[UCI]I804:24011-200000335010

2020년 8월

석사학위 논문

An Experimental Analysis on the Use of Microworld Simulator: A Comparison of Human Performances Between Operators and Students

조선대학교 대학원

원자력공학과

김 정 택

마이크로월드 시뮬레이터를 이용한 실험적 분석: 운전원과 학생 피실험자간 인적 수행도 비교

An Experimental Analysis on the Use of Microwrold Simulator: A Comparison of Human Performances Between Operators and Students

2020년 8월 28일

조선대학교 대학원

원자력공학과

김 정 택

마이크로월드 시뮬레이터를 이용한 실험적 분석: 운전원과 학생 피실험자간 인적 수행도 비교

지도교수 김 종 현

이 논문을 공학 석사학위 신청 논문으로 제출함

2020년 5월

조선대학교 대학원

원자력공학과

김 정 택

김정택의 석사학위논문을 인준함

위원장	조선대학교	교수	나 만 균	(인)
위 원	조선대학교	교수	송 종 순	(인)
위 원	조선대학교	교수	김 종 현	(인)

2020년 6월

조선대학교 대학원

CONTENTS

LIST OF TABLES

LIST OF FIGURES

ABSTRACT

마이크로월드 시뮬레이터를 이용한 실험적 분석: 운전원과 학생 피실험자간 인적 수행도 비교

조선대학교 대학원

원자력공학과

김 정 택

지도 교수 : 김 종 현

인간신뢰도분석 (Human Reliability Analysis, HRA)은 확률론적 안전성 평가에서 요구되는 운전원 행위를 평가하고 그에 대한 오류확률을 정량적으로 분석하는 방법이다. 이를 위해, HRA 연구자들은 실제 데이터, 시뮬레이터 연구 또는 전문가의 판단 등으로 인간 신뢰도 데이터를 수집하여 인적오류확률을 추정하기 위해 노력해왔다. 그러나, 인간 신뢰도 데이터 부족 문제는 아직까지 HRA의 주요 현안으로 남아있는 상태이다. 대표적으로, 현재까지 수많은 HRA 방법론들이 개발되었음에도 불구하고 대부분의 방법론들은 1970년도 초부터 1980년도 후반까지 수집된 Technique for Human Error Rate Prediction (THERP) 데이터에 의존하고 있다. 이에 따라, 미국 Idaho National Laboratory에서는 HRA 데이터를 수집하기 위한 목적으로 Rancor Microworld 시뮬레이터를 개발하였다. 기존에 수행되는 연구들의 경우, 주로 실제 발전소와 유사도가 높은 시뮬레이터와 운전원들을 대상으로 실험을 수행하여 HRA 데이터를 수집해온 반면에, Rancor Microworld 시뮬레이터의 경우 비교적 단순한 시뮬레이터로서 운전원뿐만 아니라 학생들을 대상으로 실험을 수행할 수 있어 저렴한 비용으로 비교적 많은 데이터를 수집할 수 있다는 장점을 갖는다.

본 연구는 Rancor Microworld 시뮬레이터를 HRA 데이터 수집 용도로 활용하기 위한 사전 연구로서, Rancor Microworld 시뮬레이터 운전에 대한 운전원과 학생 피실험자의 수행도를 실험적으로 분석하는 것을 목적으로 한다. 본 연구에서는 총 6개의 인적 수행도를 고려하였고 몇 가지 통계분석기법을 활용하여 운전원과 학생 피실험자의 인적 수행도를 비교하였다.

I. Introduction

Lack of data is a major challenge in human reliability analysis (HRA) [1, 2]. To date, the most broadly used HRA methods—such as the Standardized Plant Analysis Risk HRA—depend on a dataset provided by the Technique for Human Error Rate Prediction [2], which was generated from the early 1970s until the late 1980s, mostly from non-nuclear experience. Furthermore, although new technologies like digital main control rooms (MCRs) are already implemented on new or upgraded nuclear power plants, HRA methods still have been applied as is without modification to accommodate differences due to digital technologies. Accordingly, to update these data, several institutes and researchers tried to collect HRA data from different data sources, such as actual historical measurements, expert judgements, and simulator studies [3].

The majority of recent HRA studies focus on collecting data using full-scope MCR simulators with operators. For most current studies, these are predominantly concentrating on collecting data from simulator studies with full-scope simulators. The largest current efforts are led by the U.S. Nuclear Regulatory Commission (U.S. NRC) and Korea Atomic Energy Research Institute (KAERI). These efforts are collecting data from full-scope simulators using the Scenario Authoring, Characterization, and Debriefing Application (SACADA) database [4] and Human Reliability data Extraction (HuREX) [3] framework, respectively. But in fact, full-scope studies pose several intrinsic challenges for securing adequate data. A full-scope study entails high expense in securing a full-scope facility and numerous operators. Many experts in operating nuclear power plants (NPPs) also participate in the research. Because this work is relatively resource-intensive and time-consuming, in addition to presupposing utilities' cooperation in partially releasing collected data, it is strictly limited to those few organizations able to satisfy such conditions.

