of which the variable type is continuous/numerical. Fig. 2 shows the relevant data input method for logistic regression in order to investigate the relation between UA occurrence and associated TACOM score.

As shown in the input table in Fig. 2, for performing the logistic regression, as many rows as the total number of task opportunities are generated along with the associated TACOM scores. In the right column, a value of 0 is assigned when the task is successfully performed by MCR operators and a value of 1 is assigned when a UA occurs.

Table 6
Task opportunities and UA occurrences for hypothetical group tasks with respect to LOCA and SGTR accident conditions.

Group task ID	Task opportunity	TACOM score	UA occurrence
LOCA-04-13	29	6.187	23
LOCA-05-14	28	6.212	23
LOCA-06-15	22	6.220	22
LOCA-07-16	10	6.367	10
LOCA-08-17	6	6.356	6
LOCA-09-18	5	6.594	5
LOCA-10-19	5	6.230	5
LOCA-11-20	5	6.098	5
LOCA-12-21	4	6.087	4
LOCA-13-22	2	5.781	2
LOCA-14-23	1	5.769	1
SGTR-04-13	23	5.831	23
SGTR-05-14	23	5.900	23
SGTR-06-15	20	5.997	20
SGTR-07-16	15	6.101	15
SGTR-08-17	13	6.093	13
SGTR-09-18	10	5.948	10
SGTR-10-19	10	6.079	10
SGTR-11-20	9	5.959	9
SGTR-12-21	9	5.860	8
SGTR-13-22	8	5.894	8
SGTR-14-23	6	5.839	6
SGTR-15-24	6	5.763	5

4.3. Results and implications

Table 7 summarizes the results of the logistic regression analysis including the estimated coefficients, exponentiated coefficients for the independent variable, 95% confidence interval for the exponentiated coefficients, and p-values. Fig. 3 also depicts the results of the logistic regression between TACOM score and UA occurrence, depicting how the estimated probabilities of UA occurrence according to the TACOM scores and the logistic regression model are obtained. The TACOM score at 50% UA occurrence (i.e., 50% cutoff of TACOM score) is also derived.

Table 7
Coefficients resulting from logistic regression.

Term	Coefficients (B)	Standard error	Exponentiated coefficient (EXP(B))	95% confidence interval for EXP(B)		p-value
				Lower	Upper	
TACOM score	1.901	0.103	6.691	5.470	8.185	< 0.01
Constant	−9.643	0.481	0.000	—	—	< 0.01

In our logistic regression model, the constant had a coefficient of −9.643 and the TACOM score had a coefficient of 1.901, and as shown in the last column of Table 7, the constant and TACOM score both had p-values smaller than 0.01, which means that they are statistically significant. The confidence interval in the regression analysis does not include 1.00, which implies that the TACOM score influences whether or not a UA occurs. If the confidence interval contains 1.00, it means that increasing the TACOM score by one point does not have an effect on UA occurrence. Using the estimated coefficients for the constant and TACOM score in Table 7, it is possible to estimate the probabilities of UA occurrence for arbitrary TACOM scores, as depicted in Fig. 3. If we substitute the estimated coefficients for the constant and TACOM score ($\beta_0 = -9.643$ and $\beta_1 = 1.901$) respectively into the logistic regression Equation (3) introduced in Section 4.1, the following estimated regression equation for the specific problem can be made:

$$\widehat{p}(x) = \frac{e^{\beta_0+\beta_1 x}}{1+e^{\beta_0+\beta_1 x}} = \frac{e^{-9.643+1.901x}}{1+e^{-9.643+1.901x}} = \frac{1}{1+e^{-(-9.643+1.901x)}} \tag{4}$$

where $\widehat{p}(x)$ is the estimated probability that a UA occurs and x is any TACOM score to estimate the probability of UA occurrence.

Using the estimated regression Equation (4), for example, the estimated probability of UA occurrence when the TACOM score is equal to 5.00 or 6.00 can be calculated as follows.

