INL/RPT-23-71788

Solutions for Enhanced Legacy Probabilistic Risk Assessment Tools and Methodologies

Improving Efficiency of Model Development and Processing via Innovative Human Reliability Dependency Analysis

March 2023

U.S. Department of Energy Office of Nuclear Energy

INL/RPT-23-71788

Solutions for Enhanced Legacy Probabilistic Risk Assessment Tools and Methodologies: Improving Efficiency of Model Development and Processing via Innovative Human Reliability Dependency Analysis

Thomas Ulrich[1], Ronald Boring[1], Jisuk Kim[1], Anna Hall[1], Torrey Mortenson[1], Svetlana Lawrence[1], Roger Lew[2], Gregory Zucal[3], Andrew Miller[3], Erin Collins[3]

March 2023

[1]Idaho National Laboratory
[2]University of Idaho
[3]Jensen Hughes, Inc.

Prepared for the
U.S. Department of Energy
Office of Nuclear Energy
Under DOE Idaho Operations Office
Contract TBD

EXECUTIVE SUMMARY

Probabilistic risk assessments (PRAs) are integral to nuclear power plant operations, having tremendously benefitted the safety of the U.S. reactor fleet for decades. Insights from these models have provided perspectives on a variety of applications, both at the plant and for the regulator. While these models are useful, they are now being asked to represent and analyze more than they were ever envisioned to by initial PRA practitioners. Furthermore, heightened demands on PRA models have led to increased computing power requirements. As the complexity of the PRA models increased, the difficulty experienced by non-PRA experts in trying to understand these models, grasp the insights they provide, and effectively use that information has become problematic.

The need for research to address key issues regarding PRA tools and methods has never been greater. Although the nuclear power industry has been well-served by these tools and methods, the underlying science is dated, remaining mostly unchanged for over two decades. Three areas most beneficial to address to maintain and improve the usefulness of the current practice legacy PRA tools are improved quantification speed, increased ability to efficiently model multihazard models, and improved modeling of human-action dependencies in PRAs. This report is focused on the cross-area subject, improvements to modeling speed by improving the dependency analyses of human actions conducted as part of a typical human reliability assessment.

CONTENTS

FIGURES

TABLES

ACRONYMS

3D SART	Situation Awareness Rating Technique
CBP	computer-based procedure
DL	dependency level
GOMS	Goals, Operators, Methods, Selection rules
HDA	HRA dependency analysis
HEP	human event probability
HFE	human failure event
HRA	human reliability analysis
HSSL	Human Systems Simulation Laboratory
IEEE	Institute of Electrical and Electronics Engineers
INL	Idaho National Laboratory
JHEP	joint human error probability
LOFW	loss of feedwater
MSE	mean-square error
n	sample size
NPP	nuclear power plant
PORV	power-operated relief valve
PRA	probabilistic risk assessment
PSF	performance shaping factor
r	correlation
SCORE	Supervisory Control and Resilience Evaluation
SI	safety injection
SPAR-H	Standard Plant Analysis Risk-Human
THERP	Technique for Human Error Rate Prediction

1. INTRODUCTION

Probabilistic risk assessments (PRAs) have advanced the safe operation of the U.S. reactor fleet over past decades. Risk insights from these PRAs have provided information from many different perspectives, from what is most important to maintain at a facility to a better understanding of how to address new information regarding safety issues. The methods and tools that have supported the creation and enhancement of PRA models were established through multiple decades of research, from the 1970s starting with the WASH-1400 Reactor Safety Study (U.S. NRC 1975) through comprehensive plant-specific models in use today.

In August 2020, a report titled “R&D Roadmap to Enhance Industry Legacy Probabilistic Risk Assessment Methods and Tools” was published that outlined key challenges identified by PRA practitioners and detailed research and development priorities for advancing the state of PRA software (Miller, Hess, and Smith 2020). That report identified three areas with near-term strategic benefits to the PRA community:

- Quantification speed of models
- Multihazard modeling development, maintenance, and treatment
- Human-action dependency analysis.

The report concluded that there are meaningful research activities that can provide near-term benefits to PRA applications and the treatment of models could push the risk assessment community’s state of practice further along. These topics contain well known issues and are quite broad in scope. Further investigation is warranted to determine appropriate tasks and focus areas within each of the broader topics. This will allow resources to be applied to research initiatives that are most likely to have meaningful outcomes and provide tangible improvements to the state of practice.

The research continued in 2021, and a report titled “Enhancement of Industry Legacy Probabilistic Risk Assessment Methods and Tools” (Vedros et al. 2021) was published that addressed the specific focus areas identified for each broader area, as well as the research, methodology, benchmarking, and conclusions. The work described in this report represented initial research in addressing the identified high-value applications and additional follow-on work required to fully explore a topic or determine the appropriate scope to realize changes to the applicable methods and implement software systems. This report served as a first step in a broader effort to address the legacy PRA software issues and provide suggested upgrades to algorithms, methodologies, and technologies.

One of the topics that touched two strategic areas is addressing human-action dependency analyses, which also affects model quantification efficiency and speed. There two primary ways to improve quantification speed: 1) by speeding up the existing quantification process (e.g., using advanced high-speed computers) or 2) by innovating the base process (e.g., use a novel methodology to accomplish the same goal more efficiently). Improvements made to HRA dependency analysis base processes fall under the latter category offering to improve quantification speed by innovative approach to underlying processes.

This report focuses on the potential to improve model efficiencies, including quantification speed, via enhanced techniques for modeling human-action dependency analyses. In addition, this report explores opportunities to improve the underlying theoretical bases of a dependency analysis by investigating prospects for empirical data collection.

2. INCREASING PROBABILISTIC RISK ASSESSMENT QUANTIFICATION EFFICIENCY BY IMPROVING HRA DEPENDENCY ANALYSIS

2.1 Background

A dependency analysis is performed as part of a human reliability analysis (HRA) within a PRA. A PRA model is developed with human failure events (HFEs) included in the model logic, and typically, each event represents a single human action required during an accident sequence. However, accident sequence mitigation often requires multiple human actions to be successful in preventing an adverse consequence. Therefore, when evaluating the risk associated with sequences that require human intervention, it is important to take a holistic approach and look at not just the individual actions but also the multiple actions throughout a sequence.

NUREG-1792 (U.S. NRC 2005), "Good Practices for Implementing HRA," describes the importance of the dependency analysis as:

> Dependencies among the post-initiator HFEs and hence the corresponding HEPs in an accident sequence should be quantitatively accounted for in the PRA model by virtue of the joint probability used for the HEPs. This is to account for the evaluation of each sequence holistically, considering the performance of the operators throughout the sequence response and recognizing that early operator successes or failures can influence later operator judgments and subsequent actions. This is particularly important so that combined probabilities that are overly optimistic are not inadvertently assigned, potentially resulting in the inappropriate decrease in the risk-significance of human actions and related accident sequences and equipment failures. In the extreme, this could result in the inappropriate screening out of accident sequences from the model because the combined probability of occurrence of the events making up an accident sequence drops below a threshold value used in the PRA to drop sequences from the final risk results.

However, the current process for conducting an HRA dependency analysis (HDA) is a major hinderance to the effective and efficient use of plant PRA models. The predominant procedure currently in use today utilizes specially generated cutsets to determine when individual events may show up together in the same sequence, determines their dependency on each other, creates combination events, and assigns a joint human error probability (JHEP) to the combination events. While the process is thorough, it is highly labor intensive and inefficient for large models. Once the analysis is complete, the analysis results are included in the PRA model quantification through a post processing step to find and replace individual HFEs with the combination events and JHEPs.

Due to the need to produce cutsets with individual HFEs, this process could leave certain combinations of HFEs unmodified or partially modified if not all potential combinations are realized in the quantification performed. This potentially skews the results, especially when the models are designed to treat combinations conservatively. This is a particular problem when the PRA model is used for applications that consider different plant configurations as that increases the potential for new combinations to appear in results with a conservative treatment applied to them. Lastly, the process must be updated and reperformed for even minor changes to the PRA model, significantly prolonging PRA development time.

This report proposes novel concepts for performing a dependency analysis that eliminates the reliance on identifying combinations through quantification, yet still allows the user to apply structured dependency analysis rules. This approach provides the ability to define the dependency rules and values during model development instead of waiting until after cutsets have been generated specifically to identify HFE combinations. Additionally, the concepts presented here allows for a clearer representation of human actions within the context of an accident sequence scenario represented by cutsets. The aim of this approach is to enable the user to be more precise when it is warranted and more efficient when performing HDA during PRA model development.