In keeping with the need for human reliability data sources, Idaho National Laboratory (INL) began collecting HRA data via the Rancor Microworld simulator, a simplified simulator, with student participants. In this way, INL has identified to investigate whether students could be used as subjects for collecting HRA data instead of operators in nuclear power plants when using the Rancor Microworld. As a first step toward achieving this goal, this paper compares human performances between operators and students measured across benchmark experiments, so that we can understand how much differences in the performances there are between the two subject groups. A randomized factorial experiment design was developed with two independent variables: type of scenario and type of subject. Six human performance measurements—1) time, 2) error, 3) workload, 4) situational awareness, 5) attention, and 6) number of manipulations—were selected. A couple scenarios and related procedures to be simulated by Rancor were then developed. The data collected from the experiment is analyzed using several statistical analysis methods like an analysis of variance (ANOVA) test.

II. Overview of Rancor Microworld

The Rancor Microworld simulator is a simplified simulation environment that reproduces the important characteristics of real operations at NPPs [5]. It has been used to examine theoretical and practical design concepts, and provides a graphical user interface enabling researchers to generically create process control systems. The Rancor Microworld simulator was developed based on the simulation's thermo-hydraulics, which followed a gamified Rankin cycle resembling that of a small modular reactor.

Fig. 1 is a screenshot of the Rancor Microworld interface. It consists of three windows: 1) the Overview Window, 2) the Piping and Instrumentation Diagram Window, and 3) the Controls Window. The Overview Window includes general system information such as the alarm panel. The integrated design helps inform operators when certain parameters fall outside the acceptable range. The Piping and Instrumentation Diagram Window shows parameters for things such as water level and whether or not pumps or valves have been turned on and opened. Lastly, the Controls Window pertains to all controllable measures such as buttons and sliders.

Fig. 1. Rancor Microworld interface

As a simplified simulator, the Rancor Microworld has different characteristics versus full-scope simulators in terms of collectible data levels. This section compares the different analysis levels between Rancor Microworld and existing full-scope studies. The four levels considered in this comparison are 1) task level, 2) step level, 3) instruction level, and 4) execution level. The task-level corresponds to a strategy such as “feed and bleed,” an important long-term cooling strategy for maintaining core safety in emergency situations at NPPs. This level normally consists of several steps within NPP operational procedures. The step level is a procedure step level composed of several actions. Controlling reactor coolant temperature, such as by opening the atmospheric dump valve, is an example of a procedure step level. The instruction level matches the action level included in each procedure step. Lastly, the execution level, the simplest task unit, consists of actions such as looking at or reaching for an object.

Fig. 2 summarizes the various analysis levels in different simulators. Full-scope studies collect data in the task, step, and instruction levels. In the recent studies mentioned in the Introduction, the SACADA database [4] focuses on the task and step levels, whereas the HuREX framework [3] concentrates on the instruction level. The execution level is rarely considered in full-scope studies, since they already focus on so many items in the larger task units; therefore, information from items in small task units may be missed or ignored.

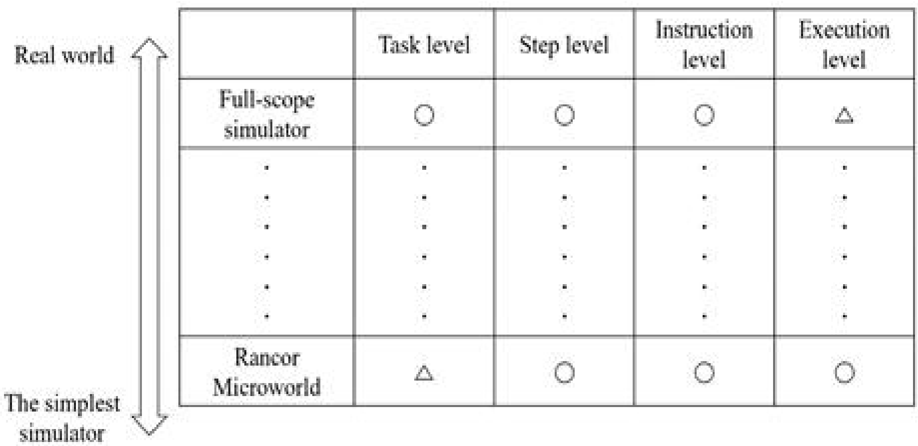

Fig. 2. A spectrum of different analysis levels in different simulators

On the other hand, for reasons of fidelity (i.e., the degree to which a simulated environment corresponds to the real world), Rancor Microworld may focus on relatively low item levels (i.e., the step, instruction, and execution levels) compared to full-scope simulators. This makes Rancor Microworld advantageous in allowing research focused on the execution level, such as how operators use perception to gather information and make decisions of confounding complexity for full-scope studies.

III. Experimental Design

In this study, a randomized factorial experiment compared human performance between operators and students. Table 1 shows the experimental design, composed of two independent variables: type of subject and type of scenario. Details of the experimental design are described in the following sections.

Table 1. Randomized factorial experiment design

Type of scenario	Type of subject	
	Operator	Student
Non-event		
Event		

A. Independent Variables

1. Type of subject

This variable is divided into two groups: operators and students. The former consists of licensed operators currently employed at Korean NPPs, while the latter is composed of undergraduate and graduate students at nuclear engineering departments. These students have at least a basic knowledge of NPP systems and operations.