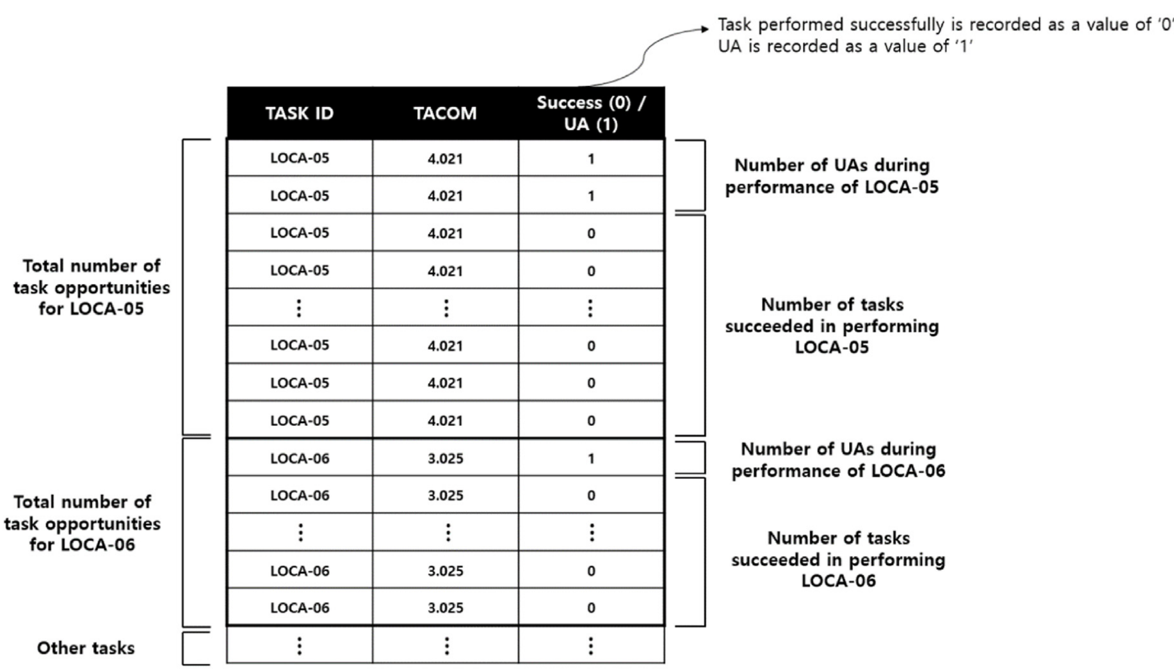

Fig. 2. Data input method for logistic regression analysis between TACOM scores and the number of UAs from Tables 5 and 6

