ESTIMATING HEPS FROM THE OPERATIONAL EXPERIENCE OF DOMESTIC NUCLEAR POWER PLANTS – A FRAMEWORK AND CASE STUDIES

Jinkyn Park[1], Yochan Kim[1], Inseok Jang[2], and Wondea Jung[1]

[1] *Integrated Safety Assessment Division: 989-111 Daedeok-daero, Yuseong-gu, Daejeon, 34057, Republic of Korea, kshpjk@kaeri.re.kr, yochankim@kaeri.re.kr, wdjung@kaeri.re.kr*
[2] *Instrumentation and Control/Human Factors Division: 989-111 Daedeok-daero, Yuseong-gu, Daejeon, 34057, Republic of Korea, isjang@kaeri.re.kr*

In this paper, a novel framework is proposed, which allows us to systematically estimate HEPs (Human Error Probabilities) from the operational experience of domestic NPPs (Nuclear Power Plants). In addition, the feasibility of the proposed framework is investigated through several case studies based on the unexpected reactor trip events of domestic NPPs, which are caused by diverse human errors. As a result, it is expected that the proposed framework could be useful for estimating more realistic HEPs compared to those from simulated environments being collected from a full-scope training simulator.

I. INTRODUCTION

One of the working definitions on a socio-technical system refers to it as a system that requires the interaction between society's complex infrastructures and human behaviors in diverse workplaces. In this regard, it can be said that typical socio-technical systems include NPPs (nuclear power plants), chemical/petro-chemical plants, railway systems, land transportation systems, and maritime transportation systems. For example, it is evident that an aviation industry belongs to the socio-technical system because the operation of commercial airplanes requires intensive interactions between infrastructures (e.g., airports, air traffic control centers, and airplanes) and stakeholders (e.g., airport managers, air traffic controllers, and cockpit pilots). This means that one of the critical factors affecting the operational risk of socio-technical systems is the reliability of operating personnel who are actually running it. Therefore, it is unavoidable that the effect of operating personnel on the risk of socio-technical systems should be properly incorporated during its assessment. For this reason, a PSA (Probabilistic Safety Assessment) or PRA (Probabilistic Risk Assessment), which has been widely used for several decades, considers a diverse spectrum of human actions for quantifying the operational risk of NPPs. To this end, various kinds of HRA (Human Reliability Analysis) techniques were developed based on HRA data including not only HEPs (Human Error Probabilities) but also other information including the effect of error-forcing contexts (e.g., PSFs; Performance Shaping Factors or PIFs; Performance Influencing Factors) on the associated HEPs should be available to HRA practitioners.

In this light, many researchers have spent huge amounts of effort in providing HRA data to HRA practitioners, of which the contents are collected from several sources of information such as (1) operational experience data based on event reports (e.g., maintenance reports, periodic test reports, near miss reports, and incident reports), (2) full-scope training simulators, (3) laboratory experiments based on partial-scope simulators or mockups, (4) expert judgments, and (5) interviews with subject matter experts [1-3]. However, since most initiating events being considered in the PSA have an extremely rare frequency, it is unrealistic to obtain sufficient HRA data from other information sources except full-scope training simulators [4-6]. Nevertheless, it is also true that HRA data obtained from the operational experience of NPPs are needed in parallel with those from full-scope training simulators. This is because the use of full-scope training simulators is an alternative solution to resolving a difficulty in extracting HRA data from operational experience data. In other words, if reliable operational experience data (e.g., near miss or incident reports) are sufficient, it is possible to secure more realistic HRA data reflecting actual working environments.

In order to address this issue, in this paper, a novel framework is proposed, which allows us to systematically estimate HEPs from the operational experience of domestic NPPs. In addition, the feasibility of the proposed framework is investigated through several case studies based on the unexpected reactor trip events of domestic NPPs, which are caused by diverse human errors. As a result, it is expected that the proposed framework could be useful for estimating more realistic HEPs compared to those from simulated environments.

II. FULL-SCOPE TRAINING SIMULATOR AND OPERATIONAL EXPERIENCE AS HRA DATA SOURCES

As briefly stated in the previous section, HRA practitioners need to have sufficient amount of HRA data that allow them to soundly estimate HEPs after understanding the characteristics of human errors. In this regard, Table I shows pros and cons in extracting HRA data from two kinds of representative information sources: operational experience (e.g., event reports) and a full-scope training simulator.

