Nuclear Power Plant Operator Performance Measurement System for Human Reliability Analysis

Kenji Mashio[a]*, Futoshi Tanaka [a]*
[a]Mitsubishi Heavy Industries Ltd., Kobe, Japan

Abstract: Collecting and evaluating operator performance data gained through simulator training is a key to understand human reliability and performance of operators under accident or abnormal conditions. In order to develop a statistically sound understating of operator performance, variety of data related to operator behavior needs to be obtained for a spectrum of accident scenarios and this can be achieved only through observation of operator training using training simulators. Current practices of obtaining data is to conduct conventional human performance measurements, such as check list, debriefing, questionaries, observations etc., to identify and classify human error modes. Classification of cognitive human errors are sometimes difficult because operators may forget their cognitive process in which human errors have occurred while the operators execute throughout the simulator scenario. The analysts also have task burden and need clear criteria to classify the error modes identified in simulator training scenario executions. To effectively and consistently evaluate human errors, an automated system to evaluate the operator's monitoring and manipulation performance in the main control room through simulator training has been development. The system acquires various data related to operator behavior obtained during the operator training using eye tracking measurement, voice recognition devices and operation logs archived with a simulator system. And based on pre-defined task patterns and task criteria, the systems classifies human error modes through computer algorithms. The prototype of the automated plant operator performance evaluation tool (POPET) shows effective and reliable role which reduces the analyst's task burden and increases reliabilities of human error mode classifications. This paper describes the POPET and its application to human reliability analysis.

Keywords: Human Reliability Analysis, Human Machine Interface, Simulator, Human Factors

1. INTRODUCTION

Collecting and evaluating operator performance data gained through simulator training is a key to understand human reliability and performance of operators under accident or abnormal conditions. In order to develop a statistically sound understating of operator performance, variety of data related to operator behavior has to be obtained for a spectrum of accident scenarios and this can be achieved only through observation of operator training using training simulators. Furthermore, a large amount of data must be objectively evaluated in a way the results are not affected by the analyst.

The analysts have task burden and need clear criteria to classify human error mode when operator errors are identified in the simulator training scenario executions. In the current practice of classifying human error modes, the analyst has to conduct conventional human performance measurements, such as check list, debriefing, questionaries, observations etc., in order to determine which human error modes the identified operator errors should be classified in. Cognitive errors can be obtained from operator's memories by operator's debriefing and questionaries after scenario executions. However, there is uncertainty in this process because operators may forget their cognitive process in which human errors have occurred while they go through simulator scenario executions.

To effectively identify human error modes identified in simulator training, an automated system that can evaluate operator's monitoring and manipulation performance in the main control room training simulator has been developed[1]. The system acquires various data related to operator behavior obtained during the operator training using eye tracking measurement, voice recognition devices and operation logs archived with a simulator system. Then, computer algorithms judge operator behaviors with pre-defined task patterns and task criteria. In that way, tasks performed by the operator can be automatically evaluated. This system can be applied to classify human errors into error modes modeled in the human reliability analysis (HRA) for the purpose of efficiently gaining data to estimate the human error probabilities.

Development of the operator performance system and its application to HRA is discussed in the following chapters as outlined in Figure 1. Chapter 2 is dedicated to the objective of the performance monitoring system. Chapter 3 describes cognitive model applied and the development of the automated performance evaluation system. The application of the automated system to classify human error into human error modes used in HRA is discussed in Chapter 4, and the conclusion is provided in Chapter 5.