2. Type of scenario

Scenarios are categorized as either non-event or event scenarios. Non-event scenarios relatively align with work performed during normal operating states such as start-up, shut-down, or full-power operations. In such a scenario, subjects may not feel the intense stress or time pressure involved in event scenarios. On the other hand, event scenarios consist of multiple critical actions needing to be finished within a limited timeframe, and that positively or negatively affect the future state of the plant. Abnormal or emergency situations are often examples of event scenarios.

This variable may help shed light on the feasibility of using simulator studies to collect event scenario-based data. To date, data from simulator studies and experimental research are considered a challenge in terms of fidelity (i.e., the degree to which experimental environments correspond to actual conditions).

B. Experiment Scenario

Several scenarios were developed for achieving the experiment goal. They are relatively simple compared to scenarios for full-scope simulators. Table 2 lists the experiment scenarios and related procedures that are tested and ready to be carried out. Non-events include start-up and shut-down scenarios, whereas events consist of abnormal cases and emergency scenarios.

Each scenario is terminated when the subjects complete a predetermined procedure or reach a specific goal. Non-event scenarios end when reactor power reaches a predetermined target (i.e., 0% or 100%). Event scenarios end when subjects successfully perform all procedural steps or instructions, and parameters such as core temperature are stably maintained.

Table 2. List of experiment scenarios and procedures

Type of scenario	Specific scenario	Procedure
Non-event	Start-up operation (0% to 100%)	OP-001 (Start-up)
	Shut-down operation (100% to 0%)	OP-002 (Shut-down)
	Manual reactor control in part of a start-up operation	OP-001 (Start-up) and OP-010 (Manual reactor control)
	Manual feedwater control in part of a start-up operation	OP-001 (Start-up) and OP-011 (Manual feedwater control)
Event	Reactor coolant pump failure during full-power operation	AOP-001 (Rapid shutdown)
	Control rod failure during full-power operation	AOP-001 (Rapid shutdown)
	Feedwater pump failure during full-power operation	AOP-001 (Rapid shutdown)
	Abnormal turbine trip during full-power operation	AOP-001 (Rapid shutdown)
	Steam generator tube rupture with an indicator failure for the steam generator level	EOP-E-3 (SGTR)
	Loss of feedwater	EOP-E-2 (LOFW)

C. Human Performance Mesurements

In this experiment, six human performance mesuremnets—1) time, 2) error, 3) workload, 4) situational awareness, 5) attention, and 6) number of manipulations—are given for each scenario. This section details each of these measurements.

1. Time

Time-related information collected in the experiment includes the time to complete a scenario, the average time to complete a step, and the average time to complete an instruction. A procedure consists of steps that can further be broken down into individual instructions, each usually including an operator action in the Microworld procedure.

2. Error

Errors indicate when the operator's task performance deviates from the procedure. This includes errors of omission and commission, and it takes into account both the number and rate of errors. The error rate is calculated by dividing the number of errors by the total number of tasks in each scenario.

3. Workload

This study considers two different approaches—Modified Cooper-Harper (MCH) rating scale [6] and an eye-tracker—to estimate workload. The MCH rating scale was originally developed by the aviation industry to estimate operators' psychological and physical workloads. Additionally, it provides design recommendations depending on the rating scale. It evaluates workloads based on responses to post-scenario questionnaires shown in Fig. 3. The second approach is to use an eye-tracker. Certain research [7, 8] indicates a relationship between blinking rate and cognitive workload.

