INL/RPT-24-78875

Light Water Reactor Sustainability Program

Use of Time Distributions to Predict Operator Procedure Performance in Dynamic Human Reliability Analysis

June 2024

U.S. Department of Energy

Office of Nuclear Energy

INL/RPT-24-78875

Use of Time Distributions to Predict Operator Procedure Performance in Dynamic Human Reliability Analysis

Jooyoung Park[1], Ronald Boring[1], Thomas Ulrich[1], Roger Lew[2]
[1]Idaho National Laboratory
[2]University of Idaho

June 2024

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Light Water Reactor Sustainability Program

Page left intentionally blank.

ABSTRACT

The Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) framework affords software capable of conducting human reliability analysis (HRA) using a dynamic approach built around operating procedures (OPs) from nuclear power plants (NPPs). Previous HUNTER reports document the development of this software tool, the coupling of HUNTER to the simulator code, the collection of operator performance data by using simulators to calibrate HUNTER models, and linking HUNTER to probabilistic risk assessment (PRA) software. The present report largely addresses two topics. The first is a new function in HUNTER called the HUNTER-Procedure Performance Predictor (P3). HUNTER-P3 uses HUNTER's built-in Monte Carlo tools featuring human performance variability to identify potential error traps in procedures. The second topic is time distribution analysis to generate time inputs for dynamic HRA. The current analysis was performed to investigate time distributions for task primitives, which are the minimum task unit of analysis used in dynamic HRA modeling. Using the time distribution data, the elapsed time for human actions in an extended loss of AC power (ELAP) scenario is then investigated. Time data and prediction are essential for modeling procedure performance.

Page left intentionally blank.

ACKNOWLEDGEMENTS

The authors wish to thank the Risk-Informed System Analysis (RISA) Pathway of the U.S. Department of Energy (DOE)'s Light Water Reactor Sustainability (LWRS) Program for supporting the research activities presented in this report. In particular, we thank Svetlana Lawrence, RISA Pathway lead, for championing the demonstrations referenced in this report.

Page left intentionally blank.

CONTENTS

FIGURES

TABLES

Page left intentionally blank.

ACRONYMS

A_C	action on the control boards
AC	alternating current
A_F	action in the field
AOP	abnormal operating procedure
API	advanced programming interface
C_C	check on the control boards
C_F	check in the field
CNS	Compact Nuclear Simulator
D_P	decision based on procedures
D_W	decision without procedures
DG	diesel generator
ELAP	extended loss of AC power
EMRALD	Event Modeling Risk Assessment Using Linked Diagrams
EOP	emergency operating procedure
FLEX	diverse and flexible coping strategies
GOMS	Goals Operator Method Selection Rules
HEP	human error probability
HFE	human failure event
HMI	human-machine interface
HRA	human reliability analysis
HUNTER	Human Unimodel for Nuclear Technology to Enhance Reliability
I_P	producing instructions
I_R	receiving instructions
LOFW	loss of feedwater
MCR	main control room
NPP	nuclear power plant
OP	operating procedure
P3	Procedure Performance Predictor
PRIMERA	Procedure-based Investigation Method of EMRALD Risk Assessment
PSF	performance-shaping factor
Rancor	Rancor Microworld Simulator
R_C	receiving information from the control boards
R_F	receiving information from the field

SGTR	steam generator tube rupture
SHEEP	Simplified Human Error Experimental Program
SPAR-H	Standard Plant Analysis Risk-HRA
S_S	selecting on the control boards
S_F	selecting in the field
TCS	turbine control system

USE OF TIME DISTRIBUTIONS TO PREDICT OPERATOR PROCEDURE PERFORMANCE IN DYNAMIC HUMAN RELIABILITY ANALYSIS

1. INTRODUCTION

The Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER), a modeling framework for performing human reliability analysis (HRA) in nuclear power plants (NPPs), was developed by researchers at Idaho National Laboratory. HRA is a discipline centered around quantifying human risk and mitigating error by measuring the impact of human decisions/actions against system performance, with the goal of maximizing safety. It uses various techniques, including static and dynamic approaches to HRA, to assess human error probabilities (HEPs).

Historically, HRA methods have used static HRA—a technique based on worksheet completion—to analyze errors and compute HEPs. Static HRA considers a "snapshot" of a human failure event (HFE) (Lew, et al., 2022), but it doesn't generally account for evolving scenarios and tasks (hence the term "static") nor human variability. The fact that static HRA cannot readily demonstrates other aspects integral to human performance (e.g., task duration, event sequencing, and operator deviations from the norm) makes it less realistic in terms of modeling. In contrast, a novel emerging method known as dynamic HRA paints a more realistic picture. It not only applies Monte Carlo sampling, time distributions, and cognitive modeling methods to HEP analyses, but also evaluates the task time—a critical parameter in HRA. Most importantly, dynamic HRA considers factors such as cognitive processes, decision making abilities, and evolving contexts surrounding human performance (Lew, et al., 2022). But compared to static HRA, its increased modeling fidelity tends to make its outcomes much more complex to generate and evaluate.

By dynamically modeling human performance under various scenarios in complex systems, dynamic HRA can more accurately account for risk and errors, thus better assisting in the mitigation of adverse events. In particular, the HUNTER framework dynamically models operator performance. To assess human error, HUNTER creates a virtual operator that researchers can use to detect and analyze HEPs and consider contributing factors. For example, the information that HUNTER provides on both time and mental workload affords valuable insights into the tasks performed by plant staff (Boring, et al., 2022). Such data can be analyzed to maximize human performance in regard to the plant, the task being performed, and the system being used (Lew, et al., 2022).

The HUNTER reports published to date are as follows:

- Initial framework and demonstration of HUNTER concepts (Boring, et al., 2016)
- Initial standalone software demonstration of HUNTER (Boring, et al., 2022)
- Data collection to support HUNTER modeling (Park, et al., 2022)
- Demonstration of new scenarios and simulator coupling (Lew, et al., 2022)
- Demonstration of linking HUNTER with the Event Modeling Risk Assessment Using Linked Diagrams (EMRALD) probabilistic risk assessment software (Lew, et al., 2023).

The present report largely covers two topics. The first is a new function added to the HUNTER software, called the HUNTER-Procedure Performance Predictor (P3). HUNTER-P3 uses HUNTER's built-in Monte Carlo tools featuring human performance variability in order to identify potential error traps in procedures. Section 2 explores HUNTER-P3 in detail. The second topic is time distribution analysis for generating time inputs for dynamic HRA. Using data collected from the Simplified Human Error Experimental Program (SHEEP) study (Park, et al., 2022), the present study investigated time distributions for task primitives, which are the minimum task units employed in dynamic HRA modeling. Elapsed time for human actions in an extended loss of alternating current (AC) power (ELAP) scenario is then investigated, based on the time distribution data (see Sections 3 and 4 for details). Time data and prediction are essential for modeling procedure performance.

2. HUNTER'S PROCEDURE PERFORMANCE PREDICTOR: SUPPORTING NEW PROCEDURE DEVELOPMENT WITH A DYNAMIC HUMAN RELIABILITY ANALYSIS METHOD[a]

2.1 Review of HUNTER

HUNTER, a dynamic HRA tool designed to be simple to use, was initially based on an effort to create a dynamic implementation of the Standard Plant Analysis Risk-HRA (SPAR-H) method (Gertman, et al., 2005). HUNTER later evolved into a standalone software package for enabling analysts to use procedures and a linked NPP model to realistically simulate human performance—in essence serving as a virtual operator.

The basic structure of HUNTER consists of three functional modules:

- *Task module*—driven by plant operating procedures (OPs)
- *Individual module*—encompassing the performance-shaping factors (PSFs) that affect the operator
- *Environment module*—virtually models the "world" of the simulation (typically a simulator).

The software implementation of HUNTER includes additional modules necessary for executing it as a standalone program. Among these are a scheduler that coordinates the interface between the task, the individual, and the environment, and also coordinates Monte Carlo runs so as to produce performance outcome distributions.

Recent versions of HUNTER (Lew, et al., 2022) incorporate the Rancor Microworld Simulator (Ulrich et al., 2017), a simplified pressurized-water reactor simulator that has been used in a variety of studies involving both students and licensed reactor operators (Park, et al., 2023). The advantages of Rancor center around its simplicity, which makes it more readily usable than full-scope/scale simulators when it comes to collecting operator-in-the-loop data. It also has fewer parameters than do full-scope training simulators, meaning it interfaces more easily with HUNTER, while also readily enabling the collection of empirical data needed to validate HRA models.

Figure 1 shows the basic framework for connecting HUNTER to a simulator. Scenarios are run by using the simulator and representative operators to collect initial human performance data. The simulator is then coupled to HUNTER, and the human performance data are applied to help refine the HUNTER model. For example, data on procedural sticking points or timing can be used to refine and calibrate the basic modeling parameters in HUNTER. A new set of scenarios is then simulated in HUNTER, and the

[a] Adapted from an article (Boring, et al., 2023): Boring, R., Ulrich, T., Lew, R., & Park, J. (2023). HUNTER procedure performance predictor: Supporting new procedure development with a dynamic human reliability analysis method. *AHFE Open Access, 117,* 29-38.

simulation outputs can once again be compared to the available human performance data generated from previous simulator runs. When human performance data are collected from two different scenarios (e.g., steam generator tube rupture [SGTR] and loss of feedwater [LOFW]), one dataset (e.g., that pertaining to LOFW) would be used to calibrate HUNTER, which could then predict performance for SGTR. The predicted and actual SGTR performance data can then be compared so as to validate the modeling effort.

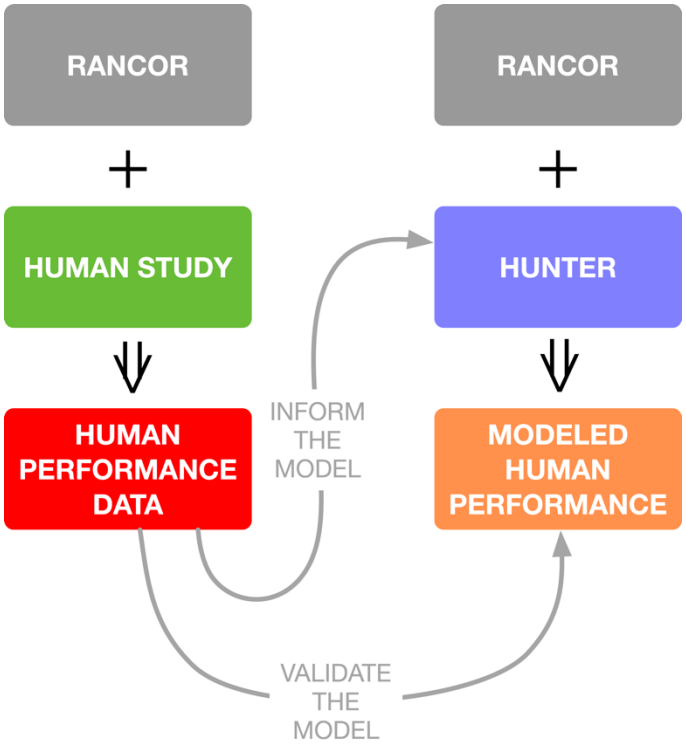


Figure 1. The relationship between HUNTER and simulators when performing dynamic HRA modeling.

Although preliminary work has focused on using Rancor in connection with HUNTER, the general approach can readily be extended to full-scope plant training simulators. In fact, Rancor's human-machine interfaces (HMIs) are based on tools for prototyping digital upgrades in plant simulators (Boring, et al., 2017), meaning that Rancor mimics the functionalities and advanced programming interfaces (APIs) of full-scope simulators. This makes scaling from Rancor to a plant-specific simulator straightforward. The process depicted in Figure 1 may be replicated for plant simulators, assuming human operational performance data are available from the simulator in order to calibrate and validate the HUNTER simulation.

2.2 Introducing HUNTER-P3

Much has been written about control room upgrades and the transition from analog to digital systems (Boring, et al., 2019), but relatively little research has specifically focused on the use of procedures with these new systems. An exception is the case of computer-based procedures, where procedures represent one of the technological systems being introduced into the modernized control room (Lew, et al., 2018). Despite minimal research having specifically been conducted on the application of procedures amid changing concepts of operations, the procedures used to operate any plant system are an important part of the plant's overall HMI.

In recent industry forums held to discuss HUNTER, a strong use case has emerged outside traditional HRA applications for risk assessment. HUNTER's focus on running procedures via a plant simulator opens up a unique and much-needed application in which HUNTER is used to evaluate new procedures. Existing plant OPs benefit from extensive operating experience, industry benchmarking, the sharing of lessons learned (e.g., through the Pressurized Water Reactor Owners Group), and continuous improvement via procedure revisions. However, this is hindered by two new developments:

- Plant upgrades that introduce new digital systems into the main control room (MCR) and require new or extensively modified procedures
- New plants that feature entirely neoteric MCRs that likewise require new procedures.

These Version Null procedures present potential safety and efficiency concerns when it comes to operator performance.

To address this challenge, HUNTER-P3 uses HUNTER's built-in Monte Carlo tools featuring human performance variability in order to identify potential error traps in procedures. In this way, HUNTER-P3 can flag problems with the procedures themselves, or issues with how they are executed by reactor operators. HUNTER-P3 can serve as a screening tool for novel procedures, helping iterate and refine them prior to deployment. Identified error traps serve to prioritize scenarios that warrant empirical evaluation.

HUNTER includes a procedure authoring system that makes it easy to input procedures for driving the Task module. A prototype tool called HUNTER-Gatherer uses natural language processing to automate the inputting of procedures from existing libraries. In this manner, HUNTER-P3 can be used in conjunction with other procedure authoring tools to simulate Version Null procedure performance.

2.3 The Importance of Simulator Coupling

As noted earlier in this section, to realize HUNTER-P3, HUNTER must be coupled to a simulator. Boring et al. (Boring, et al., 2023) explained the importance of synchronous vs. asynchronous coupling for realistic modeling of human-system interactions. Coupling refers to the link between the virtual operator (i.e., HUNTER) and the environment model (i.e., the plant simulator). Asynchronous model coupling occurs via a model code (e.g., a thermohydraulic simulation) designed to operate without evolving inputs. Asynchronous models take all inputs at the beginning and then run in a batch mode to a defined stopping point. For example, an SGTR scenario run may feature the following event sequence:

- *Initial Condition:* Normal operations at 100% power for the starting point
- *Fault:* Fault introduction in the form of a rupture
- *Mitigative Actions:* Manual responses to the rupture, such as reactor trip and safety injection
- *Termination:* Completion of the scenario at a specified time or upon achieving cooldown status.

This sequence is repeated in Monte Carlo fashion, with slight systematic variations (e.g., different times at which mitigative actions are performed) and a degree of stochasticity (e.g., normal variability in plant conditions and timings), generating a range of outcomes for parameters such as total leak time/volume.

The key distinction between synchronous and asynchronous coupling lay in the performing of mitigative actions. In asynchronous coupling (see Figure 2), the mitigative actions are prespecified to be performed at particular points during the scenario runs. In contrast, in synchronous coupling, mitigative actions are part of an iterative plant-operator feedback loop in which the operator responds to and alters the plant conditions, and then the plant proceeds from there. The value of synchronous coupling is predicated on three assumptions:

1. Human actions are responsive to emerging plant conditions and cannot be completely predicted a priori.
2. Human actions will change plant conditions in a way that meaningfully alters the course of the scenario in an evolving manner that cannot be completely predicted.
3. The timing of human actions, as well as the selection of specific actions from among many possible mitigation options, changes plant conditions in ways not fully predictable a priori.

