

Use of IDHEAS General Methodology to Incorporate Human Performance Data for Estimation of Human Error Probabilities

Jing Xing^a and Y. James Chang^a

^aU.S. Nuclear Regulatory Commission, Washington DC, USA

Abstract: The Integrated Human Event Analysis System (IDHEAS), a human reliability analysis method developed by the US Nuclear Regulatory Commission, provides a hierarchical structure to analyze and assess the reliability of human actions. The method is based on cognitive science and is capable of incorporating human performance data to support the estimation of human error probabilities. IDHEAS models human performance in five macrocognitive functions: *Detection, Understanding, Decision-making, Action execution, and Teamwork*. IDHEAS defines a set of cognitive failure modes for each function to describe the various ways of failing to perform the function. IDHEAS analyzes an event in progressively more detailed levels: event scenario, human actions, critical tasks of the actions, macrocognitive functions and cognitive failure modes of the tasks, and performance influencing factors. This structure provides an intrinsic interface to integrate various sources of human error data for human error probability estimation. We reviewed numeric data of human errors in the literature and synthesized the data in the IDHEAS structure. This paper presents the hierarchical structure along with the demonstration of using empirical and experimental human error data of various resources in the structure. The data, once sufficiently populated, can provide a basis for estimating human error probabilities.

Keywords: HRA, IDHEAS method, Human error data, Macrocognitive function

1. INTRODUCTION

Probabilistic risk assessment (PRA) results and insights support risk-informed regulatory decision making. The U. S. Nuclear Regulatory Commission continues to improve the robustness of PRA, including human reliability analysis (HRA) through many activities. Improving HRA has been a focus of the NRC's research activities. To date, there have been about fifty HRA methods developed worldwide to estimate human error probabilities (HEPs). Method-to-method variability and analyst-to-analyst variability in the estimates of human error probabilities (HEPs) have been observed in applying these methods [1]. This variability in HRA quality could affect risk-informed decisions.

Existing HRA methods were built on behavioral observations of human performance and cognitive science. Without explicitly modeling the intrinsic cognitive mechanisms underlying human errors, an HRA method may result in different interpretations of the same observed phenomena and poor understanding of the causes of human errors. As such, HRA methodologies should be enhanced to incorporate the advances in cognitive and behavioral science in the past decades. Furthermore, the use of empirical data for HEP estimation has been limited due to the lack of data and discrepancies between the formats of available data and HRA methods. Lack of a strong data basis in the methods challenge method validity and introduce additional variabilities in HEP estimation.

To tackle these variability issues, the staff in the U. S. Nuclear Regulatory Commission began development of an enhanced HRA method, referred to as the Integrated Human Event Analysis System (IDHEAS). The method was to integrate the strengths in existing HRA methods, enhance the cognitive basis for HRA, and build the capability for using human error data to improve HEP estimation. Since 2012, we have developed an IDHEAS suite that includes the following:

- The cognitive basis for HRA. The cognitive basis synthesizes the fundamentals of human cognition into a structure that supports HRA method development and HRA practices. The cognitive basis is documented in NUREG-2114, “Cognitive Basis for HRA.” [2]
- The IDHEAS general methodology (IDHEAS-G) that is independent of specific HRA applications and applicable to a wide range of nuclear HRA applications. The methodology incorporates state-of-art cognitive and behavioral sciences and integrates the strengths of existing HRA methods. IDHEAS-G lays out the fundamentals to develop application-specific HRA methods in the IDHEAS suite. IDHEAS-G is being documented in NUREG-2198 for publication in 2019 [3].
- IDHEAS internal, at-power application. This is a HRA method in the context of internal nuclear power plant events. The work is a collaboration between the US Nuclear Regulatory Commission and the Electrical Power Research Institute (EPRI) and is documented in NUREG-2199, Vol. 1 “An Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal At-Power Event Application.” [4]

IDHEAS-G includes two parts: a cognition model and its implementation in HRA. This paper introduces IDHEAS-G and demonstrates using IDHEAS-G to integrate human error data from various sources to inform HEP estimates.

2. RESULTS

2.1 IDHEAS-G Cognition Model

The cognition model includes a macrocognition model that describes the brain process of success or failure of a task, and a performance influencing factor (PIF) model that describes how various factors affect the success or failure of tasks.