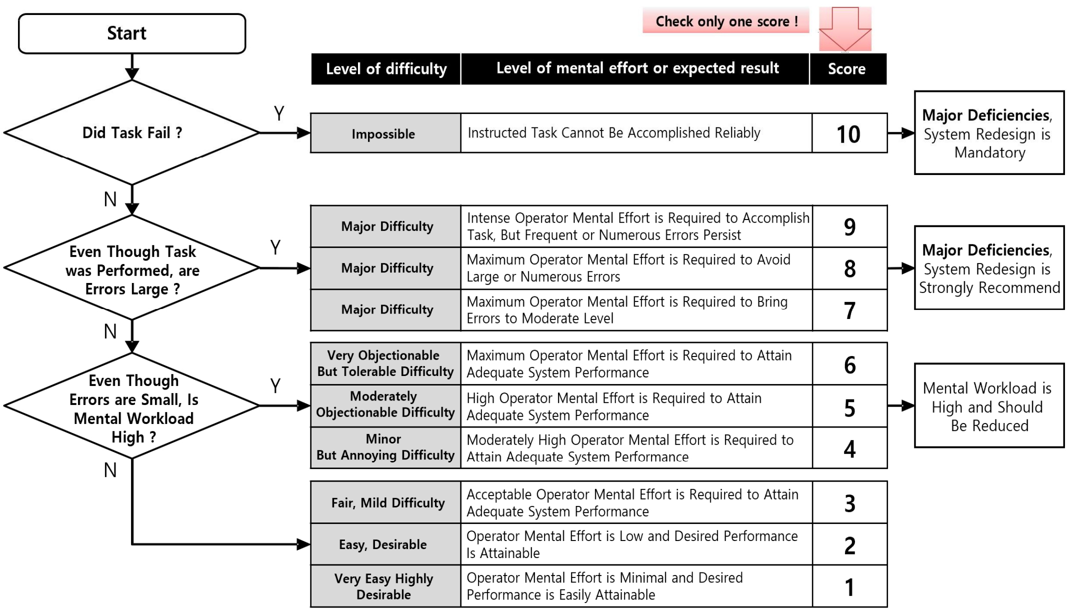

Fig. 3. The questionnaire of the MCH rating scale

4. Situation awareness

Situational awareness is the perception of elements in an environment within a volume of time and space; comprehending the meaning and projecting the status of the elements in the near future [9]. In this study, the Situation Awareness Rating Technique (SART) [10] was used to estimate subjects' situational awareness. Fig. 4 shows the questionnaire of SART rating scale.

Date : **Role :**

1. How changeable is the situation? [Instability]

stable and straightforward	1	2	3	4	5	6	7	Changing suddenly

2. How many variables are changing within the situation? [Variability]

Very few variables changing	1	2	3	4	5	6	7	A large number of factors varying

3. How complicated is the situation? [Complexity]

Simple and straightforward	1	2	3	4	5	6	7	Complex with many interrelated components

4. How aroused are you in the situation? [Arousal]

A low degree of alertness	1	2	3	4	5	6	7	Alert and ready for activity

5. How much mental capacity do you have to spare in the situation? [Spare capacity]

Nothing to spare at all	1	2	3	4	5	6	7	Sufficient to attend to many variables

6. How much are you concentrating on the situation? [Concentration]

Focusing on only one	1	2	3	4	5	6	7	Concentrating on many aspects of the situation

7. How low much is your attention divided in the situation? [Attention division]

Focusing on only one	1	2	3	4	5	6	7	Concentrating on many aspects of the situation

8. How much information have you gained about the situation? [Quantity]

Very little	1	2	3	4	5	6	7	A great deal of knowledge

9.How good information have you been accessible and usable? [Quality]

Difficult to get required operating parameters / symptoms	1	2	3	4	5	6	7	Required operating parameters / symptoms are adequately supplied

10. How familiar are you with the situation? [Familiarity]

New situation	1	2	3	4	5	6	7	A great deal of relevant experience

Fig. 4. The questionnaire of the SART rating scale

5. Attention

This performance measurement, which relies on the eye-tracker system, estimates the proportion that focuses on major information in the microworld interface: alarms, the primary system, the steam generator, the turbine system, etc.

6. Number of manipulations

The number of manipulations refers to how many times the subjects manipulate microworld interfaces. Manipulations include turning pumps, valves, and sliders on or off. It is counted by using the log data generated from the microworld as shown in Fig. 5.


```
{"EventType":"Simulation", "Tag":"Start", "Time":0, "RX":9.92348143457273E-05, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"AtmosDumpActive", "State":"Cleared", "Unit":1, "Time":0.993694, "RX":1E-05, "MW":9.01121149850884}
{"EventType": "ControlAction", "Name":"RodControl", "Command":"SetRodCtrlManualAuto", "SelectedRod":2, "RodPositions":[100, 100, 100, 100], "Time":135.3903243, "RX":1E-05, "MW":9.01121149850884}
{"EventType": "ControlAction", "Name":"RodControl", "Command":"ExecuteRodCtrlGo", "SelectedRod":2, "RodPositions":[100, 100, 100, 100], "Time":186.4149074, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"AllRodsDown", "State":"Cleared", "Unit":1, "Time":187.411882, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"Mode", "Tag":"Reactor Startup", "Unit":1, "Time":187.411882, "RX":0.151763255799332, "MW":9.01121149850884}
{"EventType": "ControlAction", "Tag":"SGcAauto", "Value":"1", "Time":203.4440899, "RX":3.27450435653858, "MW":9.01121149850884}
{"EventType": "ControlAction", "Tag":"SGcBauto", "Value":"1", "Time":207.4673477, "RX":4.92614066879076, "MW":9.01121149850884}
{"EventType":"Mode", "Tag":"Ready To Roll", "Unit":1, "Time":267.8325139, "RX":21.3265340133663, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"CoreLowTemp", "State":"Cleared", "Unit":1, "Time":279.8512014, "RX":21.4343592884394, "MW":9.01121149850884}
{"EventType": "ControlAction", "Name":"RodControl", "Command":"ExecuteRodCtrlGo", "SelectedRod":2, "RodPositions":[75, 70, 70, 70], "Time":301.8445335, "RX":26.7460852941824, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"CoreHighTemp", "State":"Alarmed", "Unit":1, "Time":320.8521435, "RX":45.2322332512789, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"CoreSafetyInterlockEngaged", "State":"Alarmed", "Unit":1, "Time":325.8399882, "RX":48.27888440554, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"SafetyInjectionActive", "State":"Alarmed", "Unit":1, "Time":325.8399882, "RX":48.27888440554, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"SafetyInjectionRunning", "State":"Alarmed", "Unit":1, "Time":325.8399882, "RX":43.4356639010997, "MW":9.01121149850884}
{"EventType":"Mode", "Tag":"Reactor Startup", "Unit":1, "Time":325.8399882, "RX":43.4356639010997, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"AllRodsDown", "State":"Alarmed", "Unit":1, "Time":326.8686777, "RX":43.4356639010997, "MW":9.01121149850884}
{"EventType":"Mode", "Tag":"Shutdown", "Unit":1, "Time":326.8686777, "RX":20.2337321863767, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"CoreSafetyInterlockEngaged", "State":"Cleared", "Unit":1, "Time":379.8851382, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"ControlAction", "Tag":"SGSGAin", "Mode":"auto", "Value":"1", "Time":446.8982159, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"ControlAction", "Tag":"SGSGBin", "Mode":"auto", "Value":"1", "Time":446.8982159, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"ControlAction", "Tag":"SGSGAin", "Mode":"manual", "Value":"-0.19889556923219", "Time":448.8957542, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"ControlAction", "Tag":"SGSGBin", "Mode":"manual", "Value":"-0.19889556923219", "Time":448.8957542, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"CoreHighTemp", "State":"Cleared", "Unit":1, "Time":450.9072393, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"SafetyInjectionActive", "State":"Cleared", "Unit":1, "Time":450.9072393, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"Alarm", "Tag":"SafetyInjectionRunning", "State":"Cleared", "Unit":1, "Time":450.9072393, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"ControlAction", "Tag":"SGSGAin", "Mode":"auto", "Value":"0.666958008067955", "Time":450.9072393, "RX":1E-05, "MW":9.01121149850884}
{"EventType":"ControlAction", "Tag":"SGSGBin", "Mode":"auto", "Value":"0.666958008067955", "Time":450.9072393, "RX":1E-05, "MW":9.01121149850884}
```