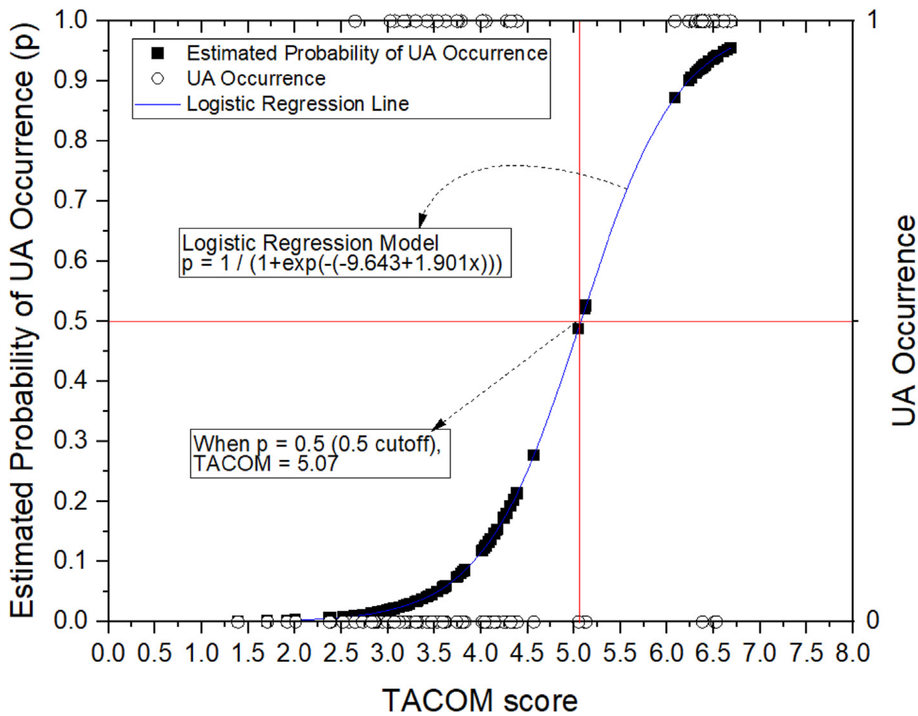

Fig. 3. Result of the logistic regression analysis between UA occurrence and TACOM score for the proceduralized tasks shown in Tables 5 and 6

$$\text{when TACOM score} = 5.00,\ \widehat{p}(5.00) = \frac{e^{-9.643+1.901*5.00}}{1+e^{-9.643+1.901*5.00}} = 0.466$$

This result shows that if the TACOM score quantifying the task complexity is 5.00, then the estimated probability that a UA occurs is about 0.466 or 46.6%.

$$\text{when TACOM score} = 6.00,\ \widehat{p}(6.00) = \frac{e^{-9.643+1.901*6.00}}{1+e^{-9.643+1.901*6.00}} = 0.853$$

This result shows that if the TACOM score is 6.00, then the estimated probability that a UA occurs is about 0.853 or 85.3% according to the model. From the above two calculation results, it can be seen that the difference between TACOM scores of 5.00 and 6.00 is quite substantial in terms of the estimated probability of UA occurrence, 46.6% versus 85.3%. Moreover, using the results of the logistic regression and Equation (2), the TACOM score when the estimated probability of UA occurrence is equal to 0.5 (i.e., even odds of UA occurrence/no UA occurrence) can be obtained as follows:

$$\ln\frac{0.5}{1-0.5} = -9.643 + 1.901x,$$

$$\therefore x(\mathit{TACOM\ score\ when}\ \widehat{p}(x) = 0.5) = 5.07$$

From this result, it is possible to say that if the TACOM score is 5.07 or more, the estimated probability of UA occurrence is more than 50%, and if it is less than 5.07, the estimated probability of UA occurrence is less than 50%. This strongly implies that it is necessary to maintain the complexity of proceduralized tasks at 5.07 or less in order to reduce the probability of UA occurrence to 50% or less.

Similar to this, the levels of task complexity may possibly be classified based on the estimated probabilities of UA occurrence according to the TACOM scores providing arbitrary task complexity ranges. However, in order to propose the TACOM measure as a baseline for classifying the levels of task complexity, it would be better to consider not only the estimated probabilities of UA occurrence according to the TACOM scores but also the change in estimated probabilities of UA occurrence when the TACOM score increases continuously. In this light, the Appendix shows the estimated probability of UA occurrence according to the TACOM score and the probability changes by increasing the TACOM score by 0.1 intervals. Based on the data in the Appendix, Fig. 4 depicts the estimated probabilities of UA occurrence and their probability changes by increasing TACOM score.

As shown in the Appendix and Fig. 4, based on TACOM scores of 3.6 or 3.7, the estimated probability of UA occurrence and its probability change are relatively stable compared to other sections with TACOM scores lower than 3.6 or 3.7, which implies that the task complexity does not significantly affect human performance at these levels. However, as easily identified in Fig. 4, the estimated probability of UA occurrence is higher than 0.05 and the change in estimated probability of UA occurrence starts to increase on the order of −2 power of 10 with TACOM scores higher than 3.6 or 3.7, which indicates that this point (i.e., a TACOM score of 3.6 or 3.7) is a possible candidate point designating where human performance starts to dramatically degrade due to task complexity, and could be therefore a criterion to identify the levels of task complexity. In addition, based on the TACOM score of 5.1, the estimated probability of UA occurrence exceeds 50%, and the probability change also steadily increases to this point. At TACOM scores over 5.1, the estimated probability of UA occurrence still increases rapidly, but after a certain point both measures are saturated, which implies