The Macrocognitive Model

The macrocognitive model elucidates the cognitive process of human performance in applied work domains where human tasks are complex and often involve multiple individuals or teams. The model is described as follows:

- Macrocognition consists of five functions: *Detection*, *Understanding*, *Decisionmaking*, *Action Execution*, and *Teamwork*. The first four functions may be performed by an individual, a group or a team, and the *Teamwork* function is performed by multiple groups or teams.
- Any human task is achieved through these functions; complex tasks typically involve all five functions;
- Each macrocognitive function is processed through a series of basic cognitive elements; failure of a cognitive element leads to the failure of the macrocognitive function;
- Each element is reliably achieved through one or more cognitive mechanisms; errors may occur in a cognitive element if the cognitive mechanisms are challenged;
- PIFs affect cognitive mechanisms.

Table 1 shows the basic cognitive elements for the macrocognitive functions. The cognitive mechanisms are not presented due to the space limit of the paper.

Table 1: Macrocognitive Functions and Their Basic Elements

Detection	Understanding	Decisionmaking	Action execution	Teamwork
D1- Initiate detection - Establish mental model and criteria D2 – Identify and attend to	U1 - Assess/select data U-2 Select / adapt / develop the mental model U-3 Integrate data with mental model	DM1 – Manage the goals DM2 – Adapt a decision model DM3 – Acquire / select information	E1 - Assess action plan E2 - Develop / modify action scripts E3 - Synchronize,	T1 - Establish or adapt teamwork infrastructure T2 - Manage information T3 - Maintain common ground

sources of information D3- Perceive, recognize, and classify information D4-Verify the information acquired D5- Communicate the acquired information	to maintain situational awareness, diagnose problems, and resolve conflicts in information U-4 Verify, revise, and iterate the understanding U-5 Communicate the understanding	DM4 - Make judgment or plans DM5 - Simulate the decision DM6 - Communicate and authorize the decision	supervise, and coordinate action implementation E4 - Implement action scripts E5 - Verify and adjust actions	T4 - Manage resources T5 - Plan inter-team collaborative activities T6 - Implement decisions/commands T7 - Verify, modify, and control the implementation
---	--	---	---	--

The Performance Influence Factor Model

PIFs affect cognitive mechanisms and increase the likelihood of macrocognitive function failure. We developed a PIF model that is independent of HRA applications and links to cognitive mechanisms. The model systematically organizes PIFs to minimize inter-dependency or overlapping of the factors. The PIF structure has four layers:

- 1) PIF category: PIFs are classified into three categories, corresponding to characteristics of systems, tasks, and personnel.
- 2) PIFs: Each category has high-level PIFs describing specific aspects of the systems, tasks, or personnel. Table 2 shows the PIFs within the three categories.

Table 2: Performance influencing factors in IDHEAS-G

System-related PIFs	Task-related PIFs	Personnel-related PIFs
<ul style="list-style-type: none"> ○ Availability and reliability of systems and instrument & control ○ Environmental factors ○ Work location accessibility and habitability ○ Tools and equipment 	<ul style="list-style-type: none"> ○ Information availability and reliability ○ Scenario familiarity ○ Multi-tasking, interruptions and distractions ○ Cognitive complexity ○ Mental fatigue and stress ○ Physical demands 	<ul style="list-style-type: none"> ○ Human-system interface (HSI) ○ Staffing ○ Training ○ Procedures / guidelines / instructions ○ Teamwork factors ○ Work process

- 3) PIF attributes: These are the specific traits of a performance influencing factor. A PIF attribute represents a poor PIF state that challenges cognitive mechanisms and increases the likelihood of errors in cognitive processes. Table 3 shows some example attributes of the PIF *Information availability and reliability*.

Table 3: Example Attributes of the PIF “*Information Availability and Reliability*”

Nominal state of <i>Information availability and reliability</i> : Information is needed for personnel to perform tasks. Information is expected to be complete, reliable, unambiguous, and available timely to personnel.
<ul style="list-style-type: none"> • Inadequate updates of information (e.g., a party receives information but fails to inform another party). • Information of different sources is not synchronized. • Conflicts in information

- There are multiple alternative explanations for the pattern of symptoms observed
- Available information contradicts with each other or does not support a coherent understanding of the situation
- Information does not match procedures/guidance.
- Sources or meanings of information are unfamiliar to personnel
- Information is ambiguous.
 - Pieces of Information change over time at different paces thus they become uncertain by the time personnel use them together
- Incomplete information, e.g., primary sources of information are not available while secondary sources of information are not readily perceived
- Information is misleading or wrong
 - Sensors or indicators may be unreliable or misleading (e.g., damaged or degraded while appearing to be working, false alarms in design, out-of-range, inherently unreliable sources, conflicting data indicating a false situation, or flaw in system state indication)
- Information is masked