TABLE I. Comparing pros and cons in extracting HRA data from event reports and full-scope training simulators; reproduced from Ref. [7]

Data source	Pros	Cons
Event reports	• Enable to secure more realistic HRA data reflecting a real task environment • Free from the fidelity problem • Enable to obtain HRA data pertaining to routine tasks	• Not easy to extract sufficient HRA data • Need a careful translation due to uneven contents and descriptions • Difficult to understand task contexts resulting in a human error
Full-scope training simulators	• Enable to simulate rare events • Enable to observe the variation of human performance with respect to diverse task contexts	• Require a huge amount of resources (e.g., budget) • Not easy to secure sufficient times for using a training simulator • Need to consider the validity of simulation results

As can be seen from Table I, full-scope simulators provide distinctive advantages in collecting HRA data, such as observing the variation of human performance with respect to diverse task contexts, which is usually not identifiable and/or accessible from event reports. Nevertheless, the collection of HRA data from event reports is very important because of at least two reasons. The first one is the reality (or fidelity) of HRA data. That is, from the point of view of supporting HRA practitioners, event reports seem to be more attractive because they tell us the actual responses of human operators who were faced with real events. Although the results of existing studies suggest that the overall tendencies of human behaviors as observed from simulated conditions are comparable to those from real conditions [8, 9], it is still evident that HRA data available from the analysis of event reports could be more realistic than those from full-scope training simulators.

Another reason is that the event reports are still valuable source for collecting HRA data for routine tasks being conducted in a normal operating condition (e.g., periodic tests or calibrations). For example, the IAEA specified three kinds of task types that are needed to be emphasized from the PRA perspective [10]. They are (1) *Type A* tasks representing a maintenance and/or testing that can degrade the availability of a given system, (2) *Type B* tasks pertaining to the direct triggering of initiating events (e.g., an unexpected reactor trip caused by a periodic test), and (3) *Type C* tasks specifying crucial human actions for responding DBAs (Design Basis Accidents), which are usually prescribed in AOPs (Abnormal Operating Procedures) and EOPs (Emergency Operating Procedures). Unfortunately, since most full-scope training simulators are not able to be used for collecting HRA data for *Types A* and *B* tasks, the analysis of event reports is crucial for securing HRA data. From the above-mentioned reasons, it is very interesting to point out that the coverage of HRA data collection seems to be clearly distinguished as depicted in Fig. 1.

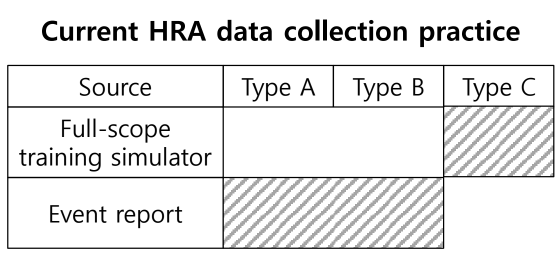

Fig. 1. Current HRA data collection practices; adopted from Ref. [7]

Under the current HRA data collection practice, it is very seldom to have HRA data pertaining to *Type C* tasks, which are extracted from event reports. This may imply that, from the point of view of HRA practitioners, this situation could become a source of uncertainties for estimating reliable HEPs. In other words, it would be careful for HRA practitioners to directly use HRA data collection from full-scope training simulators because they are not able to make sure their fidelity compared to real situations. In this regard, as shown in Fig. 2, it will be very helpful to unravel this problem if we are able to collect HRA data for *Type C* tasks. In other words, if there is a framework that allows us to crop the HEPs of *Type C* tasks from event reports, it is possible to give more reasonable guidelines to HRA practitioners in using HRA data extracted from full-scope training simulators. For this reason, a novel framework is proposed in this study.

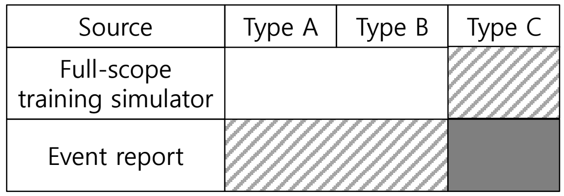

Fig. 2. Bridge between two kinds of HRA data sources, adopted from Ref. [7]

III. FRAMEWORK TO ESTIMATE AN HEP FROM THE REVIEW OF EVENT REPORTS

The definition of a probability is "how likely it is that something will happen [11]." In terms of an HEP, therefore, it is very straightforward to use Eq. (1) to quantify it.

$$HEP_i = \frac{m_i}{n_i} \qquad (1)$$

Here, HEP_i means the HEP of the i_{th} task. In addition, m_i and n_i denote the number of human errors observed and the number of trials for the performance of the i_{th} task (i.e., a demand of the i_{th} task). Actually, this formula is the direct reflection of a traditional assumption such that human operators will show similar HEPs if they have to accomplish identical tasks under a specific task environment [12, 13]. This implies that HEPs can be soundly estimated if the demand of a task can be counted in a systematic manner. In this regard, Table II shows an example about how to calculate an HEP from the review of event reports.