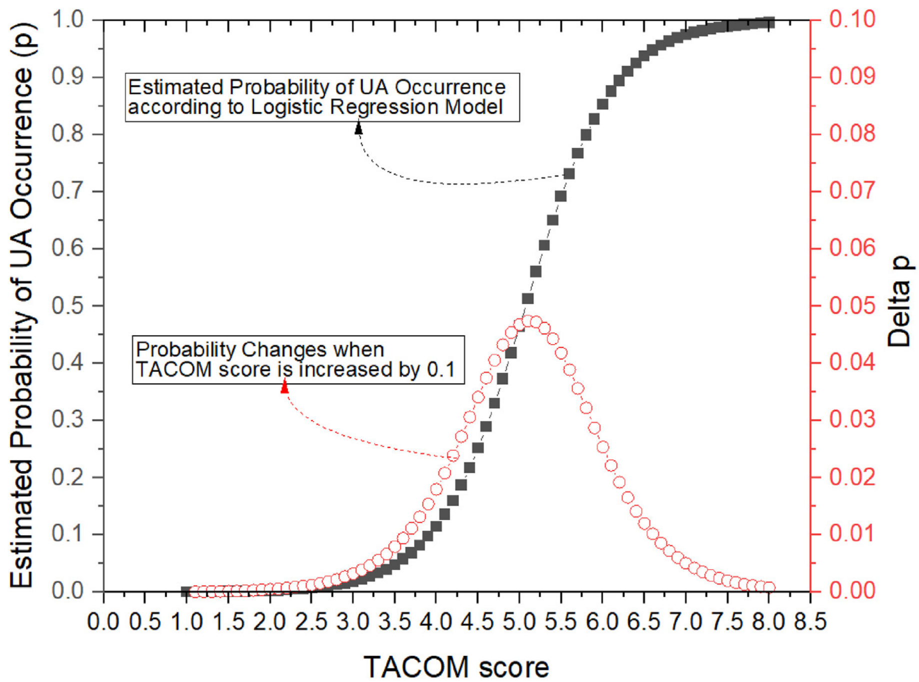

Fig. 4. Estimated probability of UA occurrence according to the derived logistic regression model and probability changes.

that this point (i.e., TACOM score of 5.1) could be another criterion to identify higher levels of task complexity.

Based on this interpretation of the results of this study, the level of task complexity is designated as 'Nominal' when the estimated probability of UA occurrence is lower than 0.05, corresponding to a TACOM score of about 3.6, and the level of task complexity is designated as 'Moderate complexity' when the estimated probability of UA occurrence is between 0.05 and 0.5, corresponding to TACOM scores between about 3.6 and 5.l. Finally, the level of task complexity is designated as 'High complexity' when the estimated probability of UA occurrence exceeds 0.5, corresponding to a TACOM score above 5.l. If such designation is applied to the SPAR-H and K-HRA methods introduced in Table 1 of Section 1, the levels of complexity could be determined in a more objective manner based on the above ranges of TACOM scores, as shown in Table 8. (Note that for the application of the TACOM measure to determine the level of task complexity in various HRA methods, the definitions of task complexity in the TACOM measure and in the target HRA method should be similar; this point is discussed in detail in Section 5.).

From these observations, it is possible to say that the TACOM measure could be a baseline for determining the levels of task complexity to be faced by MCR operators. It is also anticipated that, based on the results of this study, the levels of task complexity in various HRA methods could be determined by more objective methods, thereby reducing the subjectivity of HRA practitioners.

Table 8
Application of the results of this study to SPAR-H and K-HRA methods.