2.2 Issues and Challenges

2.2.1 Current Human Reliability Analysis Dependency Issues

The process of producing or refining a PRA can take several months due to the complexity of the various tasks, both individually and collectively. At the end of the process, the PRA model produces cutsets representative of accident sequence progression, and multiple HFEs can reside within these cutsets. The HDA update typically begins when the PRA update is nearly finished. As schedules can often get compressed near the end of a project, there is usually little time for the HDA and refinement in coordination with the PRA model.

The high-level process for performing dependency analysis in context with the PRA model update process used in the nuclear fleet today follows the steps shown in Figure 1. The detailed overview of Box 4 is shown in Figure 2, which outlines the process of identifying HFE combinations and assigning JHEPs.

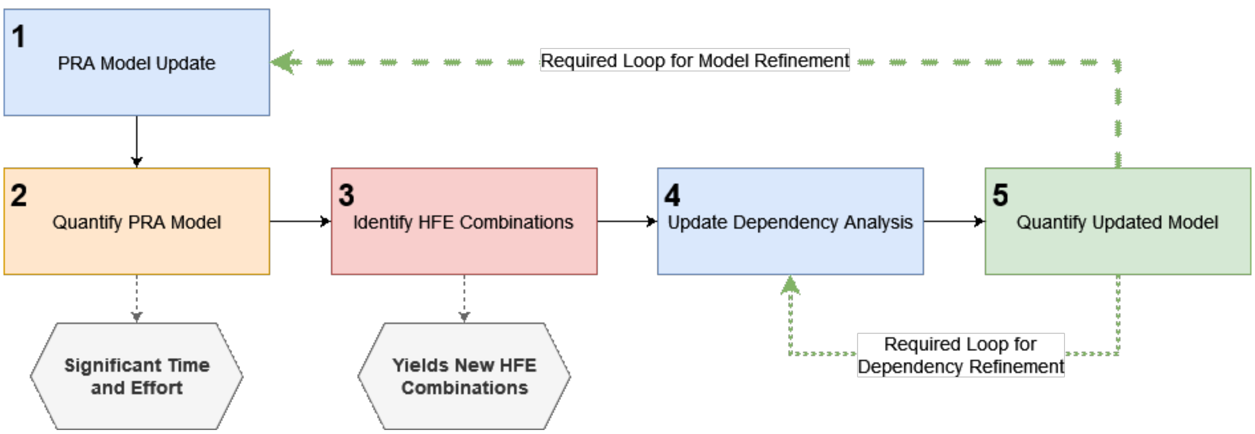

Figure 1. Current PRA model update process.

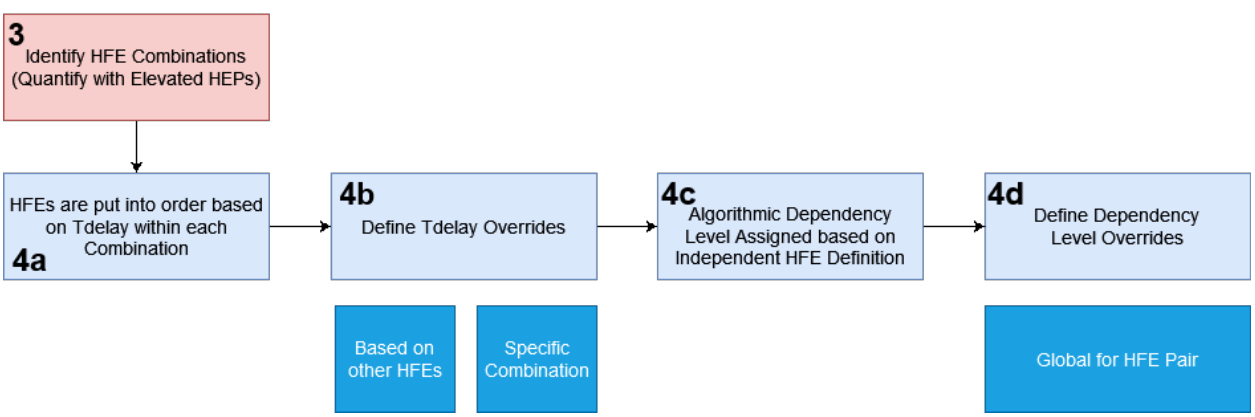

Figure 2. Primary tasks of current dependency analysis process.

Following the completion of the PRA model update (Box 1), the PRA model is first quantified with intentionally high (e.g., 0.9) HEP values (Box 2). The higher HEP values ensure cutsets contain the majority of possible HFE combinations and cutsets are not removed during model quantification due to any falling below the truncation limit. Since the HFE values will increase if a nonzero dependency level is assigned to an action, increasing the HFE values for this step in the process is critical to help identify as many HFE combinations as possible. However, elevating the HFE values drastically increases the amount of time it takes to quantify the PRA model and then start the dependency analysis.

The resulting cutsets are then processed to identify the HFE combinations and each combination is assigned a number (Box 3). The HFEs within a combination are placed in chronological order based on the parameter Tdelay (Figure 2, Box 4a). Tdelay represents the time delay or time it takes for an operator to acknowledge the cue and take an action from a common initiating time at the beginning of the scenario (T0). It is common for analysts to develop independent HFEs based on a bounding accident sequence and use them in other accident sequences where the conservatism in the HEP is not significant. However, reusing the HFE in other accident sequences may introduce undesired HFE ordering based on inaccurate Tdelay values. Therefore, an analyst must evaluate the HFE combinations to determine if any Tdelay overrides are required (Box 4b). After the HFEs are placed in the desired order, an algorithm assigns a degree of dependence between the HFEs using the independent HFE properties (Box 4c). The algorithm does not consider the context of actions (i.e., under which accident scenario these actions are taken). Therefore, an additional review (Box 4d) determines if any HFE pairs require dependency level (DL) overrides because the algorithmic approach was not sufficient.

To complete the HDA (shown in Figure 2), the conditional HFE failure probabilities are calculated using the independent HFE values and the formulas associated with the specified DLs. At the end of the dependency analysis, a JHEP is calculated for each combination. The JHEPs are incorporated into the PRA model through post-processing (Figure 1, Box 5). The impact of the dependency analysis upon the PRA results is therefore unknown until the PRA model is quantified with the JHEPs incorporated. The introduction of new probabilities and events late in the model update process is inefficient since it usually necessitates changes that require performing the entire process shown in Figure 1 from the beginning (Figure 1, Boxes 1–5).

The PRA model refinement process involves cutset review meetings subsequent to the incorporation of the updated JHEPs. These meetings are generally focused on identifying the combinations that are dominant contributors to overall plant risks and determining if the combination’s JHEP is appropriately assigned. These combinations are provided to the HRA analyst and then scrutinized to evaluate whether further DL refinement is possible. The refinement of one combination leads to cascading effects that may not be fully known until after another round of PRA model quantification. There are usually multiple rounds of combination JHEP refinement and quantification until the PRA results stabilize to an appropriate level.

2.2.1.1 Overriding Event Ordering Based on Cue Time (Tdelay)

In the current process, HFEs are initially sequentially ordered in each combination based on the timing parameter Tdelay for the independent HFEs (Figure 2, Box 4a). As stated, Tdelay represents the time delay or the time it takes for an operator to acknowledge the cue and take an action. As discussed in NUREG-1921 (EPRI 2021),

> The cue for an operator cognitive response may consist of a single parameter or multiple parameters. For example, low lubrication oil pressure for a pump is a single parameter that would actuate an alarm that would require the operator to trip the pump to protect the bearings. As an example of multiple parameters, the cue for implementing the functional restoration procedure for loss of secondary cooling on a PWR is based on multiple parameters: low steam generator feed flow and low steam generator narrow range level.

The multiple parameters for a cue may significantly alter Tdelay.

The Tdelay is assigned for each independent HFE by an HRA analyst based on the accident sequence progression and the presentation of the cue condition. However, it is common for an HFE to be used in multiple accident sequences, and as a result, the specified Tdelay may be disconnected from certain accident sequences. This leads to the need for refinement by the HRA analyst and the application of Tdelay overrides to correct the sequence of HFEs in a combination.