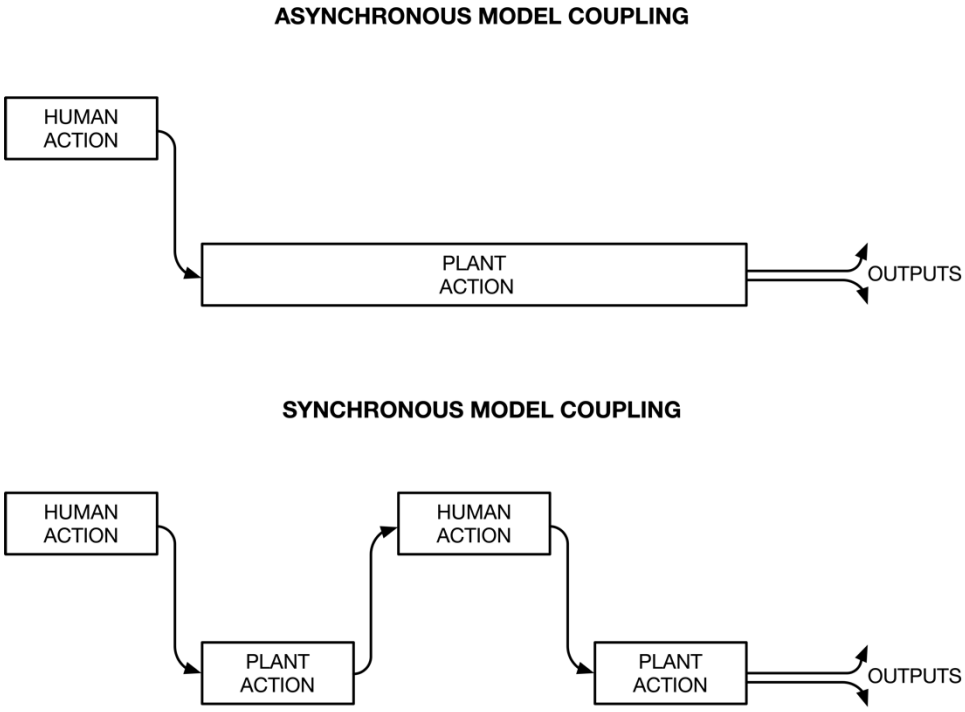


Figure 2. Human-plant interaction for asynchronous and synchronous coupling (from Boring et al., 2023).

The common theme pertaining to these assumptions is that myriad plant and human outcomes become possible as a scenario unfolds. For example, a 30-second delay in responding to an upset may change the course of that event in a manner that necessitates completely different actions. In general, asynchronous models do not fully simulate cases in which the dynamics of the evolving operational context may lead to deviations from the nominal path.

Another way to consider asynchronous vs. synchronous is in the context of normative vs. descriptive models. Bell, Raiffa, and Tversky (Bell, et al., 1988) delineate normative models as being those that predict an ideal outcome, whereas descriptive models reflect actual performance. Asynchronous modeling typically results in normative outcomes, meaning the expected normal or ideal case. Synchronous modeling results in descriptive models, meaning the actual case. The latter is essential for understanding realistic courses of operator behavior when using procedures. HUNTER-P3 synchronously couples its virtual operator representation with the plant full-scope simulator in order to predict how operators would actually perform when using procedures.

Note that HUNTER-P3 possesses all the features necessary to automate plant operations if coupled to an actual plant instead of a simulator. However, its performance would not align with the normative performance expected from automation. Rather, HUNTER-P3 would provide an operating context that

incorporates operator shortcomings, and such shortcomings are generally antithetical to the goals of automation.

2.4 Evaluating New Procedures

In this section we consider how digital upgrades affect procedures. For example, a newly modernized digital turbine control system (TCS) may largely mimic the functionality of the existing electro-hydraulic control it replaces at the plant, but its new digital interface and control system require slightly different operator actions, necessitating new procedural steps—or perhaps even whole new procedures entirely.

A thorough operating experience review (IAEA, 2018) can identify potential problem areas pertaining to the new system and the use of procedures. However, all existing experience may be based on the legacy electro-hydraulic control, meaning limited experience with novel systems. Moreover, regarding the example given in the previous paragraph, the TCS is one of the first wholly digital control subsystems installed as an upgrade at most U.S. NPPs, as it offers a potentially high return on investment through possible power uprates. With first-of-a-kind installations of digital systems, there is little operating experience to draw upon to ensure that the procedures adequately support operator usage of the new system.

In the absence of adequate operating experience to provide confidence in novel procedures, the next course of action is to perform empirical evaluations with operators in the loop. This approach is identical to the types of human factors validation activities performed as part of upgrades. Scenarios are identified that represent the range of activities performed in regard to the system or procedure, with a particular emphasis on any critical safety functions. These scenarios capture the continuum ranging from frequent and normal activities to rare and abnormal events. In a TCS, this would cover reactor startup, shutdown, and power evolutions under normal operations, along with upset conditions such as failed governor valves or grid disturbances. Operators undergo these scenarios using the new system and accompanying new procedures, and any performance deficiencies (e.g., confusion, erroneous actions, or response delays) are documented and corrected in the system/procedure. Though highly effective, this approach is costly in terms of the staffing needed to set up and carry out the studies. In addition, it is only effective to the degree that the scenarios anticipate the actual range of use, and the sample size of operators may be limited, depending on the plant's ability to deviate from operating and training schedules so as to support development and evaluation activities.

HUNTER-P3 presents a novel third approach to identifying issues pertaining to new procedures. HUNTER is coupled to the plant's training simulator with the updated control system via the simulator's available API. The API allows for monitoring and controlling all simulated plant parameters. Thus, the plant indicators that should be monitored by the virtual operator can be fed into HUNTER, and any control actions taken by the virtual operator in HUNTER can be input to the simulator, enabling HUNTER to function like an actual operator at the control panels. Typically, the API also affords control of instructor station functions, meaning it is possible to start and stop the simulator and insert faults. This functionality is used for the Monte Carlo repeated trials. In this manner, HUNTER-P3 controls the plant's new control system by following the new procedures embedded in HUNTER's Individual module.

As noted, HUNTER-P3 simulates the proceduralized activities—not in a normative or idealized manner, but in a way that incorporates a realistic level of operator fallibility. The Individual module accounts for those factors that may impinge on optimal performance. For example, the elevated stress PSF may decrease the time it takes to complete a given task, whereas the complexity introduced by simultaneously performing multiple tasks may increase the likelihood of skipping a procedure step. Hollnagel (Hollnagel, 2017) suggests that a disparity often exists between work as imagined and work as done. HUNTER-P3 captures this at two different levels of analysis:

- *Operator level:* Contextual factors hinder the operator from performing the procedures perfectly, potentially resulting in less-than-perfect plant performance.
- *Procedure level:* The procedure does not adequately cover the use context, meaning that even following the procedures perfectly will fail to result in perfect plant performance.

Operator-level issues in procedure performance result from systematic decrements in work as done. HUNTER can account for these via its Individual module. Procedure-level issues arise when work as imagined inadequately covers the operating envelope of the system. HUNTER-P3 can account for work as imagined by modeling what-if plant contexts in its Environment module (e.g., inserting faults so as to stress-test the procedure). To avoid confounding operator-level issues with procedure-level issues, the Individual module and Environment module factors can be manipulated separately.

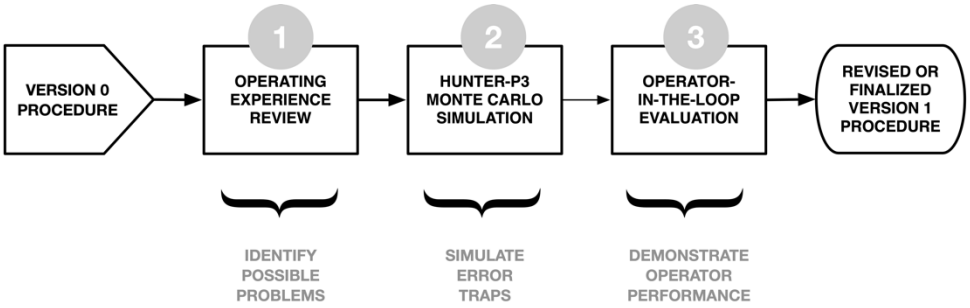


Figure 3. Three stages in validating Version Null procedures.

HUNTER-P3 works to identify overall error traps in the procedures and in how operators use them. This enables procedure writers to refine the procedures as necessary prior to deployment. This approach inserts a needed procedure analysis tool between operating experience reviews and empirical evaluations. Potential problem areas identified in the operating experience reviews can be accounted for in the HUNTER-P3 modeling. If HUNTER-P3 reveals such problem areas, this process (see Figure 3) serves as a screening tool for identifying those use cases that should be further explored through operator studies. A graded approach can be employed in which only one or two of the three phases are performed, and in some circumstances the HUNTER-P3 evaluation may be sufficient to preclude operator studies, thanks to sufficiently identifying the procedural issues to be resolved.

As illustrated back in Figure 1, HUNTER models are informed by human performance data. It may initially seem paradoxical to suggest HUNTER-P3 modeling be performed prior to the operator-in-the-loop evaluation. To forego this step, assume that the HUNTER-P3 model is mature, meaning a phase has already been completed to calibrate individual plant operational characteristics to HUNTER, and can be generalized to future scenarios. If separate human performance studies are necessary prior to running HUNTER, Stages 2 and 3 in Figure 3 may need to be switched. However, once a HUNTER model is calibrated to a given plant, ongoing human performance studies should be unnecessary for the purpose of model building.

A current limitation of this approach is that most plant training simulators cannot run at speeds greatly exceeding real time, yet their purpose is to allow for human interactions that closely follow plant changes. Timing precision stands as a roadblock for accelerated operations, and plant simulators aim to reflect actual plant response timings as closely as possible. A so-called “headless” version of Rancor can be run that is not linked to any external timing constraints but can pass along timing durations. While Rancor may run a particular plant function hundreds of times faster than real time, it logs the time required for the plant evolution, enabling HUNTER to respond as if the actual time had passed, allowing for its use in faster-than-real-time synchronous coupling. This consideration matters because using Monte

Carlo simulation runs to capture the range of human performance in HUNTER may prove dauntingly slow for a plant's actual training simulator. A 20-minute scenario that is run 500 times in Rancor faster than real time in under a minute may take 10,000 minutes (i.e., 167 hours) when performed using an unoptimized full-scope plant simulator in real time. This limitation can be overcome by parallelizing simulator installations for multi-core simultaneous execution, or by optimizing the simulator vendor's software code so as to enable it to run at faster-than-real-time speeds.

3. INVESTIGATION OF TIME DISTRIBUTIONS FOR TASK PRIMITIVES TO SUPPORT DYNAMIC HUMAN RELIABILITY ANALYSIS USING HUNTER[b]

3.1 Background

Time information on human actions is very important in HRA. HRA methods such as the Technique for Human Error Rate Prediction (Swain & Guttmann, 1983), Human Cognitive Reliability (Parry, et al., 1992), and Korean Standard HRA (Jung, et al., 2005) use time information and time response curves to estimate diagnosis HEPs. In the human factors engineering program outlined in NUREG-0711 (O'Hara, et al., 2012), HRA methods are used to investigate HFE feasibility by comparing time required against time available. "Time required" is the duration needed by operators to perform a task, while "time available" is the actual time operators are given to complete said task. If the time required for human actions exceeds the time available for an HFE, this is considered a guaranteed failure (HEP = 1.0), and the plant state is assumed to be irreversible.

To date, time windows for calculating time available have been determined via thermohydraulic analysis, which produces accurate values based on simulations. On the other hand, determining time required relies on structured interviews with instructors, operators, and other knowledgeable experts. Structured interviews would be useful for easily gathering time required estimations pertaining to human actions. However, it may be difficult to objectively explain how time required is measured, estimate time distributions or uncertainties in light of each individual and/or every trial, and judge whether human actions can be completed within the given timeframe if unexpected variables interrupt said actions.

An important feature of the HUNTER method is its focus on timing data and the overall time duration of task performance. Due to the extremely dynamic nature of human performance, this focus is critical for ensuring a robust understanding of human error in complex systems. As PSFs impact human performance to varying degrees as work on a task continues, inclusion of this timing structure can foster a more precise understanding of when human errors are more likely to occur. The HUNTER method uses Goals, Operators, Methods, and Selection Rules (GOMS)-HRA to store and manage these task timings and durations (Boring & Rasmussen, 2016; Ulrich, et al., 2017). GOMS-HRA enables each task to be broken down into subtask primitives that can then be summed at various levels to provide timing data on procedure steps or the task performance as a whole. While this allows for capturing instances in which task failure is linked to running out of time—as opposed to an error being made—it also provides a critical contextual data point for digging into human performance data and better capturing when error rates rise/fall and when PSFs trigger human errors.

Our research team developed an HRA data collection framework called the Simplified Human Error Experimental Program (SHEEP) to complement full-scope simulator studies and collect input data for

[b] Adapted from an article (Park, et al., 2023): Park, J., Yang, T., Kim, J., & Boring, R.L. (2023). An investigation of time distributions for task primitives to support the HUNTER dynamic human reliability analysis. *AHFE Open Access, 117,* 21-28.

dynamic HRA such as that afforded by HUNTER (Park, et al., 2022). The SHEEP framework infers full-scope data based on experimental data collected from simplified simulators—specifically, Rancor and the Compact Nuclear Simulator (CNS). Under the SHEEP framework, our research team experimentally collected human reliability data from 36 student operators and 36 professional operators using CNS and Rancor. The human errors and performance measurements collected from experiments to date have been analyzed, and are discussed in previous research (Park, et al., 2022; Park, et al., 2023).

Under the umbrella of the SHEEP framework, the present study aims to investigate time distributions for task primitives defined in the GOMS-HRA method (Boring & Rasmussen, 2016). GOMS-HRA was developed to provide cognition-based time and HEP information for dynamic HRA calculations performed in the HUNTER framework. Here, we investigate time distributions for GOMS-HRA task primitives, based on the SHEEP database, which includes experimental data pertaining to 20 student operators and 20 professional operators using Rancor. The experimental data were used as the foundation for investigating the time required for GOMS-HRA task primitives in order to satisfy 13 different statistical distributions. The resulting time distributions were then compared and discussed. GOMS-HRA task level primitives are mapped to procedure level primitives (Boring, et al., 2017), which allows the use of GOMS-HRA timing data to inform the task duration for procedure-based activities.

3.2 GOMS-HRA Task Primitives

GOMS-HRA (Boring & Rasmussen, 2016) was developed to provide cognition-based time and HEP information for dynamic HRA calculations performed within the HUNTER framework. GOMS-HRA has been used to model proceduralized activities and evaluate user interactions with HMIs in human factors research. As a predictive method, GOMS-HRA is well-equipped to simulate human actions under specific circumstances in a particular scenario. The basic approach of GOMS-HRA consists of three steps: (1) breaking up human actions into a series of task-level primitives, (2) allocating time and error values to each task-level primitive, and (3) predicting human actions or task durations.