HRA method	Level of task complexity	Evaluation criteria from TACOM score ranges
SPAR-H	Nominal	TACOM score $\leq$ 3.6
	Moderate complexity	3.6 $<$ TACOM score $\leq$ 5.1
	High complexity	5.1 $<$ TACOM score
K-HRA	Simple	TACOM score $\leq$ 3.6
	If-then	3.6 $<$ TACOM score $\leq$ 5.1
	Complex	5.1 $<$ TACOM score

5. Discussion and conclusion

The determination of the levels of task complexity for estimating HEPs in most HRA methods largely depends on the judgments of HRA practitioners, which may cause inconsistencies in the HEP estimations due to analyst subjectivity. To overcome this limitation and determine the levels of task complexity in an objective manner, the TACOM measure that can quantify the complexity of proceduralized tasks was applied in this study. Since it is difficult to determine the levels of task complexity only with the quantified TACOM scores, logistic regression analysis between UAs observed from a full-scope simulator of Korean domestic NPPs and the associated TACOM scores was performed to identify the human performance degradation points that could provide evidence to classify the levels of task complexity.

In the analysis, the estimated probability of UA occurrence and its probability change were calculated based on the derived logistic regression model. As a result, it was confirmed that the TACOM measure has the potential to be used as a baseline for classifying the levels of task complexity in digital MCRs of NPPs. The results indicated that the TACOM measure classified several task complexity points: (1) where the estimated probability of UA occurrence starts to dramatically increase, (2) where the estimated probability of UA occurrence increases and exceeds 50%, and (3) where the estimated probability of UA occurrence continues to increase and finally saturates. Based on these interpretations, task complexity was divided into three levels applying the SPAR-H (K-HRA) method, namely 'Nominal (Simple)', 'Moderate complexity (If-then)', and 'High complexity (Complex)', by using the likelihood of the estimated probability of UA occurrence and its probability changes. Since most HRA methods classify task complexity into two or three levels, we provided the results considering three levels of task complexity for easy an application to HRA. Based on the likelihood of the estimated probability of UA occurrence and its probability changes, task complexity could alternatively be divided into five or more levels depending on which domain it is applied to.

Before concluding this work, first, it should be discussed whether the definitions of task complexity in SPAR-H, K-HRA, and the TACOM measure are similar to each other. Fortunately, the definition of task complexity in K-HRA and TACOM seem to be similar, focusing on the complexity of proceduralized tasks, which is a key point when applying TACOM to quantify the task complexity. However, based on the description of task complexity in SPAR-H shown in Table 1, the definition of task complexity with the phrase "concurrent diagnoses" in SPAR-H seems to be wider than that of the TACOM measure if the tasks related to the concurrent diagnoses are not strictly described in the procedures. As shown in Table 3, the TACOM measure includes five sub-measures to quantify the complexity of proceduralized tasks. These five sub-measures were selected for the TACOM measure among many factors affecting the task complexity of the proceduralized tasks since they are quantifiable [8]. For this reason, the TACOM measure might not be able to quantify the tasks that do not include the characteristics of the five sub-measures, such as concurrent diagnosis tasks not described in the procedures. However, it is expected that there is still a possibility to apply the results of this study to the determination of task complexity levels in SPAR-H. That is, if the application scope of the results of this study is limited to the determination of task complexity levels for proceduralized tasks such as the tasks in the EOPs and abnormal operating procedures, tasks related to concurrent diagnoses could potentially be decomposed into several sub-tasks, especially in the EOPs. These proceduralized sub-tasks could then be used as input data to quantify the task complexity in the TACOM measure, allowing us to determine the task complexity levels in SPAR-H using the TACOM measure. Although it might not be possible to include all internal and/or external factors related to concurrent diagnosis tasks in the TACOM measure, the result of this study could possibly be applied to the determination of task complexity levels in SPAR-H if the difference in the determination of task complexity levels between SPAR-H and the TACOM measure is possibly minimized by decomposing the concurrent diagnosis tasks into several sub-tasks. For application of the TACOM measure to tasks that are not proceduralized, more discussion is necessary.

Second, the difference between nominal HEP and basic HEP should be explained. According to THERP [11], "The nominal HEP is the probability of a given human error when the effects of plant-specific PSFs have not yet been considered (p. 5–10).", while the basic HEP is the probability of human error when the effects of PSFs have been considered. In this study, the meaning of the estimated probabilities of UA occurrence by the logistic regression model is close to the basic HEP according to TACOM scores.