In the current process, Tdelay overrides are applied manually by the HRA analyst on a pair-wise basis between two HFEs (Figure 2, Box 4b). These changes for each pair of HFEs are implemented globally for all sequences where the HFE pair is present. However, the timing override appropriate in one combination may be incorrect in other combinations of three or four HFEs and may disrupt the sequence of actions in those other combinations. If more than one override is defined for the same HFE, the analyst can choose to use the shortest or longest Tdelay override, which may not always be appropriate. Another approach for overriding the Tdelay for an HFE is to define the override for the specific combination. Given the large number of HFE combinations that are typically identified, this approach is not as widely used as the pair-wise global Tdelay override.

The Tdelay manipulations for the current dependency analysis process is highly labor intensive as HFEs are often initially misordered in combinations appearing in cutsets for multiple accident sequences. Ordering HFEs using a global override for all combinations is limiting as it does not directly consider the context, i.e., accident sequence in which it may appear.

2.2.1.2 Dependency Level Overrides

In the current approach, DLs between HFEs in a PRA cutset are automatically assigned to all subsequent HFEs with the first HFE set to its independent value (Figure 2, Box 4c). The DL assignments are based on a logic structure that applies a standard set of performance shaping factors (PSFs) to assess the degree of dependence between sequential HFEs within a combination. These PSFs are consistent with the ones used to quantify individual HFEs, including the timing of the action, location of the action, and execution stress level. A DL preliminary assignment is done using the pair-wise approach (i.e., HFE2 dependent of HFE1, HFE3 dependent on HFE2, and so on) based upon PSFs and their correlation to a scale of five DLs from zero dependence to complete dependence per NUREG/CR-1278 (Swain and Guttman 1983).

The automatic assignment of DLs based on the limited information assigned to each independent HFE lacks the accident sequence information. Therefore, manual overrides to the DLs are necessary (Figure 2, Box 4d). The HFE combinations are manually reviewed by the HRA analyst to determine if the DL assigned to the HFE pairs should be changed. DL changes are implemented via a DL override, which impacts those HFE pairs globally across the analysis (meaning wherever those pairs appear sequentially in any HRA combination). However, this DL override may not be appropriate for every accident sequence in which that combination of events appears. Care must be taken in the review and application of the revised DLs. In some cases, the HRA analyst may need to create new independent events specifically for a particular set of accident sequences to ensure the appropriate DL can be assigned. This would result in restarting the HDA at Step 1 as new cutsets would need to be generated to identify the new HFE combinations. An additional challenge is maintaining the objectivity and consistency of the refinement process.

2.2.1.3 Applying Human Reliability Analysis Dependency in Quantification Process

The final step in the current HDA process is to define the HFE combination event and assign the calculated JHEP as its value. The calculated JHEP is the product of the initial HFE in the combination event multiplied by the conditional values of each subsequent HFE. The conditional values are calculated based on the HFE independent value and the DL assigned to the HFE pair involving the HFE and preceding HFE in the combination event.

The application of the HDA into the PRA model is performed after the quantification engine creates the basic events cutsets. This process is shown in Figure 3. The post-processing of the cutsets involves identifying when the HRA combinations analyzed are present in the cutsets (Figure 3, Box 5b) and replacing the basic events with the HRA combination event (Figure 3, Box 5c).

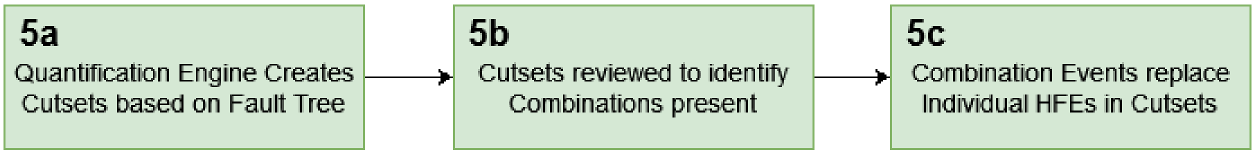

Figure 3. Current PRA model quantification steps to include HDA.

2.3 Proposed Updated Process

2.3.1 Novel Approach to Human Reliability Analysis Dependency Analysis

Based on the information about the current dependency analysis process outlined in Section 2.2, the conclusion is that the HDA could be significantly streamlined. There is also an impetus to better explain the impact of dependency analysis with respect to the PRA results, especially for time-sensitive and risk-important actions. A description of the proposed changes to the process is provided below. The main goals of the changes include expediting PRA model development, reducing the amount of rework, speeding up model quantification, and generating clearer risk insights.

Figure 4 provides the high-level overview of the proposed streamlined process. The proposed approach does not fundamentally change the dependency analysis outcomes, but rather promotes the ability to move the definition of DL between events much earlier in the model development process. The HDA should not need to wait until the model is essentially finished and cutsets are produced to start the analysis. Instead, by defining DL separately and upfront, the dependency analysis can be treated as a parallel part of the model development process instead of an add on at the end. This effectively removes Boxes 2 and 3 from Figure 1 and rearranges Boxes 1 and 4 to be in parallel before finalizing the model in Box 5 as shown in Figure 4.

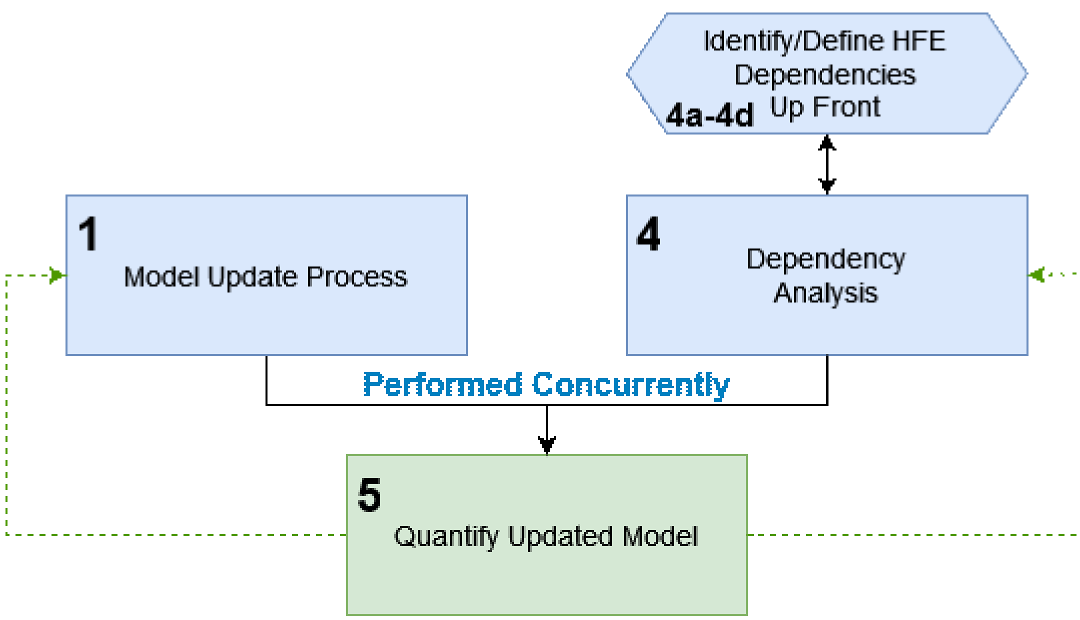

Figure 4. Proposed streamlined dependency analysis method.

To accomplish this change, the requirement for combinations from PRA model cutsets is eliminated for the steps represented in Boxes 4a–d of Figure 2. The concept is explored further with details presented in the subsections below.

The current concept of DLs and organizing individual HFEs in a sequence by Tdelay is retained. However, the sequencing of events is determined earlier in the PRA process and may be based on events other than HFEs. The intention is to utilize a more sophisticated set of post-processing rules that would allow the analyst to look at HFE attributes and functions with respect to other HFEs in the same sequence and evaluate the DLs between those HFEs. The intent is to take similar rules and apply them across the analysis in a more structured and consistent way.

Figure 5 illustrates the changes between the current prescribed process and the proposed process. The major difference is removing the requirement for model cutsets to begin the dependency analysis. The proposed method allows for the independent HFEs to be ordered and their dependency levels defined without the need for identifying combinations through cutset generation with artificially high individual HEP values. The upfront definition allows for a more predictable and efficient model quantification as the project moves toward closure.