Table 1 lists the GOMS-HRA task primitives. GOMS-HRA originally suggested 12 task primitives, performed either in control rooms or in the field. However, the present analysis focuses squarely on the five task primitives (i.e., A_C, C_C, R_C, S_C, and D_P) highlighted in grey in the table. The data were collected as part of the SHEEP experiment (see Section 3.1), in which a single operator used procedures to monitor and control a simulator in a control room environment, without any field operation. As such, the task primitives pertaining to field operations (i.e., A_F, C_F, R_F, and S_F), decision-making without procedures (i.e., D_W), and communication between operators (i.e., I_P and I_R) were excluded in the present study.

Table 1. GOMS-HRA task primitives.

Task Primitives	Description
A_C	Performing required physical actions on the control boards
A_F	Performing required physical actions in the field
C_C	Looking for required information on the control boards
C_F	Looking for required information in the field
R_C	Obtaining required information on the control boards
R_F	Obtaining required information in the field
I_P	Producing verbal or written instructions
I_R	Receiving verbal or written instructions
S_C	Selecting or setting a value on the control boards
S_F	Selecting or setting a value in the field
D_P	Making a decision based on procedures
D_W	Making a decision without available procedures

3.3 The SHEEP Experiment Data

The SHEEP data were collected from 36 student operators and 36 professional operators using Rancor (Ulrich, et al., 2017). Most of the professional operators were licensed reactor operators currently employed at NPPs. They were all operators on shift (i.e., shift supervisor, shift technical advisor, reactor operator, or turbine operator) or instructors at the training center. The student operators were all undergraduate seniors or graduate students from the Department of Nuclear Engineering at Chosun University. They were knowledgeable about NPP systems and operations, having completed a significant portion of their coursework, which included courses such as "Introduction to Nuclear Engineering," "Reactor Theory," "Reactor Control," and "Simulator Operation."

The present analysis investigated time distributions for the GOMS-HRA task primitives, which were found in several different procedures and thus pertained to different contexts. In essence, different procedures used in NPPs correspond to different goals. For example, OPs are designed to stably achieve different operating modes such as startup or hot standby, whereas emergency OPs (EOPs) are mainly instructions for rapidly cooling down reactors. Accordingly, this study delineated seven different procedure sets, listed in

Table 2. OP-03/OP-04 and OP-05/OP-06 were combined into single procedural sets because they are used in a scenario for achieving a singular goal. OP-04 and OP-06 are parts of OP-03 and OP-05.

Table 2. Procedure information used for the SHEEP experiment.

Procedure Set No.	Procedures Included	Description	Related Scenario
1	OP-01	Explains how to start up and operate Rancor in auto mode.	Scenario #1 (fully auto startup)
2	OP-02	Details the process of shutting down Rancor.	Scenario #2 (shutdown)
3	OP-03 & OP-04	Explains how to start up and operate Rancor in control rod manual operation mode.	Scenario #3 (manual rod control during startup)
4	OP-05 & OP-06	Explains how to start up and operate Rancor in feedwater manual operation mode.	Scenario #4 (manual feedwater flow control during startup)
5	Abnormal Operating Procedure (AOP)-01	Shuts down the plant in an expedient manner.	Scenarios #5–#10 (failure of a reactor coolant pump, failure of a control rod, failure of a feedwater pump, turbine failure, SGTR, and LOFW)
6	EOP-01	Provides actions to minimize reactor coolant leakage into the secondary system following a SGTR.	Scenario #9 (SGTR)
7	EOP-02	Provides actions to diagnose and mitigate LOFW.	Scenario #10 (LOFW)

3.4 Time Distribution Analysis Results

Table 3 and Table 4 show, respectively, the number of tasks used for the time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the OP01 procedure, in regard to both participant types (i.e., student operators vs. professional operators). A total of 1,591 tasks were performed by 20 student operators and 20 professional operators in executing procedure OP-01 within Rancor. The number of tasks observed for the student operators (808) exceeded that for the professional operators (783). Differences in the number of tasks when comparing the two participant types result from situations in which a participant performs additional steps that can be omitted within the procedure context, or cannot continue a scenario because the reactor has abnormally tripped. Ultimately, the task types generated statistically significant results, as shown in Table 4.

Table 3. The number of tasks used for time distribution analysis (OP-01).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	A_C	C_C	R_C	S_C	D_P		
Student operators	122	396	117	75	98	808	1,591
Professional operators	112	388	117	73	93	783	

Table 4. Time distribution analysis for the five GOMS-HRA task primitives in the OP-01 procedure, in regard to both participant types (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Normal (after Box-Cox transformation)	<0.005	<0.005	<0.005	<0.005	<0.005	0.005	<0.005	<0.005	0.022	<0.005
Lognormal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.022	<0.005
Exponential	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	0.007	<0.003	<0.003
2-parameter exponential	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Weibull	<0.010	<0.010	0.016	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
3-parameter Weibull	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Smallest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Largest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Gamma	<0.005	<0.005	0.018	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Logistic	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Loglogistic	<0.005	<0.005	<0.005	<0.005	<0.005	0.015	<0.005	<0.005	0.006	<0.005
Normal (after Johnson transformation)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 5 and Table 6 show, respectively, the number of tasks used for the time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the OP-02 procedure, in regard to both participant types. A total of 763 tasks were performed by 20 student operators and 20 professional operators in executing procedure OP-02 within Rancor. The number of tasks for the student operators (374) was lower than that for the professional operators (389). For the student operators, the task primitive elapsed time for S_C was statistically significant on the normal (after Johnson transformation) distribution. For the professional operators, on the other hand, the elapsed time for S_C proved statistically significant on the normal (after Box-Cox transformation) distribution, the largest extreme value distribution, the gamma distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution.

Figure 4 and Figure 5 summarize the most optimal time distributions representing the highest p-values from among the time distributions for each task primitive. The figures show the normal (after Johnson transformation) distributions of S_C for student operators and professional operators carrying out the OP-02 procedure. Based on these time distributions, the student operator data reflected an elapsed time of 20.68 seconds for S_C, while the professional operator data reflected 12.49 seconds.

Table 5. The number of tasks used for the time distribution analysis (OP-02).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	AC	CC	RC	SC	DP		
Student operators	70	90	56	75	83	374	763
Professional operators	75	83	60	82	89	389	

Table 6. Time distribution analysis for the five GOMS-HRA task primitives in the OP-02 procedure, in regard to both participant types (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Normal (after Box-Cox transformation)	<0.005	<0.005	<0.005	0.009	<0.005	0.005	<0.005	<0.005	0.160	<0.005
Lognormal	<0.005	<0.005	<0.005	0.009	<0.005	<0.005	<0.005	<0.005	0.009	<0.005
Exponential	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003
2-parameter exponential	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Weibull	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
3-parameter Weibull	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Smallest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Largest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	0.073	<0.010
Gamma	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.122	<0.005
Logistic	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Loglogistic	<0.005	<0.005	<0.005	0.024	<0.005	<0.005	<0.005	<0.005	0.155	<0.005
Normal (after Johnson transformation)	N/A	N/A	N/A	0.102	N/A	N/A	N/A	N/A	0.788	N/A

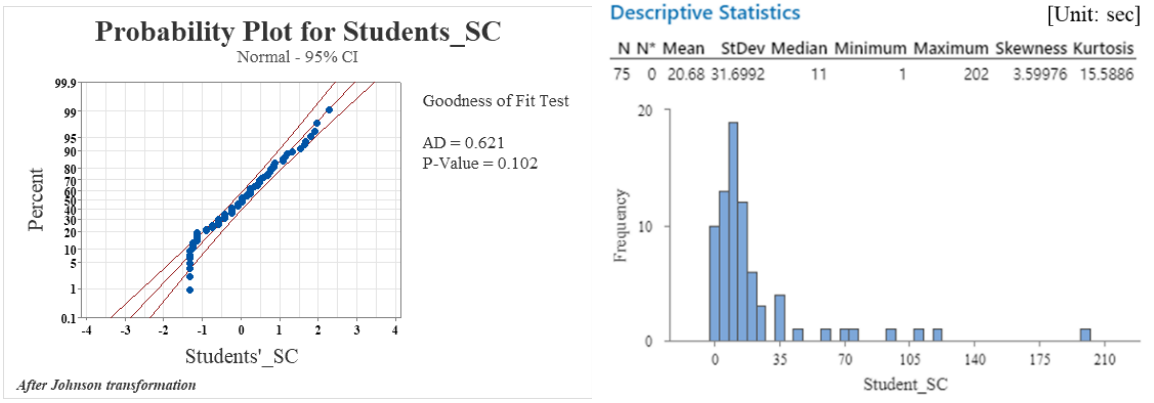


Figure 4. Normal (after Johnson transformation) distribution of S_C for student operators carrying out the OP-02 procedure.

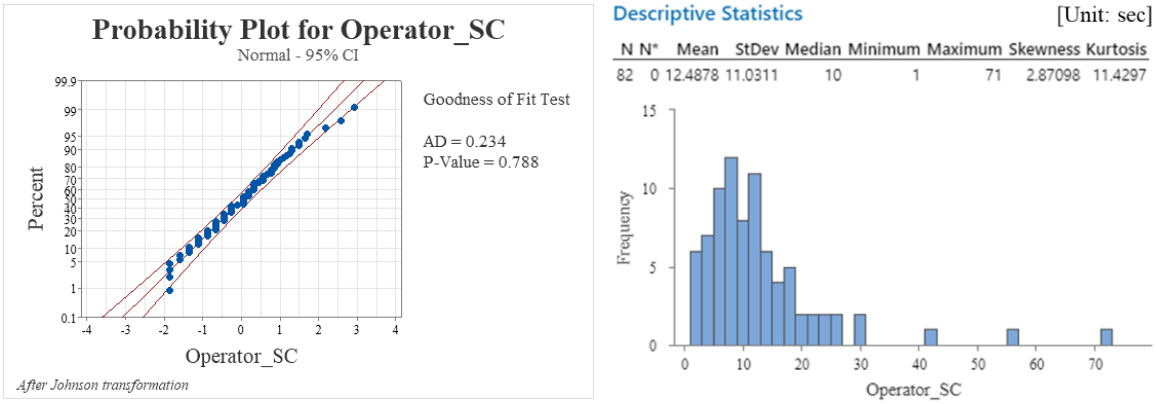


Figure 5. Normal (after Johnson transformation) distribution of S_C for professional operators carrying out the OP-02 procedure.

Table 7 and Table 8 show, respectively, the number of tasks used for time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the OP-03 and OP-04 procedures, in regard to both participant types. A total of 1,728 tasks were performed by 20 student operators and 20 professional operators in executing the OP-03 and OP-04 procedures within Rancor. The number of tasks for student operators (893) exceeded that for professional operators (835). Based on the professional operator data, the elapsed time for S_C proved statistically significant on the normal (after Johnson transformation) distribution. Figure 6 shows the normal (after Johnson transformation) distribution of S_C for professional operators carrying out the OP-03 and OP-04 procedures. Per the time distribution, the average S_C value reflected in the professional operator data was 45.16 seconds.

Table 7. The number of tasks used for the time distribution analysis (OP-03 & OP-04).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	A_C	C_C	R_C	S_C	D_P		
Student operators	182	385	112	47	167	893	1,728
Professional operators	148	368	113	45	161	835	

Table 8. Time distribution analysis for the five GOMS-HRA task primitives in the OP-03 & OP-04 procedures, in regard to both participant types (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Normal (after Box-Cox transformation)	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.023	0.023
Lognormal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.023	0.023
Exponential	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	0.003	<0.003	<0.003
2-parameter exponential	<0.010	<0.010	0.015	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Weibull	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
3-parameter Weibull	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Smallest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Largest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Gamma	<0.005	<0.005	0.009	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Logistic	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Loglogistic	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.013	0.013
Normal (after Johnson transformation)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	0.193	N/A

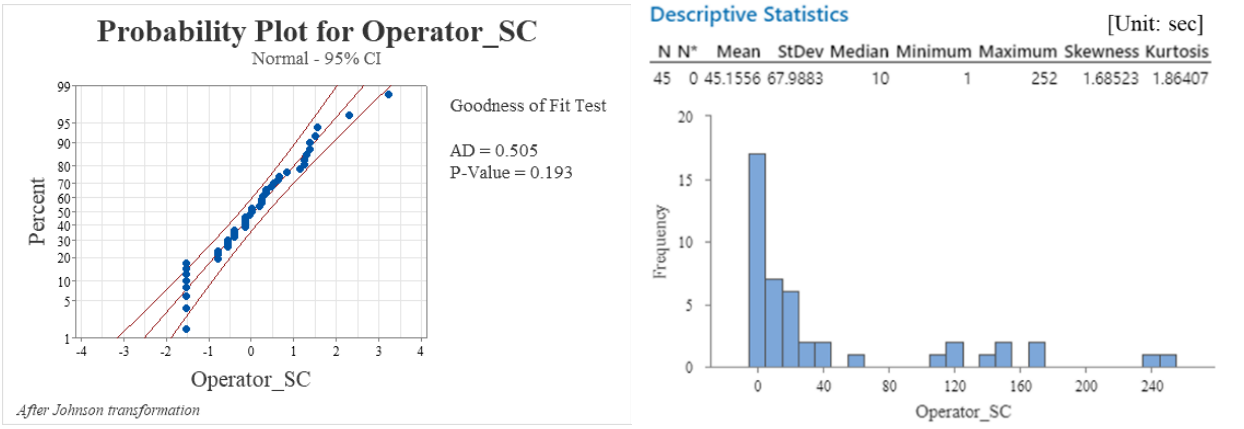


Figure 6. Normal (after Johnson transformation) distribution of S_C for professional operators carrying out the OP-03 & OP-04 procedures.

Table 9 and Table 10 show, respectively, the number of tasks used for time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the OP-05 and OP-06 procedures, in regard to both participant types. A total of 1,632 tasks were performed by 20 student operators and 20 professional operators in executing the OP-05 and OP-06 procedures within Rancor. The number of tasks for student operators (789) was lower than that for professional operators (843). No statistically significant results were uncovered by this time distribution analysis.

Table 9. The number of tasks used for the time distribution analysis (OP-05 & OP-06).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	A_C	C_C	R_C	S_C	D_P		
Student operators	130	368	105	71	115	789	1,632
Professional operators	129	397	117	78	122	843	

Table 10. Time distribution analysis for the five GOMS-HRA task primitives in the OP-05 & OP-06 procedures, in regard to both participant types (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Normal (after Box-Cox transformation)	<0.005	<0.005	<0.005	0.009	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Lognormal	<0.005	<0.005	<0.005	0.009	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Exponential	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	<0.003	0.005	<0.003	<0.003
2-parameter exponential	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Weibull	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
3-parameter Weibull	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Smallest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Largest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Gamma	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Logistic	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Loglogistic	<0.005	<0.005	<0.005	0.036	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Normal (after Johnson transformation)	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

Table 11 and Table 12 show, respectively, the number of tasks used for time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the AOP-01 procedure, in regard to both participant types. A total of 2,068 tasks were performed by 20 student operators and 20 professional operators in executing the AOP-01 procedure within Rancor. The number of tasks for student operators (1,055) exceeded that for professional operators (1,013). Per the student-operator task primitive data, the elapsed time for C_C proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution. In addition, for both the student operators and professional operators, the data showed the elapsed time for S_C to be statistically significant on the normal (after Box-Cox transformation) distribution and the lognormal distribution.

