

A Human Reliability Analysis Approach Based on the Concepts of Meta-Operation and Task Complexity

Yongping Qiu^{a*}, Dan Pan^b, Zhizhong Li^b, and Peng Liu^b

^aShanghai Nuclear Engineering Research & Design Institute, Shanghai, China

^bDepartment of Industrial Engineering, Tsinghua University, Beijing, China

Abstract: To avoid the difficulties in human error data collection or elicitation and make Human Reliability Analysis (HRA) more generic, a new HRA approach is proposed based on meta-operation identification and task complexity measurement. In this approach, a task is decomposed into meta-operations which is defined as the smallest identifiable operation or activity exhibited in the performance of a task. An eye-tracking experiment was conducted with a conclusion that the meta-operation description model of procedure tasks is reasonable. Then action complexity of these meta-operations is measured. Experiment results indicate that the complexity scores have significant correlation with and are able to explain high percentage of variations of task completion time and error rate. Thus it is possible to establish a generic model for human error probability (HEP) based on task complexity. We intend to provide HEP parametric models rather than HEP tables. For the improvement of task complexity quantification, a survey with 69 licensed NPP operators was conducted to identify complexity factors in main control rooms. Significant factors will be later included into the complexity quantification model. In the proposed method, most PSFs are treated as complexity factors and thus not used as modifiers as in many existing HRA methods.

Keywords: HRA, Meta-operation, Task Complexity, HEP model

1. INTRODUCTION

The existing human reliability analysis (HRA) methods always provide tables of human error probabilities (HEPs). These tables are the critical part of a HRA method. Unfortunately, they mostly come from expert elicitation. For each so called “generic task” (actually domain-specific) or a context situation described by a combination of several performance shaping factors (PSFs), a HEP value has to be elicited. The elicitation process is inherently subjective and in most cases the experts have to make a consensus agreement (guess). The definition of a complete set of generic tasks is not easy, due to the diversity of practical tasks.

Besides data issues that have been well known, another HRA issue relates to the analysis process. The foundation of HRA is task analysis. It not only provides the basic task information, but also influences the logic and process of human error probability calculation. Conversely, the goal of HRA determines the level of detail of task analysis. Traditional HRA methods usually defined their basic task sets. The basic task sets defined in existing HRA methods are usually different. During the process of task analysis, although analysts commonly carry out task decomposition and description based on their experience, but not referring to the existing HEP tables, however, because of the existence of predefined tasks, HRA-oriented task analysis ultimately will be decomposed into predefined tasks or task units.

HRA is also facing the challenges brought by technology advances. Compared to traditional analog Main Control Rooms (MCRs), tasks in a digital MCR is much more cognitive intensive and complex. Task definition and description based on traditional HRA method is too macro and it is not enough to distinguish the difference between various cognitive tasks. Similarly to the existing HRA methods, one may like to propose a classification method of cognitive task, and establish human error

*qiuy@snedi.com.cn

probability table or calculation method for these limited cognitive task categories. Another thinking, which is preferred by the authors of this paper, is to propose a general cognitive behavioral model which can use limited standard cognitive operations to describe the execution process of any cognitive task, and establish human error probability table or calculation method for these limited standard cognitive operations. A specific task is described by a number of standard task elements and the differences between tasks could be fully reflected. The HEP of the task can then be derived from the probability of standard task elements.

By introducing the concept of meta-operation, a new method of task analysis and description for tasks in digital MCRs is proposed. A unified calculation method for meta-operation task probability based on task complexity will then be established. This paper presents the basic idea of the proposed HRA approach and some preliminarily experimental results.

2. A NEW HRA APPROACH

The proposed HRA approach is based on decomposition of a task into meta-operations. For procedure tasks, primary meta-operations are identified. An eye-tracking experiment was conducted with a conclusion that the meta-operation description model of procedure tasks is reasonable. Task complexity measure is then adopted for the quantitative assessment of meta-operations. Lab experiment results indicate that the idea is promising. However, the proposed approach is still in its very preliminary development stage. There is still much work to do in the future.

