An Experimental Analysis on the CNS Simulator Comparing Human Performance between Operators and Students

Bumjun Park[a], Sungheon Lee[a], Taewon Yang[a], Jeong Hun Choi[b], Jooyoung Park[c], Awwal M. Arigi[a], Ronald L. Boring[c], Jonghyun Kim[a*]
[a] *Department of Nuclear Engineering, Chosun University, 309 Pilmun-daero, Dong-gu, Gwangju, 501-709, Republic of Korea*
[b] *Korea Hydro & Nuclear Power Co. LTD., 1655 Bulguk-ro, Gyeong-ju, 38120, Republic of Korea*
[c] *Idaho National Laboratory, 2525 Fremont Avenue, Idaho Falls, ID 83402, USA*

[*]*Corresponding author: jonghyun.kim@chosun.ac.kr*

1. Introduction

Nuclear power plants (NPPs) are operated or managed by operators capable of taking appropriate measures in the event of an accident. However, any human action carries the potential for error [1], which is why human reliability analysis (HRA) researchers systematically analyze and predict the potential risks due to human factors in order to eliminate or reduce such errors. In this area, much research is still needed to improve the safety of NPPs. Typically, human error probabilities (HEPs) are estimated by collecting HRA data inferred from other industries, expert judgment, or training simulator research. However, the lack of adequate data has been highlighted as a major challenge in the field of HRA [2].

Currently, to accumulate HEP data, the U.S. Nuclear Regulatory Commission and the Korea Atomic Energy Research Institute focus on collecting data via full-scope simulators. These studies are conducted using the Scenario Authoring, Characterization, and Debriefing Application (SACADA) database [3] and the Human Reliability Data Extraction (HuREX) framework [4], respectively, to collect data from full-scope digital main control room (MCR) simulations of actual NPP environments [5].

HRA data collection using a full-scope simulator is beneficial in terms of more accurate results, but there are problems with using the data collected from such simulators to support the data required for the HRA quantification process. Full-scope simulator studies conducted to collect HRA data are expensive, due to the need for a full-scope simulator facility and numerous operators, experts, and researchers. For this reason, researchers at Idaho National Laboratory (INL) developed the Simplified Human Error Experimental Program (SHEEP) as a method of overcoming the drawbacks of full-scope research. The purpose of the SHEEP study is to infer the HEPs from full-scope simulator data by comparing student and operator performance using a simplified simulator.

Similar to a previous study conducted under the SHEEP project, the current study aims to leverage experiments conducted on students and operators using the Compact Nuclear Simulator (CNS) in order to identify differences in human performance data. To achieve this goal, the study analyzes correlations between the measured human performance of students and operators, via a randomized factorial experiment design. Five human performance measures are investigated. In addition, several scenarios and related procedures were developed for the CNS experiments.

2. SHEEP Framework

As mentioned, SHEEP (See Fig. 1) was developed at Idaho National Laboratory as a method for inferring HEPs from full-scope data. The SHEEP framework infers HEPs via the following steps: (1) identify the collectible HRA data items when using a simplified simulator, (2) analyze the HRA data items collected from experiments in the simplified simulator, and (3) integrate the HRA data items obtained from the experiments into a full-scope database for developing or improving HRA methods.

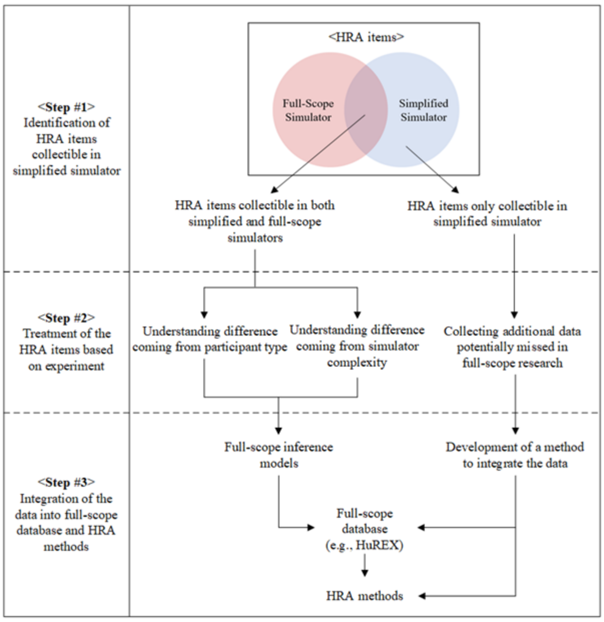

Fig. 1. SHEEP framework

Fig. 2 shows, in detail, the process of using simplified simulator data to infer a full scope of data. When using a less simplified simulator (e.g., CNS) or a more simplified one (e.g., Rancor Microworld), student/operator error data are collected via experiments.

