A Study on Definitions of Nine Dimensions for Human Performance Measures with Their Features

Hae Bin Jang [a], Jung Hwan Kim [a], Jinkyun Park [a*],
[a]*Korea Atomic Energy Research Institute, 111 Daedeok-daero 989 Beon-gil, Yuseong-gu, Daejeon, Republic of Korea*
[*]*Corresponding author: kshpjk@kaeri.re.kr*

1. Introduction

The U.S. Nuclear Regulatory Commission (NRC) defines integrated system validations (ISV) as an evaluation of an integrated system design using performance-based tests to determine whether it meets performance criteria and supports safe operations [1]. Various human performance measures are used for ISV evaluation. In the NUREG-6393, the NRC suggests nine dimensions of Construct Validity, Diagnosticity, Impartiality, Intrusiveness, Objectivity, Reliability, Resolution, Sensitivity, and Simplicity to evaluate human performance measures [2]. However, there is a lack of a clear definition of these nine dimensions. Without a clear definition of these dimensions, it is not possible to use human performance measures for ISV evaluation. Therefore, the objective of this study is to propose a definition of these dimensions and to classify human performance measures widely used in nuclear industry based on the proposed definition.

2. Methods and Results

This section describes the process of defining nine dimensions and examples of the classified human performance measures based on the nine dimensions.

2.1 Definition of nine dimensions

To define nine dimensions suggested by the NRC, this study reviewed various literature reviews from safety-critical industries such as nuclear, aviation, and railway. This study analyzed previous studies using human performance measures, measurement characteristics, Workload, Situation awareness, Teamwork, Construct Validity, Diagnosticity, Impartiality, Intrusiveness, Objectivity, Reliability, Resolution, Sensitivity, and Simplicity as keywords. Based on the literature review, about one hundred seventy studies were selected and used to define nine dimensions [3-8]. Table 1 shows the proposed definition of nine dimensions.

Table 1: Proposed definition of nine dimensions

Dimension	Definition
Diagnosticity	Measurement or metric should provide Information that can be used to identify the cause of acceptable (or unacceptable) human performance.
Sensitivity	Measurement or metric should detect the changes in human performance.
Construct Validity	Measurement or metric should represent what it claims in terms of human performance.
Intrusiveness	Data collection should not intervene participants in terms of physical and psychological aspects.
Reliability	Human performance should be similar when it repeatedly measured in an identical condition.
Objectivity	Human performance should be measured by objective information instead of subjective information.
Simplicity	Straightforward and simple measurement or metric should be used to make sure its applicability.
Resolution	Measurement or metric should reflect human performance at an appropriate level of resolution to ensure sufficient details to permit a meaningful analysis.
Impartiality	Measurement or metric should be equally capable of reflecting good as well as bad performance.

In addition, this study suggested specific features that can be used to evaluate human performance measures. As shown in Table 2, the user answers with yes/no to each dimension, and they are collected and used.

Table 2: Suggested features of nine dimensions

Dimension	Features
Diagnosticity	1. Should information related to human performance assessment be directly observable or indirectly collectible, and can be provided in value form? 2. Is it possible to provide reference information to distinguish (1) good/bad, (2) high/low, or (3) sufficient/insufficient values for the observed values?
Sensitivity	3. Is the range(scale) of measurement/metric for human performance measurement a clear way?
Construct Validity	4. Is the technical basis for human performance to be evaluated clear?
Intrusiveness	5. Is it a method to measure or evaluate human performance without physical or psychological interference?
Reliability	6. Is it a method to repeatedly measure human performance with similar values?
Objectivity	7. Is it a method of measuring human performance with direct observable physical quantity? 8. When measuring human performance through expert opinion or observation, are standards or guidelines provided to minimize subjectivity?
Simplicity	9. Is it possible to measure human performance without additional training and training, assuming that the public is an appraiser? 10. Is human performance measurement results intuitive and easy to understand without expertise or skills, assuming that the public is an appraiser?
Resolution	Properly covered by combining the features belonging to Diagnosticity, Sensitivity, Construct Validity, and Intrusiveness.
Impartiality	Properly covered by combining the features belonging to Reliability, Objectivity, and Simplicity.

