Expert Elicitation for Estimating Inter-Organizational Communication Error Probabilities in NPP operations

Ilgeun Song [a], Taewon Yang [b], Jonghyun Kim [a*]
[a]*Department of Nuclear and Quantum Engineering, Korea Advanced Institute of Science and Technology, 291 Daehak-ro, Yuseong-gu, Daejeon, Republic of Korea, 34141*
[b]*Department of Nuclear Engineering, Chosun University, 10 Chosundae 1-gil, Dong-gu, Gwangju, Republic of Korea, 61452*
[*]*Corresponding author: jonghyun.kim@kaist.ac.kr*

***Keywords :** Human Reliability Analysis (HRA), Communication Error Probability (CEP), Inter-organizational Communication

1. Introduction

In Human Reliability Analysis (HRA), communication is crucial due to the inherent need for seamless interaction between operators during the operation of a Nuclear Power Plant (NPP). This interaction is essential for ensuring that all actions are coordinated and aligned with safety protocols, which is vital for preventing accidents. The complexity of communication increases in emergency situations, particularly in multi-unit NPP scenarios, where multiple organizations such as the Emergency Operation Facility (EOF), Technical Support Center (TSC), and Operational Support Center (OSC) are involved. This heightened complexity significantly raises the risk of communication errors, which can have severe consequences for plant safety.

The importance of addressing communication-related human errors in NPPs has been well-documented. For instance, Hirotsu et al. (2001) examined 885 incidents in Japanese NPPs and identified 193 as human error cases, with 13% due to written communication failures and 5% due to verbal communication failures [1]. Similarly, Lee (2007) reported that poor communication contributed to 20 out of 27 incidents in Korean NPPs from 2001 to 2007 [2]. These findings underscore the pervasive nature of communication-related issues in the nuclear industry and highlight the need to quantify CEPs as part of HRA. However, an extensive literature review reveals that existing research does not adequately address inter-organizational CEPs in multi-unit emergency situations. This gap highlights the need for focused analysis, and this paper aims to fill that gap by quantifying these error probabilities, which is crucial for enhancing safety in multi-unit NPP scenarios. Given the absence of empirical data on inter-organizational communication errors—largely due to statistically insufficient data and the complexity of capturing such interactions in real-time—alternative methods for quantification must be employed. This study employs expert elicitation, specifically Cooke's Classical method (1991), a well-established approach for deriving probability estimates in situations where direct data is scarce [3]. By systematically collecting and aggregating expert judgments, this method provides a robust framework for estimating communication error probabilities.

In this paper, the expert elicitation method is described in detail, with a particular emphasis on Cooke's Classical method. Through this approach, the study estimates inter-organizational CEPs, offering valuable insights for mitigating communication-related risks in multi-unit emergency operations.

2. Expert Elicitation Method

According to Colson and Cooke (2018), expert elicitation approaches can be categorized into two primary types: behavioral and mathematical [4]. Behavioral methods, like the Delphi method, involve multiple rounds of expert interaction to reach consensus but are often costly, time-consuming, and susceptible to groupthink. In contrast, mathematical approaches aggregate individual expert judgments analytically, without requiring direct interaction among experts.

Cooke's Classical model, a type of mathematical approach, was chosen for this study due to its simplicity, cost-effectiveness, and practical structure. It systematically aggregates expert judgments into a single probability distribution, making it a practical choice for estimating inter-organizational CEPs in emergency operations where empirical data is limited. This method allows experts to provide their assessments independently, avoiding the logistical challenges and high costs of gathering all experts together.

This model focuses on two main types of questions: calibration and target questions. Calibration questions have known outcomes to the analyst but are unknown to the experts, while target questions require new assessments. Experts provide their 5th, 50th, and 95th percentiles for both types of questions, offering a range of uncertainty around their estimates. The reliability of their uncertainty judgments is measured by how closely these percentiles match the actual answers of the calibration questions.

The model assigns a calibration score to each expert, reflecting their accuracy in predicting the outcomes of the calibration questions. Additionally, an information score is calculated to assess how concentrated or dispersed an expert's judgments are. These individual

expert judgments are then combined into a single, aggregated estimate through a performance-based weighting system, where experts with higher calibration and information scores have more influence on the final result. This approach ensures that the combined judgment is both accurate and reliable, making it an invaluable tool for contexts like multi-unit emergency operations.