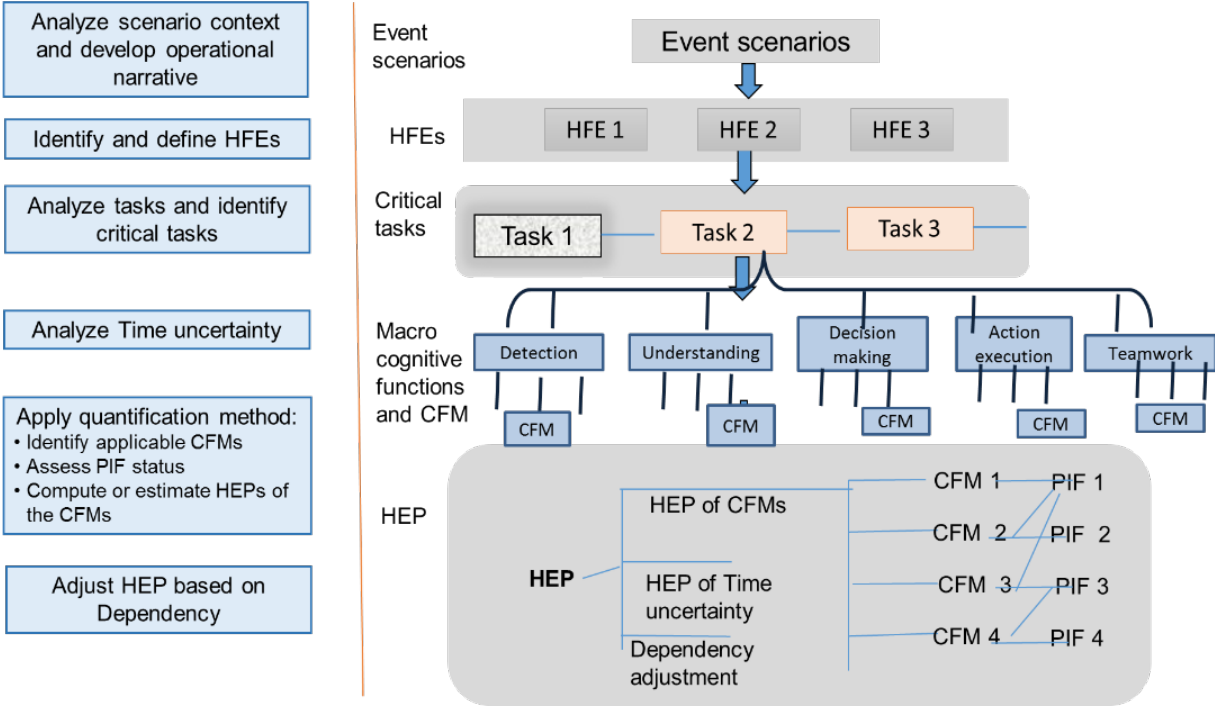
4) Every PIF attribute challenges one or several cognitive mechanisms. IDHEAS-G provides links between PIF attributes and cognitive mechanisms synthesized and inferred from the literature.

The PIF model consolidates the state-of-knowledge about PIFs. A specific HRA application may only involve a subset of PIFs in the model, and various applications may involve different subsets of PIFs and attributes. On the other hand, the subsets of PIFs for various HRA applications share a common structure, which would increase method-to-method consistency and allow comparisons of HRA results from different HRA methods.

2.2 Implementation of IDHEAS-G Cognition Model in HRA

Overview IDHEAS-G Process

Figure 1: An Overview of IDEHAS-G Process for HRA



IDHEAS-G implements the cognition model in the general HRA process, which includes qualitative analysis and HEP quantification. Figure 1 shows the overview of the process. It begins with an event and progressively analyzes more detailed elements of the event: event scenarios, human actions (referred to as human actions, i.e., HFEs, in PRA) in the scenarios, critical tasks in the human actions, macrocognitive functions, cognitive failure modes (CFMs) of the critical tasks, the states of PIFs, and HEPs. The analysis of these elements are carried out through six steps:

- Step 1 - Scenario analysis: Analyze the event and develop the operational narrative for the event scenarios
- Step 2 - Identification and definition of human actions: Identify the key human actions pertinent to the mission of the event and define the human actions
- Step 3 - Task analysis: Analyze tasks required for the human action and characterize the critical tasks for HEP quantification
- Step 4 - Time uncertainty analysis: Analyze time uncertainties in the human action and quantify the HEP attributing to time uncertainties
- Step 5 - Cognition failure analysis: Identify cognition failure modes of every critical tasks in a human action and estimate the HEP attributing to failures of macrocognitive functions for the critical tasks)
- Step 6 - Dependency analysis: Analyze dependency between human actions and adjust the Pc and Pt of a human action based on its dependency with other human action.