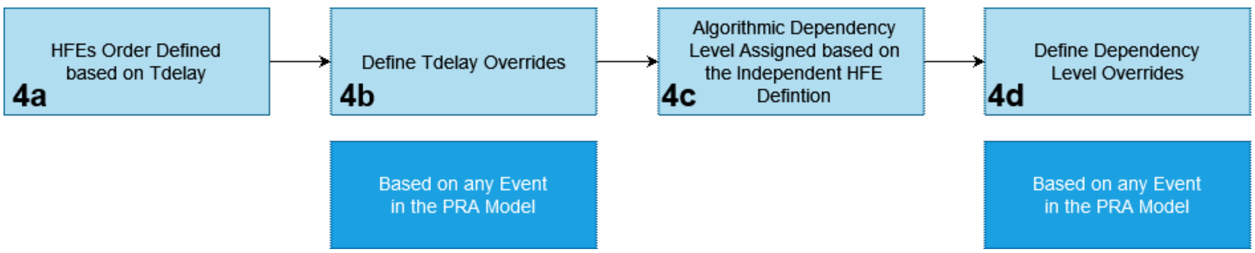

Figure 5. Streamlined dependency analysis processes.

The following sections provide description of the facets of the proposed process for streamlining the HDA.

2.3.1.1 Early Definition of Human Factors Engineering Dependencies

The critical part of dependency analysis is to identify whether HFEs must have their failure probability increased if they occur after the failure of other HFEs in the accident sequence progression. The current process relies on combinations identified from PRA cutsets as the vehicle for applying the DLs. However, the dependence between HFEs should not rely solely on the generated cutsets but primarily on the knowledge of the plant systems and operation during off-normal conditions and the postulated accident sequences and initiating events from the PRA logic model.

If “combinations” could be eliminated as a critical part of the HDA, the dependency evaluation could begin prior to the PRA model quantification and at relatively early stages of the HRA updates of individual HFEs by reviewing all the HFEs in the PRA model and identifying the plant functional states to which they relate. The event trees, which already lay out the post-initiator functional progression into which the HFEs are linked, offer a helpful mechanism for structuring the correlation between HFEs.

This process would more clearly show the progression from an initiating event to equipment failures to HFEs, which establishes the context for the HFEs and therefore provide:

- Better insights on the context of the HFEs for the HRA analyst
- Scenario-based context that can be used for operator interviews.

The proposed process permits the HRA analyst to “pre-assign” dependency levels between HFE pairs and assesses how those levels change when other events come into play. Moreover, it is likely that the HFEs and dependency between them will remain the same for several model development or update cycles, allowing the use of the same dependency analysis and to only require updates to the HFEs that have changed or a new dependency analysis on HFE pairs if they are added to the model.

2.3.1.2 Tdelay Override Enhancement

The proposed streamlined approach orders HFEs in the context of a given scenario based on other events in the sequence and cutsets, not just combinations of HFEs. The consideration of the entire sequence discussed above accounts for similar attributes and functions between HFEs.

All post-initiator HFEs are still sequenced based on their Tdelay, but the results are reviewed to determine if any Tdelay overrides are necessary to modify the original ordering. Existing cutsets could be used to help focus the analyst's attention but are not required. Overrides are prioritized and defined based on any basic event that may appear in cutsets along with the HFEs. The scenario-focused streamlined approach permits the HRA analyst to look at all events in a timeline, see all the override rules, and determine which rules should be applied.

The current process allows users to change the Tdelay for a pair of HFEs in a combination, but it is changed globally across the entire dependency analysis. The streamlined method proposes to instead use rules based on events in the sequence to control the application of the Tdelay changes. These rules can include other events in the accident sequence progression, not just HFEs (e.g., initiating event, sequence labels, equipment failures, etc.). Table 1 shows an example of rules that could be applied based on the sequence in which an HFE appears as well as how an existing Tdelay override could be included in the new process.

The HFE for failing to initiate alternate low-pressure injection (HFE-ALTINJ) in a boiling-water reactor has an initial Tdelay of 30 minutes based on the bounding accident sequence involving a failure of early high-pressure injection. If the independent failure probability does not change significantly with the extended time available, the ability to adjust the Tdelay for the HDA helps streamline the overall process.

Table 1. Example structure of rule-based Tdelay overrides.

Order	Tdelay	Rule	Comment
1	200	SEQ-TT-25	If HFE-ALTINJ appears in sequence TT-25, the Tdelay is increased from 30 to 200 minutes as initial high-pressure injection is available per the event tree structure
2	200	SEQ-TM-25	If HFE-ALTINJ appears in sequence TM-25, the Tdelay is increased from 30 to 200 minutes as initial high-pressure injection is available per the event tree structure
3	150	HFE-LPEXT	If HFE-ALTINJ appears with HFE- LPEXT, the Tdelay is increased from 30 to 150 minutes because low-pressure injection was available but not for an extended period

2.3.1.3 Dependency Level Override Enhancement

In the suggested streamlined method, the existing algorithm would still be used to assign DLs to all possible HFE pairs (HFE1 affects HFE2 and HFE2 affects HFE3) but this would be performed earlier in the PRA update process.

The analyst still reviews the HFE pairs to determine if the assigned DLs should be overridden. Overrides are prioritized and defined based on any basic event that may appear in the sequence and cutsets together. The DL assessment is stored for each event pair (HFE_N and HFE_{N-1}) and overrides are specified based on other events in a sequence like the method used for Tdelay overrides.

Table 2 shows the initial DL assignments for the HFE_N to initiate alternate low-pressure injection (HFE-ALTINJ). Table 3 provides an example of how the rule-based DL overrides are defined for an HFE pair (HFE-ESW preceding HFE-ALTINJ).

Table 2. Example HFE DL assignments for HFE-ALTINJ.

HFE_{N-1}	Description	DL for HFE_N after HFE_{N-1}	Global DL Override	Rule-Based DL Overrides
HFE-ESW	Operator fails to initiate essential service water	High	—	Two
HFE-HPI	Operator fails to initiate high-pressure injection	High	Medium	—
HFE-LPEXT	Operator fails to extend low-pressure injection duration	Low	—	—

Table 3. Example DL overrides for HFE-ESW preceding HFE-ALTINJ.

Order	DL	Rule	Comment
1	Low	SEQ-TT-25	DL should be decreased due to increased time between HFEs
2	Low	SEQ-TM-25	DL should be decreased due to increased time between HFEs

Looking at the case above, we see that DLs between HFE-ALTINJ and all the HFEs in Table 2 can be assigned and reviewed. If HFE-ALTINJ appears after HFE-ESW based on the Tdelay or Tdelay override, the DL is set to high. If the same pair appears with SEQ-TM-25, the DL drops to low based on an analysis of the timing and analyst input. Similarly, if HFE-ALTINJ comes after HFE-HPI, the normal DL would be set to high (algorithmically derived), but in the example above, the pair's DL is globally overridden through an analyst evaluation and the pair's DL can be set to medium for the entire HDA (Table 2).

Although the fundamental elements of the analysis haven't changed, the application of DLs based on each pair's ordering provides a greater fidelity of how an override is applied and ultimately how the JHEP is derived.

2.3.1.4 Applying Human Reliability Analysis Dependency in Quantification Process Enhancement

The current post-processing technique for including the HDA replaces sets of independent HEPs with a combination event that has the calculated JHEP value. Eliminating the use of HFE combinations requires revising the technique. The alternative technique (shown in Figure 6) includes the following steps performed for each cutset created by the quantification engine (Step 5a) containing more than one HFE:

1. Identify HFEs in the cutset and sequentially order them based on the Tdelay and any Tdelay overrides (Step 5b)
2. Identify the sequential HFE pairs and appropriate DL considering any DL overrides (Step 5c)
3. Replace the subsequent HFEs (HFE_N) that are part of each HFE pair with a basic event that represents the calculated conditional probabilities (Step 5d).

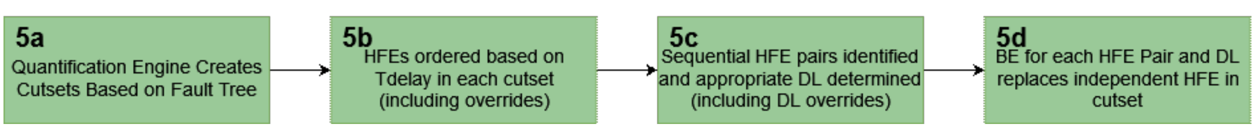

Figure 6. Proposed PRA model quantification steps to include HDA.

The application of the HFE conditional probabilities ideally provides a naming convention that enables the user to quickly identify the sequential HFEs and assigned DL; however, a unique ID that can be cross-referenced back to this information would also be acceptable. Table 4 displays a representative set of cutsets that may be produced by a quantification engine with the specific random failure events not shown. The cutsets involve four HFEs with their independent Tdelay values shown in minutes. Table 5 shows that, after applying the Tdelay overrides, the order of the HFEs changes in Cutset 4.

