Updating the Dependency Equations Proposed in THERP Using Experimental Data

Taewon Yang[a], and Jooyoung Park[b*]
[a] Korea Advanced Institute of Science and Technology, Daejeon, Republic of Korea,
Taewon.Yang@kaist.ac.kr
[b] Korea National University of Transportation, Uiwang, Republic of Korea,
Jooyoung.Park@ut.ac.kr
[*] Corresponding Author

Abstract: Dependency analysis in human reliability analysis (HRA) adjusts the failure probability of a subsequent human action according to the influence of a preceding action. The Technique for Human Error Rate Prediction (THERP) characterizes this relationship through five discrete levels of dependence—zero, low, moderate, high, and complete—each associated with a closed-form equation for the conditional human error probability. However, the degree-of-dependence values underlying these equations were assigned without a documented technical basis, and the THERP handbook itself recommended that they be revised as supporting data became available. Despite this, the equations have been applied without modification for over four decades. This study presents a preliminary analysis aimed at updating the THERP dependency equations using experimental data collected under the Simplified Human Error Experimental Program (SHEEP) framework developed by Idaho National Laboratory. The five THERP failure equations are first unified into a single generalized expression in which the degree of dependence appears as a continuous interpolation weight between the zero-dependence and complete-dependence bounds. An estimation formula is then derived that computes the degree of dependence directly from experimental counts of joint and preceding-task failures. The approach is demonstrated using human error data collected from student operators performing a startup operation scenario on the Compact Nuclear Simulator. Degrees of dependence are estimated for combinations of consecutive major human actions, and six candidate dependency factors—area of interest, paper procedure page, cognitive activity, task type, component type, and error type—are coded for each action pair and regressed against the estimated degrees of dependence. The regression results indicate that area of interest, cognitive activity, task type, and error type are statistically significant at the 0.05 level, suggesting that factors beyond those traditionally considered in THERP may influence dependence between human actions. The findings demonstrate the feasibility of empirically estimating degrees of dependence from simulator data and provide a basis for refining the THERP dependency model. Further analyses with additional data and candidate factors are ongoing, and more conclusive results will be presented at the conference.

1. INTRODUCTION

Dependency analysis is a method used in human reliability analysis (HRA) to adjust the failure probability of a subsequent action according to the influence of a preceding action. The Technique for Human Error Rate Prediction (THERP) [1], one of the most widely used HRA methods, originally introduced a dependency analysis approach based on five distinct levels of dependence: Zero Dependence (ZD), Low Dependence (LD), Moderate Dependence (MD), High Dependence (HD), and Complete Dependence (CD). Each level is associated with a specific equation that adjusts the failure probability of the subsequent action. For example, the LD equation assumes a degree of dependence of 5%, interpolating between the basic human error probability (the ZD value) and 1.0 (the CD value) to obtain the conditional failure probability. However, the values underlying these dependency equations were assumed without any supporting empirical or theoretical basis, and the THERP handbook itself recommended that the values be updated in future work. Despite this, the THERP equations have continued to be applied without modification, retaining their original assumptions regarding the degrees of dependence. Although many HRA researchers have sought to develop improved dependency quantification methods, the field still relies on the THERP equations; the alternative approaches proposed to date have remained largely at the research stage and have yet to replace the original formulation.

This study aims to update the THERP dependency equations using experimental data. The data were collected through the Simplified Human Error Experimental Program (SHEEP) framework [2], an HRA data collection framework proposed by Idaho National Laboratory (INL). The SHEEP framework focuses on simplified simulators and non-expert operators in order to complement full-scope HRA data collection studies. The SHEEP database comprises human error and performance data collected from 36 student operators and 36 professional operators, obtained using the Rancor Microworld simulator developed by INL and the Compact Nuclear Simulator (CNS) developed by the Korea Atomic Energy Research Institute (KAERI). Through this analysis, the study ultimately seeks to answer the following questions:

- How can degrees of dependence be estimated, and the THERP dependency equations updated, using experimental data?
- Can the number of dependency levels be reduced?
- Do the dependency factors currently used in HRA meaningfully influence the dependence between human actions?
- Can new dependency factors be identified from the experimental data?