2.1. Task Analysis Methods

Tasks can be classified into two broad categories: Physical and mental (or cognitive). The most famous method for the decomposition and description of physical tasks is motion study, which was originated in the earlier 20th century by Gilbreth and his wife (Niegel&Freivalds, 2003). The other task analysis methods often consider both physical and cognitive tasks. Currently, there are various task analysis methods used in human factors field for the description (and evaluation) of human-machine or human-human interactions.

The description of a task includes all actions and cognitive activities necessary for the achievement of task goals, and information and facilities involved in the performance process (Kirwan & Ainsworth, 1992). It can provide helpful information for function allocation, operator selection and training, work organization, interface design, and performance prediction.

Unlike the other task analysis methods which focus on observable behaviors, Cognitive Task Analysis (CTA) methods focus on the cognitive process behind the observable behaviors. As Cooke (1994) pointed out, there exist more than 100 CTA methods, and can be classified into three categories: (a) observations and interviews, (b) process tracing, and (c) conceptual techniques. Wei & Salvendy (2004) adds the fourth one: (d) formal models. In this paper, the fourth category is renamed or extended to *task description models*. CTA methods emphasize the micro-level analysis of internal cognitive process and are definitely useful for HRA. However, they are often tedious and time consuming. Thus the level of details has to be carefully determined. Task description models aims to describe the process and logic of task performance. Qualitative or quantitative analysis can then be done based on these models.

One basis of the proposed HRA approach is to establish a task description model with meta-operations. The meta-operation concept actually borrows the idea from Gilbreth's work that 17 generic therbligs (motion elements) are used to describe any motions.

2.2. Task Decomposition and Description with Meta-operations

A meta-operation is defined as the smallest identifiable operation or activity exhibited in the performance of a task (Zhao, 1992). That is, if we try to further decompose a meta-operation, we

cannot obtain operations or actions with specific and complete description. For procedure tasks, seven primary meta-operations have been identified, as listed in Table 1. These meta-operations are commonly performed in the execution of a procedure. There are some other secondary meta-operations in the execution of a procedure, mainly related to interface management and use, such as page switching, zooming, window movement, link and menu selection, list dropping, mouse movement and positioning to popup instructions or tips, reading interface tips and help information, and various confirmations. The secondary meta-operations are mainly introduced by interface design. Except for safety considerations, they should be avoided if possible. Currently we have not defined meta-operations related to teamwork.

By defining the above meta-operations, a formal description of a procedure task can be established. Such a description model is obvious a discrete model. Different from the real-time and parallel controls in a driving task, tasks in MCRs are often not real-time, and can be performed in a serial manner. From the viewpoint of reliability, actions performed parallel by humans are always much lower than being performed in order. Thus even if operators have the ability to perform parallel actions, they are not encouraged to do so. So we more like the proposed model to be a serial model.

Table 1: Primary Meta-operations

Notation	Meta-operation	Resource	Example
RT	Reading procedure text	Visual & Cognitive	Read text on the procedure "Check if primary pump is running"
L	Locating system/equipment status indicator	Visual	Search for primary pump on displays
RV	Reading system/equipment status indicator values	Visual	Read the value of tube pressure or status of a pump
J	Judgment	Cognitive	Judge whether the pressure is higher than 15 MPa
C	Calculation	Cognitive	Calculate the temperature difference between the hot and cold legs.
M	Manual operation	Psychomotor	Open/close a pump
S	Selecting a branch in procedure to go ahead	Cognitive & Psychomotor	Select the "Yes", "No", "Next", or "Go to XXX" branches

When developing a description model for a task, not only primary meta-operations but also secondary ones should be identified since secondary meta-operations provide helpful information for interface evaluation. However, because secondary meta-operations mainly influence task efficiency and user satisfaction (they may also have some influences on attention and situation awareness), if the interface design has conform to human factors requirements, we can often ignore their influence on human reliability. In most existing HRA methods, quality of interface design is normally considered as a Performance Shaping Factor (PSF). So when determining the basic HEP for a task, we can mainly consider the seven primary meta-operations, by which task description can be simplified accordingly. To be brief, if not specified, meta-operations in the rest of this paper refer to primary meta-operations.