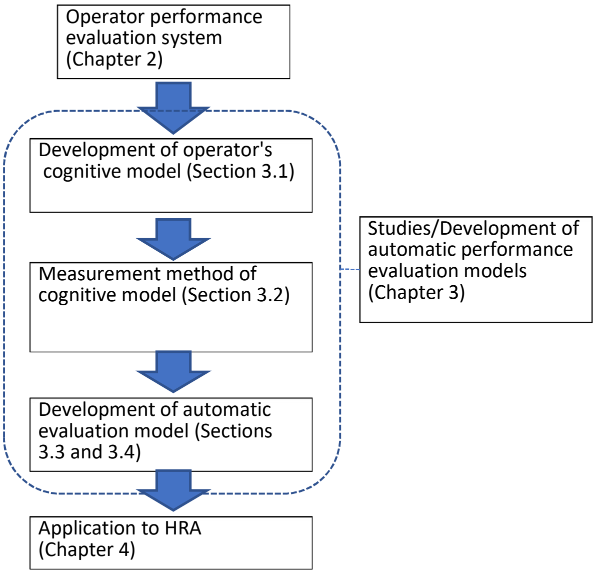

Figure 1. Outline of operator performance system development

2. OPERATOR PERFORMANCE EVALUATION SYSTEM

2.1. System Development

It is important to measure operator's situation awareness and workload to evaluate human system interactions. Mental workload caused by cognitive tasks are dominant rather than physical workload, particularly in digital main control rooms. In addition, for HRA, it is crucial to identify cognitive levels in addition to procedural human errors. Therefore, the performance evaluation system was developed to identify operator's cognitive errors.

2.2. Plant system of interest

The evaluation objectives are selected from human system interaction aspects as follows.

Human Machine Interface (HMI) objective

The latest pressurized water reactor (PWR) digital control room is selected as HMI objectives. The digital control room consists of operator console, supervisor console and large display panel. Operators perform plant control and monitoring at operator console. A supervisor sits down at supervisor console, which is located behind the operator console, and directs operators based on operating procedures and follow/supervise operator's actions as well as plant status/parameters monitoring at the video display units (VDUs) in the supervisor console. A large display panel is placed in front of operator console and supervisor console and provides plant overview displays to operators and supervisor. Operating procedures are provided with either paper-based or computer-based types. In case of computer-based procedure (CBP), the CBP is embedded in the operator console.

Human objective

Control room operators in PWR plants are selected as human objectives. Control room operators consist of operators and a supervisor. Two operators are typically assigned and working at operator console, then split their work for nuclear and turbine systems. A supervisor sits down at a supervisor console which is located behind operators and directs plant operations to operators based on operating procedures (Figure 2).

Figure 2. Digital Control Room

3. Automation of Operator Performance Measurement

3.1. Cognitive process model

Operator's cognitive model applied is settled in Figure 3. This model is a minor modification of the model discussed in Reference [2]. Inter-team coordination is not an independent process and affects understanding and decision-making of operator's cognitive process. Therefore, inter-team coordination is located in parallel with understanding and decision-making. The arrow between inter-team coordination and understanding/decision-making indicated with dual process represents bi-lateral because this process is interactive before moving to action execution.

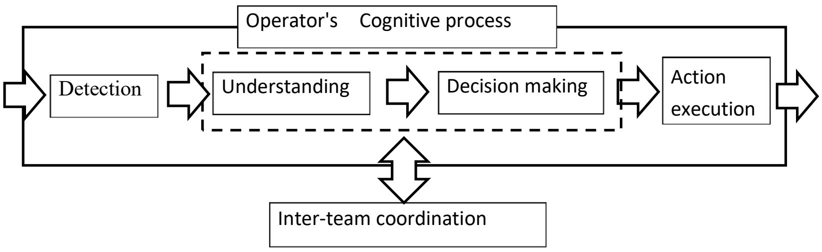

Figure 3. Operator's cognitive process model

3.2. Plant diagnosis performance measurement

Measurement method for operator's cognitive process is shown in Figure 4. Among the senses of human (i.e., sight, sound, smell and touch) used for detection, sight and sound are the major senses of detection in the control room except for particular events, such as fire incidents. Sound sense compensates vision by proactively notifying operator by alarm sounds. In case of emergency or accident situation, many alarms are initiated, and visual sense takes a major role because operations focus on procedural operations. It is important to identify which display areas the operators are looking at in the digital control room, because plant operations are performed thorough various displays in the digital control room.

To measure detection focuses on visual sense, eye pointer measurement system is applied to monitor and record eye pointers during operations for measuring what operators understand and/or make decisions in understanding and decision-making process. Verbal information by operators recorded in conversation log system is used to measure understanding and decision-making process through communications between operators and supervisor.

Action execution is identified using display touch logs which are equipped within a training simulator. The display touch log system can track plant component status (e.g., valve open/close) and display touch logs (e.g., touch record on navigation button for display switching).