Figure 7, Figure 8, and Figure 9 summarize the most optimal time distributions, as indicated by the time distribution analysis results. Figure 7 and Figure 8 include the normal (after Johnson transformation) distribution of C_C and the lognormal distribution of S_C, as generated based on the student operator tasks performed in accordance with the AOP-01 procedure. Figure 9 shows the lognormal distribution of S_S that was generated based on professional operator tasks performed in accordance with the AOP-01 procedure. Per the time distributions, the student operator data showed an average of 17.19 seconds for C_C and 19.58 seconds for S_C, while the professional operator data showed an average of 8.00 seconds for S_C.

Table 11. The number of tasks used for the time distribution analysis (AOP-01).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	A_C	C_C	R_C	S_C	D_P		
Student operators	222	54	212	143	424	1,055	2,068
Professional operators	222	58	200	154	379	1,013	

Table 12. Time distribution analysis for the five GOMS-HRA task primitives in the AOP procedure, in regard to both participant types (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Normal (after Box-Cox transformation)	<0.005	0.205	<0.005	0.062	<0.005	<0.005	<0.005	<0.005	0.068	<0.005
Lognormal	<0.005	0.205	<0.005	0.062	<0.005	<0.005	<0.005	<0.005	0.068	<0.005
Exponential	<0.003	0.009	<0.003	<0.003	<0.003	<0.003	0.007	<0.003	<0.003	<0.003
2-parameter exponential	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	0.013	<0.010
Weibull	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
3-parameter Weibull	<0.005	0.019	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	0.014	<0.005
Smallest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Largest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010
Gamma	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Logistic	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005	<0.005
Loglogistic	<0.005	0.174	<0.005	0.013	<0.005	<0.005	<0.005	<0.005	0.015	<0.005
Normal (after Johnson transformation)	N/A	0.658	N/A	N/A	N/A	N/A	N/A	N/A	N/A	N/A

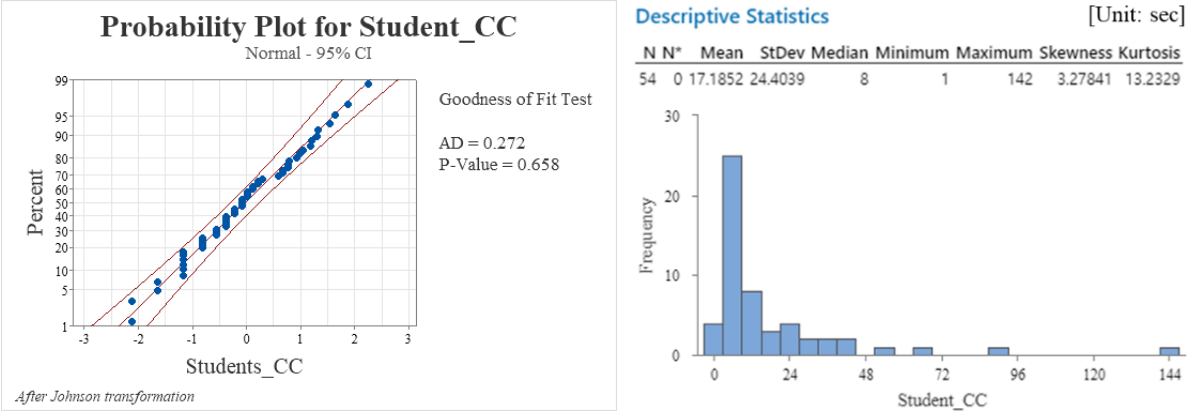


Figure 7. Normal (after Johnson transformation) distribution of C_C for student operators carrying out the AOP-01 procedure.

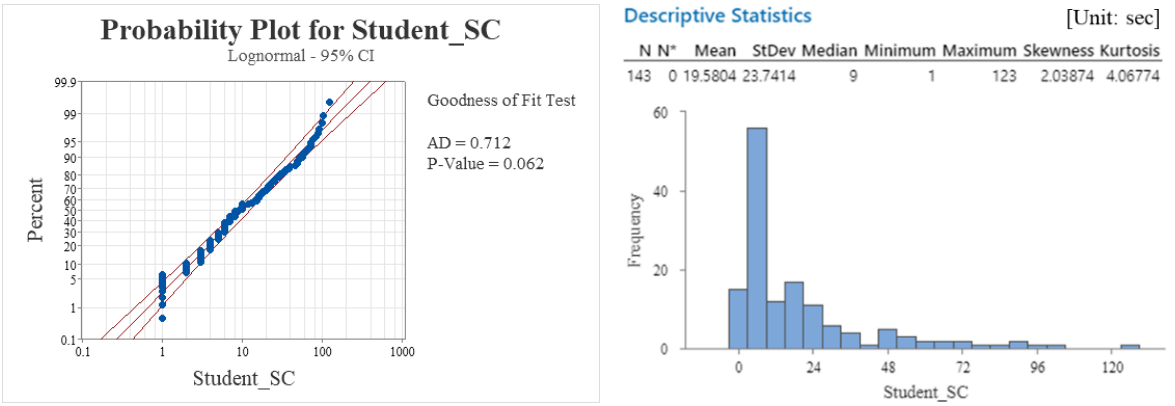


Figure 8. Lognormal distribution of S_C for student operators carrying out the AOP-01 procedure.

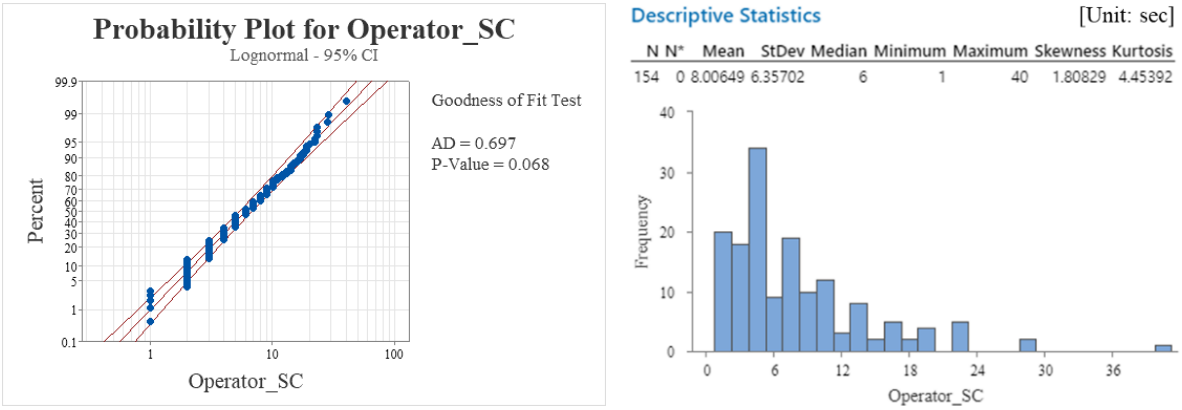


Figure 9. Lognormal distribution of S_C for professional operators carrying out the AOP-01 procedure.

Table 13 and Table 14 show, respectively, the number of tasks used for the time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the EOP-01 procedure, in regard to both participant types. A total of 490 tasks were performed by 20 student operators and 20 professional operators in executing the EOP-01 procedure within Rancor. The number of tasks for student operators (248) exceeded that for professional operators (242). Per the student-operator task primitive data, the elapsed time for A_C proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the 2-parameter exponential distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution. And the elapsed time for S_C proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the exponential distribution, the 2-parameter exponential distribution, the Weibull distribution, the 3-parameter Weibull distribution, the gamma distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution. In addition, the elapsed time for D_P proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, and the exponential distribution. On the other hand, per the professional operator data, the elapsed time for A_C proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the 3-parameter Weibull distribution, the largest extreme value distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution. The elapsed time for S_C proved statistically significant on the normal distribution, the normal (after Box-Cox transformation) distribution, the Weibull distribution, the 3-parameter Weibull distribution, the smallest and largest extreme value distributions, the gamma distribution, the logistic distribution, and the loglogistic distribution.

Figure 10, Figure 11, Figure 12, Figure 13, and Figure 14 reflect the most optimal time distributions from the time distribution analysis results. More specifically, Figure 10, Figure 11, and Figure 12 show, respectively, the normal (after Johnson transformation) distributions of A_C and S_C and the lognormal distribution of D_P for student operators carrying out the EOP-01 procedure. The average elapsed time from the time distributions are 11.92 seconds for A_C, 9.80 seconds for S_C, and 6.14 seconds for D_P. Figure 13 and Figure 14 show, respectively, the normal distribution (after Johnson transformation) of A_C and the 3-parameter Weibull distribution of S_C for professional operators carrying out the EOP-E0 procedure. The average elapsed times from these time distributions are 7.21 seconds for A_C and 5.30 seconds for S_C.

Table 13. The number of tasks used for the time distribution analysis (EOP-01).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	A_C	C_C	R_C	S_C	D_P		
Student operators	59	80	49	10	50	248	490
Professional operators	57	78	49	10	48	242	

Table 14. Time distribution analysis for the five GOMS-HRA task primitives in the EOP-01 procedure, in regard to both participant types (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	<0.005	0.014	<0.005	<0.005	<0.005	<0.005	0.237	<0.005
Normal (after Box-Cox transformation)	0.374	<0.005	0.010	0.653	0.070	0.340	<0.005	<0.005	0.237	0.041
Lognormal	0.374	<0.005	0.010	0.404	0.070	0.340	<0.005	<0.005	0.031	0.041
Exponential	0.023	<0.003	0.018	0.486	0.051	<0.003	<0.003	<0.003	0.021	<0.003

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
2-parameter exponential	0.083	<0.010	<0.010	>0.250	0.011	<0.010	<0.010	<0.010	0.012	0.034
Weibull	<0.010	<0.010	<0.010	0.189	<0.010	0.015	<0.010	<0.010	0.236	0.022
3-parameter Weibull	0.013	<0.005	<0.005	0.404	<0.005	0.084	<0.005	<0.005	0.254	0.006
Smallest extreme value	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	<0.010	0.092	<0.010
Largest extreme value	<0.010	<0.010	<0.010	0.037	<0.010	0.088	<0.010	<0.010	0.227	0.016
Gamma	<0.005	<0.005	0.007	0.208	0.006	0.167	<0.005	<0.005	0.182	0.047
Logistic	<0.005	<0.005	<0.005	0.016	<0.005	<0.005	<0.005	<0.005	0.235	<0.005
Loglogistic	>0.250	<0.005	<0.005	>0.250	0.032	0.233	<0.005	<0.005	0.104	0.017
Normal (after Johnson transformation)	0.563	N/A	N/A	0.763	N/A	0.364	N/A	N/A	N/A	N/A

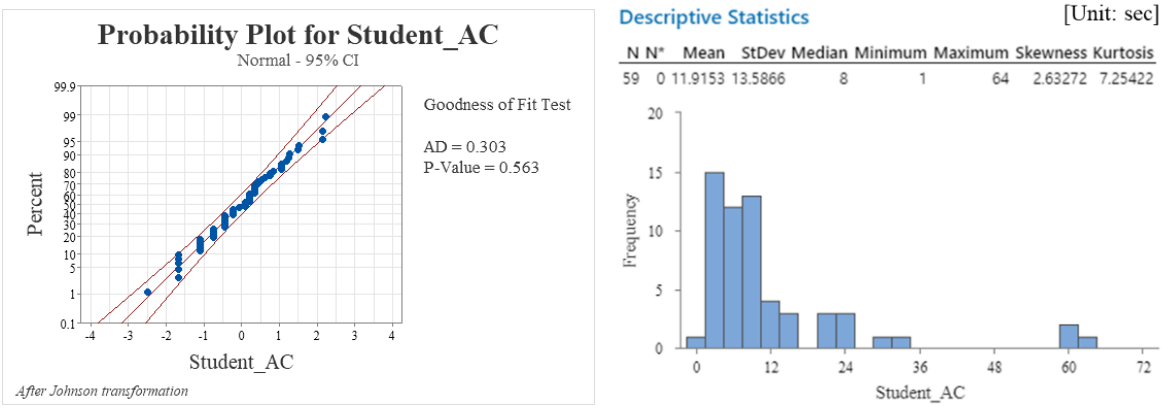


Figure 10. Normal (after Johnson transformation) distribution of A_C for student operators carrying out the EOP-01 procedure.

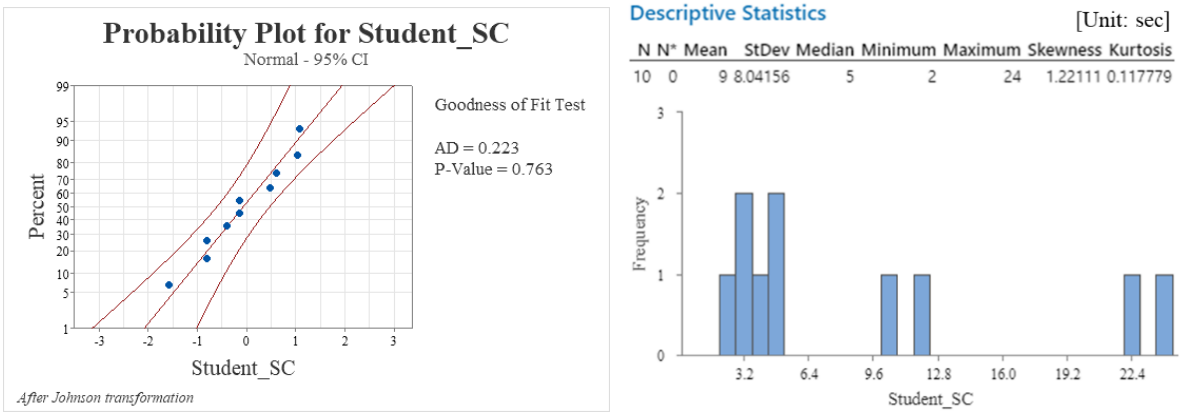


Figure 11. Normal (after Johnson transformation) distribution of S_C for student operators carrying out the EOP-01 procedure.

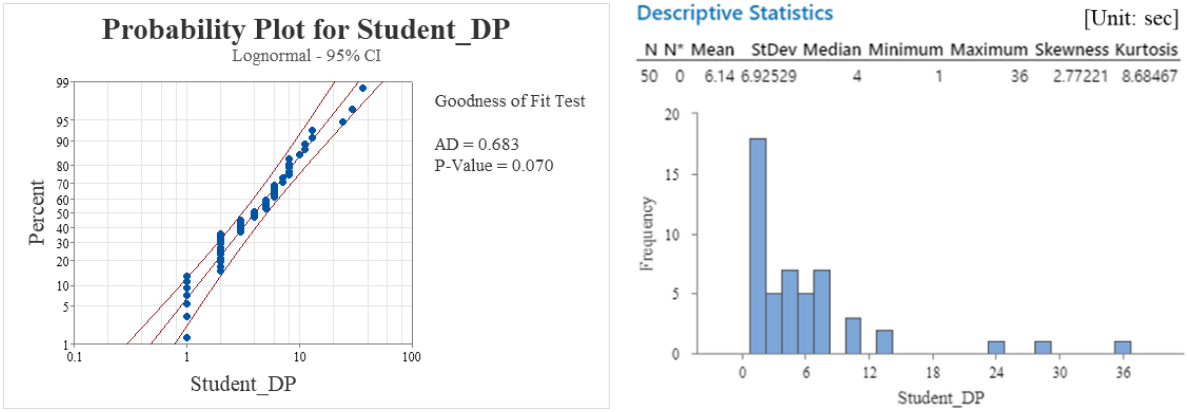


Figure 12. Lognormal distribution of D_P for student operators carrying out the EOP-01 procedure.