Below are some examples of procedure task description using meta-operations. For a simple procedure step, say, "Manually start safety injection", its description could be: RT (read procedure text "Manually start safety injection") →M (perform the start action)→S (select "Next step" if needed). For a typical step, say, "Check if primary pump I is running", the task description could be: RT (read procedure text "Check if primary pump I is running ") →L (locate the status of primary pump I, i.e. find the corresponding parameter in displays)→RV (read the status)→J (make judge whether the condition is satisfied)→S (select a procedure branch according to the judgment). Some procedure steps involve more than one parameter, then L and RV meta-operations will be repeated, and C (calculation) meta-operation is often needed. For example, "Check if temperature difference between cold and hot legs of RCS less than 10°C?", the description would be: RT→L (locate temperature of the cold leg)→RV (read temperature of the cold leg) →L (locate temperature of the hot leg)→RV (read

temperature of the hot leg) → C (calculate the temperature difference) → J (make judgment whether the difference is less than 10°C) → S.

It can be seen that the above task description model based on meta-operations is simpler than the existing formal models, very convenient for practice, and more importantly, generic for the task description of various procedures in digital MCRs. If quantitative assessment models can be developed for each meta-operation, we will be able to innovate a new HRA method suitable for digital MCRs. We have to point out that the above model is applicable for diagnosis tasks without procedure aids. Knowledge-based diagnosis process is mainly internal mental process that could not be observed from outside. It also varies largely among operators. Thus it is very difficult to create a formal model for such a task. Reliability assessment for such diagnosis tasks has been a tough HRA problem for very long time. The current practice is to simplify the assessment. Fortunately, procedures are prepared for "standard accidents". Even if "non-standard" accidents happen, operators are not required to determine the failures, but use prepared procedures to transfer the system to safe conditions first.

2.3. Task Complexity Measurement

Task complexity is acknowledged to be an important PSF in HRA. Several researchers tried to understand the effects of task complexity on operator/crew performance in NPP MCRs and to integrate task complexity into HRA. Researchers (Braarud, 2000; Braarud et al., 2004; Laumann et al., 2004; Laumann et al., 2005) in the OECD Halden Reactor Project conducted several studies to characterize task complexity that making MCR task difficult and to explore how specific complexity factors influence crew performance. Recently, several researchers (Liu et al., 2012; Podofillini et al., 2013; Liu and Li, 2014) tried to quantitatively measure task complexity of emergency operating procedures and build the empirical relationship between task complexity and human error probability (HEP). Gertman (2012) proposed a rating scale to more clearly allocate levels of complexity for HRA. These studies highlight the strong interest of HRA practitioner on complexity.

That how to integrate task complexity into HRA, however, is still a big challenge. Several critical issues should be touched. First, task complexity is a rich concept with several dimensions (Liu and Li, 2012). There are a variety of complexity factors in MCRs (Sasangohar and Cummings, 2010). They are possible PSFs. What complexity factors should be considered in HRA? The number of complexity factors should be parsimonious in HRA, as suggested by Galyean (2006). Second, how to prioritize various complexity factors? This question is associated with the first one. While applying HRA, highly salient complexity factors or factors that have high effect on operator performance should be given high priority. Third, how do we quantify complexity factors and what properties of complexity factors should be quantified? These two questions are critically important for building the quantitative relationship between complexity and HEP. They are also related to the first and second questions. Fourth, do we just need to consider complexity as a regular PSF in HRA or can we predict HEP from the complexity perspective?

In Liu et al. (2012), a task (action) complexity measure is proposed based on meta-operations. Task complexity is operationalized as the amount of required resources in Human Information Processing (HIP) activities. The equation for calculating the task complexity (TC) score of a step is written as follows:

$$TC = \sum_{ns=1}^{NS} \sum_{no=1}^{NO} \sum_{nr=1}^{NR} C_{ns,no,nr} \quad (1)$$

where

NS: the number of sub-steps in the execution sequence with maximum required resources in a step;

NO: the number of meta-operations in a sub-step;

NR: the number of types of resources used;

$C_{ns,no,nr}$: the demand level (or complexity score) of resource nr of the meta-operation no in the sub-step ns . The score is determined based on the VACP (Visual, Auditory, Cognitive, and Psychomotor) scales updated by Aldrich et al. (1989). The resource demanded by the meta-operations are given in Table 1.