(Experiments using the Rancor Microworld simulator were conducted in a previous study [5].) Then, by developing (1) a method of identifying differences in the human performance of students and operators, as well as (2) a method of defining the differences between simplified and full-scope simulators, the operator data for the full-scope environment can be inferred using the student data from the simplified simulators.

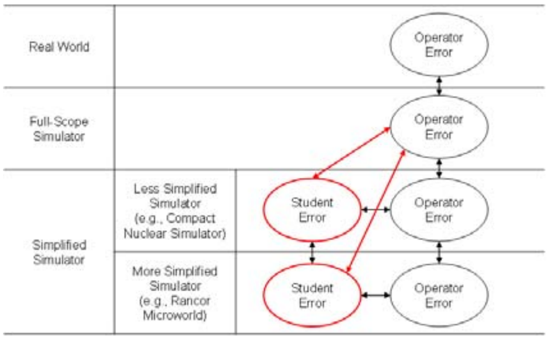

Fig. 2. Process of inferring full-scope operator error data

3. Experimental Design

In this study, which was based around a randomized factorial experimental design, we used the CNS to compare the human performance measures of operators and students. As shown in Table I, the experimental design consisted of two independent variables — "type of subject" (i.e., operator vs. student) and "type of event" (i.e., non-event vs. event)—with each subject performing two normal and two emergency scenarios.

Table I: Randomized factorial experiment design

Type of Event	Type of Subject		Scenario
	Operator	Student	
Non-event			- Startup operation (2 to 50%) - Shutdown operation (100% to hot standby)
Event			- Steam generator tube rupture (SGTR) - Loss of feedwater (LOFW)

3.1 Simulator

The CNS used in this experiment (see Fig. 3) is a small, pressurized-water reactor (PWR)-type NPP simulator developed by the Korea Atomic Energy Research Institute, and is based on the Westinghouse 930 MWe and 3-Loop plant models, as embodied by the Hanul 1 and 2 plants.

This simulator is modeled as having primary and secondary systems, as well as a containment, an electrical system, and alarms, enabling simplified simulation of emergency accidents and power increase/decrease operations.

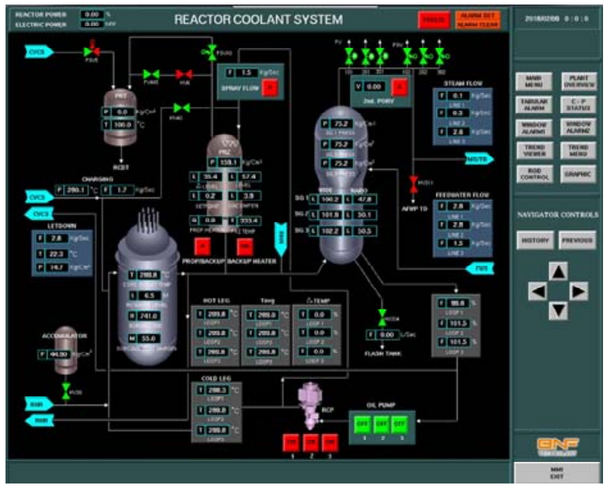

Fig. 3. CNS interface

3.2 Type of Subject

The experiment involved 32 test subjects in total: 16 operators and 16 students. The former group consisted of licensed operators currently employed at a Korean NPP, while the latter group consisted of students enrolled in graduate or undergraduate classes at Chosun University's Department of Nuclear Engineering. Moreover, these students had at least a basic knowledge of PWR-type NPP systems and how to operate them.

3.3 Type of Scenario

There were two scenario types: non-events and events. Non-event scenarios usually resemble normal operation conditions, including startup, shutdown, or full-power operations. In these scenarios, participants may experience less mental pressure and have more time to complete actions than in event scenarios. Conversely, event scenarios involve emergency or abnormal situations (e.g., malfunction or failure of NPP components). These event scenarios imposed greater mental stress and required high levels of concentration.