TABLE II. Calculating an HEP from event reports; reproduced from Ref. [14]

Information content	Description from an event report
Error description	During an in-cave operation to load active material into waste flasks, a piece of highly active waste was places in the wrong flask.
Operating history	4 years
Task frequency	20 loading tasks were carried out per a week; Loading tasks have been done for 26 weeks per a year.

As can be seen from Table II, an event report issued from a nuclear reprocessing facility described that a human error was observed during an in-cave operation because a human operator put radioactive materials into a wrong flask (e.g., a wrong flask selection). Accordingly, it is evident that $m_i$is 1. In addition, the operation history revealed that the task demand of a waste flask loading (i.e., n_i) is 2,080 because of: 20 (loading tasks/week) × 26 (weeks/year) × 4 (years). Therefore, the HEP of the wrong flask selection becomes 4.81E-4 (=1/2080). This means that there are two kinds of key challenges in calculating HEPs from event reports: (1) specifying the catalog of generic task types applicable to the analysis of event reports, and (2) determining the task demand of a given generic task type.

III.A. Generic task types

The first challenge is to develop the catalog of generic task types that can be used to properly distinguish the nature of a task to be carried out by human operators. To this end, task types and the associated human error modes proposed in Ref. [15] are adopted in this study (Table III). As can be seen from Table III, in total 23 task types belonging to 7 cognitive activities are summarized with the associated human error modes, such as EOO (Error of Omission) and EOC (Error of Commission). It is to be noted that an EOC can be subdivided into detailed categories including WDEV (Wrong Device selection), WDIR (Wrong Direction), and WQNT (Wrong Quantity), if necessary. That is, the WDEV and WDIR represent human errors caused by selecting a wrong device and selecting an inappropriate control direction, respectively. In addition, the WQNT denotes a situation when human operators put a wrong control input to a control device even if these selection were correct (e.g., adjust the openness of a valve to 20% instead of a full close position).

TABLE III. Generic Task types and the associated human error modes; modified from Ref. [15]

Cognitive activity	Task type	Error mode
Information gathering and reporting – checking discrete state	Verifying alarm occurrence	EOO, EOC
	Verifying state of indicator	EOO, EOC
	Synthetically verifying information	EOO, EOC
Information gathering and reporting – measuring parameter	Reading simple value	EOO, EOC
	Comparing parameter	EOO, EOC
	Comparing in graph constraint	EOO, EOC
	Comparing for abnormality	EOO, EOC
	Evaluating trend	EOO, EOC
Response planning and instruction	Entering step in procedure	EOO
	Transferring procedure	EOO, EOC
	Transferring step in procedure	EOO, EOC
	Directing information gathering	EOO, EOC
	Directing manipulation	EOO, EOC
	Directing notification	EOO, EOC
Situation interpreting without explicit guide of document	Diagnosing	EOO, EOC
	Identifying overall status	EOO, EOC
	Predicting	EOO, EOC
Action – manipulation	Manipulating simple (push button) control	EOO, EOC (WDEV, WDIR)
	Manipulating simple (rotary) control	EOO, EOC (WDEV, WDIR, WQNT)
	Manipulating dynamically	EOO, EOC (WDEV, WDIR, WQNT)
Action – notifying/requesting to the outside of the MCR	Notifying to external agent	EOO, EOC
Other	Unauthorized control – unguided response planning and instruction	EOC
	Unauthorized control – unguided manipulation	EOC