By using the task complexity measure, quantitative assessment of meta-operations and procedure task is realized.

Because our task description model is a serial model, similarly to the serial model in classic reliability engineering, the failure of one meta-operation will lead to failure of the whole task. Suppose a task consists of two meta-operations, namely A with error probability of $P(A)$ and B with error probability of $P(B)$, then the error probability of the task is $P(A \cup B) = P(A) + P(B) - P(AB) \approx P(A) + P(B)$, since the probability that two meta-operations both fails is much lower than that of the failure of only one meta-operation. This is the reason that we can use a cumulative form in Equation (1). Such a conservative calculation is acceptable in HRA.

3. SOME EXPERIMENTAL SUPPORTS

3.1. Examination of Task Description Model

This experiment has two goals. One is laboratory observation to confirm whether the execution process of a procedure task is consistent with its task description model based on meta-operations. The other is to preliminarily examine whether the VACP scale is suitable for assessing meta-operations quantitatively according to objective behavior indices.

3.1.1. Operating Procedure

Based on the existing emergency operating procedures (EOPs) of a NPP, Wang (2012) made up a dummy procedure with 15 steps each with a different complexity score (Table 2). This EOP was used in this experiment. TC score is the independent variable for later data analysis.

3.1.2. Dependent Variables

Dependent variables include subjective assessment and objective eye movement criteria. Subjective assessment was done by observing the execution process of the procedure task during the experiment and after experiment by watching videos. Objective eye movement criteria were recorded by eye tracking system during the experiment. These criteria included:

- Average blink duration
- Average blink times (counts)
- Average fixation
- Average fixation times (counts)
- Average saccade duration
- Average saccade times (counts)

These criteria were calculated for each step in the procedure.

3.1.3. Participants

A total of 15 male university students aged from 18 to 25 (mean: 22.0; SD: 2.0) with engineering education background at Tsinghua University were recruited as the participants. They were right handedness, had more than 5 years' computer experience, and had a normal vision ($< 300^\circ$ after correction). They didn't know much about NPPs, digital control systems, and digital EOPs, so that they would not have any doubts at the dummy procedure.

3.1.4. Experiment Platform

The experiment platform included a digital parameter monitoring display and another display for the computerized dummy procedure.

In this experiment, the participants wore the eye tracker iView XTM HED to record the participants' eye movements. Then, the eye movement data was exported from iViewX program by BeGaze program. The data could be used to analyze the fixation point movement, fixation time and the residence time of region of interest. The data acquisition frequency is 50 Hz.

During the experiment, the participants were asked to perform the dummy procedure accurately and timely.

3.1.5. Experiment Procedure

Before the experiment, every participant read the experiment introductions, filled an informed consent form and personal information form, and answered a knowledge background questionnaire.

The experiment consisted of two stages: the practice stage and the formal experimental stage. In the first stage, each participant performed three trials (each trial has 15 steps) under the experimenter's instruction. In this stage, if an error occurred or the time was out, the participant could continue to finish the EOP. In the second stage, the participant should finish more than 20 trials without the experimenter's help and the last three trials should be successful. In this stage, if an error occurred or the time was out, the participant had to restart another trial.

3.1.6. Results

The experimental observation didn't find out any other primary meta-operations which should be added into the proposed model. The execution process of the procedure task was consistent with the description based on meta-operations. However, it was found that *Reading Value* (RV) and Judgment (J) were often difficult to be separated from each other. Because both of them involved the internal cognitive process and *Judgment* always followed *Reading Value* immediately, it was difficult to determine the finish time of *Reading Value* and the start time of *Judgment*. Similarly, *Judgment* (J) followed *Calculation* (C). In spite of this, the meta-operations proposed in this study were based on expert knowledge of procedure execution process, and indeed existed in the task performance process, but merely they were difficult to be segmented in time. Because of the difficulty in time segmentation, it was unable to conduct quantitative validation for each meta-operation, but the validity can be inferred through analysis on the combination of two or more meta-operations, such as at the procedure step level.