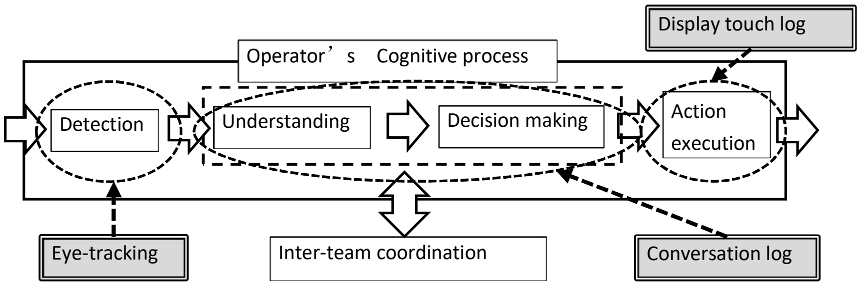

Figure 4. Measurement methods for operator's cognitive process

3.3. Automatic performance measurement process

Automatic performance evaluation algorithms are developed based on operator's cognitive process model and performance measurement data, respectively described in Sections 3.1 and 3.2. The overall process of the automated performance measurement is depicted in Figure 5. Performance measurement is performed in the following two steps.

<u>Step 1: Data collection and database construction</u>

The first step is to organize different performance measurement data source collected from each measurement device, then integrate those data in one database. Eye point data, conversation data and display touch data are organized with the same data formula and aligned in chronological order. In this way, measurement record along with cognitive process are clarified and utilized for automatic evaluation algorithm discussed in Step 2. Conversation data is converted to text to enable to search and identify key words during data evaluation.

<u>Step 2: Data evaluation</u>

The second step is to apply automatic performance evaluation algorithms for each type of task at operator's cognitive process level, and to provide evaluation results in real time. Automatic performance evaluation algorithms are developed are created based on how evaluators of the performance measurement data, hereafter called evaluator, do their evaluations. Operator's tasks are categorized into following three types of tasks and four sub-categorized tasks.

a. Plant diagnosis
b. Plant monitoring
    - Plant parameters and component status monitoring
    - Alarm detection & monitoring
c. Plant operation
    - On-off operations
    - Process control by controllers

Indirect plant monitoring and control tasks (e.g., plant administrative tasks, on-duty/job reporting, shift transfer and plant surveillance) are not within the scope of this evaluation. Automatic performance evaluation algorithms are developed through investigating how evaluators evaluate each type of task and allocate performance measurement data to each cognitive process.

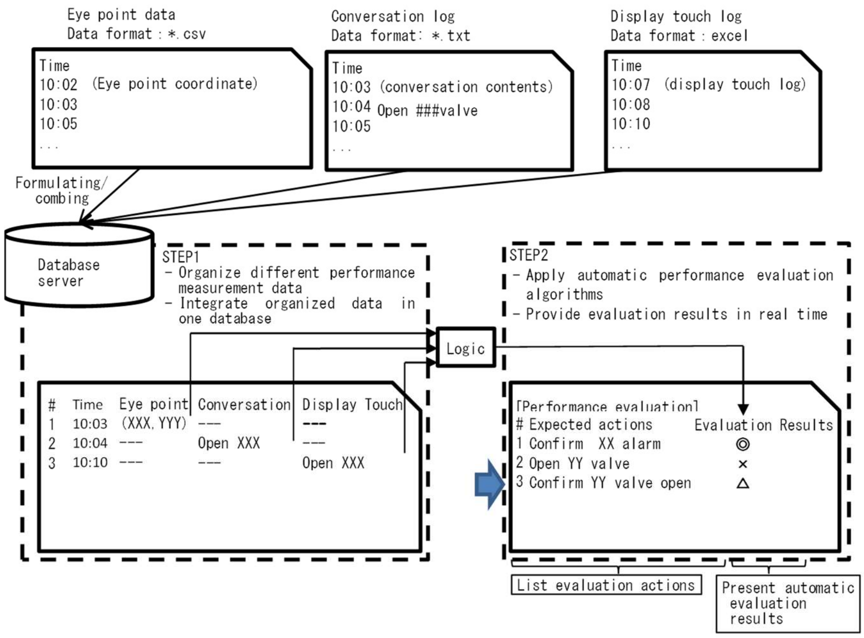

Figure 5. Overall evaluation process

3.4. Performance Evaluation Algorithm for Diagnosis

Evaluator needs to evaluate whether operators have correctly identified plant event/incident with associate parameters/alarms indicating the symptom, through observation of operator's behaviours, such as operator's actions and conversations. The cognitive process of "detection-understanding-decision-making", described in Section 3.2 is applied to develop performance evaluation algorithm for plant diagnosis. The automated evaluation methodology for the cognitive process is described below.