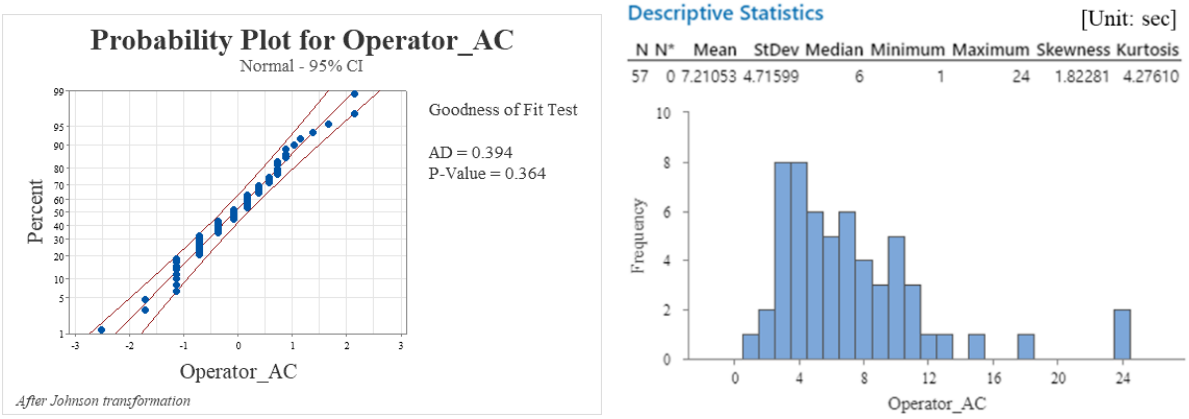


Figure 13. Normal (after Johnson transformation) distribution of A_C for professional operators carrying out the EOP-01 procedure.

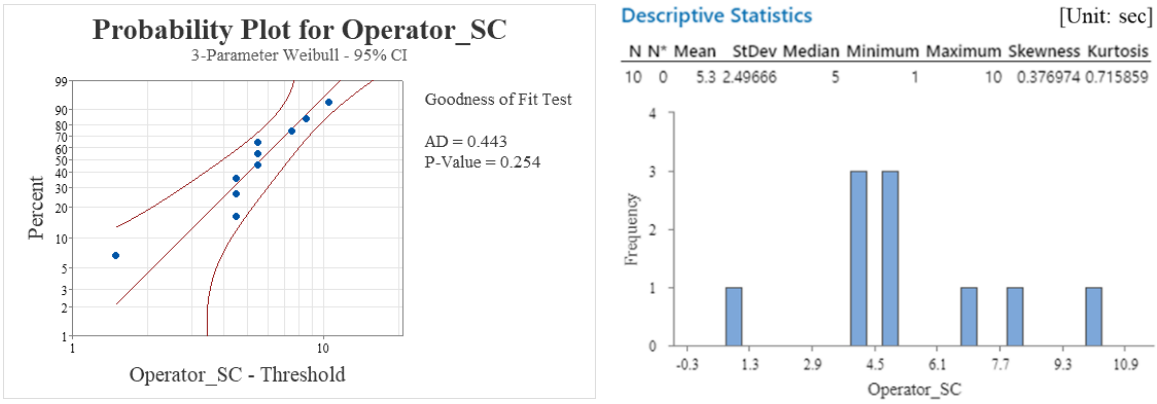


Figure 14. 3-parameter Weibull distribution of S_C for professional operators carrying out the EOP-01 procedure.

Table 15 and Table 16 show, respectively, the number of tasks used for time distribution analysis and the goodness-of-fit test results for the 13 statistical elapsed time distributions pertaining to the five GOMS-HRA task primitives in the EOP-02 procedure, in regard to both participant types. A total of 195 tasks were performed by 20 student operators and 20 professional operators in executing the EOP-02 procedure within Rancor. The number of tasks for the student operators (114) exceeded that for the professional operators (81). Per the student-operator task primitive data, the elapsed time for A_C proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the 2-parameter exponential distribution, the Weibull distribution, the 3-parameter Weibull distribution, the gamma distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution. The elapsed time for S_C proved statistically significant on the normal distribution, the normal (after Box-Cox transformation) distribution, the lognormal distribution, the exponential distribution, the 2-parameter exponential distribution, the Weibull distribution, the 3-parameter Weibull distribution, the largest extreme value distribution, the gamma distribution, the logistic distribution, and the loglogistic distribution. On the other hand, the professional operator data showed the elapsed time for A_C to be statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the exponential distribution, the 2-parameter exponential distribution, the Weibull distribution, the 3-parameter Weibull distribution, the gamma distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution. The elapsed time for D_P proved statistically significant on the normal (after Box-Cox transformation) distribution, the lognormal distribution, the 2-parameter exponential distribution, the 3-parameter Weibull distribution, the loglogistic distribution, and the normal (after Johnson transformation) distribution.

Figure 15, Figure 16, Figure 17, and Figure 18 list the most optimal time distributions from the time distribution analysis results. More specifically, Figure 15 and Figure 16 show, respectively, the normal (after Johnson transformation) distribution of A_C and the lognormal distribution of S_C for student operators carrying out the EOP-02 procedure. Per these distributions, the average elapsed times are 8.10 seconds for A_C and 2.67 seconds for S_C. Figure 17 and Figure 18 indicate the normal (after Johnson transformation) distributions of A_C and D_P for professional operators carrying out the EOP-02 procedure. Per these distributions, the average elapsed times are 4.41 seconds for A_C and 5.96 seconds for D_P.

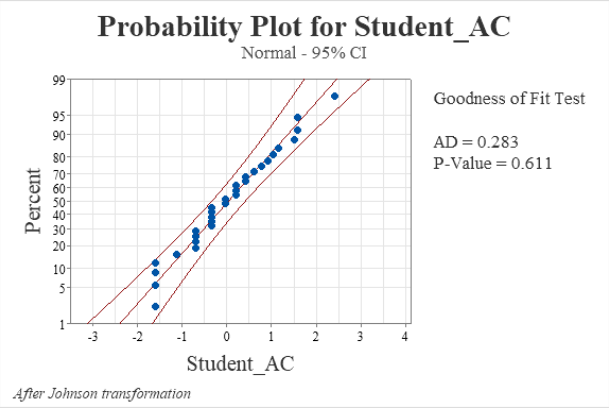
Table 15. The number of tasks used for the time distribution analysis (EOP-02).

Participant Type	GOMS-HRA Task Primitive					Number of Tasks per Participant Type	Total Number of Tasks
	A_C	C_C	R_C	S_C	D_P		
Student operators	30	36	N/A	9	39	114	195
Professional operators	22	33	N/A	2	24	81	

Table 16. Time distribution analysis for the five GOMS-HRA task primitives in the EOP-02 procedure, in regard to participant type (student vs. operator).

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Normal	<0.005	<0.005	N/A	0.200	<0.005	<0.005	<0.005	N/A	N/A	<0.005
Normal (after Box-Cox transformation)	0.536	<0.005	N/A	0.488	<0.005	0.211	<0.005	N/A	N/A	0.422
Lognormal	0.536	<0.005	N/A	0.488	<0.005	0.211	<0.005	N/A	N/A	0.081
Exponential	0.016	<0.003	N/A	0.094	0.035	0.151	<0.003	N/A	N/A	0.005
2-parameter exponential	>0.250	<0.010	N/A	>0.250	<0.010	>0.250	<0.010	N/A	N/A	0.213

Distribution	P-value of Goodness-of-Fit Test									
	Student					Operator				
	A_C	C_C	R_C	S_C	D_P	A_C	C_C	R_C	S_C	D_P
Weibull	0.052	<0.010	N/A	>0.250	<0.010	0.144	<0.010	N/A	N/A	<0.010
3-parameter Weibull	0.204	<0.005	N/A	>0.500	<0.005	0.140	<0.005	N/A	N/A	0.135
Smallest extreme value	<0.010	<0.010	N/A	0.047	<0.010	<0.010	<0.010	N/A	N/A	<0.010
Largest extreme value	0.049	<0.010	N/A	>0.250	<0.010	0.028	<0.010	N/A	N/A	<0.010
Gamma	0.175	<0.005	N/A	>0.250	<0.005	0.175	<0.005	N/A	N/A	<0.005
Logistic	<0.005	<0.005	N/A	>0.250	<0.005	<0.005	<0.005	N/A	N/A	<0.005
Loglogistic	>0.250	<0.005	N/A	>0.250	<0.005	0.158	<0.005	N/A	N/A	0.185
Normal (after Johnson transformation)	0.611	N/A	N/A	N/A	N/A	0.297	N/A	N/A	N/A	0.440



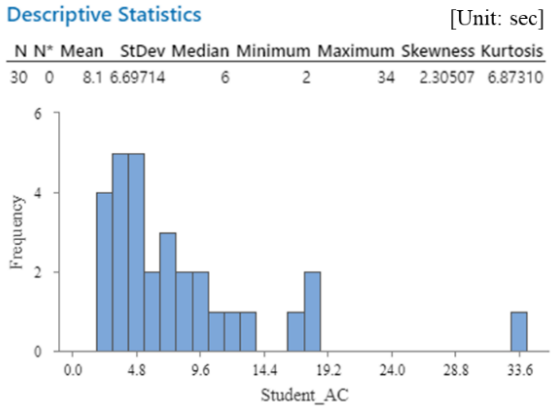
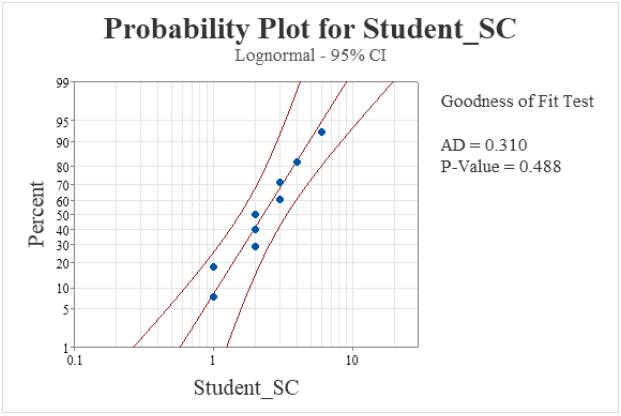


Figure 15. Normal (after Johnson transformation) distribution of A_C for student operators carrying out the EOP-02 procedure.



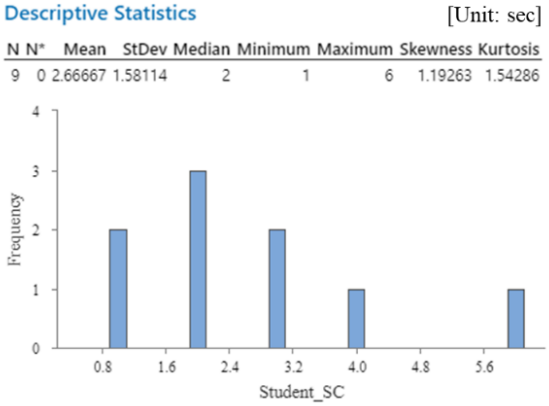


Figure 16. Lognormal distribution of S_C for student operators carrying out the EOP-02 procedure.

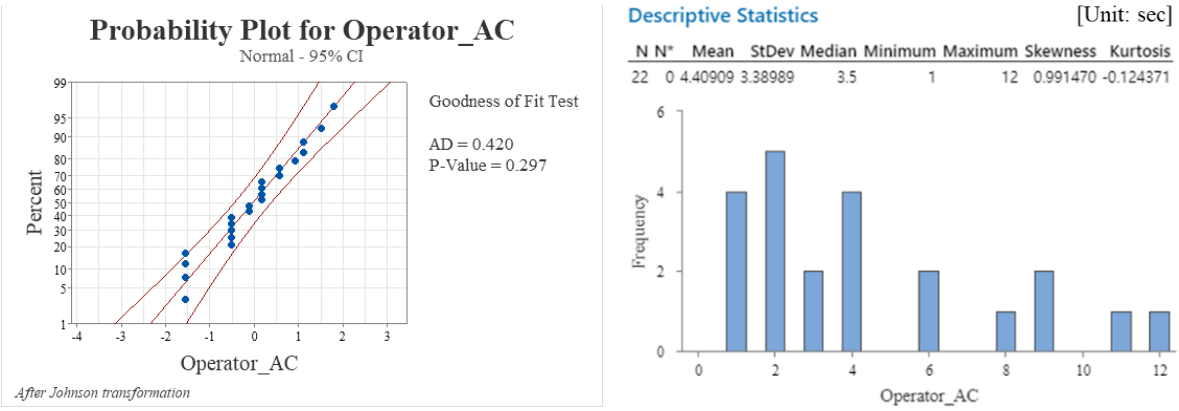


Figure 17. Normal (after Johnson transformation) distribution of A_C for professional operators carrying out the EOP-02 procedure.

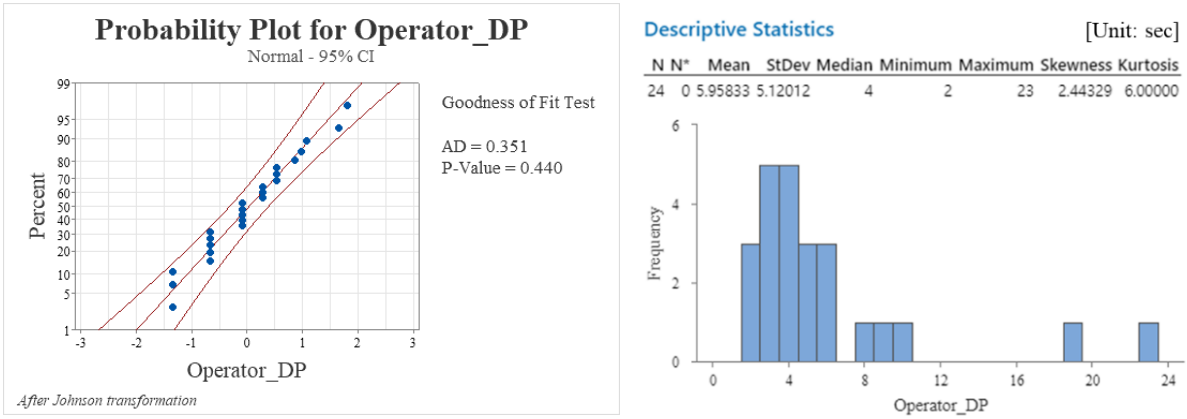


Figure 18. Normal (after Johnson transformation) distribution of D_P for professional operators carrying out the EOP-02 procedure.