Quantification of Human Error Probabilities in IDHEAS-G

This section describes HEP quantification in Step 5. IDHEAS-G states that the HEP of a human action consists of two parts: the error probability caused by variability and uncertainties in time available to perform the human action, and the error probability caused by failures of the macrocognitive functions under the assumption that the time available for performing the action is adequate. Estimating the cognitive HEP includes three parts:

- 1) identify applicable CFMs for every critical task
- 2) assess PIF attributes relevant to the CFMs
- 3) estimate HEPs of the CFMs

Identify applicable CFMs for every critical task

Based on the macrocognition model, we developed a basic set of cognitive failure modes in three levels of detail. The first level is failure of macrocognitive functions, the next level is failure of the basic cognitive elements for every macrocognitive function, and we further break each basic element failure mode into several detailed, behaviorally observable failure modes. The layered structure of the CFMs is to provide flexibility in the level of detail of an HRA and to adapt to available human error data that serve as the basis for HEP estimation. Table 4 shows the full set of the CFMs. An application-specific IDHEAS method may only include a subset of these CFMs. A critical task may have one or several applicable CFMs.

Table 4: The Basic Set of Cognitive Failure Modes in IDHEAS-G

Failure of Detection	
Failure of the basic elements	Detailed cognitive failure modes for Detection
Fail to Initiate detection	D1-1 Detection is not intended (e.g., skip steps of procedures for detection, no for detection) D1-2 Wrong mental model for detection (e.g., incorrect planning on when, how, or what to detect) D1-3 Fail to prioritize information to be detected
Fail to identify and attend to sources of information	D2-1 Unable to access the source of information D2-2 Attend to wrong source of information
Incorrectly perceive	D3-1 Key alarm not perceived D3-2 Key alarm incorrectly perceived

information	D3-3 Cues not perceived D3-4 Cues misperceived (e.g., information incorrectly perceived, fail to perceive weak signals, reading errors, incorrectly interpret, organize, or classify information etc.) D3-5 Fail to monitor parameters (e.g., information or parameters not monitored at proper frequency or for an adequate period of time, fail to monitor all the key parameters, and incorrectly perceiving the trend of a parameter)
Incorrectly recognize information	D4-1 Fail to recognize that primary cue is incorrect or misleading D4-2 Incorrectly verify the perceived information against the detection criteria
Fail to communicate the acquired information	D5-1 The detected information not retained or incorrectly retained (e.g., mark wrong items, wrong recording, and wrong data entry) D5-2 The detected information not communicated or miscommunicated

Failure of Understanding	
Failure of the basic elements	Detailed cognitive failure modes for Understanding
Fail to Assess or select data	U1-1 Incomplete data selected (e.g., critical data dismissed, critical data omitted) U1-2 Incorrect or inappropriate data selected (e.g., fail to recognize the applicable data range, and not recognize the information is outdated)
Incorrect mental model	U2-1 No mental model exists for understanding the situation U2-2 Incorrect mental model selected U2-3 Fail to adapt the mental model (e.g., fail to recognize and adapt mismatched procedures)
Incorrect integration of data and mental model	U3-1 Incorrectly assess situation (e.g., situational awareness not maintained, and incorrect prediction of the system evolution or upcoming events) U3-2 Incorrectly diagnose problems (e.g., conflicts in data not resolved, under-diagnosis, fail to use guidance outside main procedure steps for diagnosis)
Fail to iterate the understanding	U4-1 Premature termination of data collection (e.g., not seeking additional data to reconcile gaps, discrepancies, or conflicts, or fail to revise the outcomes based on new data, mental models, or viewpoints) U4-2 Fail to generate coherent team understanding (e.g., assessment or diagnosis not verified or confirmed by the team, and lack of confirmation and verification of the results).
Fail to communicate the outcome	U5-1 Outcomes of understanding miscommunicated or inadequately communicated

Failure of Decisionmaking	
Failure of the basic elements	Detailed cognitive failure modes for Decisionmaking
Incorrect goals or priorities	DM1-2 Unable to prioritize multiple conflicting goals DM1-1 Incorrect goal selected
Inappropriate decision model	DM2-1 Incorrect decision model or decision-making process (e.g., incorrect on who, how, or when to make decision, decision goal is not supported by the decision model or process). DM2-2 Incorrect decision criteria
Information is under-represented	DM3-1 Critical information not selected or only partially selected (e.g., biased, under-sampling of information) DM3-2 Selected information is not appropriate or not applicable for the situation DM3-3 Misinterpret or misuse selected information
Incorrect judgment or	DM4-1 Misinterpret procedure DM4-2 Choose inappropriate strategy or options

planning	DM4-3 Incorrect or inadequate planning or developing solutions (e.g., plan wrong or infeasible responses, plan the right response actions at wrong times, not plan configuration changes when needed, plan wrong or infeasible configuration changes) DM4-5 Decide to interfere or override automatic or passive safety-critical systems that would lead to undesirable consequences
Fail to simulate or evaluate the decision / strategy /plan	DM5-1 Unable to simulate or evaluate the decision's effects (e.g., fail to assess negative impacts or unable to evaluate the pros and cons) DM5-2 Incorrectly simulate or evaluate the decision (e.g., fail to evaluate the side effects or components, or fail to consider all key factors) DM5-3 Incorrect dynamic decision-making
Fail to communicate or authorize the decision	DM6-1 Decision is incorrectly communicated DM6-2 Decision not authorized DM6-3 Decision is delayed in authorization