<u>Detection</u>

Eye point detection areas representing each plant parameter, component status and/or alarm, are mapped on displays (Figure 6). Thresholds are set for the period of time eye point stands in the detection area. If eye point stays in the detection area for a certain period of time, then it is judged that detection of the parameter or status within the area has been made. Similarly, view area detection mapping is also applied to a large display panel.

In accordance with the model human processor [3], perceptual processor was considered to be around 100 msec., so the threshold is set as 100 msec. This means it is judged that detection of the parameter or status is made if eye points stay within specific area for more than 100 msec.

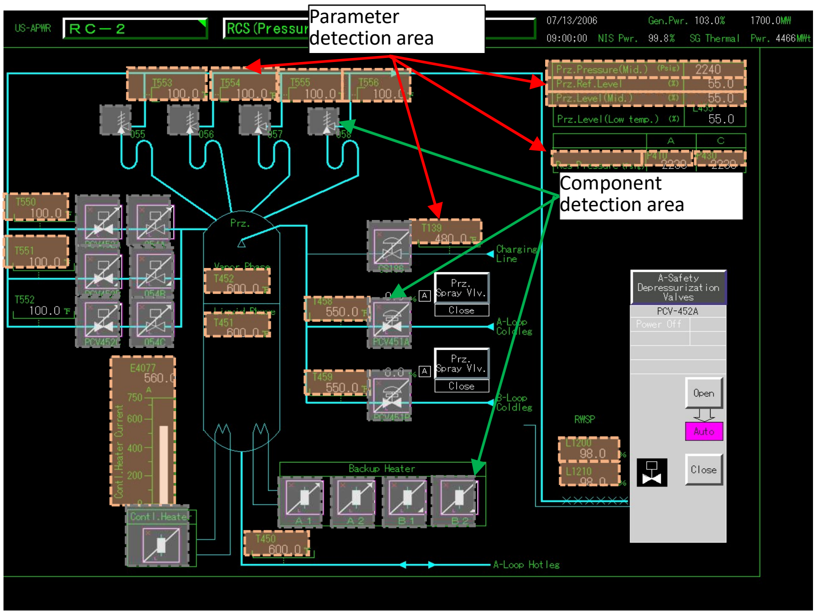

Figure 6. View area detection mapping (operation display[1])

Understanding and decision-making

As for understanding and decision-making, information sets, which are necessary for plant diagnosis, are pre-identified, then speaking information in the conversation log are compared with the information set. If the key verbal phrase in the information set is found in the conversation log, then understanding/decision-making is assume to be successfully made.

The pre-determined information set includes the name of plant incident, component names, tagging numbers, plant parameter names and alarms defined in the plant diagnosis part of the operating procedure. The parameter set of the pre-defined information set is basically identical to the parameter set for detection.

There is no execution action (e.g., manual controls of components and valves) during plant diagnosis. Display touch operations (e.g., switching other display through navigation functions) are performed for monitoring plant parameters, components and/or alarms so that display touch log are used for detection process in combination with view area information. So, for digital MCR plants with computer-based procedure system (CBP), the logic should be simplified to check display touch log for indication of actuation of a specific procedure instead of comparing key words in conversation log.

In performance evaluation, achievement level of diagnosis should also be specified depending on completeness of the cognitive process. If the cognitive process is fully achieved, completeness of cognitive process is marked as "◎". If understanding and decision-making process was achieved but detection process was incomplete, completeness of cognitive process is marked as "○". Completeness of cognitive process is marked as "X" if all cognitive processes are insufficient. Figure 6 shows the overall automatic performance evaluation process for plant diagnosis.

Plant diagnosis relatively relies on operator's skill and knowledge rather than operating procedure. In plant diagnosis, operator actions are not always aligned to procedural steps but are performed based on operator's skills and knowledges. In this case, operators are trained by monitoring key parameters, component status and alarms described in the event diagnosis procedure. Key parameters can be pre-determined and performance measurement data can be selected as actions performed within a certain period of time after malfunction initiation (e.g., several ten minutes from simulator malfunction initiation) or operator selecting and entering a particular event mitigation procedure. Operator's speaking information from the conversation log is compared with pre-determined key words crucial to situation awareness for plant event diagnosis and specific event identifications to determine the completeness of diagnosis.