Table 4. Example cutsets before dependency analysis.

Cutset	Accident Sequence	Non-HFE Failures	HFE-1 (Tdelay)	HFE-2 (Tdelay)	HFE-3 (Tdelay)
1	SEQ-TT-25	—	HFE-ESW (10)	HFE-ALTINJ (30)	—
2	SEQ-TT-35	—	HFE-ESW (10)	HFE-ALTINJ (30)	—
3	SEQ-TT-35	—	HFE-HPI (25)	HFE-ALTINJ (30)	—
4	SEQ-TT-35	—	HFE-HPI (25)	HFE-ALTINJ (30)	HFE-LPEXT (30)

Table 5. Example cutsets after Tdelay overrides dependency analysis.

Cutset	Accident Sequence	Non-HFE Failures	HFE-1 (Tdelay)	HFE-2 (Tdelay)	HFE-3 (Tdelay)
1	SEQ-TT-25	—	HFE-ESW (10)	HFE-ALTINJ (200)	—
2	SEQ-TT-35	—	HFE-ESW (10)	HFE-ALTINJ (30)	—
3	SEQ-TT-35	—	HFE-HPI (25)	HFE-ALTINJ (30)	—
4	SEQ-TT-35	—	HFE-HPI (25)	HFE-LPEXT (30)	HFE-ALTINJ (150)

By reviewing the data in Table 2, the DLs for each sequential pair can be identified. Table 6 displays how the updated cutsets would look with each subsequent HFE replaced with an event that represents the HFE, the preceding HFE, and the DL assigned between the pair. This also considers any DL overrides defined in Table 3. Cutsets 2 and 4 have basic events added based on the standard dependency levels for HFE-ALTINJ of High and Low with the preceding events HFE-ESW and HFE-LPEXT respectively. Cutset 1 takes advantage of the Rule Based DL override for HFE-ALTINJ and HFE-ESW to Low based the accident sequence label. The dependency level of Medium applied in Cutset 3 is based on the global override for HFE-HPI and HFE-ALTINJ. Cutset 4 also provides an example on how this approach would apply when more than two HFEs are present by including an additional HDA event representing the dependency between HFE-LPEXT and HFE-HPI with a DL of High.

Table 6. Example cutsets after applying proposed technique.

Cutset	Accident Sequence	Non-HFE Ilures	HFE-1	HFE-2	HFE-3
1	SEQ-TT-25	IFE-ESW	HDA-ESW:ALTINJ-LOW	—	
2	SEQ-TT-35	—	HFE-ESW	HDA-ESW:ALTINJ-HIGH	—
3	SEQ-TT-35	—	HFE-HPI	HDA-HPILTINJ-MED	—
4	SEQ-TT-35	—	HFE-HPI	HDA-HPI:LPEXT-HIGH	HDA-LPEXT:ALTINJ-LOW

Table 7 displays the results of applying the current technique for the four example cutsets. The same HFE combination event (JHEP-123) is applied in the first two cutsets, as the rule-based DL override was not possible in the previous technique.

Table 7. Example cutsets after applying current technique.

Cutset	Initiating Event	Accident Sequence	Non-HFE Failures	HRA combination
1	IE-TT	SEQ-TT-25	—	JHEP-123
2	IE-TT	SEQ-TT-35	—	JHEP-123

Cutset	Initiating Event	Accident Sequence	Non-HFE Failures	HRA combination
3	IE-TT	SEQ-TT-35	—	JHEP-75
4	IE-TT	SEQ-TT-35	—	JHEP-456

Comparing Table 6 and Table 7 illustrates how the proposed updated technique provides more resolution when obtaining basic event importance measures directly from the final cutsets as the HFE-related event counts are four for the current technique and seven in the proposed update. Being able to quickly obtain the risk importance values for the specific HFE pairs that may be common across several HRA combinations (as they are currently defined) can help focus where refinements to the dependency analysis would be most effective at reducing overall model risk.

In addition to the existing PRA importance measures, a new importance measure could show how sensitive the quantified results are to the assigned DLs. Similar to how the Birnbaum value measures the change in the quantified result when an event value is set to 1 or 0, a new risk metric could measure the change in the quantified result when the DLs for individual HFE pairs are increased or decreased. Under the current methodology, this type of sensitivity would be extremely time consuming to perform. Understanding the parts of the model to which the results are sensitive allows the analyst to focus efforts on ensuring the most accurate DLs are assigned.

2.4 Proposed Research Roadmap

This report introduces enhancements to the HDA and a novel technique for integration of HRA dependency results into the PRA model quantification. The proposed approach improves efficiency of PRA by moving the HDA to the early steps of model development instead of postponing it all the way to the end. The approach is also expected to improve model quantification speed by reducing the number of HFE combinations that must be processed at each quantification run. Lastly, the proposed approach improves understanding of the overall system performance and more specifically operators' performance during the upset plant conditions.

However, there are some details of the process that need to be explored further. Additional research is necessary to investigate how to address the HRA dependency minimal JHEP value (i.e., the JHEP floor value). The application of a JHEP floor value is expected by the PRA standard (ASME 2013); however, current practices for doing so impact the ability to measure the risk significance of HRA dependencies and relies on the current HFE combination approach. While not discussed in this paper, techniques are available to permit an application of the JHEP floor value within the proposed new process.

As the proposed process eliminates the reliance on HFE combinations, one of the first priorities is to verify that the proposed post-processing technique to apply HFE dependencies can be performed efficiently with current computer resources used by commercial nuclear power plants (NPPs). This task would require developing the post-processing tool software as well as a way to translate existing Tdelay and DL overrides into the new format. A pilot effort of the process would then be possible.

3. HUMAN DATA COLLECTION ON DEPENDENCY

3.1 General Introduction

Error dependency is often mentioned in the human reliability literature. Error dependency is concisely captured by the notion that "error begets error." Less concisely, once a human commits an initial error, they are more likely to commit subsequent errors. For the sake of example, imagine making an unprotected left turn and accidentally not seeing an oncoming vehicle that must take evasive action to avoid collision. Subsequent to that event, one might be frazzled and more likely to commit further driving errors (e.g., speeding, poor navigation, distractedness).

The concept of error dependence was first described by Swain and Guttman (1983) in NUREG/CR-1278 as part of the Technique for Human Error Rate Prediction (THERP) method. Error dependence is most often thought of as positive dependence where an error in one task increases the probability of error in subsequent tasks. However, error dependence can also be negative, where failure during a task decreases the chance of error during subsequent tasks. Error awareness can serve as a learning cue that improves subsequent performance. Our focus here is on positive dependence, where an initial error triggers a greater likelihood of subsequent errors.

Swain and Guttman describe error independent tasks as tasks where one task does not impact another and vice versa. Tasks are dependent if subsequent tasks result from previous tasks (direct dependence) or if performance on multiple tasks are influenced by shared PSFs (see Figure 7).

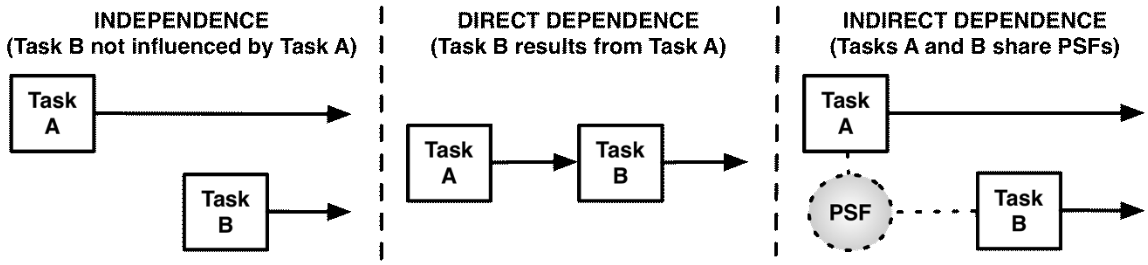

Figure 7. Task dependence diagrams (Boring 2015).

We are now 40 years past Swain and Guttman's first description of error dependence. Despite the decades that have passed, we still do not have a concrete understanding of error dependence to support HRA (Mortenson and Boring 2021). Notably, task batteries treat tasks as independent and do not account for task dependency. Mortenson and Boring go as far as to question whether task dependency is a real phenomenon or merely a theoretical concept. If task dependency is real, why does it exist? Is it an evolutionary mechanism, or the symptom of underlying confusion?