Failure of Action Execution	
Failure of the basic elements	Detailed cognitive failure modes for Action Execution
Fail to assess action plan	E1-1 Action is not initiated E1-2 Incorrectly interpret the action plan (e.g., wrong equipment / tool preparation, or coordination) E1-3 Wrong action criteria E1-4 Delayed implementation E1-5 Incorrectly add actions or action steps to manipulate safety systems outside action plans (e.g., error of commission)
Fail to develop / modify action scripts	E2-1 Incorrectly modify or develop action scripts for the action plan
Failed to coordinate action implementation	E3-1 Fail to coordinate the action implementation (e.g., fail to coordinate team members, errors in personnel allocation) E3-2 Fail to initiate action E3-3 Fail to perform status checking required for initiating actions
Failed to take planned action	E4-1 Fail to follow procedures (e.g., skip steps in procedures) E4-2 Fail to execute simple action E4-3 Fail to execute complex action (e.g., execute a complex action with incorrect timing or sequence, execute actions that do not meet the entry conditions) E4-3A Fail to execute control actions E4-3B Fail to execute skill-of-craft actions E4-3C Fail to execute long-lasting actions E4-4 Fail to execute physically demanding actions E4-5 Fail to execute fine-motor actions
Fail to verify or adjust action	E5-1 Fail to adjust action by monitoring, measuring, and assessing outcomes E5-2 Fail to complete the entire action scripts or procedures (e.g., omit steps after the action criteria are met) E5-3 Fail to record, report or communicate action status or outcomes

Failure of Teamwork	
T1	Fail to establish or adapt the teamwork infrastructure
T2	Fail to manage information
T3	Fail to maintain common ground
T4	Inappropriately manage resources
T5	Fail to make inter-team decisions or generate commands

T6 Fail to implement decisions/commands

T7 Fail to control the implementation

Assess PIF attributes relevant to the CFMs

PIF attributes challenge one or more cognitive mechanisms, which leads to errors in macrocognitive functions. Each CFM represents one type of macrocognitive error. Thus, a CFM is associated with a specific set of PIF attributes. IDHEAS-G provides the links between PIF attributes, cognitive mechanisms, and CFMs. Table 5 presents several examples of relevant PIF attributes to CFMs.

Table 5: Example PIF Attributes for Two CFMs

Example PIF attributes for CFM “U3 – <i>Incorrectly integrate data and mental model for understanding</i> ”

- Cognitive complexity - Cognitive complexity in Understanding
- Scenario familiarity - Unfamiliar scenario
- Work process – lack of or ineffective process reconciling different viewpoints in a team.
- Procedures - Sequential presentation of guidelines requires the crew to go through several loops before finding the correct indications to diagnose the plant status.
- Procedures - Multiple guidance documents are needed simultaneously.

Example PIF attributes for CFM “E1 – <i>Fail to assess action plan</i> ”
--

- Reluctance to execute the action plan (e.g., adverse economic impact, and personnel injury)
- Inadequate leadership to initiate assessment of action scripts
- Unable to verify the plan because of inadequate communication (of the goals, negative impacts, deviations) with decision-makers
- Inadequate training on verifying and evaluating action plans
- Inappropriate crew assignment, e.g., under-staffing, lack of skills, and limited access to the action sites

Estimate human error probabilities of CFMs

The HEP of a CFM is determined by states of the relevant PIFs, that is,

$$P_c = f(w_1, w_2, w_3, w_4, \dots) \quad (1)$$

Where P_c is the HEP of a CFM, and w models the quantitative effects of a PIF state on the HEP.

At present, there is no adequate data allowing calculation of the HEPs of all CFMs for any given combination of PIF states, nor have cognitive studies clearly elucidated the mathematic relation between PIFs and HEPs. We can only estimate HEPs from sparse human error data and knowledge available. The estimation can be based on the simplest linear function as the following:

$$P_c = P_0 \times (w_1 + w_2 + w_3 + w_4, \dots) \times R, \quad (2)$$

Where P_0 is the base HEP of a CFM when relevant PIF states are nominal; w is the weight of a PIF state representing the increment of the HEP caused by the PIF poor state compared to the nominal state; R is a numeric factor representing possible interaction between the PIFs, and R can be set to 1 if no interaction between PIFs is assumed.