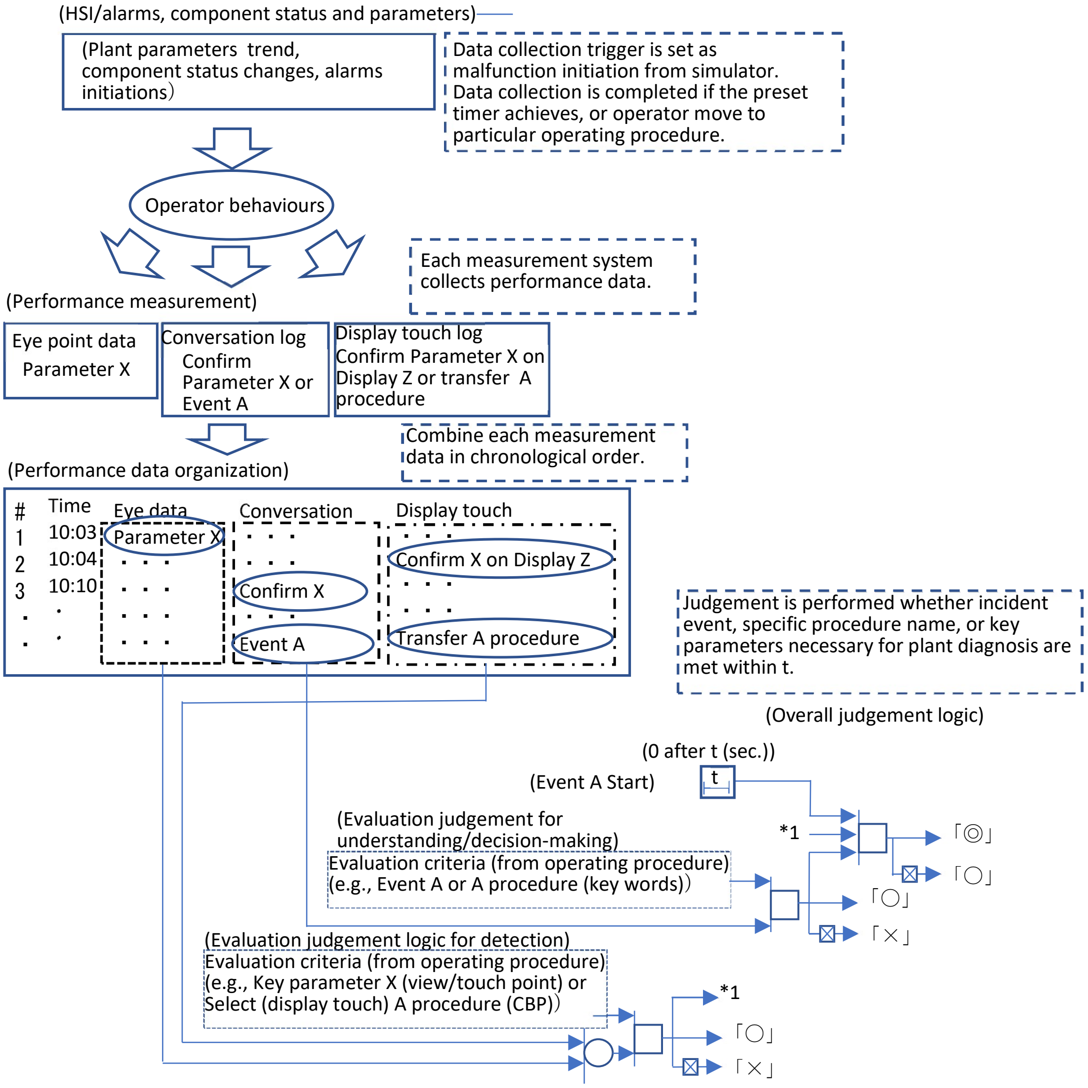

Figure 6. Overall automatic performance evaluation process for plant diagnosis

4. APPLICATION TO HUMAN RELIABITY ANALYSIS

4.1. Human reliability analysis

Human reliability is conducted through identification and characterization of the relevant performance shaping factors (PSFs), and evaluation of the likelihood of the failure mode/mechanism characterized by PSFs. Human performance data collected through the automated plant operator performance evaluation tool (POPET) can be used to evaluate the likelihood of the operator failure modes/mechanisms.