This paper presents the results of a preliminary analysis addressing these questions.

2. METHODOLOGY

The THERP HRA method [1] is one of the most widely used first-generation methods in HRA. THERP quantifies the probability of human error over a sequence of operator tasks and supports the identification and reduction of error-prone actions in safety-critical systems such as nuclear power plants. The method represents a task sequence as an HRA event tree, assigns a human error probability (HEP) to each task, and combines these values to estimate the overall probability of human failure.

A central feature of THERP is its treatment of dependence between successive human actions. When two tasks are not independent, the failure of an earlier task increases the failure probability of a subsequent one; neglecting this effect can substantially underestimate the overall HEP. THERP characterizes this relationship through five discrete levels of dependence—ZD, LD, MD, HD, and CD—which are selected on the basis of qualitative factors such as the functional and spatial relatedness of the tasks and their closeness in time.

For each level, THERP provides a closed-form expression for the conditional human error probability (CHEP) of $Task_N$ given failure of the preceding $Task_{N-1}$. The failure equations are:

$$P(F_N|F_{N-1}|ZD) = k \tag{1}$$

$$P(F_N|F_{N-1}|LD) = \frac{1+19k}{20} \tag{2}$$

$$P(F_N|F_{N-1}|MD) = \frac{1+6k}{7} \tag{3}$$

$$P(F_N|F_{N-1}|HD) = \frac{1+k}{2} \tag{4}$$

$$P(F_N|F_{N-1}|CD) = 1.0 \tag{5}$$

where F_N and F_{N-1} denote the failure of $Task_N$ and of the preceding $Task_{N-1}$, respectively; $P(F_N \mid F_{N-1})$ is the CHEP of $Task_N$ given that the preceding task has failed; and k is the basic human error probability (BHEP) of $Task_N$, i.e., its failure probability under the assumption of independence. Equations (1)–(5) correspond to Eqs. (10-14)–(10-18) of the THERP handbook [1].

As the level of dependence increases, the CHEP rises monotonically from the independent value k under zero dependence to unity under complete dependence. The zero-dependence case assumes the BHEP without dependency effects, whereas complete dependence forces the subsequent failure with certainty. The intermediate levels (LD–HD) can be interpreted as a weighted average of the independent failure probability k and the complete-dependence bound of 1.0, with progressively greater weight assigned to the latter.

The five failure equations can be unified into a single generalized expression, as given in Equation (6):

$$P(F_N|F_{N-1}) = \frac{a \times P(F_N|F_{N-1}|CD) + (100-a) \times P(F_N|F_{N-1}|ZD)}{100} = \frac{a + (100-a) \times k}{100} \quad (6)$$

where a is the degree of dependence (%). When $a = 100\%$, the conditional failure probability equals 1.0, corresponding to CD; when $a = 0\%$, it reduces to the basic HEP k, indicating the absence of any dependency effect. Setting a to 5% (1/20), 14.3% (1/7), and 50% (1/2) recovers the conditional failure probability equations for LD, MD, and HD dependence, respectively.

In the original THERP formulation, these degree-of-dependence values were assigned without a documented technical basis, and the handbook itself recommended that they be revised as supporting data became available. Accordingly, this study conducts a preliminary analysis of how the value of a in these equations can be estimated from experimental data, and whether the dependency factors governing a can be identified from such data.

To estimate the value of a from the experimental data, this study first defines the input variables obtainable from the data and then reformulates the conditional probability in Eq. (6) so that it is expressed in terms of these variables. Using the definition of conditional probability, $P(F_N \mid F_{N-1})$ can be rewritten as:

$$P(F_N|F_{N-1}) = \frac{P(F_N \cap F_{N-1})}{P(F_{N-1})} = \frac{a + (100-a) \times k}{100} \quad (7)$$

where the joint and marginal failure probabilities are estimated from the experimental data as:

$$P(F_N \cap F_{N-1}) = \frac{m}{s} \quad (8)$$

$$P(F_{N-1}) = \frac{w}{s} \quad (9)$$

Substituting Eqs. (8) and (9) into Eq. (7) and solving for a yields:

$$a = \frac{100(m - wk)}{w(1-k)} \quad (10)$$

where m is the number of scenario trials in which both human actions failed; s is the total number of scenario trials, each conducted with a different participant, under the assumption that every participant performed both human actions; w is the number of scenario trials in which $Task_{N-1}$(the preceding action) failed; and $P(F_{N-1})$ is the failure probability of $Task_{N-1}$, which incorporates the dependency effects of any human actions performed earlier. Equation (10) is defined for $k \neq 1$ and $w \neq 0$.