Fig. 5. Log data generated from Rancor Microworld

D. Subjects

Table 3 summarizes the experiment subjects. Basically, the Rancor Microworld simulator was designed so a single subject could operate the system. The current plan calls for twenty subjects (both operators and students) to participate in the experiment. The operators will be licensed operators employed at Korean NPPs. The students will be knowledgeable about NPPs and their operation, and will mostly come from universities in Korea.

Table 3. Summary of experiment subjects

Type of subject	Number of subjects	Description
Operator	Twenty	Licensed operators employed at Korean NPPs
Student	Twenty	Knowledgeable about NPPs and their operations, or having participated in the undergraduate class, "Reactor Operation and Simulator Training"

E. Facility

The Rancor Microworld simulator is installed on a laptop dedicated to Microworld experiments only. This experiment can be performed wherever a desk, chair, and power source are available. Also, the laptop enables subjects to operate microworld via the touch screen.

F. Data Acquisition

In this study, the majority of data is collected via the aforementioned questionnaires and eye-tracker. Table 4 summarizes the data acquisition methods, their collectible items, and human performance measurements. All items collected from each method are directly linked to human performance data or additional data potentially helpful for understanding analysis results and compiling alternative methods for identifying other significant results.

Table 4. Summary of data acquisition methods, their collectible items, and human performance measurements

Method	All Items collected	Human performance
Questionnaires	General information, Situational awareness scores from SART, Workload from MCH	Situational awareness, Workload
Eye-tracker	Video record, Area of interest, Gaze, Workload (based on blinking data)	Time, Error, Attention, Workload
Microworld	Microworld log data	Number of manipulations

G. Experiment Procedure

Each subject conducts three different scenarios consisting of non-events or events. Scenarios are randomly selected from among those introduced in Table 2. Termination conditions for each scenario are described in Section 2.2. The time it takes for each subject to participate in an experiment by completing the six scenarios is approximately 1.5 hours.

Before conducting the scenarios, an introductory presentation will be shown to give the subjects an overview of the experiment. The subjects will have enough time to become familiar with the Microworld interfaces.

H. Data Analysis

Statistical analysis methods are applied for the randomized factorial experiment design introduced in Section 2. An analysis of variance (ANOVA) test is performed to identify significant results between items in each independent variable. Then, A correlation analysis is performed to identify how much human performance measures are correlated one another.

IV. Result

This section mainly introduces analysis results on 1) how much differences in the human performance measures there are depending on two independent variables, i.e., type of subject and type of scenario, and 2) how much human performance measures are correlated one another. The two statistical analysis methods, i.e., 1) ANOVA test and 2) correlation analysis, have been applied for the data collected from the experiment.

In addition, the normality test and homogeneity test have been additionally applied for the data collected from the experiment. The normality test aims to determine if a data set is well-modeled by a normal distribution and to compute how likely it is for a random variable underlying the data set to be normally distributed. The homogeneity test determines if two or more populations have the same distribution of a single categorical variable. The normality and homogeneity tests are required to guarantee the quality of the result of the statistical methods. In this study, the Kolmogorov-Smirnov method and the Levene's test are used for investigating the normality and the homogeneity, respectively.

A. The Result of ANOVA Test

Table 5 shows a summary of results from the ANOVA test for the independent variables on the six human performance measures. The detail on the result is described in the following sections.