3.4 Experiment Scenario

Prior to the experiment, several scenarios and procedures were developed to achieve the project goal. These were relatively simple compared to those used for full-scope simulators. In addition, the CNS was modified based on the Westinghouse model (e.g., in terms of procedures and initial conditions) in order to more accurately measure human performance. Table II shows experimental scenarios and relevant procedures that have been validated and are ready for application in the CNS experiments. Each scenario is terminated when the subject completes the procedure or achieves a specified goal. A non-event scenario is terminated when the reactor power reaches a predetermined target (i.e., 2 to 50%) via the procedure, without necessitating a reactor shutdown. An event scenario is terminated when the subject transfers to the wrong procedure, or reaches a predetermined target step after correctly moving through all the right procedures.

Table II: Experiment scenarios and procedures list

Type of Event	Scenario	Procedure
Non-event	Startup operation (2% to 50%)	GOP-01 (startup)
	Shutdown (100% to hot standby)	GOP-02 (shutdown)
Event	Steam generator tube rupture (SGTR) with the failure indicator for the steam generator level	EOP-00 (reactor trip or Safety Injection) EOP-03 (SGTR)
	Loss of feedwater (LOFW)	EOP-00 (reactor trip or Safety Injection) ES-1.1 (Safety Injection termination)

3.5 Human Performance

For each scenario, this experiment collected data pertaining to five different human performance categories: situation awareness, workload, time, error, and eye movement. Table III summarizes the measurement techniques applied to each of these five human performance categories.

Table III: Summary of each human performance measure

Human Performance	Description
Situation awareness	- Perception of environmental elements within a volume of time and space, comprehension of their contextual meaning, and prediction of the future status of those elements. - The situation awareness rating technique (SART) was adopted in the questionnaires given at the end of each scenario
Workload	- The task-imposed demand on limited mental resources when performance is to be maintained at a desired level - The experiment adopted a modified Cooper-Harper (MCH) using a one-dimensional scale, and a questionnaire was filled out after completing each scenario
Time	- Average time to complete a step - Average time to complete an instruction - Average time to complete a task
Error	- The error rate is calculated by dividing the number of errors by the total number of operations in each scenario - An error occurs when the subject's performance deviates from the correct procedure - Errors include both errors of omission and commission - To determine the errors, this experiment applied the same rules and analysis categories as proposed by the HuREX project.
Eye movements	- Eye movement data were collected using the Tobii Pro Glasses 2 eye-tracking system. - Most eye-tracking systems record blink frequency, eye closure fraction, blink duration, fixations, pupil diameter, and saccades. - The experiment used fixation data from 75 milliseconds. - A subject's eye fixations an also be measured using a heatmap.

3.6 Apparatus

The apparatus used in this experiment consisted of a laptop with CNS, an eye tracker capable of recording eye movements, a video recording system to collect the images needed for error analysis, and a procedure that related to the experimental scenario.

3.7 Data Acquisition

In this experiment, most of the data were collected via video recorders, questionnaires, and eye trackers. Table IV summarizes the data collection methods, measurements of collectible data items, and human performance measures. All data items collected through each method are directly linked to human performance data or additional data. These data are used to decipher the analysis results and derive ways of identifying other important inferences.

Table IV: Summary of data collection methods, collection items, and human performance

Method	Data Item Collected	Human Performance
Video Recording	- Time to completion for each step, instruction, and task - Number of errors	- Error - Time
Questionnaire	- General information on each participant - SART - MCH	- Situation awareness - Workload
Eye Tracker	- Number of fixations - Fixation duration - Number of blinks - Heatmap over (areas of interest) AOIs	- Time - Error - Eye movements

4. Analysis of Human Performance Differences

The human performance data collected via the experiment were applied using two statistical analysis methods: Analysis of Variance (ANOVA) tests and correlation analysis using IBM's Statistical Package for the Social Sciences statistics program.

Statistical analysis methods were performed in three steps. The first was to check the normality of the human performance data. Next, two-group (e.g., operator vs. student) ANOVA tests were performed, the results of which confirmed the significance between groups. Lastly, a correlation analysis was performed on those variables that represented correlations.

In this experiment, statistical analysis methods were performed according to the above sequence. Figure 4 shows the results for four of the human performance measures (i.e., SART, MCH, average task completion time, and eye movements [fixation count and blink rate]) in the ANOVA test—the one exception being error.