In contemporary practice, Standard Plant Analysis Risk-Human (SPAR-H; Gertman et al., 2005) provides a dependency determination table (see Table 8) that allows an analyst to identify a qualitative dependency classification that can be used to increase the human error probability (HEP) in proportion to the level of dependency (Blackman and Boring 2017). The authors also caution that simply following the table may not yield accurate results.

Within traditional HRA, the dependency calculation is the last step for HEP quantification, and dependence can essentially override previously estimated HEPs to the extent where Boring, Park, and Mortenson (2021) suggest conducting a dependency screening to examine whether non-dependency factors (nominal HEP and basic HEP) even need to be considered.

Table 8. Dependency determination table from SPAR-H.

Condition Number	Crew (same or different)	Time (close in time or not close in time)	Location (same or different)	Cues (additional or no additional)	Dependency	Number of Human Action Failures Rule ☐ - Not Applicable. Why?________________
1	s	c	s	na	complete	When considering recovery in a series e.g., 2nd, 3rd, or 4th checker If this error is the **3rd error in the sequence**, then the dependency is at least **moderate**. If this error is the **4th error in the sequence**, then the dependency is at least **high**.
2				a	complete	
3			d	na	high	
4				a	high	
5		nc	s	na	high	
6				a	moderate	
7			d	na	moderate	
8				a	low	
9	d	c	s	na	moderate	
10				a	moderate	
11			d	na	moderate	
12				a	moderate	
13		nc	s	na	low	
14				a	low	
15			d	na	low	
16				a	low	
17					zero	

In the HRA domain, the contention is that discarding dependency could lead to a critical underestimation of human error probabilities. However, strong empirical evidence for dependency does not exist. Paglioni and Groth (2021) have gone as far as stating a "need for a revolution in HRA dependency treatment by showing that there does not exist an adequate, robust definition of dependency as a general concept in any authoritative literature."

Paglioni and Groth (2021) provide a revised general definition for dependency in HRA:

> A dependency exists between HRA variables if they are connected by a direct or indirect causal relationship which changes the conditional probabilities of the variables, regardless of whether the existence or utility of the variables is acknowledged within HRA.

It is a common understanding amongst HRA researchers that foundational work is needed to examine human error dependency. The primary domain of HRA is nuclear power, where the tasks under decomposition are complex nuclear operations tasks. Licensed operators are a rare commodity, and gathering operator data in research contexts to support an examination of error dependence has not happened in the last 40 years and is perhaps equally likely to not happen in the next 40. Scenario event trees are complex with dozens or more branch points. When the probability of error for any given cutset can be on the order of one out of 1,000, it becomes unfathomable to acquire the millions of iterations needed to statistically examine error dependence within an event or fault tree.

Dependency in HRA is a critical yet ill-defined concept. Traditional HRA does not function without it. Despite several decades of research, a formal definition and understanding of dependency does not exist, yet dependency remains a large driver in overall HEPs in probabilistic risk models. Practitioners are left with heuristics and rules of thumb that yield estimations that seem reasonable, but researchers are unsatisfied (Blackman and Boring 2017). As noted, some have begun more formal theoretical treatments of dependency (Paglioni and Groth 2021), but foundational work is still needed.

In this report, we describe an initial effort to investigate error dependency using empirical data derived from a simulator study. The original intent of the study, described in the next section, was not to investigate dependency but rather human performance using computer-based procedures. However, due to the way human actions were recorded, it proved a particularly fruitful way to capture successful and erroneous human actions. After a review of the study, the subsequent sections describe data extraction and dependency findings.

3.2 Study on Three Types of Computer-Based Procedures

One upgrade being undertaken at NPPs involves a shift away from paper-based procedures towards computer-based procedures (CBPs). The purpose of a procedure is to guide operators through a particular task—they are expected to follow it precisely and without deviation. A 1995 report from the U.S. Nuclear Regulatory Commission revealed that problems following paper-based procedures contributed to the majority of reportable events (U.S. NRC 1995). The report highlighted several challenges with paper-based procedures, including identifying the correct procedure to follow, the stress involved with following multiple nested procedures in emergency scenarios, and issues with divided attention.

Modernizing procedures includes integrating electronic devices and machines, such as computers, into an environment where tasks were originally executed by humans (Parasuraman and Riley 1997). Using digital tools can help lower the number of human errors by alleviating some of the operator's mental workload, which can directly impact performance (Hwang et al. 2008). The Institute of Electrical and Electronics Engineers (IEEE) Standard 1786 classifies CBPs into three types, depending on the amount of digitalization integrated into the procedure (IEEE 2022):

- Type 1 closely resembles traditional paper-based procedures in that it displays the procedure on a computer screen
- Type 2 has more capabilities and can additionally display process data and step logic, display results, and provide access links to displays and soft controls that reside on a separate system
- Type 3 has Type 2's capabilities and can automatically carry out sequences in the procedure and has embedded soft control features.

Unlike Types 1 and 2, Type 3 procedures can manipulate the plant directly from within the procedure instructions.

A between-subjects experimental design was used with one independent variable, the CBP type with three levels: Type 1 (sample size [n]=10), Type 2 (n=8), and Type 3 (n=9). Participants each completed two different scenarios: non-event (start-up) and event (loss of feedwater [LOFW]). The start-up scenario required participants to use a procedure that would start the reactor. To successfully complete this scenario, the reactor must not be tripped (shut down). The LOFW scenario required participants to carry out a fault diagnosis and then shut down the reactor. To successfully complete this scenario, the reactor must safely be shut down within the time limit.

The Rancor Microworld Simulator (Rancor) was used in the study. Rancor is a simplified simulator that mimics both everyday operations and emergency situations in an NPP. It was developed at INL and the University of Idaho (Ulrich et al. 2017) by human factors scientists and contains essential NPP components and systems, such as the reactor core, control rods, reactor cooling and feedwater pumps, bypass and load valves, steam generators, and turbine. The interface consists of three areas (see Figure 8). The piping and instrumentation diagram depicts the relationship between piping, process equipment, instrumentation, and control devices. Rancor displays the plant state. The overview area collates much of the important information from the piping and instrumentation diagram and puts it in one convenient location. This visual area also contains the alarm panel. The control area allows the user to control the plant. These interface elements may be joined as a single screen or split across separate windows or screens.

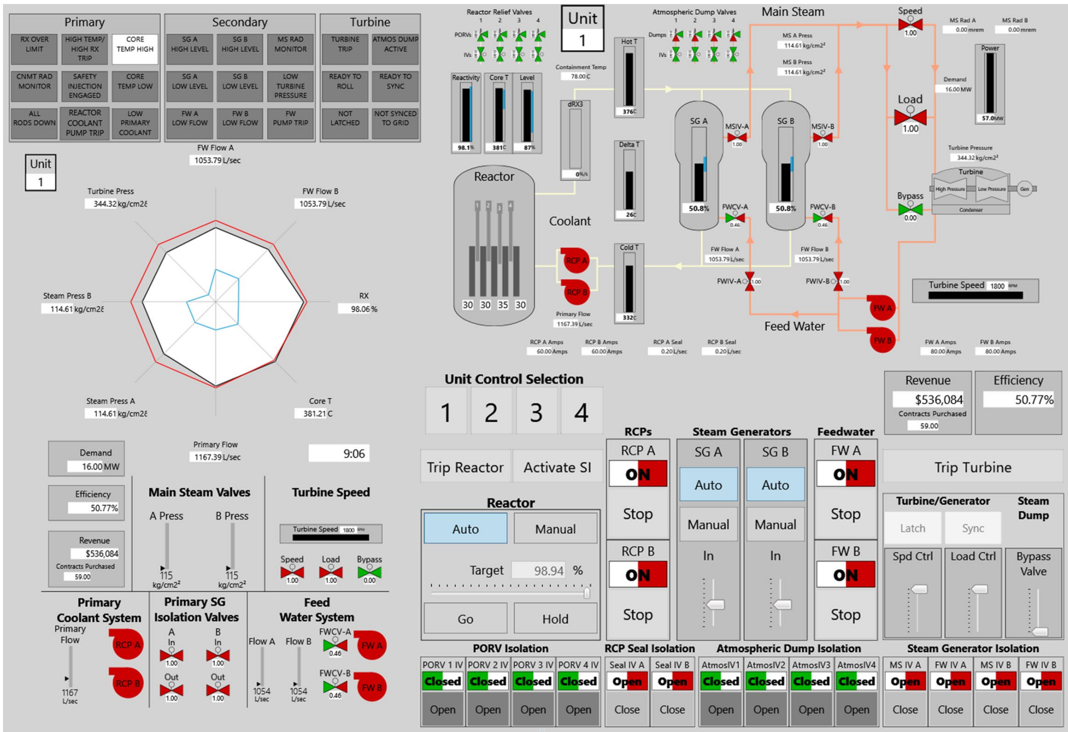

Figure 8. Rancor display as presented to participants in its default configuration.