III.B. Task demand

Once the type of a task being related to the human error of a given event report is identified, the next step is to determine the number of trials for a given task type (i.e., task demand). To this end, it is necessary to consider two kinds of opportunities, such as a procedure opportunity and task opportunity. That is, in the case of *Type C* tasks, human operators have to carry out a series of tasks to cope with off-normal conditions in a timely manner. However, since the off-normal conditions can be caused by diverse failures (e.g., signal or component failures) and DBAs (Design Basis Accidents), human operators are apt to feel a significant workload in conducting required tasks. For this reason, most *Type C* tasks are prescribed in various kinds of procedures, such as AOPs and EOPs. This means that the determination of a task demand for a specific task type should start from how many times a procedure that contains it. For example, let us assume that a human error has occurred in verifying the initiation of an *Alarm A* that is included in an AOP_1. According to Table III, the type of this task is 'Verifying alarm occurrence,' and its error mode is EOO. In this case, the task demand of 'Verifying alarm occurrence' can be estimated through multiplying the number of previous successes in conducting AOP_1 without any human errors by the number of identical task types involved in the AOP_1. More detailed explanations in calculating task demands for *Type C* tasks can be found from Ref. [7].

IV. CASE STUDY

In order to investigate the applicability of the proposed framework, event reports being archived in NEED (Nuclear Event Evaluation Database) are reviewed in detail. NEED is one of the working databases operated by the nuclear regulatory body of Republic of Korea (KINS; Korea Institute of Nuclear Safety). NEED aims to provide valuable information that can be used to prevent the recurrence of similar incidents in NPPs [16]. From this purpose, when a safety significant incident has occurred, the KINS creates a special inspection team comprised of several subject matter experts along with the nature and importance of the incident. The primary role of the special inspection team is to identify the underlying causes of the incident, from which remedial actions (or countermeasures) for preventing the recurrence of similar incidents can be effectively suppressed. Since 2002, when an investigation is completed, the KINS has uploaded all kinds of investigation results to NEED, which is connected to the Internet. Thus, anyone can access detailed investigation results including the name of investigators, the initiation and progression of an incident, and the catalog of remedial actions being proposed by the investigators.

Here, it is very important to point out that NEED contains detailed event reports that could be used to estimate the task demand of a *Type C* task. In this vein, in total 193 incidents that have occurred in the period from January of 2002 to December of 2013 are reviewed. As a preliminary result, it is identified that ten human errors have occurred during the performance of *Type C* tasks. Table IV summarizes the characteristics of the human errors with respect to the associated task types and error modes.

TABLE IV. Human errors identified during the performance of *Type C* tasks

ID	Task type	Error mode
1	Manipulating dynamically	EOC (WDEV)
2	Manipulating dynamically	EOC (WQNT)
3	Manipulating dynamically	EOC (WQNT)
4	Manipulating simple (discrete) control	EOC (WDIR)
5	Manipulating simple (discrete) control	EOC (WQNT)
6	Manipulating simple (discrete) control	EOO
7	Unguided manipulation	EOC
8	Unguided manipulation;	EOC
9	Unguided response planning and instruction	EOC
10	Unguided response planning and instruction	EOC

For example, according to NEED, a significant incident happened in *Hanul Unit* 4 in October 10, 2003, which was initiated by the abnormal event of *Turbine generator trip* [16]. Unfortunately, human operators working in the main control room (MCR) failed to properly control the feedwater flow of steam generators (SGs), which is one of the required tasks institutionalized in an AOP to be carried out under the situation of the *Turbine generator trip*. As a result, from this event report, it is possible to say that an EOC has occurred in the course of conducting a proceduralized task belonging to the category of *Manipulating dynamically*. In addition, the total number of task demands for *Manipulating dynamically* can be counted along with the proposed framework. As a result, Table V shows a part of preliminary HEPs for *Type C* tasks estimated from the analysis of domestic event reports. It is to be noted that the HEPs of EOOs and EOCs imply their 95 percentile calculated by Bayesian update technique with Jeffery's non-informative prior for Beta distribution.

TABLE V. A part of preliminary HEPs for Type C tasks estimated from domestic event reports

Subtask type	Task demand (n_i)	$m_i(EOO)$	$m_i(EOC)$	$p(EOO)$	$p(EOC)$
Verifying alarm occurrence	195	0	0	9.900E-03	9.900E-03
Verifying state of indicator	3598	0	0	6.001E-04	6.001E-04
Comparing parameter	674	0	0	2.900E-03	2.900E-03
Evaluating trend	316	0	0	6.101E-03	6.101E-03
Entering step in procedure	713	0	0	2.800E-03	2.800E-03
Manipulating simple (discrete) control	1034	1	2	3.700E-03	5.301E-03
Manipulating dynamically	35	0	3	5.371E-02	1.883E-01

V. DISCUSSION AND CONCLUSION

It is evident that the contribution of human errors to the safety of socio-technical systems is very critical. For this reason, it is important for HRA practitioners to provide reliable HRA data including HEPs. Although a full-scope simulator can be used to collect valuable HRA data, it is still necessary to extract HRA data from the review of operational experience. If so, it is possible to expect several benefits, such as the use of HRA data gathered from the operational experience of domestic NPPs as reference information to clarify the appropriateness of those collected from full-scope simulators.