Table 5. Summary of results from the ANOVA test for the independent variables on the human performance measurements

Human performance	Measurement	Independent variable	
		Type of subject	Type of scenario
Time	Average time to complete a step	-	★
	Average time to complete an instruction	-	★
	Average time to complete a task	-	★
Error	The number of errors	★	-
	Error rate	★	-
Workload	MCH scale	★	-
Situation awareness	SART scale	★	-
Attention	Average duration per a fixation for the entire interface	-	★
	Average duration per a fixation for alarm display	-	-
	Average duration per a fixation for controls window	-	★
	Average duration per a fixation for overview window	-	★
	Average duration per a fixation for PID window	-	-
Number of manipulations	The total number of manipulations	-	★
	The total number of manipulations per a step	-	★
	The total number of manipulations per an instruction	-	★
	The total number of manipulations per a task	-	★
	The total number of manipulations per the scenario completion time	★	★

Note:

'★' denotes that a result from the ANOVA test shows a statistical difference with respect to the independent variable within the 95% confidence level ($p<0.05$).

'-' denotes that a result from the ANOVA test shows no statistical difference with respect to the independent variable ($p>0.05$).

1. Time

The human performance measures for time include 1) average time to complete a step, 2) average time to complete an instruction, 3) average time to complete a task. The result of ANOVA test for all the measures indicates significant differences only for the type of scenario, while there is no statistical significance result for the type of subject.

This study compares average values for the measures having a significant relation with the independent variable. It is identified that the non-event scenarios in the type of scenario have averagely higher time measures than the event scenario. Table 6 indicates a summary of average values for the time measures depending on categories in the type of scenario, i.e., non-event and event scenarios. The ratios between non-event and event scenarios for the average time to complete a step, the average time to complete an instruction and average time to complete a task are 2.31, 1.27 and 1.54, respectively.

Table 6. Summary of average values for time measures depending on categories in the type of scenario

Type of scenario	Human performance measures - Time		
	Average time to complete a step	Average time to complete an instruction	Average time to complete a task
Non-event	50.07	8.68	7.25
Event	21.72	6.84	4.72
The ratio between non-event and event scenarios	2.31	1.27	1.54

2. Error

The error consists of two human performance measures; 1) the number of errors and 2) the error rate. The ANOVA test for the number of errors produces a statistical significance on the type of subject, while the test for the error rate does not show any significant result with the type of scenario. In addition, the relationships between 1) the type of subject and the error rate and 2) the type of scenario and the number of errors do not satisfy a significant level, but indicate a p-value close to the 95% confidence level.

Table 7 shows a summary of average values for the number of errors depending on categories in the type of subject. The student group averagely shows 2.19 times higher error numbers than the operator group.

Table 7. Summary of average values for the number of errors depending on categories in the type of subject

Type of subjects	The number of errors
Student	0.68
Operator	0.31
The ratio between student and operator	2.19

3. Workload

The workload is measured by the MCH questionnaire. The result of ANOVA test for the MCH scale shows a significant difference with the type of subject, while there is no significant result with the type of scenario.

Table 8 shows a summary of average values for the MCH scale depending on categories in the type of subject. The student group averagely shows 1.33 times higher value than the operator group.

Table 8. Summary of average values for the MCH scale depending on categories in the type of subject

Type of subjects	MCH scale
Student	4.15
Operator	3.12
The ratio between student and operator	1.33

4. Situation Awareness

The situation awareness is measured by the SART questionnaire. The result of ANOVA test for the SART scale shows a significant difference with the type of subject, while there is no significant result with the type of scenario.

Table 9 shows a summary of average values for the SART scale depending on categories in the type of subject. The student group averagely shows 0.89 times lower value than the operator group.

Table 9. Summary of average values for the SART scale depending on categories in the type of subject

Type of subjects	SART scale
Student	18.18
Operator	20.21
The ratio between student and operator	0.89

5. Attention

The attention consists of five human performance measures; 1) average duration per a fixation for the entire interface, 2) average duration per a fixation for alarm display, 3) average duration per a fixation for controls window, 4) average duration per a fixation for overview window and 5) average duration per a fixation for PID window. For the average duration per a fixation for the entire interface, controls window and overview window, those have significant differences on the type of scenario, while the others do not include any significant result on the both of the independent variables.

Table 10 shows a summary of average values for attention measures depending on categories in the type of scenario. In the table, the non-event scenarios in the type of scenario have averagely higher attention measures than the event scenario.

Table 10. Summary of average values for attention measures depending on categories in the type of scenario

Type of scenario	Human performance measures - Attention		
	Average duration per a fixation for the entire interface	Average duration per a fixation for controls window	Average duration per a fixation for overview window
Non-event	0.36	0.37	0.35
Event	0.31	0.31	0.32
The ratio between non-event and event scenarios	1.16	1.19	1.09

6. Number of manipulations

The number of manipulations has five measures; 1) the total number of manipulations, 2) the total number of manipulations per a step, 3) the total number of manipulations per an instruction, 4) the total number of manipulations per a task and 5) the total number of manipulations per the scenario completion time. All the measures have statistically significant relation with the type of scenario. For the type of subject, the total number of manipulations per the scenario completion time has a significant difference, while the other measures are not satisfying the confidence level.