Third, for the improvement of the results of this study, it would be better to compare the estimated probabilities of UA occurrence by the logistic regression model shown in Equation (4) to the data in Table 5. However, before this comparison, it should be noted that since the regression analysis is a set of statistical processes *to find the line that most closely fits the data*, the comparison results between the estimated probabilities of UA occurrence by the logistic regression model and the data might not be exactly the same. Since the number of data in Table 5 is too small, such as the task opportunities at the low TACOM scores, the comparison is performed by selecting temporary TACOM ranges, as shown in Table 9.

As shown in Table 9, although the averaged probabilities of UA occurrence estimated by the logistic regression model are not exactly same as the probabilities of UA occurrence calculated by the data in Table 5, it can be seen that the trend of probability changes according to changes in TACOM range is similar. In a way, this comparison result is natural since the logistic regression model was derived using the currently available (limited) data in Table 5. However, it is expected that if the number of data such as task opportunity and UA occurrence at a specific TACOM score is large enough, the two probabilities estimated by the logistic model and the data would be closer.

Table 9
Comparison of probabilities estimated by the logistic regression model and the data in Table 5.

TACOM range	Task opportunity (m)	UA occurrence	Probabilities of UA occurrence calculated by the data in Table 5	Averaged probabilities of UA occurrence estimated by the logistic regression model
0–2.5	72	0	0.0096*	0.0033
2.5–3.0	185	1	0.0054	0.0131
3.0–3.5	244	18	0.0738	0.0300
3.5–4.0	162	12	0.0741	0.0653
4.0–4.5	261	40	0.1543	0.1535

* Zero failure estimation method is used ($p = 1 - 0.5^{\frac{1}{m}}$).

Fourth, the relationship between the results of the current study and the previous study performed by Park et al. should be pointed out [21]. The previous study investigating HEPs based on the complexity of proceduralized tasks under an analog environment [21] showed that the estimated probability of UA occurrence was divided into two categories, the estimated probability of error of omission (EOO) occurrence and error of commission (EOC) occurrence, since it was observed that the occurrence of EOOs largely depends on the dynamic characteristics of the accident condition rather than on the complexity of the proceduralized tasks. Accordingly, the previous study was only able to derive the estimated probabilities of EOOs and EOCs. On the other hand, the human performance data used in this study was collected from a full-scope simulator of a fully digitalized MCR equipped with a computerized procedure system (CPS), where the function of the CPS automatically checks the trigger condition of non-sequential procedural steps and thus reduces or eliminates the effect of dynamic characteristics on the occurrence of UAs (both EOCs and EOOs). Based on this, it is expected that the logistic regression model estimating the probability of UAs in this study is mainly affected by the TACOM scores quantifying the proceduralized tasks. Considering the abovementioned points, it is evident that the results of the current study imply that the TACOM measure can be utilized as a baseline for classifying the levels of task complexity.

The following topics require more research in future studies. (1) Since our results support the identification of task complexity levels in a digital environment-based MCR, the likelihood of the estimated probability of UA occurrence to classify the levels of task complexity in an analog environment-based MCR should be derived and compared with those of the digital MCR under similar internal and external conditions. (2) The levels of task complexity determined based on the TACOM measure (main results of this study) should be applied to various HRA methods, such as SPAR-H and K-HRA as considered here, to investigate the impact of task complexity level determination by the TACOM measure on final HEP estimations for comparison with those by subjective decisions.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

This work was supported by a Nuclear Safety Research Program grant from the Korea Foundation of Nuclear Safety (KOFONS), funded by the Nuclear Safety and Security Commission of the Republic of Korea (No. 2105029-0121-CG100).