In the formal experiment stage, the participants were required to finish at least 20 trials and succeed three consecutive trials. For most participants, less than 30 trials were performed before they reach the termination criteria, except the No. 6 (65 trials), No. 9(43 trials), No. 11(57 trials), and No. 13(42 trials) participants. These four participants were treated as outliers in later analysis on the eye movement data of the last 10 trials for each participant as given in Table 2.

Table 2: Task Complexity Scores and Eye Movement Data

Step No.	TC Score	Average Blink Duration (s)	Average Blink Times	Average Fixation Duration (s)	Average Fixation Times	Average Saccade Duration (s)	Average Saccade Times
1	35.2	0.11	0.68	1.67	4.35	0.27	4.54
2	47.7	0.20	1.31	3.67	9.95	0.60	10.43
3	65	0.36	1.86	4.31	11.75	0.66	12.39
4	75	0.35	2.23	5.32	14.22	0.90	15.23
5	89.8	0.42	2.61	5.32	15.14	0.91	16.10
6	92.1	0.40	2.42	5.30	15.24	0.95	16.18
7	102	0.25	1.61	3.29	9.06	0.62	9.68
8	108.5	0.44	2.53	5.99	16.49	1.02	17.20
9	119.4	0.37	2.25	5.11	13.64	0.86	14.60
10	129.9	0.50	2.93	5.85	16.40	1.05	17.42
11	134.7	0.45	2.49	6.05	15.99	0.98	17.04
12	142.9	0.55	3.17	7.06	19.71	1.33	21.19
13	154.8	0.67	4.17	9.06	24.48	1.45	26.19
14	155.3	0.58	3.49	7.70	20.45	1.33	21.65
15	187.2	0.65	3.86	9.01	24.38	1.51	25.65

As shown in Figures 1 & 2, strong linear relationships between average blink duration and TC score, as well as between average blink times and TC score can be seen (Blink Duration=0.003*TC+0.054, $R^2=0.822$; Blink Times=0.02*TC+0.35, $R^2=0.812$). This implies that TC score can explain 82.2% variance of eye blink duration and 81.2% variance of eye blink times. The ANOVA suggests that the regression model between eye blink duration/times and TC score is significant ($F(1,13)=60.216$, $p<0.001$).