Since the degree of dependence is a percentage, the estimate is then constrained to its admissible range.

$$\hat{a} = \min\{100,\ \max\{0,\ a\}\} \quad (11)$$

Because every trial in which both actions failed is also a trial in which the preceding action failed ($m \leq w$), the observed conditional probability m/w cannot exceed unity; thus $a \leq 100$ always holds, and the upper bound in Eq. (11) serves as a formal safeguard. The lower bound is applied when the observed conditional failure probability falls below k, which is interpreted as the absence of positive dependence (i.e., zero dependence).

In this study, $P(F_{N-1})$ is estimated from the experimental data, whereas k is set to the nominal action HEP of the SPAR-H method [3]. It is assumed that every human action employed in the experiments is already affected by the dependency carried over from preceding human actions. Accordingly, although $P(F_{N-1})$ can be computed from the experimental data, k must be a basic HEP that is free of any dependency effect. Because this assumption makes it impossible to obtain a dependency-free HEP from the experimental data, the SPAR-H action value of 0.001 was adopted for k in this study.

3. EXPERIMENTAL DATA

INL has proposed the SHEEP framework [2], which utilizes human performance data collected from simplified (or partial-scope) simulators and student operators to support full-scope simulator studies such as the Human Reliability data EXtraction (HuREX) [4] and Scenario Authoring, Characterization, And Debriefing Application (SACADA) [5] data collections. To date, INL has collected human performance data—including task completion time, errors, workload, and situation awareness—from 36 professional operators and 36 student operators using Rancor developed by INL and the CNS developed by KAERI. Based on these data, INL examined differences in human performance arising from expertise (professional vs. student operators) and from simulator complexity (Rancor [most simplified], CNS [intermediate], and HuREX data [full scope]). Drawing on these performance differences, INL then proposed a method for inferring full-scope data, thereby supporting the related data collection studies.

This study analyzed the error data collected while student operators operated the CNS. The student data were used because professional operators rarely committed human errors on such a simplified simulator in the experiment. As shown in Table 1, four scenarios were conducted on CNS; among them, this paper analyzes only the human actions included in the startup operation scenario. The startup operation is primarily governed by procedure OP-01, and the major human actions in this study were defined on the basis of this procedure. Specifically, the major human actions were selected as those that either failed at least once during the experiments or whose failure could contribute to overall scenario failure; they are summarized in Table 2. Table 2 also lists the procedural step in which each major human action appears, together with the cognitive activity type of each action.

Table 1: Experimental Scenarios Using CNS.

No.	Name	Description	Success Criteria
1	Startup operation (2%–50%)	Increase reactor power from 2% to 50% in fully automatic mode.	• Reactor power = 50% • No reactor trip during the operation
2	Shutdown operation (100% to hot standby)	Shut down the reactor from 100% to 0% in fully automatic mode.	• Reactor power = 2% • No unintended reactor trip during the shutdown
3	Steam generator tube rupture	When a steam generator tube ruptures, it is required that the damaged steam generator be isolated, the safety functions be maintained, and the reactor coolant system temperature be reduced to 200°C.	• Diagnosis of an initiating event • Reactor coolant system temperature under 200°C • Isolation of damaged steam generator
4	Loss of all feedwater	When loss of all feedwater occurs, it is required that safety	• Diagnosis of an initiating event

		functions be maintained and the reactor coolant system temperature be reduced to 200°C.	• Reactor coolant system temperature under 200°C • Attempted recovery of the feedwater pump

Table 2: Major Human Actions in the OP-01 Procedure.