5. Conclusion

As a preliminary study based on the SHEEP project, this paper described an experiment on inferring operator HEPs in a full-scope simulator by using the CNS to examine the human performance differences between operators and students. The results reflect the statistical analysis of the collected data, except as regards the error human measure category.

The results of the current experiment showed no significant differences in situation awareness, workload, and blink rate. However, they do reveal significant differences in terms of fixation and average task completion time. More specifically, the operators had relatively fewer fixations than the students, whereas the students, on average, needed more time to complete tasks.

The statistical results (see Fig. 4) do not include error and correlation analysis, making them seemingly insufficient for discussion. Thus, further analysis (e.g., error and correlation analysis) are being conducted and will be discussed at this conference.

Tests of Between-Subjects Effects

Source	Dependent Variable	Type III Sum of Squares	df	Mean Square	F	Sig.
Corrected Model	SART	2466.062[a]	3	822.021	15.159	.000
	MCH	130.281[b]	3	43.427	15.466	.000
	Average time to complete a task	14742.56[e]	3	4914.188	113.90	.000
	Fixation count	8399566[f]	3	2799855.4	7.489	.000
	Blink rate	.074[h]	3	.025	.191	.902
Type of Subject (Operator vs Student)	SART	9.031	1	9.031	.167	.684
	MCH	9.031	1	9.031	3.216	.075
	Average time to complete a task	538.453	1	538.453	12.480	.001
	Fixation count	3029491.1	1	3029491.1	8.103	.005
	Blink rate	5.000E-5	1	5.000E-5	.000	.984
Type of Event (Non-event vs Event)	SART	2450.000	1	2450.000	45.182	.000
	MCH	120.125	1	120.125	42.780	.000
	Average time to complete a task	13820.570	1	13820.570	320.33	.000
	Fixation count	4336512.5	1	4336512.5	11.599	.001
	Blink rate	.060	1	.060	.468	.495
Type of Subject * Type of Event	SART	7.031	1	7.031	.130	.719
	MCH	1.125	1	1.125	.401	.528
	Average time to complete a task	383.541	1	383.541	8.890	.003
	Fixation count	1033562.5	1	1033562.5	2.765	.099
	Blink rate	.014	1	.014	.105	.746
Total	SART	96552.000	128			
	MCH	1950.000	128			
	Average time to complete a task	84277.018	128			
	Fixation count	4.187E+8	128			
	Blink rate	29.384	128			

a. R Squared = .268 (Adjusted R Squared = .251)
b. R Squared = .272 (Adjusted R Squared = .255)
e. R Squared = .734 (Adjusted R Squared = .727)
f. R Squared = .153 (Adjusted R Squared = .133)
h. R Squared = .005 (Adjusted R Squared = -.019)

Fig. 4. ANOVA test results for four human performance measures

ACKNOWLEDGMENTS

This work of authorship was prepared as an account of work sponsored by Idaho National Laboratory (under Contract DE-AC07-05ID14517), an agency of the U.S. Government. Neither the U.S. Government, nor any agency thereof, nor any of their employees makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed or represents that its use would not infringe privately owned rights.

REFERENCES

[1] Byeon, S. N., et al. 2011. An Human Reliability Analysis to Identify Human Error Mechanisms for Reducing the Risks Associated with Human Errors in a Main Control Room of the SMART. No. KAERI/CM-1427/2010. Korea Atomic Energy Research Institute.

[2]. Park, J., Arigi, A. M., and Kim, J. 2019. A comparison of the quantification aspects of human reliability analysis methods in nuclear power plants. Annals of Nuclear Energy, 133: 297-312.

[3] Chang, Y. J., et-al 2014. The SACADA database for human reliability and human performance. Reliability Engineering & System Safety, 125, 117-133.

[4] Kim, S., Kim, Y., Choi, S. Y., Jung, W., and Park, J. 2018. Design and Implementation of HuREX Analysis Supporting Interface for HRA Data Extraction. Nuclear Technology, 202(2-3), 259-277.

[5] Kim, Jeongtaeg, et al. 2019. "An experimental design on the use of rancor microworld simulator: a comparison of human performances between actual operators and students." Transactions of the Korean Nuclear Society Autumn Meeting, Goyang, Korea.