4. EVALUATING TASK PRIMITIVE TIME DISTRIBUTIONS FOR AN ELAP SCENARIO

The present analysis attempted to apply the time distribution data from Section 3 to a dynamic HRA model that implements an ELAP scenario within the EMRALD software program. The dynamic HRA model was developed based on the Procedure-based Investigation Method of EMRALD Risk Assessment – HRA (PRIMERA-HRA) (Park, 2024). To provide initial estimates, HRA elements were modeled in EMRALD as a precursor to formal modeling in HUNTER. EMRALD is a PRA software tool that is well suited for modeling systems of components for dynamic simulation; however, it was not developed to address certain human actions and thus entails certain limitations. Additionally, in prior work, a HUNTER module was coupled with EMRALD (i.e., EMRALD-HUNTER) (Lew, et al., 2023) to provide more simplified HUNTER functionality. In summary, EMRALD can make overall scenario time estimations based on the time distributions derived in Section 3.

This section mainly describes (1) the ELAP scenario, (2) a dynamic HRA model for application to the ELAP scenario, (3) the time input data used in the model, and (4) the simulation results. We compare the time required for HFEs when using the existing time information proposed in GOMS-HRA, against using the time information introduced in Section 3 of this report.

4.1 ELAP Scenario

ELAP is a station blackout scenario in which offsite power, emergency diesel generators (DGs), and alternate AC DGs are all rendered unavailable (Gunther, 2015). In the United States, diverse and flexible coping strategies (FLEX) (NEI, 2016) have been suggested, using emergency (i.e., mobile) mitigative equipment as a backup for fixed equipment so as to provide plant coping capability to present core damage, even under scenarios involving simultaneous ELAP and loss of normal access to the ultimate heat sink. In such scenarios, FLEX DGs provide AC power and aid in the reactor cooldown. For the present study, we specifically developed an ELAP scenario in which FLEX DGs were deployed and connected to the plant. This scenario was developed based on observations made during stress tests (Park, et al., 2019). It assumes that upon the occurrence of the initiating event, the MCR panel indicators suddenly become unavailable due to a blackout. The operators are assumed to experience a high level of disorientation and stress, and are not equipped with any flashlights. The battery power connection is delayed for 15 minutes. (In other words, the battery power associated with the MCR indicators and emergency light functionality is automatically restored after 15 minutes.) Operators find flashlights at a location outside the MCR and bring them inside the MCR. Once some of the indicators have been restored and the flashlights are made available, the MCR operators can begin to perform procedures. First, they must diagnose the initiating event. Specifically, they follow procedures to evaluate whether the AC power sources will be difficult to restore. The evaluation outcome may be to declare an ELAP scenario, at which point two operator actions must be performed almost simultaneously. First, the MCR operators must perform Direct Current load shedding in collaboration with the local operators. In this scenario, although the local operators are required to complete all their tasks onsite, they overlook a couple of manipulations. They notice the fault after coming back and communicating with the MCR operators, then leave to finalize the manipulations. It is a local recovery process. Second, the MCR operators communicate with subcontractors to deploy the FLEX DGs. At this point, the subcontractor personnel move to the mobile equipment garage and deploy all relevant equipment to the designated place for connecting them with the plant. During deployment under this scenario, the subcontractor personnel may probabilistically face three different road conditions: (1) a good road, causing no delay [50%]; (2) a damaged road, causing a 30-minute delay [30%]; and (3) a bad road blocked by debris, causing a 60-minute delay [20%]. With one of these different road conditions, the subcontractor personnel continue to deploy the equipment and connect the FLEX DGs to the plant. The scenario concludes when both operator actions are successfully carried out within the time window and are successfully reported to the MCR operators.

This ELAP scenario involves three HFEs that are considered critical events in static FLEX HRA (Park, et al., 2019; NEI, 2016). Table 17, Table 18, and Table 19 summarize PSF information based on SPAR-H PSFs; time information such as time window, delay time, diagnosis time, and execution time; and HEP calculations for three HFEs modeled in the ELAP scenario. These were assumed based on the relevant literature (Park, et al., 2019; NEI, 2016).

Table 17. PSF information pertaining to three HFEs modeled in the ELAP scenario.

No.	HFE	SPAR-H PSFs								
		Task type	Available time	Stress /stressor	Complexity	Experience /training	Procedures	Ergonomics /HSI	Fitness for duty	Work process
1	Operator fails to declare ELAP.	Diagnosis	Expansive time (x0.01)	High (x2)	Moderately complex (x2)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)
		Execution	Time available >= 5 x the time required (x0.1)	High (x2)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)
2	Operator fails to perform FLEX DG load shed.	Diagnosis	Nominal time (x1)	High (x2)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)
		Execution	Nominal time (x1)	High (x2)	Moderately complex (x2)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)
3	Operator fails to deploy and install FLEX DG.	Diagnosis	Expansive time (x0.01)	High (x2)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)
		Execution	Time available >= 5 x the time required (x0.1)	Extremely High (x5)	Moderately complex (x2)	Low (x3)	Nominal (x1)	Nominal (x1)	Nominal (x1)	Nominal (x1)

Table 18. Time information pertaining to three HFEs modeled in the ELAP scenario.

No.	HFE	Time Window	Delay Time	Diagnosis Time	Execution Time
1	Operator fails to declare ELAP	1 hr	15 min	1 min	1 min
2	Operator fails to perform FLEX DG load shed	1.5 hrs	60 min	5 min	15 min
3	Operator fails to deploy and install FLEX DG	6 hrs	90 min	5 min	30 min

Table 19. HEP calculations pertaining to three HFEs modeled in the ELAP scenario.

No.	HFE	Task Type	HEP for Each Task Type	Final HEP
1	Operator fails to declare ELAP	Diagnosis	4.00e-4	6.00e-4
		Execution	2.00e-4	
2	Operator fails to perform FLEX DG load shed	Diagnosis	2.00e-2	2.40e-2
		Execution	4.00e-3	
3	Operator fails to deploy and install FLEX DG	Diagnosis	2.00e-4	3.20e-3
		Execution	3.00e-3	

4.2 Dynamic HRA Model for the ELAP Scenario

The dynamic HRA model for the ELAP scenario modeled within EMRALD-HUNTER consists of three smaller models: the (1) main model, (2) heading model, and (3) procedure model. The main model gives an overview of the scenario, along with heading events. Figure 19 shows the main model for the ELAP scenario. The heading model includes logic for determining the success or failure of heading events. If a heading does not split into branches, it need not be modeled. Figure 20 and Figure 21 show the heading models for Headings #1 and #2. The procedure model reflects all the information obtained from the task primitive analysis (Park, 2024). Figure 22, Figure 23, and Figure 24 show the procedure models for Procedure Paths #1, #2, and #3. Details on these models are provided in (Park, 2024).

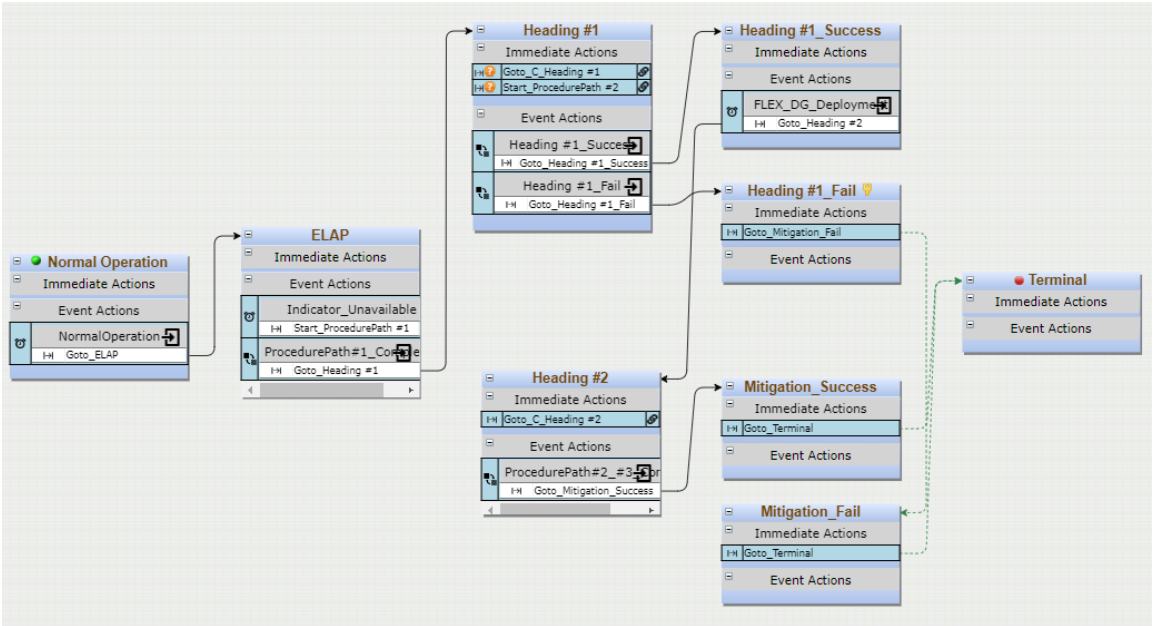


Figure 19. Main model in EMRALD-HUNTER for the ELAP scenario.

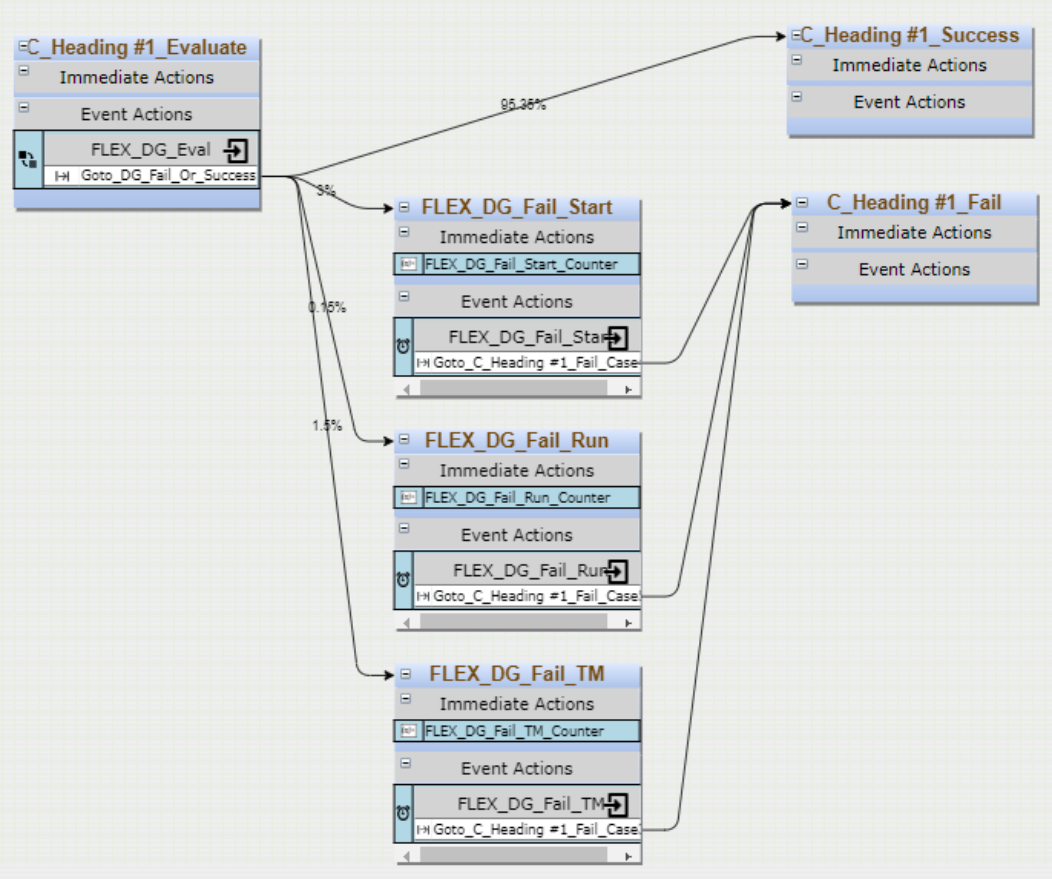


Figure 20. Heading model for Heading #1 in EMRALD-HUNTER.

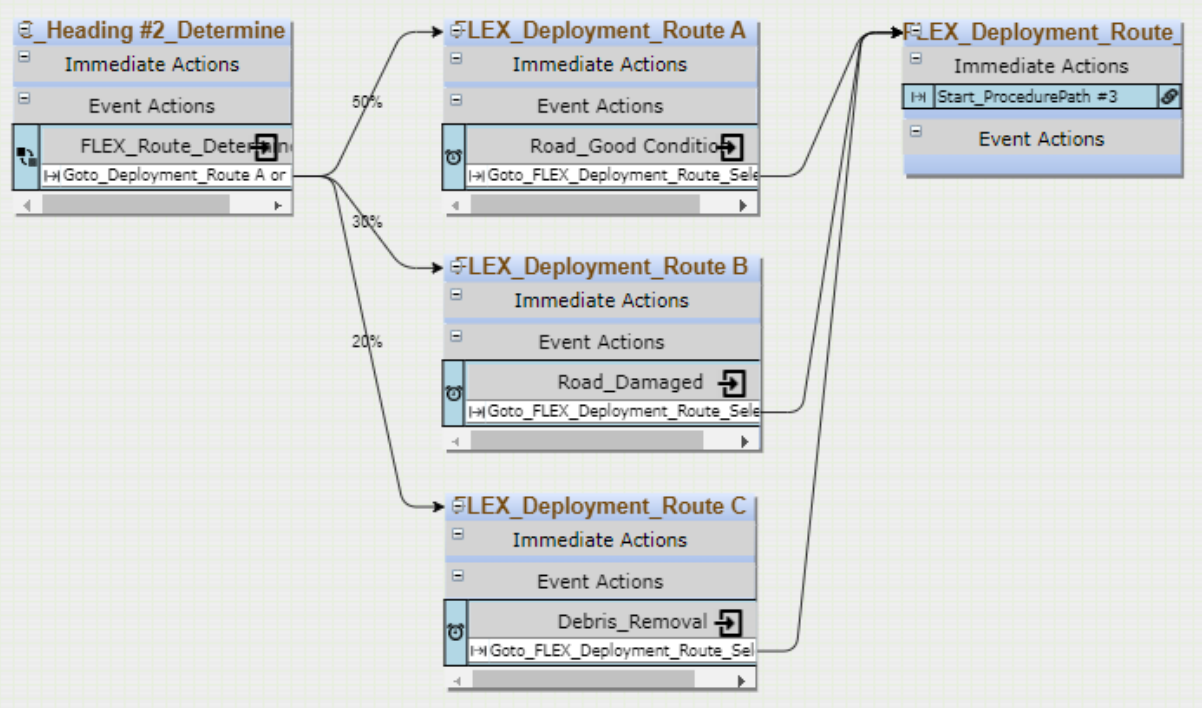


Figure 21. Heading model for Heading #2 in EMRALD-HUNTER.