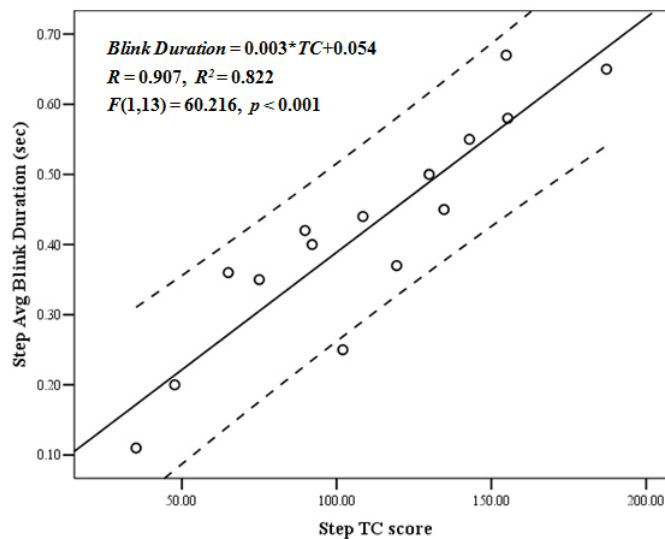


Figure 1: Relationship between TC Score and Average Blink Duration

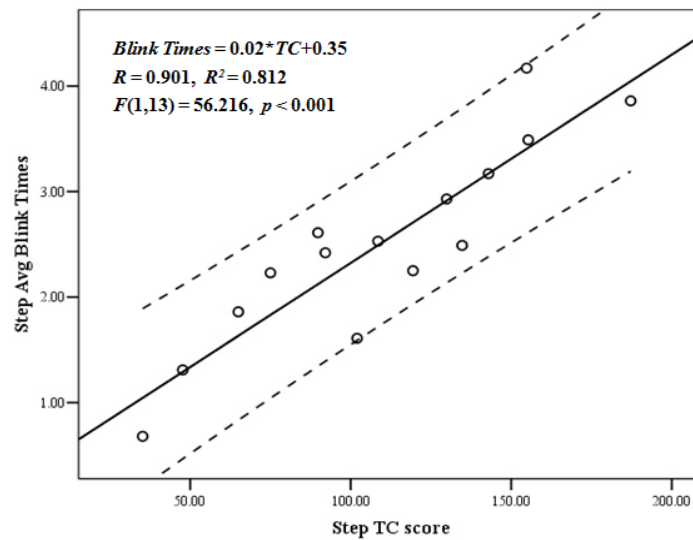


Figure 2: Relationship between TC Score and Average Blink Times

Figures 3&4 show the strong linear relationship between average fixation duration and TC score, as well as between average fixation times and TC score (Fixation Duration=0.042*TC+1.037, $R^2=0.800$; Fixation Times =0.113*TC+3.013, $R^2=0.789$), meaning that TC score can explain 80.0% variance of eye fixation duration and 78.9% variance of eye fixation times. The ANOVA suggests that the regression model between eye fixation duration/times and TC score is significant ($F(1,13) = 51.953, p < 0.001$; $F(1,13)=48.623, p < 0.001$).

From Figures 5 & 6, it can be seen that there is a strong linear relationship between average saccade duration and TC score, as well as between average saccade times and TC score (Saccade duration=0.007*TC+0.163, $R^2=0.832, F(1,13) = 64.282, p < 0.001$; Saccade Times =0.121*TC+3.145, $R^2=0.79, F(1,13)=48.868, p < 0.001$). The variances explained by TC score are also high and significant.

The above results mean that if the VACP scales are used to quantitatively assess procedure steps consisting of meta-operations, the complexity score of a task is a good predictor for eye movement indices.