Table 11 shows a summary of average values for the total number of manipulations per the scenario completion time depending on categories in the type of subject. In the table, the student scenarios in the type of scenario have 0.8 times lower value than the event scenario.

Table 11. Summary of average values for the total number of manipulations per the scenario completion time depending on categories in the type of subject

Type of subjects	The total number of manipulations per the scenario completion time
Student	0.0615
Operator	0.0772
The ratio between student and operator	0.80

Table 12 shows a summary of average values for the total number of manipulations per the scenario completion time depending on categories in the type of scenario.

Table 12. Summary of average values for the number of manipulations measures depending on categories in the type of scenario

Type of scenario	Human performance measures − Number of manipulations				
	The total number of manipulations	The total number of manipulations per a step	The total number of manipulations per an instruction	The total number of manipulations per a task	The total number of manipulations per the scenario completion time
Non-event	40.60	1.68	0.33	0.26	0.04
Event	13.23	1.87	0.59	0.41	0.09
The ratio between non-event and event scenarios	3.07	0.90	0.56	0.63	0.44

B. The Result of Correlation Analysis

Table 13, Table 14 and Table 15 show a summary of results from the correlation analysis on the five human performance measures with all the subjects, 20 operators and 20 students. First, in the Table 13, the relationships 1) between the situation awareness and the workload, 2) between the error rate and the workload, and 3) between the number of manipulations and the time, show the moderately high correlations with significant levels. Second, in the Table 14, there are the significant relationships 1) between the workload and the situation awareness and 2) between the time and the number of manipulations. Third, in the Table 15, the workload and the situation awareness have correlated with others except for the number of manipulations. There is also a correlation relationship between the time and the number of manipulations.

Table 13. A summary of results from the correlation analysis on the five human performance measures with 40 subjects

	Workload	Situation Awareness	Error	Time	Manipulation
Workload	1.00				
Situation awareness	-0.548*	1.00			
Error	0.431*	-0.06	1.00		
Time	0.12	-0.09	0.00	1.00	
Manipulation	0.00	0.02	-0.06	0.488*	1.00

‘*’ denotes that a result from the correlation analysis shows a statistical difference with respect to the independent variable within the 95% confidence level ($p<0.05$).

Table 14. A summary of results from the correlation analysis on the five human performance measures with 20 operators

	Workload	Situation Awareness	Error	Time	Manipulation
Workload	1.00				
Situation awareness	-0.375*	1.00			
Error	0.17	0.02	1.00		
Time	0.00	0.02	-0.07	1.00	
Manipulation	0.03	0.01	0.10	0.637*	1.00

'*' denotes that a result from the correlation analysis shows a statistical difference with respect to the independent variable within the 95% confidence level ($p<0.05$).

Table 15. A summary of results from the correlation analysis on the five human performance measures with 20 students

	Workload	Situation Awareness	Error	Time	Manipulation
Workload	1.00				
Situation awareness	-0.629*	1.00			
Error	0.480*	-0.365*	1.00		
Time	0.204*	-0.208*	0.03	1.00	
Manipulation	0.01	0.02	-0.11	0.342*	1.00

'*' denotes that a result from the correlation analysis shows a statistical difference with respect to the independent variable within the 95% confidence level ($p<0.05$).

V. Discussion

This study basically attempts to understand how much differences in the human performance measures there are within independent variables, i.e., type of subject and type of scenario. The details are described as below.

The human performance measures estimated in this study could be classified into two different categories, i.e., 1) primary human performance measures and 2) secondary human performance measures. As shown in Fig. 6 and Table 16, the former measures refer to the human performance measures estimated from human-system collaboration. The measures for time, attention and number of manipulations correspond to the primary human performance measures, which are highly related to humans' basic ability on system manipulation. On the other hand, the measures for error, workload and situation awareness are the latter measures which are estimated after the cognitive processes and highly employ humans' previous knowledge and experience.

As shown in Table 5, the result of ANOVA test indicates that the secondary measures i.e., human performance measures for error, workload and situation awareness, are significantly different, while the primary measures, i.e., human performance measures for time, attention and number of manipulations do not show any significant result depending on the type of subject. Through this result, it is inferred that the definitive difference between operator and student relies on the human cognitive processes that are dominantly depend on subjects' previous experience or knowledge rather than humans' basic ability. Regardless of the subject types, the result may indicate that the humans' basic ability is similar.

The human performance measures on the type of scenario indicate opposite results with those on the type of subject. First, the secondary human performance measures show significant differences on whether event or non-event scenarios. It indicates that the difference between the two scenario types mainly affects human-system collaboration, not human cognitive processes employing humans' previous knowledge and experience. Second, for primary measures, there is no significant result on the type of scenarios. As described in Section 3.1.2, the non-event scenarios relatively align with work performed during normal operating states and subjects in these scenarios may not feel the intense stress or time pressure involved in event scenarios. Whereas, event scenarios such as abnormal or emergency situations consist of multiple critical actions needing to be finished within a limited timeframe. When considering these aspects, this result may make an issue on whether simulator studies are applicable to collect event scenario-based HRA data. At least in the Rancor Microworld simulator, which is a simplified simulator, there may be a limitation to equip with the similar experimental environment for task complexity in comparison with full-scope studies.

