Transactions of the Korean Nuclear Society Spring Meeting
Jeju, Korea, May 22-23, 2025

An Approach to Effective Pairwise Comparison of NASA-TLX for the Human Factors Experiment on Nuclear Power Plant Operators

Hyun-Chul Lee*
Korea Atomic Energy Research Institute, Advanced I&C Research Division, 111 deadeok-daero 989 beon-gil, Yuseong-gu, Daejon, South Korea 34057
**Corresponding author: leehc@kaeri.re.kr*

***Keywords :** pairwise comparison, NASA-TLX, workload measurement, operator, human factors experiment

1. Introduction

In human factors researches, subjective measurement methods are widely used because they are simple to use and have the advantage of being less time-consuming for measurement and analysis than objective measurement methods. On the other hand, since subjective measurement methods are relatively vulnerable to bias or contamination of measured data, careful preparation is required for data collection and analysis. Considering these advantages and disadvantages, it can be assumed that there are not many cases in which a single measurement method is continuously used frequently over a long period of time, and NASA-TLX (NASA Task Load Index) exceeds these usual expectations.

In the nuclear domain, NASA-TLX is also often used to measure the operator's workload. However, since the weighting process of NASA-TLX is not easy and cumbersome to apply, data from other experiments may be used or omitted. Using data from other experiments is only acceptable if both experimental setups, such as participants, tasks/context and environment, are (almost) same. There has been no thorough verification of skipping weighting for the nuclear domain, so it must be decided carefully. In this paper, an approach to reduce the burden of weighting process is proposed.

2. Methods and Results

2,1 NASA-TLX

NASA-TLX is a multi-dimensional scale for estimating workload, created in 1996 by Sandra Hart of NASA Ames Research Center and Lowell Staveland of San Jose State University.

NASA-TLX is used according to a procedure that consists of two steps: weighting and rating. NASA-TLX estimates the workload by subjective scoring (referred to as rating) on the following six subscales/elements: Mental Demand, Physical Demand, Temporal Demand, Performance, Effort, and Frustration.

Since each element may have different importance depending on the job, the weighting is used to compensate it. Determining the weight also depends on the subjective decision from the subject. The weight should be determined for each group of tasks that require the subject to perform in the experiment. In the experiment, the rating must be performed first and then the weighting is next. For the same experimental task group, one weighting may be performed and applied for all experimental tasks of the group.

The weighting is completed through pairwise comparison of six elements. Relative importance assessments are made for all 15 pairs that can be made from six elements. The highest score of weighting does not exceed 5 because for each comparison an element that is determined to be more important gets 1 point and another element gets no point.

Since NASA-TLX is based on a post-experiment self-rating technique, it is conducted at the end of each experimental run. The rating uses a Likert scale with bipolar descriptions such as low/high. The total 21 vertical ticks on each subscale divide the scale from 0 to 100 in increments of 5.

The total workload is defined as the weight and rating of each element multiplied and divided by 15.

2.2 Demanding Weighting Process

Pairwise comparisons (weighting) are challenging. Subjects should perform importance assessments on all 15 comparison pairs for one experimental task group.

In the case of ergonomic experiments using NASA-TLX to find out whether the design improvement of the main control room HSI (Human System Interface) of a nuclear power plant is effective in reducing the operator's workload in both abnormal and emergency situations, the pairwise comparison may have to be made twice for abnormal and emergency situations. This is because if the spatial and contextual situations follow and the operator's task characteristics differ, a new weight profile should be defined.

It is necessary to be cautious about omitting the pairwise comparison. Comparative studies related to RTLX (Raw Task Load Index) which eliminate the weighting process of NASA-TLX have not shown consistent results. In addition, there are comparative studies on vehicle drivers, but it is difficult to find a comparative study on nuclear power plant operators. Thus, it is difficult to give a lot of trust in the use of RTLX in the nuclear field.

Therefore, before securing reliable research results on RTLX in the nuclear field, the use of RTLX needs to be cautious in experiments targeting nuclear power plant operators.

2.3 Approach to Get Weighting Profile

The weighting should be performed for each task group and for each subject, respectively. This requirement can put an unnecessary burden on the subject during the experiment. In addition, it is difficult to say that the burden on the experimenter to collect weight data from all subjects is small. During the experiment, all influencing factors must be well controlled and weight determination can sometimes make this difficult. As an approach to solve this problem, author would like to propose determining the weight profile from a pilot experiment.

In the human factors experiment a preliminary experiment or pilot test is carried out usually before the main experiment is conducted. This is to secure the feasibility of the data and results that can be obtained from the experiment, and is an indispensable and important step in predicting the value of the experimentation.

In general, the pilot experiment is performed on the representative tasks of the task group. In this pilot experiment, author would like to propose a method of creating a weight profile for the experiment.

Although a small number of subjects are attended the pilot experiment, it can be proposed that the weighting results and expert opinions are combined to create a weighting profile for each task group. This profile is viewed as a representative value of the weights of all subjects for one task group. The background of this approach is the assumption that the characteristics of the subjects are very consistent from the workload point of view. In the case of nuclear power plant operators, it can be considered that they have secured considerable proficiency and possess consistent and equivalent capabilities through periodic education and training. Therefore, this approach should be applied to situations in which all subjects are expected to suffer from the equivalent level of workload under a given experimental environment through the same level of education and training.

The weight profile collected in this way can be used as data to determine the weight in future similar experiments. From a long-term perspective, these weighting profile data will make it possible to determine representative workload weights for tasks. Importance information can be also usefully used even when analyzing by subscale.

3. Conclusions

In this paper, author argues the following for the pairwise comparison of NASA-TLX.

- RTLX should be used in a limited way before securing sufficient sensitivity analysis and reliability in the nuclear field.
- Under a specific experimental environment (equivalent capability between subjects, working context and environment), creating a weight profile that can represent all subjects in the preliminary experimental stage and using it consistently in determining the workload of all subjects is expected to reduce the burden on the subject and the experimenter.

REFERENCES

[1] NASA Task Load Index (TLX) v. 1.0 Manual, NASA, 1986.
[2] S. G. Hart, L. E. Staveland, Development of NASA-TLX (Task Load Index): Results of Empirical and Theoretical Research, In Peter A. Hancock, Najmedin Meshkati, Human Mental Workload. Advances in Psychology, Vol. 52. Amsterdam: North Holland. pp. 139–183.
[3] S. G. Hart, NASA-Task Load Index (NASA-TLX); 20 Years Later, Proceedings of the Human Factors and Ergonomics Society Annual Meeting. 50 (9): 904–908. 2006. San Francisco, CA.