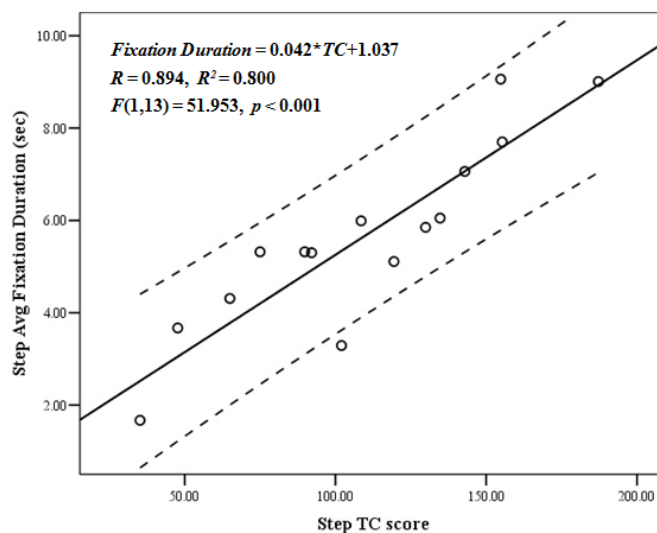


Figure 3: Relationship between TC Score and Average Fixation Duration

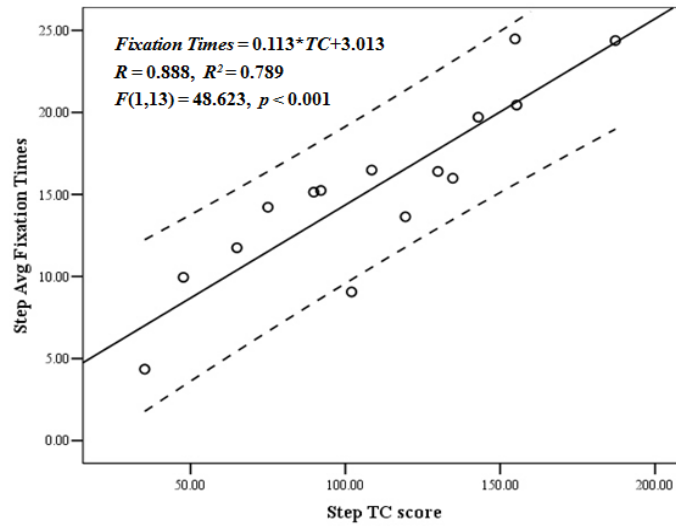


Figure 4: Relationship between TC Score and Average Fixation Times

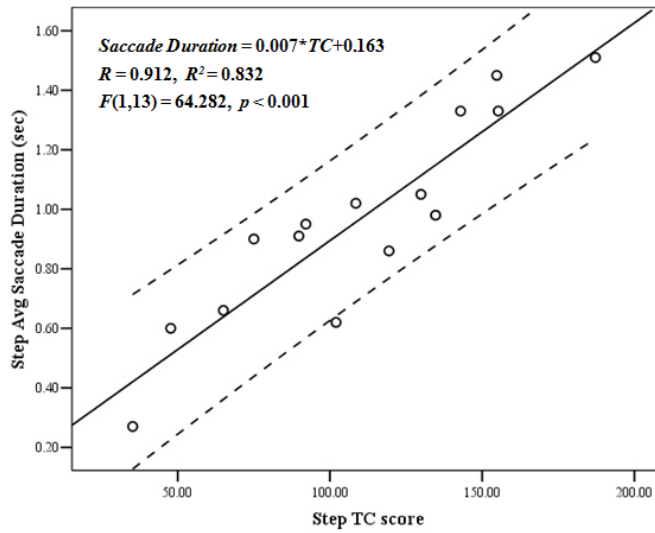


Figure 5: Relationship between TC Score and Average Saccade Duration

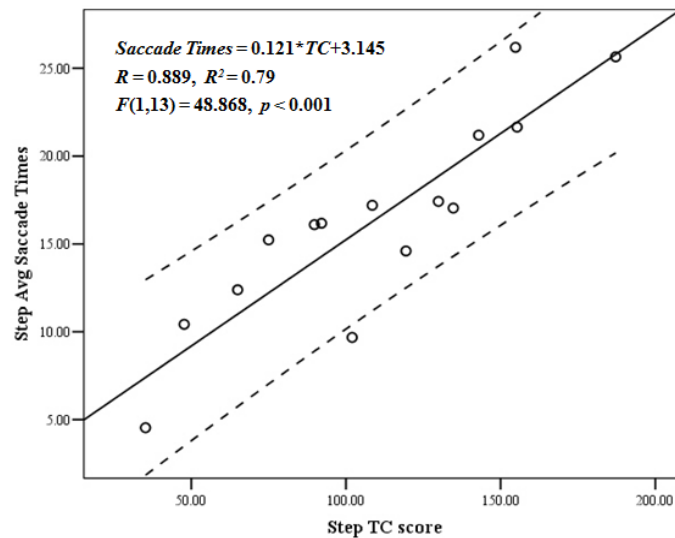


Figure 6: Relationship between TC Score and Average Saccade Times

3.2. Examination of Task Complexity Model

According to Liu et al. (2012)'s experiment results, the complexity scores have significant correlation with and are able to explain high percentage of variations of completion time and HEP of procedure tasks. Thus it could be possible to establish a generic HEP model based on task complexity. We will try to provide HEP parametric models for meta-operations rather than HEP tables. Although the proposed method originally aimed to be capable for HRA analysis for tasks in digital MCRs of NPPs, we can see that the concept of meta-operation and task complexity measurement can be very generic. Thus we can expect its broad applications.

The above task complexity measure mainly considers the dimension of action complexity. For extending task complexity quantification to other complexity dimensions, a survey with 69 licensed NPP operators was conducted to identify complexity factors in main control rooms. Significant factors will be later included into the complexity quantification model. In the proposed method, PSFs are treated as complexity factors and thus not used as modifiers as in many existing HRA methods. Another point distinguishing from other HRA methods is that the proposed method will introduce a concept of task parameter (such as the length of procedure text) into the quantification models, making HEPs adaptive to difference of tasks belong to the same type. All these following studies will be reported in the future.

