

Comparing TACOM scores with subjective workload scores measured by NASA-TLX technique

Jinkyun Park* and Wondea Jung

*kshpjk@kaeri.re.kr

Integrated Safety Assessment Division

Korea Atomic Energy Research Institute

P.O.Box 105, Duckjin-Dong, Yusong-Ku, Taejeon, 305-600, Korea

1. Introduction

It is a well-known fact that a large portion of human performance related problems was attributed to the complexity of tasks. Therefore, managing the complexity of tasks is a prerequisite for safety-critical systems such as nuclear power plants (NPPs), because the consequence of a degraded human performance could be more severe than in other systems [1]. From this concern, it is necessary to quantify the complexity of emergency tasks that are stipulated in procedures, because most tasks of NPPs have been specified in the form of procedures [2].

For this reason, Park et al. developed a task complexity measure called TACOM [3]. In this study, in order to confirm the validity of the TACOM measure, subjective workload scores that were measured by the NASA-TLX technique were compared with the associated TACOM scores. To do this, 23 emergency tasks of the reference NPPs were selected, and then subjective workload scores for these emergency tasks were quantified by 18 operators who had a sufficient knowledge about emergency operations.

2. Background

The TACOM measure consists of five sub-measures that characterize five kinds of remarkable factors making the performance of procedural steps complicated. More detailed explanations about the quantification of TACOM scores can be found in Ref. [3].

The validity of the TACOM measure was preliminarily examined by comparing the estimated TACOM scores with task performance time data that were extracted from OPERA database. OPERA database has been developed by KAERI, and it contains many kinds of plant- and domain-specific operator performance data obtained under simulated emergencies [4]. As a result, it was observed that TACOM scores were strongly correlated with averaged task performance time data that denote an elapsed time from the commencement of a given task to the accomplishment of it [3]. This result was insightful because it provided crucial evidence for supporting the fact that the TACOM measure can properly quantify the complexity of emergency tasks.

3. Rating subjective workload scores

As explained in the previous section, it is expected that the TACOM measure is serviceable for quantifying the complexity of emergency tasks. However, in order to confirm this expectation, an additional validation is indispensable because the response time (i.e., the task performance time) is one of the representative measures, which can reflect a facet of human performance [5]. In this regard, it is desirable that a subjective workload could be used as reference information, because the subjective workload is susceptible to the level of cognitive demands [6]. From this perspective, subjective workload scores were collected through the cooperation of the reference NPPs.

3.1. Subject

18 SROs who are working in the MCR of the reference NPPs were selected to quantify the subjective workload that they felt under simulated emergencies, since they had sufficient experience on emergency tasks.

3.2 The NASA-TLX technique

Of many subjective evaluation techniques, the NASA-TLX technique is selected to measure the level of a workload perceived by SROs, since this is known as one of the most suitable techniques for evaluating an experienced workload [7]. The NASA-TLX technique consists of six dimensions, such as a mental demand, physical demand, temporal demand, performance, effort and frustration [8], and the overall workload score can be quantified by the average of subjective ratings on them [7]. Therefore, based on their previous experience from the periodic training sessions, SROs were asked to rate six dimensions by using an arbitrary scale from 0 to 100.

3.3. Collecting subjective workload scores

In total 23 emergency tasks that were extracted from six kinds of EOPs were selected as the tasks to be subjectively evaluated. On the basis of the selected emergency tasks, eight tasks were randomly assigned to each SRO with written descriptions about the tasks.

4. Result

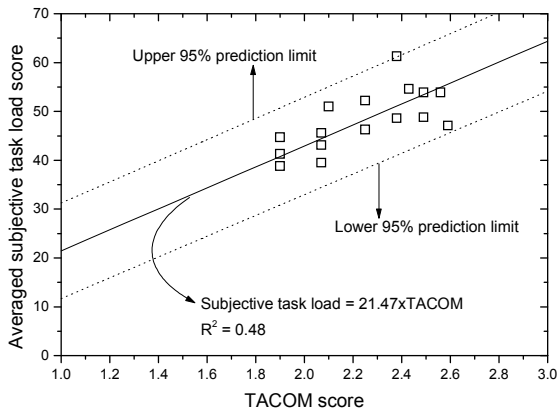
Subjective workload scores for 23 emergency tasks were successfully secured from 18 SROs. However, before comparing them with the associated TACOM scores, it is indispensable to check the reliability of subjective workload scores. To this end, the consistency (or the agreement) of subjective workload scores was

investigated by using the intra-class correlation (ICC) coefficient [9].