Importantly, Rancor allows both operations experts and novices without operations experience to effectively monitor and control a simulated NPP to a proficient level. It monitors and records the status of all plant parameters as well as any actions taken by participants in each procedure step. Several studies have employed Rancor to successfully collect human factors and human reliability data involving nuclear operations. Recent studies have compared student and licensed operators and demonstrated a good generalizability of findings from student participants (Park et al. 2022; Park et al. 2023).

The study was conducted in the Human Systems Simulation Laboratory (HSSL; see Figure 9) at INL, where up to three participants working on separate simulator bays participated at one time. Participants were randomly assigned to one CBP type (Type 1, 2, or 3). They provided informed consent, completed a brief demographic survey, and watched a Rancor training video specific to their CBP-type group assignment. The video explained how nuclear power is generated and introduced the Rancor interface, the purpose of using procedures in nuclear operations, and the study's goals. The video then went through the start-up scenario using the procedure in Rancor. Participants had 3 minutes to practice. The same occurred for the feedwater scenario. Participants then completed up to three repeated trials of the same type during a 20-minute timed scenario. Rancor logged plant and human performance parameters.

Figure 9. The HSSL at INL in 2022.

3.3 Simulator Log Analysis for Dependency

3.3.1 The Data Problem

The nuclear industry at large has experienced challenges collecting sufficient data due to difficulties in acquiring simulator access and licensed operators as participants in studies (Ulrich, Boring, and Lew 2019). Simulator access has been expanded with generic plant models that are within the ability of well-funded university laboratories but still require substantial expertise to install and maintain. Lastly, the type of researchers needed to conduct studies on human performance typically do not necessarily have nuclear engineering knowledge or operational expertise required to develop scenarios with appropriate operational aspects to maintain a realistic nuclear operations environment and conditions to elicit useful human performance data. Typical full-scope studies require an interdisciplinary team to achieve the overall level of expertise required to perform a scenario-based study. Even with an experienced team, operators are expensive and have demanding schedules occupied with their scheduled shifts to operate the plant and the remaining filled by demanding and extensive training. Ironically, commercial plants record a lot of human performance data; however, these data are proprietary. Consequently, the U.S. nuclear industry has suffered from a dearth of data on human performance, leaving the HRA community struggling to identify effective methods to develop accurate models of human performance.

The simple solution to obtaining human performance data within a nuclear context is a platform that is accessible to human performance researchers and can generate large volumes of data. Rancor is a simplified nuclear process control simulator developed specifically to serve as a platform to evaluate human performance in a nuclear process control room setting (Ulrich et al. 2017). There are several key characteristics that distinguish Rancor from the traditional simulators used for human factors and HRA research within a nuclear context. The simplicity in the design eliminates much of the overly challenging complexity of an actual nuclear control room while maintaining the core tasks and functions that operators perform. This allows the tool to be used by human factors or psychology researchers without extensive nuclear operations expertise to develop meaningful scenarios. Additionally, the simplicity allows for an inexperienced test subject pool outside of the typical licensed or formerly licensed nuclear power plant operators used in traditional full-scope simulators. This last capability is perhaps one of the most powerful for this platform, since one of the greatest challenges for human factors researchers in this field is obtaining and funding operators to act as test subjects. This barrier has limited the research to large institutions that can afford operator participants and the compliment of simulator developers and operations experts needed to perform experiments. Since this approach is using students as a surrogate for operators, the key critic is the generalizability of inexperienced participants to that of seasoned nuclear operators.

The initial development and continued refinement of Rancor focused on ensuring the simulation functionally represents the actual nuclear process control tasks in a simplified manner (Ulrich, Boring, and Lew 2019). Continuing validation research compares students and operators using full-scope and Rancor simulators to establish equivalency and identify discrepancies. The goal is to map the data collected from Rancor to actual operations to understand which constructs directly translate, which simply do not, and which must be adjusted to generalize appropriately. This validation work has thus far demonstrated promise, as evidence to validate the generalizability has been found in several experiments (Ulrich et al. 2021, Park et al. 2022). The remainder of this chapter describes using Rancor with an integrated CBP system to support human performance data collection to illustrate how the unique capabilities of Rancor provide the ability to use CBPs as an experimental framework to collect large amounts of task-level human performance data to fill the HRA data modeling gaps.

3.3.2 Computer-Based Procedures for Experimental Data Collection

As noted, a version of Rancor was developed with a selectable CBP system module with the three levels of CBPs. The CBP system module is integrated into the overall Rancor simulation, which affords data connections to support all the functionality required to represent Type 3 CBPs with embedded indications and soft controls. Thus, Type 1 and 2 CBPs are simply variants of the Type 3 system (see Figure 10). All procedure configurations allow users to select the appropriate procedure to expand its contents to reveal the first step. Users are required to manually execute all control actions and manually mark each procedure step complete.

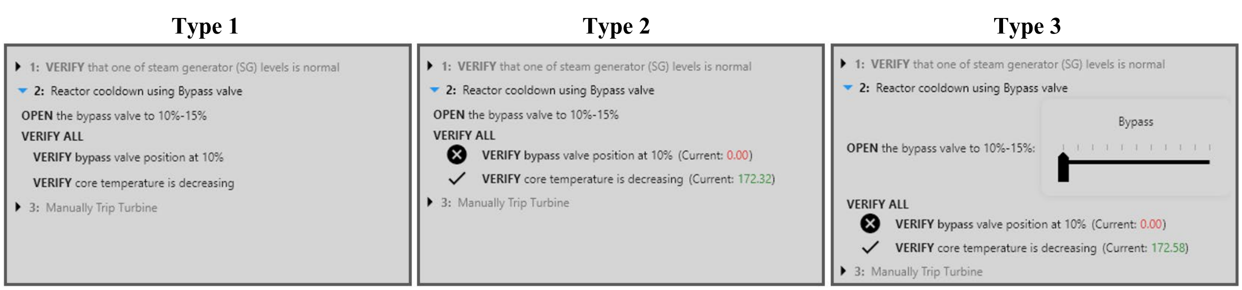

Figure 10. Three types of CBPs in Rancor.

To the best of our knowledge, prior experiments focused solely on the human-factors-related constructs when investigating CBPs. Here a novel and central aim is to illustrate the use of a CBP system as a research tool itself to unlock new possibilities for human performance evaluation. CBPs can provide a fine-grained analysis of human performance that is otherwise not possible without tedious manual event coding. To illustrate the analysis power afforded by the CBP system in Rancor, it is first necessary to explain how human performance is typically analyzed in full-scope simulator studies, then with Rancor without CBP, and finally with CBP enabled in Rancor.

Most full-scope simulators have limited data recording capabilities as they were originally intended to train operators rated by simulator trainers through observation. In an experimental setting, a subject matter expert, often the simulator trainer, acts as an expert observer. To eliminate subjectivity the experimenter may provide a scoring rubric and a list of key tasks to rate the operator(s) as they complete the scenario.

Within the training realm, the rubric may simply be a binary success or fail metric, while in some experimental methods, such as Supervisory Control and Resilience Evaluation (SCORE), a Likert scale captures more nuance (Braarud, Eitrheim, and Fernandes 2015). The observer may also record the time each key activity occurred, both when it was initiated and completed. The time to complete tasks is a powerful and useful metric, though it is typically used only in a gross expert estimation approach in which the length of time is deemed to be appropriate or insufficient. This is typically the extent of the human performance evaluation and, as should be evident, it offers very limited data because the entire scenario is collapsed into a small number of data points associated with a human error. In a small minority of experiments, based on this authors' involvement from prior full-scope studies performing the analysis, some process parameters are used to serve as an additional type of metric for performance. A limited set of process parameters are recorded throughout the scenario so that a running mean-square error (MSE) can be calculated. By referencing the timesteps recorded by the expert observer and, as needed, viewing video recordings to verify recorded timestamps, portions of the MSE for relevant process parameters can be used to correlate with the rated performance. This becomes a tedious process as it requires significant manual effort. As such, much of the data recorded during full-scope simulator studies goes unused since the hypothesis can typically be supported by lower resolution methods. Indeed, the debrief following each scenario in which operators report their experience and area probed by observers to understand any challenges they witnessed is typically sufficient to address much of the basic human factors issues often targeted during these experiments.