4. CONCLUSION

This study proposed a description model for procedure tasks based on meta-operations. The subjective observation and objective eye movement data collect from the procedure task experiment seem to support that the task description model based on meta-operations is appropriate. The task complexity validation shows that quantitative assessment of meta-operations is possible and promising. However, this new approach is still at its earlier development stage and much work needs to be done in the future.

Acknowledgements

This study was supported by the National Natural Science Foundation of China (Project No. 70931003).

References

- [1] B.Niebeland A.Freivalds. "*Methods, standards and work design*," Tsinghua University Press, 2003, Beijing.
- [2] B. Kirwan and L. K. Ainsworth. "*A guide to task analysis*," Taylor & Francis, 1992, London.
- [3] N. J. Cooke. "*Varieties of knowledge elicitation techniques*," International Journal of Human-Computer Studies, 41, pp.801-849, (1994).
- [4] J. Wei and G. Salvendy. "*The cognitive task analysis methods for job and task design: review and reappraisal*," Behaviour and Information Technology, 23, pp.273-299, (2004).
- [5] B. Zhao. "*A Structured analysis and quantitative measurement of task complexity in human-computer interaction*," Purdue University, 1992, West Lafayette.
- [6] P.Ø.Braarud. "*Subjective task complexity in the control room (HWR-621)*," OECD Halden Reactor Project, 2000, Halden.
- [7] P.Ø.Braarud, K. Laumann and H. Svengren. "*Task complexity as a factor for human reliability assessment*," Proceedings of PSAM7 & ESREL2004, 2004, Berlin.
- [8] K. Laumann, H. Svengren, P.Ø.Braarud, J. T. Ludvigsen and C. Smith. "*Effects of task complexity on operators' performance in eight medium loss of coolant accident scenarios (recovery experiment 2002/03) (HWR-757)*," OECD Halden Reactor Project, 2004, Halden.
- [9] K. Laumann, P.Ø.Braarud and H. Svengren. "*The task complexity experiment 2003/2004 (HWR-758)*," OECD Halden Reactor Project, 2005, Halden.
- [10] P. Liu, Z. Z. Li and Z. Wang. "*Task complexity measure for emergency operating procedures based on resource requirements in human information processing*," Proceedings of PSAM11&ESREL2012, 2012, Helsinki.
- [11] L. Podofillini, J. Park and V.N.Dang. "*Measuring the influence of task complexity on human error probability: an empirical evaluation*," Nuclear Engineering and Design, 45, pp. 151-164, (2013).
- [12] P. Liu and Z. Z. Li. "*Comparison of task complexity measures for emergency operating procedures: convergent validity and predictive validity*," Reliability Engineering and System Safety 121, pp. 289-293, (2014).
- [13] D. I. Gertman. "*Complexity: application to human performance modeling and HRA for dynamic environments*," Proceedings of the 5th International Symposium on Resilient Control Systems (ISRCs), 2012, Salt Lake City.
- [14] P. Liu and Z. Z. Li. "*Task complexity: a review and conceptualization framework*," International Journal of Industrial Ergonomics, 42, pp.553-568, (2012).
- [15] F. Sasangohar and M. L. Cummings. "*Human-system interface complexity and opacity part ii: methods and tools to assess HSI complexity (HAL2010-03)*," Massachusetts Institute of Technology, 2010, Cambridge.
- [16] W. J. Galyean. "*Orthogonal PSF taxonomy for human reliability analysis*," Proceedings of PSAM8, 2006, New Orleans.
- [17] T. B. Aldrich, S. M. Szabo and C. R. Bierbaum. "*The development and application of models of predict operator workload during system design*," In G. R. MaMillan, D. Beevis, E. Salas, M. H. Strub, R. Sutton, and L. van Breda (Eds.), Applications of human performance models to system design, pp. 65-80, 1989, New York.
- [18] Z. Wang. "*Effects of task complexity and time pressure on operation performance of computerized procedures*," Tsinghua University, 2012.