As a result, subjective workload scores for 16 emergency tasks were determined as reliable data. Therefore, TACOM scores for 16 emergency tasks were quantified. Table 1 summarizes averaged subjective workload scores for 16 emergency tasks with the associated TACOM scores. In addition, Fig. 1 shows the result of a regression analysis as well as ANOVA table by using the data shown in Table 1.

Table 1. Averaged workload scores with the associated TACOM scores

Emergency task ID	Averaged NASA-TLX scores	TACOM scores
2	41.3	1.90
3	44.7	1.90
4	45.6	2.07
5	46.3	2.25
6	38.8	1.90
7	53.9	2.56
8	52.2	2.25
10	54.6	2.43
12	43.1	2.07
13	48.6	2.38
14	53.9	2.49
16	39.5	2.07
17	47.1	2.59
18	48.8	2.49
22	61.3	2.38
23	51.0	2.10



Source of variance	DF	SS	MS	F	p-value
Treatment	1	250.56	250.56	12.01	< 0.01
Residual	15	312.83	20.86		
Total	16	563.39			

Figure 1. The result of comparisons between averaged subjective workload scores and the associated TACOM scores

Fig. 1 indicates that there is a significant correlation between averaged subjective workload scores and TACOM scores ($R^2 = 0.48$). In addition, the p-value indicates that the changes of subjective workload scores are attributable to the changes of TACOM scores. Therefore, it is reasonable to expect that TACOM measures are meaningful in explaining workload scores perceived by SROs.

5. Discussion and conclusion

In this study, the appropriateness of the TACOM measure is scrutinized by comparing subjective workload scores that were measured by the NASA-TLX technique with the associated TACOM scores. As a result, it was observed that TACOM scores are significantly correlated with subjective workload scores that are representative of cognitive demands placed on human operators. This observation strongly supports the notion that the TACOM measure could be useful for quantifying the complexity of emergency tasks. Therefore, although additional activities may be indispensable for a more decisive conclusion, it is expected that the TACOM measure could be used to quantify the complexity of emergency tasks.

Acknowledgement

This research was supported by ‘The Mid- and Long Term Nuclear R&D Program’ of Ministry of Science and Technology (MOST), Korea. The authors would like to express deep appreciation to all the training instructors as well as the SROs of the reference NPPs for their sincere support.

Reference

- [1] E. M. Roth, R. J. Mumaw, and P. M. Lewis, “An empirical investigation of operator performance in cognitively demanding simulated emergencies,” U.S. Nuclear Regulatory Commission, NUREG/CR-6208, Washington D.C., 1994.
- [2] Y. S. Bang, K. W. Seul, and H. J. Min, “Evaluation of total loss of feedwater accident – Recovery phase and investigation of the associated EOP,” Journal of the Korean Nuclear Society, vol. 25, no. 1, p. 37-50, 1993
- [3] J. K. Park, and W. D. Jung, “A study on the development of a task complexity measure for emergency operating procedures of nuclear power plants,” Reliability Engineering and System Safety, to be published.
- [4] J. K. Park, and W. D. Jung, “A database for human performance under simulated emergencies of nuclear power plants,” Nuclear Engineering and Technology, vol. 37, no. 5, pp. 491-502, 2005.
- [5] R. L. Henneman, and W. B. Rouse, “Measures of human problem solving performance in fault diagnosis tasks,” IEEE Transactions on Systems, Man and Cybernetics, SMC-14, no. 1, pp. 99-112, 1982.
- [6] D. J. Campbell, “Task complexity: A review and analysis,” Academy of Management Review, vol. 13, no. 1, pp. 40-52, 1988.
- [7] T. E. Nygren, “Psychometric properties of subjective workload measurement techniques: Implications for their use in the assessment of perceived mental workload,” Human Factors, vol. 33, no. 1, pp. 17-33, 1991.
- [8] S. G. Hart, and L. E. Staveland, “Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research,” in Human Mental Workload, P. A. Hancock and N. Meshkati, Ed. North-Holland: Elsevier Science Publisher B. V., 1988, p. 139-183.
- [9] J. J. Bartko, “The intraclass correlation coefficient as a measure of reliability,” Psychological Reports, vol. 19, pp. 3-11, 1966.