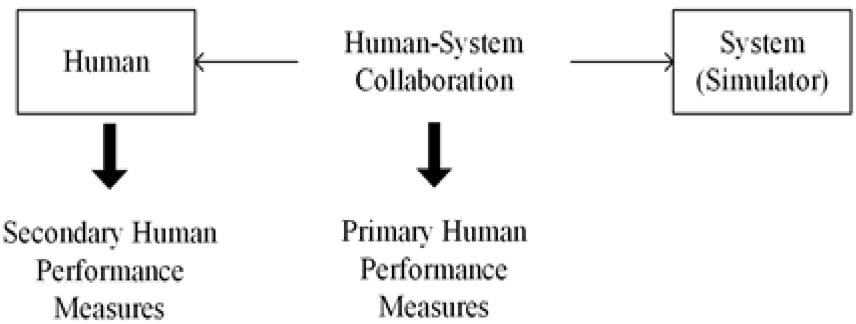

Fig. 6. The relationship between primary and secondary human performance measures

Table 16. Classification of human performance measures depending on their characteristics

Type	Definition	Human performance measures in this study
Primary human performance measures	• Human performance measures estimated from human-system collaboration • Related to humans' basic ability on system manipulation	• Time • Attention pattern • Number of manipulations

Secondary human performance measures	• Human performance measures estimated from humans after human cognitive processes • Influential on humans’ previous knowledge and experience	• Errors • Workload • Situation awareness

For operators’ human performance (see Table 14), there is no correlation between error and the others, while the error for students is correlated with workload and situation awareness (see Table 15). Operators have trained and experienced in a variety of NPP operation environment, therefore, they are relatively familiar with the circumstance. This result may say that operators are the more tolerant to the circumstance rather than students.

VI. Conclusion

This paper compares human performances between actual operators and students measured across benchmark experiments, so that we can understand how much differences in the performances there are between the two subject groups. A randomized factorial experiment design was developed with two independent variables: type of scenario and type of subject. Six human performance measurements—1) time, 2) error, 3) workload, 4) situational awareness, 5) attention, and 6) number of manipulations—were selected. A couple scenarios and related procedures to be simulated by Rancor were then developed. The data collected from the experiment is analyzed using several statistical analysis methods like an analysis of variance (ANOVA) test and correlation analysis.

This study represents an on-going effort to validate a simplified simulator, i.e., the Rancor Microworld for collecting a variety and a number of human reliability data, because the existing full-scope studies have a couple of limitations as below.

- ✓ The full-scope simulators and experiment subjects are rare and costly.
- ✓ Projecting a full-scope study is relatively re-source-intensive and time-consuming.
- ✓ It presupposes utilities' cooperation in partially releasing collected data.
- ✓ It is strictly limited to few organizations that can satisfy the conditions above.

Nevertheless, for the use of the Rancor Microworld simulator, there are still several challenges needed to be overcome like how to treat limitations of the simulator coming from simulator simplicity and whether we can collect the data using student subjects. As one of the efforts, this study conducted to compare human performance measures to understand differences in the measures from operator and student. In the future, differences using the Rancor and full-scope simulators will be investigated. It is assumed that the Rancor Microworld is a complement – not a replacement – for full-scope studies. With understanding the limitations that are difficult to be treated in the Rancor Microworld simulator, how to support the existing full-scope data collection studies will be suggested in the upcoming research.

VII. Reference

[1] Park, J., et al., An experimental investigation on relationship between PSFs and operator performances in the digital main control room. Annals of Nuclear Energy, 2017. 101: p. 58-68.

[2] Swain, A.D. and H.E. Guttmann, Handbook of human-reliability analysis with emphasis on nuclear power plant applications. Final report. 1983, Sandia National Labs.

[3] Jung, W., et al., HuREX–A framework of HRA data collection from simulators in nuclear power plants. Reliability Engineering & System Safety, 2020. 194: p. 106235.

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

[5] Ulrich, T.A., The Development and Evaluation of Attention and Situation Awareness Measures in Nuclear Process Control Using the Rancor Microworld Environment. 2017, University of Idaho.

[6] Cummings, M., K. Myers, and S.D. Scott. Modified Cooper Harper evaluation tool for unmanned vehicle displays. in Proceedings of UVS Canada: Conference on Unmanned Vehicle Systems Canada. 2006.

[7] Coral, M.P., Analyzing cognitive workload through eye-related measurements: A meta-analysis. 2016.

[8] Telford, C. and N. Thompson, Some factors influencing voluntary and reflex eyelid responses. Journal of Experimental Psychology, 1933. 16(4): p. 524.

[9] Endsley, M.R. and D.J. Garland, Situation awareness analysis and measurement. 2000: CRC Press.

[10] Taylor, R.M., Situational awareness rating technique (SART): The development of a tool for aircrew systems design, in Situational awareness. 2017, Routledge. p. 111-128.