In this section, application of POPET described in chapter 2 to HRA is demonstrated. POPET can widely be applied to existing HRA methods including those that require cognitive failure mode (CFM) identifications. Human error event detection of the failure modes used in cognition error estimations of CBDTM (Caused Based Decision Tree Method)[4] and those used for omission and commission error estimations of THERP (Technique for Human Error Prediction) [5] is discussed.

For each operational task in the accident scenario, POPET collects information relevant to each step within the cognitive process as well as the consequence of cognition (i.e., operator's action recorded in action logs). When a human error occurs, which is detected as deviation of display touch logs (action logs) from the expected actions described in the procedure, the human error event and its failure mechanism/mode can be automatically identified using human error identification logics. The automated human error identification logics for failure mechanism/mode used in CBDTM and THERP, and logics to measure response times used in HCR/OCR are described in Section 4.2.

4.2. Automatic human error identification and error mode classification logics

Human error detection logics to automatically identify human errors and to classify them into failure mechanism/mode defined in CBDTM/THERP have been investigated. Based on the predefined human error detection logic, human errors can be automatically detected and identified operator's performance data obtained by POPET. Human error detection logics for failure mechanism/mode defined in CBDTM and THERP are respectively shown in Table 1 and Table 2.

Table 1. Human error identification and error mechanism classification logics for CBDTM

Failure mode/ mechanism	**Failure mechanism definition**	**Failure detection logic**	**Example (SGTR, SGTL scenario case)**
Pca	Data not available	Monitoring parameters are not physically available in displays	N16 parameter or annunciator not available in HMI
Pcb	Data available but are not attended to	Eye-tracking data show operator has not attended to the key parameter	No evidence of operator watching N16 for certain period
Pcc	Data misread or miscommunicated	Conversation log does not match parameters required to be monitored per procedures	Conversation log of N16 monitoring confirmation step in the procedure does not match with N16 parameter value.
Pcd	Available information is misleading	Conversation log indicates conversation on degraded cue state. Cue or parameter values encountered are not stated in procedure.	Cue (operator's action logs) or conversation logs are different from those stated in procedure and expected cues. *Note: Expected cues besides those in the procedure are pre-defined and installed in the POPET.
Pce	Relevant steps in procedures are skipped	Action logs are inconsistent with the procedure steps	Action logs for SG isolations mismatch with the procedural steps lack several steps which are defined in the SG isolation procedures.
Pcf	Errors made in interpreting instructions	Conversations log do not match with the procedure instructions	Conversation log includes instructions to partially close valves (no fully close in key words are found in conversation logs) while the procedure step requires those valves fully closed.
Pcg	Errors made in interpreting diagnostic logic	Conversation log or action log show incorrect procedure or step has been selected	Instruction of Valve C closed after closure of Valve A was found in the conversation logs while the closure of Valve C shall be closed followed by the closure of Valve B per the procedure steps / logics.
Pch	Crew deliberately violate procedure	Both of conversation log and action logs show attempt of alternative means deviating from procedure.	Different valves openings (e.g., auxiliary spray valves) in the conversation logs are included in the depressurizing operation steps defined in the procedure.

Table 2. Human error identification and failure mode classification logic for THERP

Failure mode	**Failure mode definition**	**Failure detection logic**	**Example (SGTR, SGTL scenario case)**
Omission errors	Omit actions required in the procedure	Action logs are inconsistent with those in the procedure steps	Action logs for SG isolations mismatch with the procedural steps in the procedure.
Commission errors	Errors in selecting displays	Action logs of touch displays mismatch those in the procedures.	The different displays for confirming N16 radiation alarm are selected from those required in the procedure steps.
	Errors in reading quantitative information from displays	Conversation logs do not match parameters required to be monitored per procedure. And Eye-tracking data show operator attended to the key parameter	Conversation log does not match with N16 parameter value. Or Eye-tracking data detected operator has watched N16 parameter.
	Errors in manual control actions	Action logs are inconsistent with the procedure steps.	Action logs for SG isolations mismatch with the defined procedural steps.