3. Elicitation Process

3.1. Target Questions

To classify the communication acts in our study, we applied speech act theory, which traditionally categorizes communication into four types: Request, Report, Declaration, and Acknowledgment [5]. However, for the purposes of our analysis, we chose to include an additional category—Recovery Failure—given its significance in inter-organizational communication contexts. These five categories were then used to structure the target questions, with a focus on estimating the CEPs for each speech act in inter-organizational settings.

3.2. Calibration Questions

To ensure the relevance and accuracy of expert assessments, the calibration questions were based on well-established intra-organizational CEPs [6] and human error probabilities from THERP [7]. The calibration set consisted of 8 questions designed to observe substantial differences in calibration. Of these, 5 questions were derived from inter-organizational CEPs, while the remaining 3 were based on THERP human error probabilities, specifically covering error of commission, error of omission, and selection error. Additionally, to familiarize the experts with our speech act concept, we provided examples of each speech act category, ensuring that the participants had a clear understanding of the framework being used. The intra-organizational CEPs, detailed in Table I, served as benchmarks for assessing communication errors within single organizations, making them ideal for verifying the expertise of the participants.

Table I. Intra-organizational CEPs

Speech act	Error rate		
	5%	Mean	95%
Request	1.10E-02	1.59E-02	2.14E-02
Report	8.77E-03	1.31E-02	1.82E-02
Declaration	6.44E-06	1.63E-03	6.27E-03
Acknowledgment	5.00E-06	1.27E-03	4.87E-03
Recovery failure	4.01E-03	3.33E-02	8.54E-02

3.3. Selection of Experts

The next step involved selecting survey participants. To ensure a broad range of perspectives, we recruited 12 experts with experience in HRA from various fields. Guidance from the Expert Judgment Policy Symposium and Technical Workshop suggests that including between 6 and 12 experts strikes an effective balance between capturing diverse expert opinions and maintaining the manageability of the elicitation process [8]. Our expert panel consisted of 2 participants from regulatory agencies, 3 from research groups, 3 from academic institutions, and 4 field operators. Their experience ranged from 7 to 40 years, with an average of 23 years.

3.4. Application of Elicitation

The elicitation was conducted face-to-face with the experts to enhance both reliability and accountability. During the session, if experts had any questions about how to complete the questionnaire, the analyst provided clarification to ensure that responses were accurate and consistent. The procedure began with an introduction covering research ethics, key terminologies, and detailed instructions on how to provide answers. Potential biases, such as common expert mistakes, were also discussed to minimize their impact on the results. To contextualize the responses, experts were asked to provide information about their professional background, including their group affiliation and years of experience in the field. After providing this information, they began by answering the calibration questions, followed by the target questions. Figure 1 shows an example of the questionnaire provided to the experts, which they used to record their assessments.

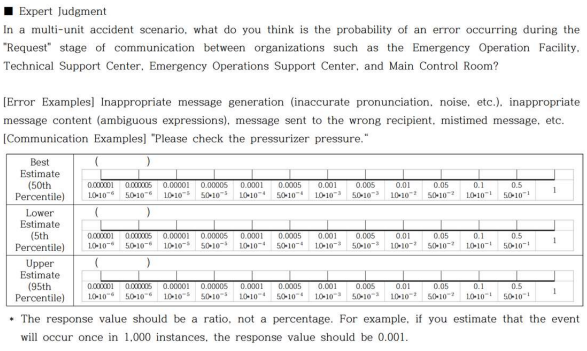

Fig. 1. Example of questionnaire

4. Results and Discussions

The evaluation of the calibration questions revealed an average of 1.92 errors per expert, particularly in questions related to intra-organizational CEPs, where the actual values did not fall within the 5% to 95% range. These errors were primarily attributed to the experts' lack of familiarity with CEPs, despite their extensive experience in HRA. Based on these results, the inter-organizational CEPs for each speech act were calculated using Vincent's R SW expert package [9]. The estimated

CEPs, derived from the expert input, are presented in Table II. The results suggest that inter-organizational CEPs are expected to increase compared to intra-organizational CEPs, with a particularly notable increase observed in the Acknowledgment. This suggests that communication between organizations is more prone to errors, likely due to differences in organizational culture, procedures, and possibly less frequent interaction compared to intra-organizational communication.