No.	Procedure Step Including Major Human Action	Cognitive Activity
1	Step 4.1	Execution
2	Step 4.4.1	Execution
3	Step 4.4.2	Execution
4	Step 5.2	Information gathering and reporting
5	Step 6.3.1	Information gathering and reporting
6	Step 6.3.2	Information gathering and reporting
7	Step 8.3	Response planning and instruction
8	Step 8.4	Information gathering and reporting / Execution
9	Step 8.7.1	Information gathering and reporting
10	Step 8.7.2	Information gathering and reporting
11	Step 9.2	Response planning and instruction

4. DATA ANALYSIS

Using Equations (7)–(11) and the experimental data, this study calculated degrees of dependence as shown in Table 3.

Table 3: Example Calculation of the Degree of Dependence Using SHEEP Data.

No.	Human Action Combination	m	s	$P(F_{N-1})$	$P(F_N \cap F_{N-1}) = \frac{m}{s}$	$P(F_N \mid F_{N-1}) = \frac{P(F_N \cap F_{N-1})}{P(F_{N-1})}$	a
1	4.1+8.4	2	15	0.25	0.13	0.53	53.29
2	4.1+9.2	1	12	0.25	0.08	0.33	33.27
3	4.4.1+4.4.2	1	16	0.06	0.06	1.00	100.00
4	5.2+6.3.1	1	15	0.13	0.07	0.53	53.29
5	6.3.1+6.3.2	2	15	0.13	0.13	1.00	100.00
6	6.3.2+8.3	1	15	0.13	0.07	0.50	49.95
7	8.4+8.7.1	1	15	0.33	0.07	0.20	19.92
8	8.4+9.2	1	12	0.33	0.08	0.25	24.93

Then, this study defined several dependency factors that may contribute to the degree of dependence, such as area of interest (AOI), paper procedure page (PPP), cognitive activity (CA), task type (TT), component type (CT), and error type (ET), and investigated whether $Task_N$ and $Task_{N-1}$ were under the same or different conditions, as shown in Table 4. For example, for the first human action combination in Table 4, the human action in Step 4.1 of OP-01 is performed in the Chemical Volume Control System (CVCS) window, while the subsequent human action in Step 8.4 of OP-01 is carried out in the Reactivity Control System window. Therefore, the area of interest was coded as "Different".

Table 4: Dependency Factors Derived from Human Action Combinations (Preliminary Analysis).

No.	Human Action Combination	AOI	PPP	CA	TT	CT	ET
1	4.1+8.4	Different	Different	Different	Same	Different	Same
2	4.1+9.2	Different	Different	Same	Same	Different	Different
3	4.4.1+4.4.2	Same	Same	Same	Same	Same	Same

4	5.2+6.3.1	Different	Different	Same	Same	Different	Same
5	6.3.1+6.3.2	Same	Same	Same	Same	Same	Same
6	6.3.2+8.3	Different	Different	Different	Same	Same	Different
7	8.4+8.71	Different	Same	Same	Same	Same	Same
8	8.4+9.2	Different	Different	Different	Different	Different	Different

This study regressed the estimated degrees of dependence on the six dependency factors. Each dependency factor was coded as a binary variable: 0 for "Same" and 1 for "Different." Table 5 presents the regression results, with an R-square value of 0.365. Among the six factors, "Area of Interest," "Cognitive Activity," "Task Type," and "Error Type" were statistically significant at the 0.05 level. The results in Table 5 can be expressed as the following regression equation:

$$a = 93.566 - 26.056x_{AOI} - 5.786x_{PPP} + 27.947x_{CA} - 39.588x_{TT} + 9.054x_{CT} - 22.094x_{ET} \quad (12)$$

where x_{AOI}, x_{PPP}, x_{CA}, x_{TT}, x_{CT}, and x_{ET} denote the binary indicators for Area of Interest, Paper Procedure Page, Cognitive Activity, Task Type, Component Type, and Error Type, respectively (0 = Same, 1 = Different).

For example, when the Area of Interest of the two human actions is "Different," substituting 1 into Eq. (12) decreases the estimated degree of dependence, whereas "Same" (coded as 0) leaves the estimate unaffected. This suggests that when two consecutive actions are performed in different areas of the human–system interface, the indirect dependence arising from the shared use of the same interface may be weakened.