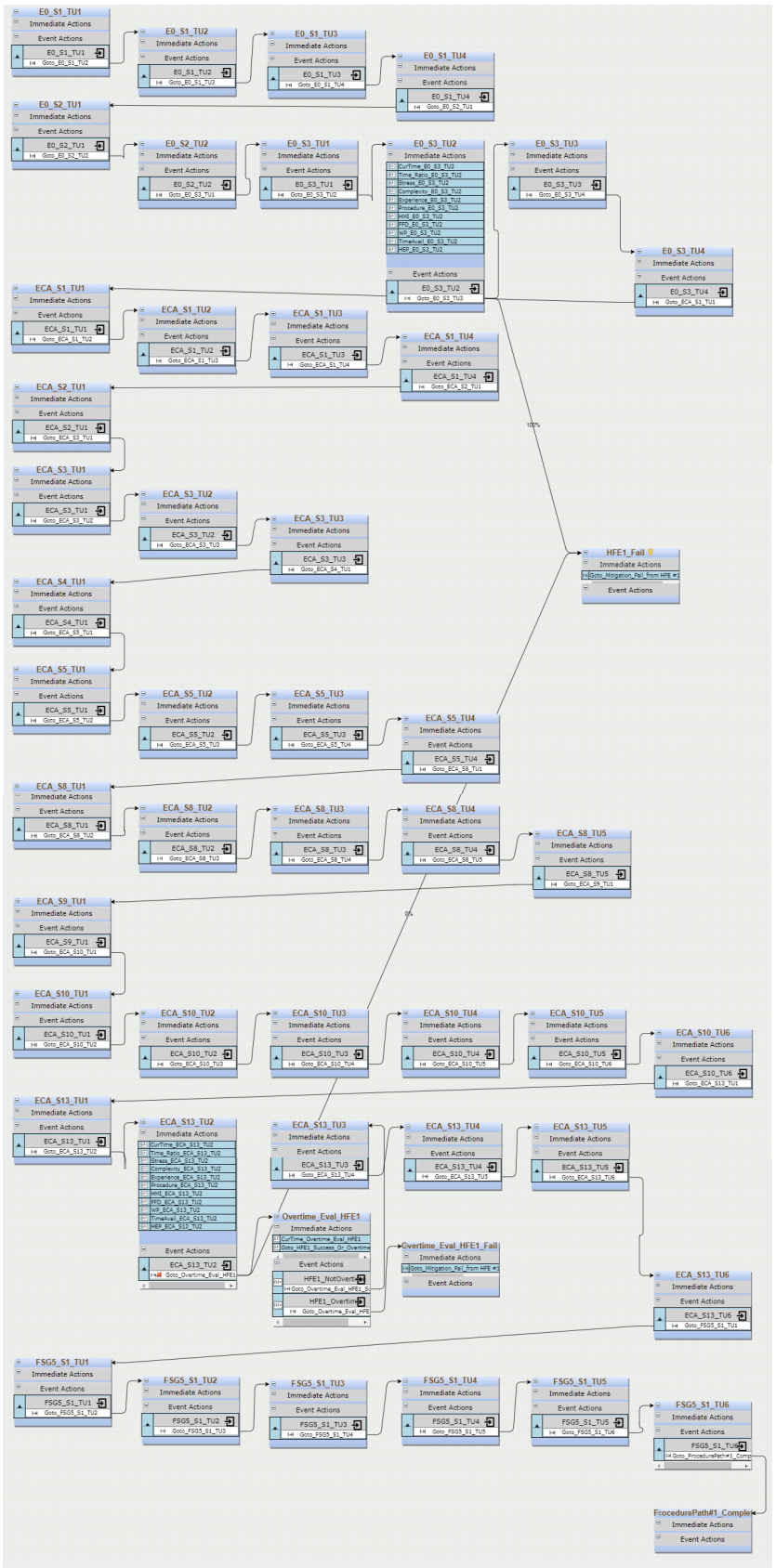


Figure 22. Procedure model for Procedure Path #1 in EMRALD-HUNTER.

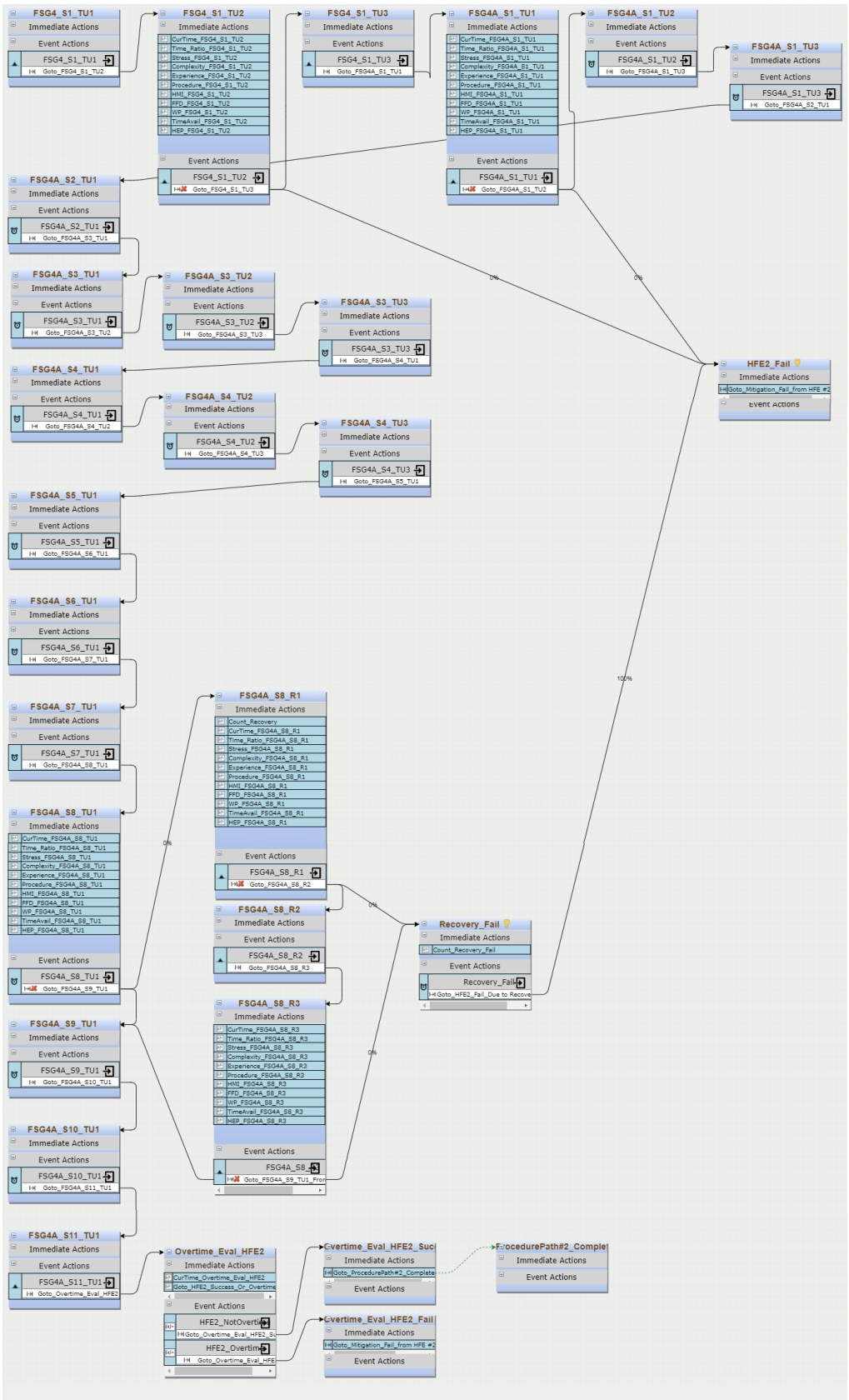


Figure 23. Procedure model for Procedure Path #2 in EMRALD-HUNTER.

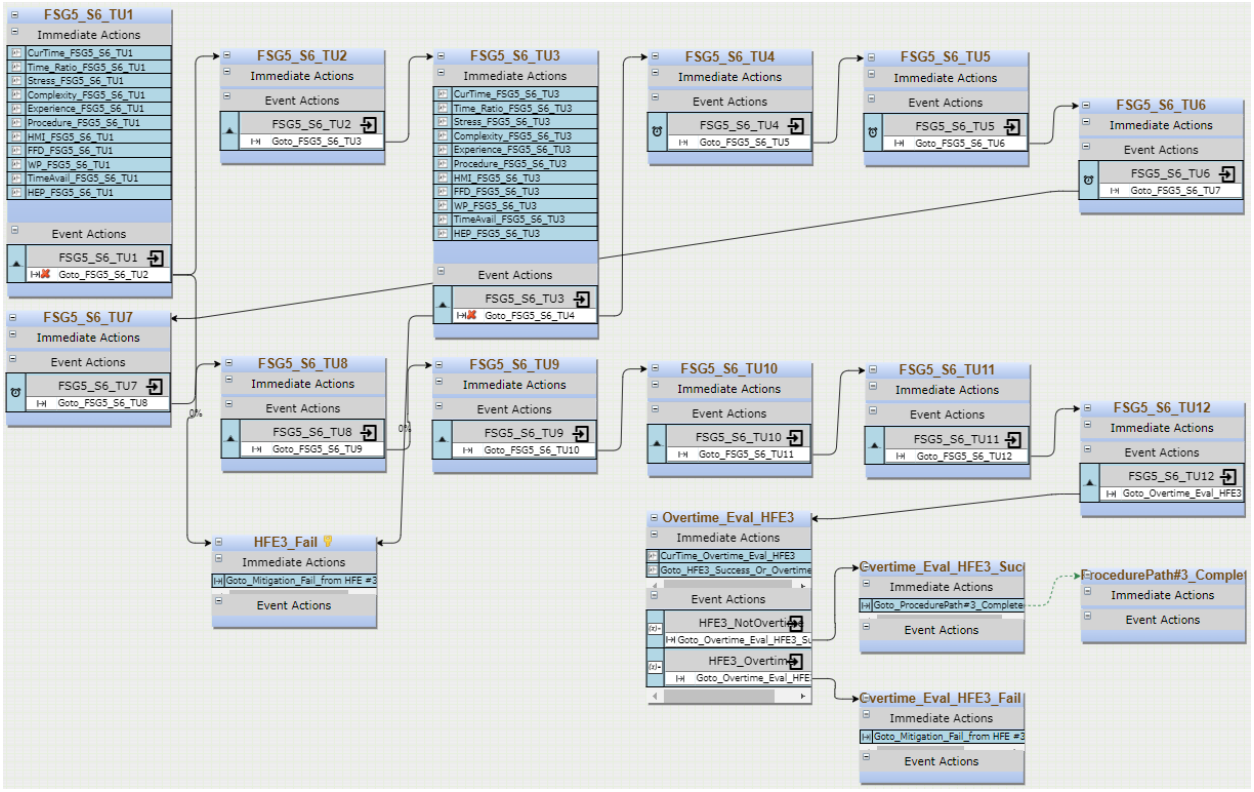


Figure 24. Procedure model for Procedure Path #3 in EMRALD-HUNTER.

4.3 Time Input Data

Table 20 summarizes time information updates for task primitives, per each procedure set. The time information, which is based on operator data, includes statistically significant results marked with an asterisk and introduced in Section 3. For task primitives that proved statistically insignificant on the time data, the present study assumed a normal distribution with corresponding time information. In this ELAP scenario, the time information for EOP-01 was mainly used as the input for tasks performed in MCRs.

Table 20. Summary of time information updates for task primitives, per each procedure set.

Procedure Set	Task Primitive	Distribution	Mean [sec]	Standard Deviation [sec]	Minimum [sec]	Maximum [sec]
OP-01	A_C	Normal	14.79	79.41	1.00	844.00
	C_C	Normal	3.26	4.75	1.00	44.00
	R_C	Normal	6.42	6.40	1.00	30.00
	S_C	Normal	62.03	88.11	1.00	393.00
	D_P	Normal	10.61	19.52	1.00	124.00
OP-02	A_C	Normal	13.48	20.68	1.00	83.00
	C_C	Normal	2.48	2.21	1.00	12.00
	R_C	Normal	3.15	2.89	1.00	18.00
	S_C	Normal (after Johnson Transformation)*	12.49	11.03	1.00	71.00
	D_P	Normal	5.79	11.22	1.00	106.00

Procedure Set	Task Primitive	Distribution	Mean [sec]	Standard Deviation [sec]	Minimum [sec]	Maximum [sec]
OP-03 & OP-04	A_C	Normal	29.95	52.46	1.00	252.00
	C_C	Normal	2.58	2.97	1.00	26.00
	R_C	Normal	5.85	8.38	1.00	81.00
	S_C	Normal (after Johnson Transformation)*	45.16	67.99	1.00	252.00
	D_P	Normal	13.42	31.69	1.00	193.00
OP-05 & OP-06	A_C	Normal	8.33	11.05	1.00	71.00
	C_C	Normal	3.79	7.35	1.00	87.00
	R_C	Normal	6.47	8.23	1.00	78.00
	S_C	Normal	71.12	119.41	1.00	599.00
	D_P	Normal	16.18	38.26	1.00	345.00
AOP-01	A_C	Normal	6.31	10.29	1.00	130.00
	C_C	Normal	5.59	6.50	1.00	38.00
	R_C	Normal	3.39	4.64	1.00	32.00
	S_C	Normal (after Johnson Transformation)*	8.01	6.36	1.00	40.00
	D_P	Normal	5.50	8.02	1.00	65.00
EOP-01	A_C	Normal (after Johnson Transformation)*	7.21	4.72	1.00	24.00
	C_C	Normal	2.37	2.30	1.00	19.00
	R_C	Normal	3.67	6.16	1.00	42.00
	S_C	3-Parameter Weibull*	5.30	2.50	1.00	10.00
	D_P	Normal	4.13	2.60	1.00	13.00
EOP-02	A_C	Normal (after Johnson Transformation)*	4.41	3.39	1.00	12.00
	C_C	Normal	2.12	1.80	1.00	10.00
	R_C	Normal	N/A	N/A	N/A	N/A
	S_C	Normal	3.5	2.12	2.00	5.00
	D_P	Normal (after Johnson Transformation)*	5.96	5.12	2.00	23.00

Note: * = $p > 0.05$

The ELAP scenario included local actions. The time information reflected in Table 20 may not be applicable to these actions, as it specifically pertains to actions performed in MCRs. As a general rule, time information regarding local actions will vary depending on the NPP and mitigation strategies involved. In the present study, the times required for these local actions were all assumed (see Table 21). The mean values were assumed based on observations made during stress tests or relevant reports (Park, et al., 2019; NEI, 2016). The standard deviation, minimum, and maximum can be calculated in the manner shown below. These assume very wide distributions with high levels of uncertainty.

- Standard Deviation = Mean / 3
- Minimum = Mean – 2 × Standard Deviation
- Maximum = Mean + 2 × Standard Deviation.

Table 21. Assumptions regarding the time required for local actions.

Task Primitive Type	Mean (sec)	Standard Deviation (sec)	Minimum (sec)	Maximum (sec)	Distribution
Local action	300	100	100	500	Normal
Access to fixed or local equipment	600	200	200	1,000	Normal
Deployment of mobile equipment	9,000	3,000	3,000	15,000	Normal

4.4 Simulation Results

The present study compared the elapsed time for HFEs when using the existing time information in GOMS-HRA, as opposed to using the revised time information introduced in Section 3. Figure 25, Figure 26, and Figure 27 compare the resulting time required for HFEs #1, #2, and #3. Overall, the time required for each HFE decreased with the updated time information. In comparison with HFEs #1 and #2, HFE #3 indicated the smallest mean time difference between using the original GOMS-HRA data and using the updated data. As a point of fact, HFEs #1 and #2 are actions performed in MCRs, whereas HFE #3 includes many local actions. Thus, it was expected that the updated time information would more strongly influence HFEs #1 and #2 and cause a greater mean time difference.