Note that IDHEAS-G does not provide numeric HEP values. It provides a basic set of CFMs, a PIF model, and a simplification of the quantitative relation between PIF states and the HEP of a CFM. This structures allows for the synthesis of human error data from various sources, at various levels of details, and in various formats to inform HEP estimation.

2.2 Use IDHEAS-G to Synthesize Human Error Data for HEP Estimation

IDHEAS-G presents a basic set of CFMs, a PIF model, and a simple linear formula to combine the contributions of PIFs to the HEP of a CFM. The basic set of CFMs represents failure modes at three

levels of granularity, i.e., failures of macrocognitive functions, failures of the basic elements in each function, and behaviourally observable failure modes. Similarly, the PIF model represents PIFs in two levels of granularity: PIFs and attributes. The cognitive mechanisms can link CFMs and PIFs at any level of granularity. The structured data can inform expert judgment or Bayesian estimates of HEPs. The NRC staff recently synthesized a variety of sources of human error data to inform expert judgment of HEPs for nuclear power plant ex-control room actions.

Next we use several example types of human error data to demonstrate how the data can be used to inform HEP estimation. The data sources include i) quantification of unsatisfied task performance in nuclear power plant operator simulator training (as collected in the Scenario Authoring, Characterization, and Debriefing Applications (SACADA) database by the US Nuclear Regulatory Commission); ii) human error rates in nuclear power plant operational tasks as well as tasks in other domains (such as aviation, assembly industry, offshore operation); iii) Human error rates of cognitive tasks in controlled experiments; iv) quantitative effects of PIFs in the literature. These data play different roles in estimating HEPs.

1) *Baseline HEPs or HEPs with known states of PIFs*

Some sources of data present statistical human error rates of certain types of tasks with various contexts and scenarios. Such data can inform the baseline HEPs for the CFMs applicable to the tasks. Below are two examples:

- Quantification of unsatisfied task performance in nuclear power plant operator simulator training, as collected in the Scenario Authoring, Characterization, and Debriefing Applications (SACADA) database by the US Nuclear Regulatory Commission [5]. The SACADA database was built with the same macrocognitive model as that in IDEHAS-G; SACADA collects operator unsatisfied task performance in different types of failures under various contexts; the types of failures can be mapped to the detailed level CFMs in IDEHAS-G, and the context can be mapped to IDEHAS-G PIF attributes. Thus, the SACADA database can inform baseline HEPs of IDEHAS-G CFMs and the quantitative effects of some PIF attributes.
- The analysis of human errors in maintenance operations of German nuclear power plants. Preischl and Hellmich [6, 7] studied human error rates of various basic tasks in maintenance operations. Below are some example human error rates they reported:
 - o 1/490 for operating a circuit breaker in a switchgear cabinet under normal conditions;
 - o 1/33 for connecting a cable between an external test facility and a control cabinet;
 - o 1/36 for reassembly of component elements;
 - o 1/7 for transporting fuel assemblies

These error rates can inform base HEPs of the CFMs for action execution.

2) *Quantification of PIF effects*

Some data sources present the changes in human error rates when varying one or more PIF from nominal to a poor state. Such data can inform PIF contribution factor W and interaction factor R in equation (2) above. Below are several examples:

- NUREG/CR-5572 [8] estimated the effects of local control station design configurations on human performance and nuclear power plants. It estimated that the HEP = $2E-2$ for ideal conditions and HEP = 0.57 for challenging conditions with poor human-system-interfaces and distributed work locations.
- Prinzo et al [9] analyzed aircraft pilot communication errors and found that the error rate increased nonlinearly with the complexity of the message communicated. The error rate was around 4% for the information complexity index of 4 (i.e., the number of messages transmitted per communication), 30% for the index of 12, and greater than 50% for indices greater than 20.
- Patten et al [10] studied the effect of task complexity and experience on driver performance. The PIF states of the tasks manipulated in the experiment were low experience vs high experience, and low complexity vs. high complexity. The mean error rates were 12%, 21%,

25%, and 32% respectively for the four combinations of PIF states: low complexity and high experience, low complexity and low experience, high complexity and low experience, high complexity and low experience. The data in this experiment suggests nearly no interaction between the two PIFs.