Appendix. Estimated probability of UA occurrence ($\hat{p}$) and probability changes ($\Delta\hat{p}$) by increasing TACOM score by 0.1 intervals

TACOM score*	Estimated probability of UA occurrence ($\hat{p}$)	$\Delta\hat{p}$
1.0	0.000434014	—
1.1	0.000524836	0.000090822
1.2	0.000634651	0.000109815
1.3	0.000767427	0.000132775
1.4	0.000927954	0.000160527
1.5	0.001122022	0.000194068
1.6	0.001356622	0.000234600
1.7	0.001640193	0.000283571
1.8	0.001982920	0.000342727
1.9	0.002397090	0.000414170
2.0	0.002897515	0.000500426
2.1	0.003502045	0.000604530
2.2	0.004232167	0.000730122
2.3	0.005113727	0.000881560
2.4	0.006177776	0.001064049
2.5	0.007461569	0.001283793
2.6	0.009009727	0.001548157
2.7	0.010875582	0.001865856
2.8	0.013122728	0.002247146
2.9	0.015826756	0.002704028
3.0	0.019077197	0.003250441
3.1	0.022979615	0.003902418
3.2	0.027657800	0.004678185
3.3	0.033255952	0.005598152
3.4	0.039940669	0.006684717
3.5	0.047902489	0.007961819
3.6	**0.057356607**	**0.009454119**
3.7	0.068542281	0.011185674
3.8	0.081720269	0.013177988
3.9	0.097167566	0.015447297
4.0	0.115168598	0.018001032
4.1	0.136002101	0.020833504
4.2	0.159923148	0.023921047
4.3	0.187140288	0.027217140
4.4	0.217788629	0.030648341
4.5	0.251900843	0.034112215
4.6	0.289379409	0.037478565
4.7	0.329974587	0.040595178
4.8	0.373273230	0.043298643
4.9	0.418702994	0.045429764
5.0	0.465554647	0.046851653
5.1	**0.513022055**	**0.047467407**
5.2	0.560255733	0.047233679
5.3	0.606422753	0.046167020
5.4	0.650764194	0.044341441
5.5	0.692641983	0.041877789
5.6	0.731569462	0.038927479
5.7	0.767223609	0.035654147
5.8	0.799440316	0.032216707
5.9	0.828196633	0.028756317
6.0	0.853584991	0.025388358
6.1	0.875784272	0.022199281
6.2	0.895031631	0.019247359
6.3	0.911597637	0.016566006
6.4	0.925766052	0.014168416
6.5	0.937818564	0.012052511
6.6	0.948024111	0.010205547
6.7	0.956632127	0.008608016
6.8	0.963868855	0.007236728
6.9	0.969935952	0.006067097
7.0	0.975010683	0.005074731
7.1	0.979247140	0.004236458
7.2	0.982778074	0.003530934
7.3	0.985717010	0.002938936
7.4	0.988160456	0.002443446
7.5	0.990190053	0.002029596
7.6	0.991874586	0.001684533
7.7	0.993271822	0.001397236
7.8	0.994430140	0.001158318
7.9	0.995389969	0.000959829
8.0	0.996185029	0.000795060

(continued on next page)

(continued)

*Bold indicates the candidate points to determine the complexity level of proceduralized tasks in a digitalized main control room using the TACOM measure.

References

[1] A.M. Whaley, D.L. Kelly, R.L. Boring, W.J. Galyean, SPAR-H Step-by-step Guidance. INL/EXT-10-18533, Idaho National Laboratory, Idaho Falls & U.S., 2011.

[2] W. Jung, D. I Kang, J.W. Kim, Development of a Standard Method for Human Reliability Analysis (HRA) of Nuclear Power Plants: Level I PSA Full Power Internal HRA, KAERI/TR-2961, Korea Atomic Energy Research Institute, Daejeon, Korea, 2005.

[3] J. Park, W. Jung, A study on the development of a task complexity measure for emergency operating procedures of nuclear power plants, Reliability Engineering and System Safety 92 (2007) 1102–1116.

[4] J. Park, W. Jung, A study on the revision of the TACOM measure, IEEE Transaction on Nuclear Science 54 (2007) 2666–2676.

[5] J. Park, Investigating the TACOM measure as a general tool for quantifying the complexity of procedure guided tasks, Reliability Engineering and System Safety 129 (2014) 66–75.