POPET can also be applied to evaluate parameters used in the HCR/ORE method[4]. HCR/ORE method predicts human error probability of cognition based on the median response time $T_{1/2}$, logarithmic standard deviation of the response time σ, and time window for operators to respond T_w. $T_{1/2}$ and σ are specific to types of accident scenarios and need to be estimated from operator performance. Using the performance evaluation system describe in chapter 3, response time can be measured as the time between the plant disturbance has occurred and the time when successful diagnosis has been confirmed through the conversation log. After response time data has been accumulated for categories of accident scenarios, the median response time $T_{1/2}$ and it's deviation σ can be statistically evaluated or used to update the HRC/ORE data.

4.3. Challenge and discussions

There are following challenges for implementation.

(1) Calibrations of performance measurement
Reference[1] demonstrated prototype test and presented test results showed good agreement with results obtained from conventional evaluation method (manual evaluations). Reference[1] also discussed about the further improvement points to measure human performance behaviors such as how veteran operators fast monitoring activities were detected with more accurate manners. There were no significant challenges for those improvements.

(2) Enhancement of automatic classification algorithms
Algorithms to automatically classify cognitive human error mode/mechanisms of CBDTM and THERP has been studied. Most human error modes can be classified solely depended on operator's performance logs in accordance with human error mode definitions, and for those human error modes, algorisms are straightforward to develop. Such human error modes include skipping, misreading and not addended to plant parameters. On the other hand, some human error modes that involve judgments or interpretations within the operator's mind (i.e., misleading information (Pcd) and deliberately violate procedure (Pch)) could be difficult to classify solely from measured data and existing procedure steps. In general, these human error modes are difficult to detect through detection logics and therefore error probabilities could be underestimated. For such human error modes, additional algorithms such as checking operator's conversation logs with alternative expected operation sequences/keywords, as analysts typically do when classifying through operator interviews and debriefings, are necessary for concrete judgements.
The automatic classifications algorithms would be enhanced by accumulating success path besides operating procedures as alternative expected cues and adding them as success criteria in the automatic classification algorithms.

(3) Estimations of human error probabilities
The estimation of human error probabilities (HEPs) used in HRA using automatic classification algorithms has not yet been implemented. Comparison of the proposed method against the conventional HEP estimation method for a series of training scenarios remains as future work.

(4) Dependency analysis
Dependency analysis needs to be executed when HEP estimations are performed. POPET can collect operator's cognitive process errors so several aspect of dependency types, such as consequential and context dependencies should be evaluated.

5. CONCLUSION

A study and development of an automated Operator Performance Measurement Tool (POPET) to evaluate operators performance in a simulator training is described. POPET acquires various data related to operator's cognitive and action behaviors using eye tracking measurement, voice recognition devices and operation logs that are archived with a simulator system, then automatically evaluates operator performances by comparing those collected data against pre-defined task patterns and task criteria.

POPET can also be applied to the automatic human error identifications and error mode classifications used in HRA using the collected operator's performance data and the automatic judgement algorithms. Especially as for cognitive errors, of which the error mode evaluation normally has to rely on operator's memory and could be subjective, POPET utilizes multidimensional cognitive performance data (i.e., eye tracking data, verbal conversational data, and action logs) to objectively evaluate cognitive error modes.

In summary, POPET is capable of automatically identifying operator's human errors and classifying error modes used in CBDTM and THERP. POPET can support error mode classifications which are currently performed in manually, and reduce analysts workload as well as makes human error rate more reliable based on objective data and transparent judgement criteria.

References

[1] K. Mashio and T. Hotta et al. Development of Online Performance Evaluation System of Nuclear Power Plant System. Journal of Nuclear Science and Technology. 59, 518-533, 2022.

[2] J. Xing, Y.J. Chang and J. DeJesus Segarra. Integrated Human Event Analysis System for Event and Condition Assessment (IDHEAS-ECA). NUREG-2256, U.S. Nuclear Regulatory Commission, 2022.

[3] S.K. Card, T.P. Moran and A. Newell. The Psychology of Human-Computer Interaction. Lawrence Eribaum Associates, Inc, Mahwah, NJ, 1983

[4] A. Swain and H. Guttmann. Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Application. NUREG/CR-1278, U.S. Nuclear Regulatory Commission, 1983

[5] G.W. Parry et.al. An Approach to the Analysis of Operator Actions in Probabilistic Risk Assessment. EPRI TR-100259, Electric Power Research Institute, 1992