3) *The significance of PIFs for certain types of tasks*

Studies in human error analysis and root causal analysis typically classify and rank the frequencies of various PIFs in reported human events. Some studies correlate PIFs with various types of human errors. Those studies only analyze the relative human error data without reporting how many times personnel performed the kind of tasks. The data from such studies cannot directly inform HEPs, but they can inform which PIFs or attributes are more relevant to the CFMs of the reported human errors. Below are several examples:

- Virovac et al [11] analyzed human errors in airplane maintenance and found that the prevalent factors with frequent occurrence in human errors are communication 16%, equipment and tools 12%, work environment 12%, and complexity 6.5%.
- Kyriakidis et al [12] analyzed UK railway accidents caused by human errors and calculated proportions of PIFs in the accidents. They reported that the most frequent PIFs in the accidents were safety culture 19%, familiarity 15%, and distraction 13%.

The above examples are just a few of a large body of human error data we have documented so far. We performed a meta-analysis of a subset of the documented data [13] and noticed that the error rate data were generally convergent across different studies. For example, most studies on dual-tasks showed that the error rate in dual-tasks was between 1 to 2 times higher than that in a single task. We also observed the consistency between the results obtained in controlled cognitive experiments and those from complex scenario simulation. The observation suggests that human error rates measured from cognitive experiments could serve as a baseline reference for estimating HEPs in more complex, real life scenarios.

3. DISCUSSION

Assessment of PIF states

The effect of a PIF on human error probabilities typically varies continuously from the nominal state to the extremely poor state of the PIF. Our preliminary meta-analysis of human error data suggests that the effects of PIFs on human error rates follow a logarithmic function. For simplification, most HRA methods typically model PIFs with binary states (e.g., good vs. poor) or several discrete states (e.g., low, medium, or high). When modeling a PIF as a binary variable, the model needs to clearly define the meaning of the states. Because the effects of PIFs on human error probabilities follow a logarithmic function, the “good” state of a PIF usually corresponds to the range within which the PIF has little effect on the human error probability. However, the “poor” state of a PIF can represent any place on the rising section of the logarithmic function. As a result, the human error probability for the poor state can vary greatly if the state is not clearly specified. Therefore, for modeling PIFs with binary or a few discrete states, the definitions of each state must be specified and used consistently in HRA,

Some PIFs may affect human error probabilities significantly more than other factors. For example, the factors “information reliability,” “cognitive complexity,” or “intermingled multitasking” can result in very high HEPs. Thus, human error probabilities are very sensitive to changes in the states of those PIFs. Ideally, such high-impact PIFs should be modeled with continuous variables.

Estimation of human error probabilities through a Bayesian approach

A human error probability can be interpreted as the number of errors in performing a human action divided by the number of times the action is performed. In the real world, there is not adequate data to precisely compute the human error probabilities of rare events. As a common HRA practice, the human error probabilities of human failure modes in an HRA method have been estimated through a Bayesian approach, which characterizes what is known about the parameter in terms of a probability

distribution that measures the current state of belief in the possible values of the parameter. A Bayesian approach can be implemented through Bayesian network computation, formal expert judgment, or the combination of both.

- When numerical data is available in the form of the number of failures in a given number of demands, the human error probability distribution can be estimated through Bayesian computation. NUREG-2122 [14] describes Bayesian estimation as, “Bayesian analysis is commonly used in the computation of the frequencies and failure probabilities in which an initial estimation about a parameter value (e.g., event probability) is modified based on actual occurrences of the event. The parameter value may have a probability distribution associated with it. Thus, the event probability to be determined is based on a belief, rather than on occurrence ratios.”
- When numerical data is not available or is sparse, expert judgment is used to estimate the human error probability distribution. The expert judgment approach relies on the knowledge of experts in the specific technical field who arrive at "best estimates" of the distribution of the probability of a parameter or basic event. This approach is typically used when detailed analyses or evidence concerning the event are very limited or unavailable. Such a situation is usual in studying rare events. Ideally, this approach provides a mathematical probability distribution that represents the expert or "best available" knowledge about the probability of the parameter or basic event. The process of obtaining these estimates typically is called "expert judgment elicitation," or simply "expert judgment" or "expert elicitation.” The US Nuclear Regulatory Commission has developed several guidance documents on expert judgment and applied them in HEP estimation [4, 15, 16].

If there are no known experiential data to evaluate the parameter, one must rely on expert elicitation to develop an uncertainty distribution about the parameter of interest (referred to as expert information). As new experiential or empirical data becomes available, the data can be used to verify or modify the expert information, or the experts can use the new data to update their judgment. As additional information becomes available, the Bayesian approach provides a methodology to account for new information, without having to go back through the expert elicitation process again. As the evidence becomes stronger and stronger, the influence of the expert elicitation becomes less and less, to the limit of that infinitely available information.