However, despite the increase in overall error probabilities, the relative ranking of error types remained consistent across both contexts: Recovery Failure > Request > Report > Declaration > Acknowledgement for intra-organizational, and Recovery Failure > Request > Report > Acknowledgement > Declaration for inter-organizational communication. This suggests that while the likelihood of errors may be higher in inter-organizational settings, the underlying patterns of vulnerability in communication remain similar.

Moreover, the wider statistical ranges observed—from the 5% to 95% confidence intervals—further highlight the variability and uncertainty in inter-organizational communication error probabilities. This greater variability could be attributed to external factors such as differing levels of trust, information sharing protocols, and the inherent challenges of coordinating between organizations. These factors introduce additional layers of complexity, making inter-organizational communication more susceptible to errors.

Table II. Inter-organizational communication error probability

Speech act	Error rate		
	5%	Mean	95%
Request	7.66E-04	6.84E-02	1.90E-01
Report	1.56E-03	3.18E-02	7.54E-02
Declaration	4.78E-04	2.42E-02	7.54E-02
Acknowledgment	6.13E-04	3.15E-02	7.54E-02
Recovery failure	1.06E-02	7.79E-02	1.56E-01

5. Conclusion

The analysis of inter-organizational CEPs reveals key insights essential for improving communication reliability in multi-unit NPP emergency operations. Inter-organizational communication is found to be more prone to errors than intra-organizational communication, likely due to differences in organizational culture, procedures, and the infrequency of interactions. However, the relative ranking of error types remains consistent, with Recovery Failures and Requests being the most error-prone acts in both settings.

Given these findings, prioritizing high-risk communication acts, particularly Requests and Recovery Failures, is crucial. Implementing more rigorous checks for Requests and developing shared recovery protocols could significantly reduce the likelihood of errors. The analysis also indicates a shift in the significance of Acknowledgment errors in inter-organizational contexts, pointing to specific vulnerabilities in message confirmation. This underscores the need for enhanced validation mechanisms to ensure accurate communication between organizations.

In conclusion, addressing these high-risk areas and tailoring strategies to the specific challenges of inter-organizational communication in multi-unit NPP emergency situations can significantly enhance the resilience and safety of these plants during critical scenarios.

6. Acknowledgments

This study was supported by the National Research Foundation of Korea (NRF) grant No. RS-2022-00144172.

REFERENCES

[1] Y. Hirotsu, K. Suzuki, M. Kojima, and K. Takano, "Multivariate analysis of human error incidents occurring at nuclear power plants: several occurrence patterns of observed human errors, " Cognition, Technology & Work, Vol. 3, No. 1, pp. 82–91, 2001.
[2] Y. H. Lee, The casebook on human error in nuclear power plant, Korea Atomic Energy Research Institute, 2007. (in Korean).
[3] R. M. Cooke, Experts in Uncertainty, Oxford University Press, 1991.
[4] A. R. Colson, and R. M. Cooke, "Cross validation for the classical model of structured expert judgment," Review of Environmental Economics and Policy, Vol. 12, No. 1, pp. 119-132, 2018.
[5] J. A. Goguen, and C. Linde, Linguistic methodology for the analysis of aviation accidents, No. NASA-CR-3741, NASA, 1983.
[6] J. Yun, T. Yang, J. Kim, and J. Kim, "Estimation of Communication Reliability for Human Reliability Analysis at Nuclear Power Plants," Transactions of the Korean Nuclear Society Autumn Meeting, Gyeongju, Korea, pp. 1-5, Oct. 26-27, 2023.
[7] A. D. Swain, and H. E. Guttmann, Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR-1278, U.S. Nuclear Regulatory Commission, 1983.
[8] R. M. Cooke, and K. N. Probst, Highlights of the Expert Judgment Policy Symposium and Technical Workshop, Resources for the Future, Mar. 13-14, 2006.