[6] L. Podofillini, J. Park, V.N. Dang, Measuring the influence of task complexity on human error probability: an empirical evaluation, Nuclear Engineering and Technology 45 (2) (2013) 151–164.

[7] I. Jang, Y. Kim, J. Park, Investigating the effect of task complexity on the occurrence of human errors observed in a nuclear power plant full-scope simulator, Reliability Engineering and System Safety 214 (2021), 107704.

[8] J. Park, The Complexity of Proceduralized Tasks, Springer-Verlag, London, 2009.

[9] S.W. Lee, A.R. Kim, J.S. Ha, P.H. Seong, Development of a qualitative evaluation framework for performance shaping factors (PSFs) in advanced MCR HRA, Annals of Nuclear Energy 38 (2011) 1751–1759.

[10] J.W. Kim, W. Jung, A taxonomy of performance influencing factors for human reliability analysis of emergency tasks, Journal of Loss Prevention in the process industries 16 (2003) 479–795.

[11] A.D. Swain, H.E. Guttmann, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR-1278, U.S. NRC, Washington D.C., 1983.

[12] D.I. Gertman, H.S. Blackman, L.N. Haney, K.S. Seidle, H.A. Hahn, INTENT: a method for estimating human error probabilities for decision based errors, Reliability Engineering and System Safety 35 (1992) 127–136.

[13] B. Kirwan, The development of a nuclear chemical plant human reliability management approach: HRMS and JHEDI, Reliability Engineering and System Safety 56 (1997) 107–133.

[14] E. Lois, V.N. Dang, J. Forester, H. Broberg, S. Massaiu, M. Hildebrandt, P.Ø. Braarud, G. Parry, J. Julius, R. Boring, I. Männistö, A. Bye, International HRA Empirical Study - Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Performance Data, NUREG/IA-2016, vol.1, U.S. NRC, Washington D.C., 2009, Halden, Norway, 2008 (also issued as HWR-844, OECD Halden Reactor Project Work Report.

[15] A. Bye, E. Lois, V.N. Dang, G. Parry, J. Forester, S. Massaiu, R. Boring, P.Ø. Braarud, H. Broberg, J. Julius, I. Männistö, P. Nelson International, HRA Empirical Study - Results from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios, NUREG/IA-2016, vol.2, U.S. NRC, Washington D.C., 2011, Halden, Norway, 2010 (also issued as HWR-915 OECD Halden Reactor Project Work Report.

[16] J. Marble, L. Huafei, M. Presley, J. Forester, A. Bye, V.N. Dang, E. Lois, Results and insights derived from the intra-method comparisons of the US HRA empirical study, in: Proceeding on 11th Probabilistic Safety Assessment and Management / European Safety and Reliability 2012 (PSAM11/ESREL2012). Helsinki, Finland, June 25-29, 2012.

[17] J. Park, Scrutinizing inter-relations between performance influencing factors and the performance of human operators pertaining to the emergency tasks of nuclear power plants – an explanatory study, Annals of Nuclear Energy 38 (2011) 2521–2532.

[18] W. Jung, J. Park, Y. Kim, S.Y. Choi, S. Kim, HuREX – a framework of HRA data collection from simulators in nuclear power plants, Reliability Engineering and System Safety 194 (2020), 106235.

[19] J. Park, Y. Kim, S.Y. Choi, W. Jung, H.E. Kim, Review on the Applicability and Appropriateness of HRA Data Extraction Based on the HuREX Framework, KAERI/TR-8038, Korea Atomic Energy Research Institute, Daejeon, Korea, 2020.
[20] C.R. Shalizi, Advanced data analysis from an elementary point of view, Preprint of book found (2013) at, http://www.stat.cmu.edu/~cshalizi/ADAfaEPoV.
[21] J. Park, H. Kim, I. Jang, Empirical estimation of human error probabilities based on the complexity of proceduralized tasks under an analog environment, Nuclear Engineering and Technology 54 (2022) 2037–2047, https://doi.org/10.1016/j.net.2021.12.025.