Assessment of data/evidence

In a Bayesian approach, regardless of the amount of data available (even massive amounts of data), engineering judgment is still needed to consider the applicability of the data, whether there are gaps in the data, and where there may be uncertainties in the data. Regardless of whether Bayesian computation or expert elicitation is used, the data/evidence used for estimating human error probabilities should describe the human errors associated with the tasks or cognitive failure modes at the same level as those in the IDHEAS-G quantification model. Because HRA data are rare, estimating human error probabilities often requires using available data from different sources. First, the data need to be assessed to understand the tasks represented by the data and their applicability to the generic tasks and cognitive failure modes in IDHEAS-G. In addition, the context of the data needs to be assessed to ensure that it is used appropriately for corresponding combinations of performance influencing factors.

4. CONCLUSIONS

IDHEAS-G is a general HRA methodology built on cognitive science and existing HRA technologies. It can be adapted to various HRA applications and can be used as a basis to develop application-specific HRA methods. Its layered structure allows for the synchronization of human error data of different formats and various levels of detail to inform HEP estimates. The NRC staff reviewed a large body of human error data from the literature and available human error databases, synchronized the data into the IDHEAS-G structure, and used the synchronized dataset to inform expert judgment of HEPs for nuclear power plant human actions outside the control rooms. The effort demonstrates the promising of data-informed and data-based human error probabilities in human reliability analysis.

References

1. Lois, E., Dang, V. N., Forester, J., Broberg, H., Massaiu, S., Hildebrandt, M., Braarud, P. Ø., Parry, G., Julius, J., Boring, R., Männistö, I., & Bye, A. *International HRA Empirical Study – Phase I Report: Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Data*. NUREG/IA-0216, Vol. 1, US Nuclear Regulatory Commission, Washington, DC, (2009).
2. US Nuclear Regulatory Commission. *Cognitive Basis for Human Reliability Analysis*, NUREG-2114, (2016).
3. US Nuclear Regulatory Commission. *An Integrated Human Event Analysis System – Internal At-Power Application*, NUREG-2198, (2019, in preparation).
4. US Nuclear Regulatory Commission, *An Integrated Human Event Analysis System – Internal At-Power Application*, US Nuclear Regulatory Commission, NUREG-2199, Vol.1. (2017)
5. Chang Y. J., et. al., *The SACADA database for human reliability and human performance*, Reliability Engineering and System Safety 125:117–133. (2014)
6. Preischl W, Hellmich M, *Human error probabilities from operational experience of German nuclear power plants, Part II*, Reliability Engineering and System Safety 109: 150–159, (2013)
7. Preischl W, Hellmich M, *Human error probabilities from operational experience of German nuclear power plants, Part II*, Reliability Engineering and System Safety 148: 44–56, 2016
8. US Nuclear Regulatory Commission. *An Evaluation of the Effects of Local Control Station Design Configurations on Human Performance and Nuclear Power Plant Risk*. NUREG/CR-5572 (1990)
9. O.V. Prinzo, A. M. Hendrix, and R.Hendrix. *Outcome of ATC (Air Traffic Control) Message Complexity on Pilot Readback Performance*. Technical report by Federal Aviation Administration, Washington, DC. Office Of Aerospace Medicine, (2007)
10. C.J.D. Patten, A. Kircher, J. O’Stlund L. Nilsson, O. Svenson. *Driver experience and cognitive workload in different traffic environments*, Accident Analysis and Prevention, 38: 887–894, (2006)
11. D. Virovac, A. Domitrović, E. Bazijanac, *The influence of human factor in aircraft maintenance*, Promet – Traffic&Transportation, 29: 257-266 , (2017)
12. Kyriakidis M, Pak, KT, Majumdar A. *Railway Accidents Caused by Human Error - Historic Analysis of UK Railways 1945 to 2012* Transportation Research Record Journal of the Transportation Research Board 2476:126-136, (2015)
13. Xing J, Chang YJ, and Siu N, 2015 *Insights on human error probability from cognitive experiment literature*. International Topical Meeting on Probabilistic Safety Assessment (PSA-15), Sun Valley, Idaho, USA, (2015).
14. US Nuclear Regulatory Commission. *Glossary of Risk-Related Terms in Support of Risk-Informed Decisionmaking*, NUREG-2122, (2013)
15. US Nuclear Regulatory Commission. *Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and the Use of Experts*, NUREG/CR-6372, (1997)
16. Xing J, Morrow S, *Practical Insights and Lessons Learned on Implementing Expert Elicitation*, US Nuclear Regulatory Commission White Paper, ADAMS ML16287A734, accessed via <https://www.nrc.gov/docs/ML1628/ML16287A734.pdf>, (2016)