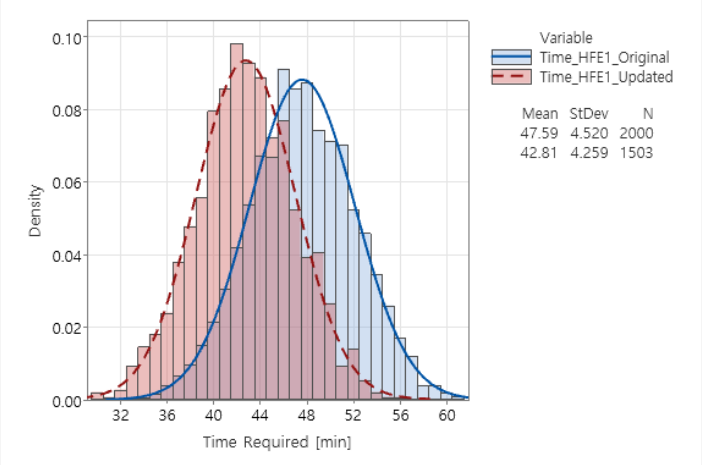


Figure 25. Comparison of time required for HFE #1.

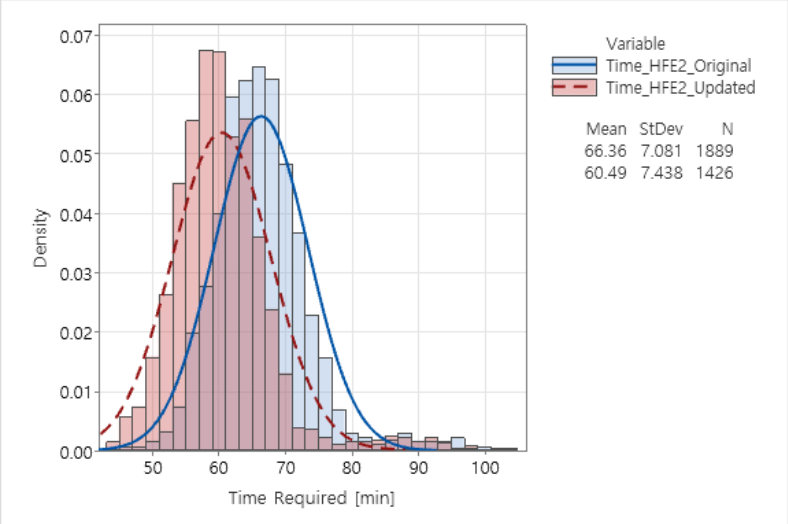


Figure 26. Comparison of time required for HFE #2.

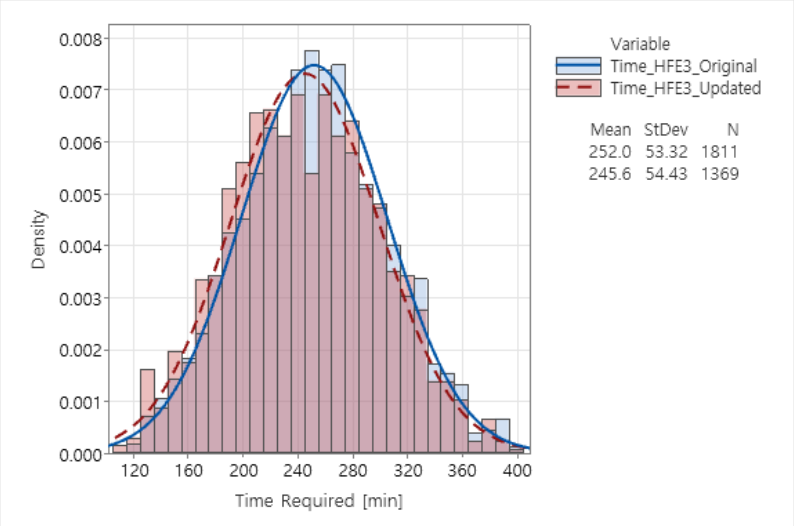


Figure 27. Comparison of time required for HFE #3.

5. DISCUSSION AND CONCLUSION

5.1 Discussion on HUNTER-P3 Development

While HUNTER was developed generally as a tool for dynamic HRA—meaning a tool to support risk analysis—applying HUNTER-P3 to procedures illustrates its strong potential to benefit non-quantitative risk uses. Procedures are essential in the safe operation of NPPs (and, indeed, in many other industries). As more digital control systems are introduced into existing control rooms, the concepts of operations at these plants will likely change, requiring new or updated procedures. These new/updated procedures can be considered via quantitative HRA methods, but many such methods lack nuance in differentiating the consequences of changes to important human actions prescribed by those procedures. The HEPs predicted for procedures may perhaps go unaltered, as the risk related to general actions and safety impacts do not necessarily change just because the HMI has evolved. On a task execution level, the procedures do change, and operator reliability in performing procedural tasks may also change in ways perhaps not fully reflected in the risk analysis. HUNTER-P3 augments existing methods of evaluating novel procedures. By anticipating the types of error traps that can occur at both the procedure and operator levels, it represents a unique solution to vetting and optimizing procedures. The outputs of HUNTER-P3 include traditional error measures, but being a dynamic modeling tool, it can also capture issues with procedure flow—issues that may not escalate into overt errors but could hinder optimal operations nonetheless. It also estimates required time, as covered in this report. Required time can have a significant impact on procedure execution and therefore be a determiner in success or failure in procedure use.

5.2 Discussion on Time Distribution Analysis

This study investigated time distributions for five GOMS-HRA task primitives taken from the seven different sets of procedures (each with different goals), comparing them in terms of participant type. Several time distributions pertaining to the five GOMS-HRA task primitives were found to be statistically significant. Specifically, a greater number of time distributions were found in AOP-01, EOP-01 and EOP-02 than in the OPs. Manipulation-related task primitives (i.e., A_C [performing required physical actions on the control boards] and S_C [selecting or setting a value on the control boards]) satisfied a relatively large amount of statistical distributions with high confidence levels in comparison with other task primitives.

A relatively smaller number of time distributions proved statistically significant in the OPs. Actually, some tasks in the OPs necessitated elapsed time for plant parameters to reach certain values. These may make it difficult to obtain time distributions that reach statistical significance. Furthermore, no time distributions were found for R_C (obtaining required information on the control boards), regardless of participant type. In the SHEEP experiment, each participant carried out six scenarios per experimental session. Accordingly, learning effects may have come into play when participants worked to obtain information from the Rancor interface. Such learning effects may inhibit the satisfying of time distributions.

Our research team continues to analyze the experimental data. Further analyses will be performed to clarify these issues and arrive at better time distributions applicable to future implementations of dynamic HRA. Already, these data show promise for using GOMS-HRA task-level primitives to arrive at time distributions. And such time distributions may eventually prove useful as outright HEP estimations in future HRA applications.

5.3 Discussion on Applying Timing Data to an ELAP Scenario

In the ELAP model, this study identified that the mean time for three HFEs when using the time information updated from the SHEEP data was shorter than when the original GOMS-HRA data were used. Based on this result, two assumptions can be made. First, the timing data in the original GOMS-HRA method may be uncertain. Specifically, the time values in GOMS-HRA were derived in limited fashion from the small amount of samples by mapping procedure-step-level primitives (not task-level primitives) into simulator logs. Second, the updated time information from the SHEEP data may be relatively shorter due to simplicity of the simulator. For Rancor, scenarios progress relatively faster than for simulators with higher complexity, such as CNS or full-scope simulators. Furthermore, its systems and interfaces are simpler as well, potentially decreasing the time required for task primitives.

Using time information that is experimentally collected and analyzed within statistical significance level enables more realistic modeling of human actions in dynamic HRA. Dynamic HRA modeling based on timing data can be useful when developing new procedures and investigating the feasibility of human actions to execute them. Here we find that the ELAP scenario can be completed quicker than earlier HUNTER modeling suggested. This gives higher confidence in the corresponding procedures as written leading to successful mitigation of ELAP. Further experimental research is required to acquire timing data that are statistically significant. In the future, more experiments are proposed under the SHEEP framework by using more complex simulators (e.g., CNS) for comparison against the timing data obtained via Rancor, then to explore time distribution differences stemming from expertise and simulator complexity. The future study would help guarantee ample sample sizes and aid in acquiring more adequate timing data that are statistically significant. Additionally, future modeling in HUNTER will seek to demonstrate the HUNTER-P3 concept for control room upgrades, providing a screening tool for benchmarking new procedures against existing ones.

5.4 Concluding Remarks

HUNTER continues to evolve as a modeling tool for human performance, with our research team having used it to explore new analyses and applications of dynamic HRA. These efforts pave the way for HUNTER applications that do not duplicate existing static HRA, but instead unlock new HRA capabilities and benefits. As development of HUNTER continues, these features will be demonstrated using plant-specific examples. Whereas HUNTER can complement existing HRA applications, its greatest value lay in helping answer emerging risk analysis topics for which legacy HRA may not be optimal. Applications such as evaluating operator performance with digital systems, evaluating procedure performance, and more accurately capturing the various facets of balance-of-plant activities lead to new methods and grant added value to the tools available to human reliability analysts.

6. REFERENCES

Bell, D., Raiffa, H. & Tversky, A., 1988. Decision Making: Descriptive, Normative, and Prescriptive Interactions: Cambridge University Press.

Boring, R., Lew, R. & Ulrich, T., 2017. Advanced nuclear interface modeling environment (ANIME): A tool for developing human-computer interfaces for experimental process control systems: Lecture Notes in Computer Science.

Boring, R. et al., 2016. Integration of Human Reliability Analysis Models into the Simulation-Based Framework for the Risk-Informed Safety Margin Characterization Toolkit: Idaho National Laboratory.

Boring, R.L., & Rasmussen, M., 2016. GOMS-HRA: A method for treating subtasks in dynamic Human Reliability Analysis. Risk, Reliability and Safety: Innovating Theory and Practice, Proceedings of the European Safety and Reliability Conference, pp. 956-963.

Boring, R. L., Rasmussen, M., Ulrich, T., Ewing, S., & Mandelli, D., 2017. Task and procedure level primitives for modeling human error. Advances in Intelligent Systems and Computing, 589, 30-40.

Boring, R. et al., 2022. Software Implementation and Demonstration of the Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER), INL/RPT-22-66564: Idaho National Laboratory.

Boring, R. et al., 2019. A comparison of operator preference and performance for analog versus digital turbine control systems in control room modernization. Nuclear Technology, Volume 205, pp. 507-523.

Boring, R., Ulrich, T., Lew, R. & Park, J., 2023. HUNTER procedure performance predictor: Supporting new procedure development with a dynamic human reliability analysis method, Proceedings of the Applied Human Factors and Ergonomics Conference, Hawaii Edition.

Boring, R., Ulrich, T., Lew, R. & Park, J., 2023. Synchronous vs. Asynchronous Coupling in the HUNTER Dynamic Human Reliability Analysis Framework. 14th International Conference on Applied Human Factors and Ergonomics (AHFE 2023) and the Affiliated Conferences.

Gertman, D. et al., 2005. The SPAR-H human reliability analysis method, NUREG/CR-6883: US Nuclear Regulatory Commission.

Gunther, W., 2015. Testing to Evaluate Extended Battery Operation in Nuclear Power Plants, NUREG/CR-7188: US Nuclear Regulatory Commission.

Hollnagel, E., 2017. Why is work-as-imagined different from work-as-done?. CRC Press ed.:Resilient Health Care, Volume 2.

IAEA, 2018. Operating Experience Feedback for Nuclear Installations, SSG-50: International Atomic Energy Agency (IAEA).

Jung, W. D., Kang, D. I. & Kim, J. W., 2005. Development of a standard method for human reliability analysis (HRA) of nuclear power plants - Level 1 PSA full power internal HRA, KAERI/TR-2961/2005, Daejeon, Republic of Korea: Korea Atomic Energy Research Institute.

Lew, R., Boring, R. & Ulrich, T., 2018. Transitioning nuclear power plant main control room from paper based procedures to computer based procedures, Proceedings of the Human Factors and Ergonomics Society.

Lew, R. et al., 2023. EMRALD-HUNTER: An Embedded Dynamic Human Reliability Analysis Module for Probabilistic Risk Assessment, Idaho Falls: Idaho National Laboratory.

Lew, R., Ulrich, T. & Boring, R., 2022. Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) Demonstration: Part 2, Model Runs of Operational Scenarios, INL/RPT-22-70076, Idaho Falls: Idaho National Laboratory.

NEI, 2016. Crediting Mitigating Strategies in Risk-Informed Decision Making, NET-16-06: Nuclear Energy Institute (NEI).

NEI, 2016. Diverse and Flexible Coping Strategies (FLEX) Implementation Guide, NEI-12-06: Nuclear Energy Institute (NEI).

O'Hara, J., Higgins, J., Fleger, S. & Pieringer, P., 2012. Human Factors Engineering Program Review Model (NUREG-0711, Revision 3): U.S. Nuclear Regulatory Commission.

Park, J., 2024. An Approach to Dynamic Human Reliability Analysis and Its Data Collection Framework, Ph.D Dissertation. Pocatello: Idaho State University.

Park, J., Arigi, A. & Kim, J., 2019. Treatment of human and organizational factors for multi-unit HRA: Application of SPAR-H method. Annals of Nuclear Energy, Volume 132, pp. 656-678.

Park, J. et al., 2022. A framework to collect human reliability analysis data for nuclear power plants using a simplified simulator and student operators. Reliability Engineering & System Safety, Volume 221, p. 108326.

Park, J. et al., 2022. Human Unimodel for Nuclear Technology to Enhance Reliability (HUNTER) Demonstration: Part 1, Empirical Data Collection of Operational Scenarios (INL/RPT-22-69167), Idaho Falls: Idaho National Laboratory.

Park, J. et al., 2023. Analysis of human performance differences between students and operators when using the Rancor Microworld simulator. Annals of Nuclear Energy, Volume 180, p. 109502.

Park, J., Yang, T., Kim, J. & Boring, R., 2023. An Investigation of Time Distributions for Task Primitives to Support the HUNTER Dynamic Human Reliability Analysis. Proceedings of the Applied Human Factors and Ergonomics Conference, Hawaii Edition.

Parry, G., Beare, A., Spurgin, A. & Moieni, P., 1992. An approach to the analysis of operator actions in probabilistic risk assessment, EPRI Rep. TR-100259: Electric Power Research Institute.

Swain, A. & Guttmann, H., 1983. Handbook of human-reliability analysis with emphasis on nuclear power plant applications, NUREG/CR-1278, Albuquerque, NM, USA: Sandia National Laboratory.

Ulrich, T., Boring, R., L., Ewing, S., & Rasmussen, M, 2017. Operator timing of task level primitives for use in computation-based human reliability analysis. Advances in Intelligent Systems and Computing, 589, 41-49.

Ulrich, T., Lew, R., Werner, S. & Boring, R., 2017. Rancor: A Gamified Microworld Nuclear Power Plant Simulation for Engineering Psychology Research and Process Control Applications. Proceedings of the Human Factors and Ergonomics Society 2017 Annual Meeting, 61, 398-402.