It is true that the framework explained in this study still has several limitations. For example, as can be seen from Table IV, four human errors related to the category of the *unauthorized control* are identified from the analysis of event reports. This implies that human operators carried out inappropriate actions that are not prescribed in procedures (e.g., AOPs or EOPs). Unfortunately, in the case of the *unauthorized control*, it is not possible to determine the associated task demand because of its nature (i.e., unauthorized controls can happen whenever human operators want to do something based on their own decisions). Accordingly, it is necessary to come up with alternative solutions such as considering the occurrence rate of unauthorized controls (e.g., 2.00E-4/hr).

However, it is strongly anticipated that the uncertainty in determining the annual frequency of an abnormal event could become tolerable if we are able to use a reliable component failure database. In addition, the comparison of two kinds of HEPs, one from operational experience data and the other from full-scope simulators, could become a good source of information that allows us to clarify how to use simulator-based HRA data for conducting a practical HRA. From this perspective, the results of this study seem to be meaningful because they could be a technical basis for securing more reliable HEPs from event reports.

ACKNOWLEDGMENTS

This work was supported by Nuclear Research & Development Program of the National Research Foundation of Korea (NRF) grant, funded by the Korean government, Ministry of Science, ICT & Future Planning (Grant Code: 2012M2A8A4025991).

REFERENCES

1 International Atomic Energy Agency, *Collection and classification of human reliability data for use in probabilistic safety assessment*, IAEA-TECDOC-1048, Vienna, Austria (1995)
2 A. Isaac, S. T. Shorrock, and B. Kirwan, "Human error in European air traffic management: the HERA project," *Reliability Engineering and System Safety*, **75(2)**, 257 (2002)
3 Nuclear Energy Agency, *HRA data and recommended actions to support the collection and exchange of HRA data*, Report of Working Group on Risk Assessment, NEA/CSNI/R(2008)9, Vienna, Austria (2008)
4 J. Chang, and E. Lois, "Overview of the NRC's HRA data program and current activities," PSAM 11, Helsinki, Finland (2012)
5 L. Lederman, "Accident sequences sensitive to human errors," *Reliability Engineering and System Safety*, **22**, 269 (1988)
6 N. Stanton, "Simulator: a review of research and practice," *Human Factors in Nuclear Safety* N. STANTON, Ed., Taylor & Francis, London (2005)
7 J. Park, Y. Kim, and W. Jung, "A framework to estimate task opportunities from the operational experience of domestic nuclear power plants," *Safety Science*, **88**, 146 (2016)
8 W. H. Gibson, B. Hickling, and B. Kirwan, "Feasibility study into the collection of human error probability data," Eurocontrol Experimental Center, EEC Note No. 02/06 (2006)
9 J. Park, and W. Jung, 2007. "OPERA - a human performance database under simulated emergencies of nuclear power plants," *Reliability Engineering and System Safety*, **92**, 503 (2007)
10 International Atomic Energy Agency, "Human reliability analysis in probabilistic safety assessment for nuclear power plants," IAEA-SS-50-P-10, Vienna, Austria (1995)
11 www.dictionary.cambridge.org
12 E. A. Fleishman, and L. C. Buffardi, "Predicting human error probabilities from the ability requirements of jobs in nuclear power plants," *Nuclear Safety: A Human Factors Perspective,* J.(MISUMI, B. WILPERT, and R. MILLER Ed., Taylor & Francis Ltd. (1999)
13 J. Reason, 2000. "Human error: models and management," *British Medical Journal*, **30**, 768 (2000)
14 S. Taylor-Adams, and B. Kirwan, "Human reliability data requirements," *International Journal of Quality and Reliability Management,* **12(1)**, 24 (1995)
15 J. Park, Y. Kim, and W. Jung, "A framework to estimate HEPs from the full-scope simulators of NPPs: unsafe act definition, identification and quantification," Korea Atomic Energy Research Institute, Daejeon, Korea (2016)
16 Nuclear Event Evaluation Database (2016). Available at: http://opis.kins.re.kr/opis?act=KEOBA1100R