In contrast, Cognitive Activity exhibited the opposite trend to that of Area of Interest: the degree of dependence increased when the cognitive activities of the two human actions differed. A plausible explanation is that operator tasks are typically performed as a sequence of distinct cognitive activities—for example, a diagnosis followed by an execution. Accordingly, the action pairs with different cognitive activities may correspond to a single continuous task performance, which could account for the higher degrees of dependence observed. To substantiate these findings, we are currently conducting further analyses of the data and exploring additional dependency factors. More conclusive results will be presented at the conference.

Table 5: Results of the Regression Analysis (Preliminary Analysis).

	B	Std. Error	β (standardized)	t	p-value
Constant	93.566	8.539	-	10.958	<0.001
AOI	-26.056	11.574	-.408	-2.251	0.030
PPP	-5.786	9.066	-.095	-.638	0.527
CA	27.947	13.262	.453	2.107	0.042
TT	-39.588	14.249	-.591	-2.778	0.008
CT	9.054	10.802	.144	.838	0.407
ET	-22.094	10.178	-.330	-2.171	0.036

5. SUMMARY

This study presented a preliminary analysis toward updating the dependency equations of THERP using experimental data. The five THERP failure equations were unified into a single generalized expression in which the degree of dependence serves as a continuous interpolation weight between the zero-dependence and complete-dependence bounds, and an estimation formula was derived to compute the degree of dependence from experimentally observed failure counts. The approach was demonstrated using human error data collected from student operators performing a startup operation scenario on the CNS under the SHEEP framework. The estimated degrees of dependence for the analyzed human action combinations varied widely, indicating that a single discrete dependency level cannot adequately represent the dependence observed between human actions. A regression analysis of six candidate dependency factors showed that area of interest, cognitive activity, task type, and error type significantly influenced the estimated degrees of dependence, whereas paper procedure page and component type did not. Notably, the cognitive activity factor exhibited a positive effect, suggesting that consecutive actions involving different cognitive activities may represent successive phases of a single task and therefore exhibit stronger dependence.

Several limitations should be noted. The analysis relied on a limited number of human action combinations from a single scenario, the basic human error probability was assumed to be the SPAR-H nominal action value rather than estimated from data, and the dependency factors were coded as binary variables. To address these limitations, we are currently expanding the analysis with additional experimental data, refining the estimation of the basic human error probability, and exploring additional dependency factors. Future work will also examine whether the number of dependency levels can be reduced, and whether other data sources, such as licensee event reports, can be used to update the dependency equations. More conclusive results will be presented at the conference.

Acknowledgement

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (Ministry of Science and ICT) (No. RS-2026-25541562).

References

[1] A. Swain and H. Guttmann. "*Handbook of human reliability analysis with emphasis on nuclear power plant applications*", NUREG/CR-1278, Sandia National Laboratories, Albuquerque, NM, USA, (1983).
[2] J. Park, R. Boring, T. Ulrich, R. Lew, S. Lee, B. Park and J. Kim. "*A framework to collect human reliability analysis data for nuclear power plants using a simplified simulator and student operators*", Reliability Engineering & System Safety, vol. 221, p. 108326, (2022).
[3] D. Gertman, H. Blackman, J. Marble, J. Byers and C. Smith. "*The SPAR-H human reliability analysis method*", NUREG/CR-6883, U.S. Nuclear Regulatory Commission, Washington, DC, USA, (2005).
[4] W. Jung, J. Park, Y. Kim, S. Kim and S. Choi. "*HuREX – A framework of HRA data collection from simulators in nuclear power plants*", Reliability Engineering & System Safety, vol. 194, p. 106235, (2020).
[5] Y. J. Chang, D. Bley, L. Criscione, B. Kirwan, A. Mosleh, T. Madary, R. Nowell, R. Richards, E. Roth, S. Sieben and A. Zoulis. "*The SACADA database for human reliability and human performance*", Reliability Engineering & System Safety, vol. 125, pp. 117-133, (2014).