From Table 2, it should be noted that the features of Resolution and Impartiality seem to be largely overlapped with those of other dimensions. For example, it is possible to say that a certain human performance measure has a sufficient Resolution if it can evaluate the performance of human operators without any intervenes, which is not only representative but also very sensitive in terms of providing diagnostic information. Similarly, it is expected that the satisfaction of Impartiality can be determined by considering the features of Reliability, Objectivity, and Simplicity.

2.2 Examples of classified human performance measures

Plant performance, Personnel task, Situation awareness, Workload, Teamwork, Anthropometric/Physiological factors are keys for human performance assessment [9]. Among these factors, workload, situation awareness, and teamwork were selected as examples of human classification of human performance measures. Based on table 2, this study classified the NASA-task load index (NASA-TLX), situation awareness global assessment technique (SAGAT), and behaviorally anchored rating scales (BARS). Table 3 shows the classification results [10-12].

Table 3: Classification results of three human performance measures

Dimension	Human performance measures		
	Workload	Situation awareness	Teamwork
	NASA-TLX	SAGAT	BARS
Diagnosticity	Yes	No	Yes
Sensitivity	Yes	No	Yes
Construct Validity	Yes	Yes	Yes
Intrusiveness	No	No	No
Reliability	Yes	Yes	Yes
Objectivity	No	No	No
Simplicity	Yes	No	No

It is worth noting that, as mentioned at the end of the previous section, the decisions of two dimensions (Resolution and Impartiality) are not explicitly included in Table 3. Instead, it can be implicitly decided by considering the decisions of relevant dimensions. For example, in case of NASA-TLX, it is reasonable to expect that Resolution would be ‘Yes’ because three of the four dimensions related to it are assessed as ‘Yes.’

3. Conclusions

Although human performance measures are keys for ISV evaluation, there is a lack of a clear definition of nine dimensions suggested by the NRC. This study defined the nine dimensions through literature reviews and classified human performance measures based on the proposed definitions and features. The suggested definitions and features could be used to identify human performance measures, and to evaluate the ISV of the system. Future studies will suggest a framework to develop human performance selection tool.

Acknowledgment

This work was supported by a 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).

REFERENCES

[1] O’Hara, John, James Higgins, and Stephen Fleger. "Human factors engineering program review model (NUREG-0711) revision 3: update methodology and key revisions." proceedings of 8th international topical meeting on nuclear plant instrumentation and control and human-machine interface technologies. 2012.
[2] O’Hara, J., et al. Integrated system validation: methodology and review criteria. NUREG/CR-6393, US NRC, 1997.
[3] Rehmann, Albert J. Handbook of human performance measures and crew requirements for flightdeck research. Crew System Ergonomics Information Analysis Center Wright-Patterson AFB OH, 1995.
[4] Eggemeier, F. T., et al. "Workload assessment of workload in multiple-task environments." Multiple-Task Performance, 1991.
[5] Nickel, Peter, and Friedhelm Nachreiner. "Psychometric properties of the 0.1 Hz component of HRV as an indicator of mental strain." Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Sage Ca: Los Angeles, Ca: Sage Publications, 2000.
[6] Stein, Earl S. The measurement of pilot performance: A master-journeyman approach. Federal Aviation Administration Technical Center Atlantic City Nj, 1984.
[7] Endsley, Mica R. "A methodology for the objective measurement of pilot situation awareness." AGARD, Situational Awareness in Aerospace Operations, 1990.
[8] O’Hara, J., W. Stubler, and J. Wachtel. Methodological issues in the validation of complex human-machine systems. No. BNL-NUREG-61712; CONF-9506175-3. Brookhaven National Lab., Upton, NY,1995.
[9] Ha, Jun-Su, and Poong-Hyun Seong. "Development of human performance measures for human factors validation in advanced nuclear power plants." Journal of the Ergonomics Society of Korea, 2006.
[10] Rubio, Susana, et al. "Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and Workload Profile Methods." Applied Psychology, 2004.
[11] Endsley, Mica R., et al. "A comparative analysis of SAGAT and SART for evaluations of situation awareness." Proceedings of the Human Factors and Ergonomics Society Annual Meeting. Vol. 42. No. 1. Sage Ca: Los Angeles, Ca: Sage Publications, 1998.
[12] Braarud, P. Ø., and H. Brendryen. "Task demand, task management, and teamwork (HWR-657)." Halden, Norway: OECD Halden Reactor Project, 2001.