Rancor was developed to augment full-scope studies and focuses explicitly on methods to efficiently output human performance data and plant process data to support higher resolution analyses than is typical in full-scope studies. Due to its simplistic model, the entire model can be output at each timestep such that an exact history of all process parameters can be used for analysis. Full-scope simulators with 10,000 to 100,000 parameters do not have this luxury. Additionally, Rancor records every action an operator takes to a subsecond accuracy. A set of Python-based analysis tools are freely available and provide aggregated data outputs that integrate the process parameters with operator actions to aid analyses. This essentially automates the tedious process required to integrate, time-synchronize, and code the disparate data sources as must be done in a full-scope experiment. It certainly is a step in the right direction, since it is much easier to develop analysis scripts that rely on action logs to focus on key time windows that may be informative to human performance, but it is still necessary for the experimenter to manually define these. The following paragraphs will describe how CBP, as implemented in Rancor, can be used to support human performance evaluation at a detailed task resolution.

Here, the initial data analysis required developing a post-analysis Python application to integrate the data files recording plant state parameters at 1 second intervals, participant action logs, and procedure logs. This analysis became the impetus for this report, as it revealed the potential to systematically collect large amounts of task-level human performance data with Rancor. Second, the analysis revealed the data outputs were largely adequate, but with a few simple additional data outputs, much of the post study analysis required to extract the task-level details from the various data sources could be automated to further reduce the barriers to evaluating task-level data.

The potential power of CBPs as an experimental tool is due to the hierarchical structure of the procedures. This structure provides the framework that supports detailed task-level resolution and overall scenario-level resolution analysis. The suite of procedures represents a collection of possible goals operators can achieve with the system. Each procedure represents a single high-level goal and is comprised of steps that provide methods to achieve the overall goal. Based on the plant state, the primary procedure path that provides the optimal method to achieve the procedure goal may not be feasible. Therefore, the procedure contains contingencies that guide operators to identify the system state and then choose an appropriate path to maneuver the plant to the desired goal state. As such, a well-designed procedures system is deemed closed loop if it prescribes methods to accommodate all known plant states, though it is rare or impossible to truly achieve a closed loop procedure system. However, through the continued use and revision of the procedures, the gamut of plant states gradually becomes sufficiently covered such that the procedures effectively cover all possible states, as is the case in nuclear process control for the existing U.S. fleet of commercial reactors.

Armed with this hierarchy, the CBP system serves as an automatic trial management system that is able to mark time intervals with explicit goals. In traditional full-scope studies, an experimenter must manually code the time intervals, which is often simply not feasible to support experiments aiming to achieve thousands of samples of individual step time intervals. The CBP system divides the scenario into discrete time intervals surrounding explicit goals. Since the CBP system is integrated into Rancor, it can query the model throughout each procedure step's time interval. Several key elements are needed to evaluate human performance at this basic task level, including initial plant state, actions, and postcondition plant state. First, the initial plant state as the participant entered the step must be determined. Then the logic of the procedure step must be evaluated against the initial plant state to determine the appropriate actions needed to satisfy the step's goal. The participant executes the actions they deem necessary based on their understanding of the step instructions as they relate to their mental model of the plant state. When the step is marked by the participant, the CBP evaluates a postcondition for the step to determine if the desired state has been achieved. A miniature trial of sorts has now been defined for a task-level human action for qualitative and quantitative characterization suitable for HRA methods.

There are many useful metrics that can be used to evaluate human performance within the task level, which will be dictated by the experiment goals. Since Rancor was developed to augment existing data collection methods for HRA purposes, metrics with direct implications for HRA model construction will be highlighted. Dynamic HRA methods, such as HUNTER (Boring et al. 2022) need well-defined timing distributions for the basic tasks. HUNTER uses a Goals, Operators, Methods, Selection Rules (GOMS)-HRA framework, which consists of a dictionary of basic human actions, for example, task-level primitives that can be combined to represent more complicated actions found within a procedure step (Boring and Rasmussen 2016). The elapsed time to complete the action from entering the procedure step is the key element needed to populate the various GOMS-HRA primitives. Thus far, only manual actions have been considered since the actions are logged and available to analyze against the procedure logs. Information tasks, such as reading indicators or making a decision, are more challenging since they do not have a specific element recording a timestamp and value. However, the CBP system does record the time a step is completed, which serves as a more crudely defined timestamp since determining the step was complete and marking it are captured in the elapsed time. Eye tracking provides another method for acquiring accurate timestamps for identifying required plant parameter information. With this information, the CBP framework to identify the relevant time interval can then function in an identical manner to the actions but instead using the eye tracking events to evaluate against.

To evaluate performance in terms of human error probabilities, actions completed within a procedure step time interval can be evaluated. Each action can be evaluated to determine if the correct plant state was achieved, which represents subtask-level analyses. The sequence of tasks can be evaluated to determine if they were executed in the order prescribed by the procedure. Missing steps provide a metric for errors of omission, while extra unprescribed steps provide a metric for errors of commission. Quantitatively, actions that entail a manipulation occurring across a continuum can be evaluated in terms of error from target optimal value similarly to how MSE is calculated for process parameters at the scenario level in full-scope scenarios. Lastly, constructs such as dependency can be evaluated by examining subsequent errors to identify proximity in time and potentially system (i.e., the same or independent system).

This work merely represents the initial development of a CBP system and evaluation of a small set of data to serve as a proof of concept and demonstration of using a CBP system as an experimental tool to target detailed task-level analyses. Future work will add additional validation logic to automatically generate the data in a synthesized format for timing and human error performance analyses.

3.3.3 Dependency Data Extraction

We configured Rancor to output several files containing different parameters recorded during each trial. Each of the files contains either time series data recorded at set intervals or event data captured at the moment of the event. This data output structure followed a typical database approach in which a primary and secondary key were used to link data between different files. In this case, the timestamp and simulator timestep variables were used as linking keys to aggregate the data since they support the chronological ordering of various data objects. The analysis compiled action and procedure events recorded during the simulation. An action event is simply a control actuation of some component within the simulator while a procedure event is opening a particular CBP and place keeping within the procedure by marking each step complete after performing the prescribed task. Some procedure tasks did require participants to perform verification activities, but these were excluded from the analysis, since there was no way to discern between the procedure acknowledgment associated with the verification tasks and the verification task itself. In some cases, the verification precedes an action, and therefore the verification task is dependent on any other tasks as well as the acknowledgment of the step itself. Without eye tracking to identify when participants acquired or examined the appropriate indication for the verification task, it is not currently feasible to carry out a conclusive analysis of these data.

The time-based analysis of each action and procedure event followed the following process for each trial:

1. Extract the list of action events
2. Extract the list of procedure events
3. Iterate over the list of action objects to categorize each action object with a procedure identifier and procedure step number identifier based on timestamps within the procedure objects.

The rows represent sequential time-based procedure and action events. Each row only contains a single procedure or action event. As the action events were categorized, time-based metrics were calculated as well. The time-based metrics included elapsed time since the last procedure event, elapsed time since the last action event, remaining time between the event and completion of the procedure step, and the duration of the time spent manipulating the action for continuous items (i.e., action events that required multiple interactions, such as modulating the bypass valve to maintain a particular core temperature as prescribed in the procedure). While study participants completed both start-up and LOFW scenarios, only the LOFW scenarios are analyzed here. A LOFW represents a fault or abnormal operation scenario, which is generally of higher interest in HRA than normal operations like start-up.

Table 9. Example of a single procedure sequential set of recorded events classified by the prescribed procedure script.

Timestamp	Step Number	Event	Correct Execution	Error Commission	Error Omission
15:03:56.551	1	open_procedure	1	0	0
15:05:16.464	1	mark_procedure	1	0	0
15:05:30.233	2.1	rc_pump_a	1	0	0
15:05:31.232	2.1	rc_pump_b	1	0	0
15:05:37.803	2.1	mark_procedure	1	0	0
15:06:02.246	2.2	mark_procedure	1	0	0
15:06:06.934	2	mark_procedure	1	0	0
15:06:18.829	3.1	reactor_control_mode	1	0	0
15:06:22.828	3.1	reactivity_rx_target	1	0	0

15:06:33.228	3.1	rod_control_mode	1	0	0
	3.1	bypass_demand	0	0	1
15:07:04.849	3.1	mark_procedure	0	0	0
15:07:10.771	3	mark_procedure	1	0	0
15:07:37.425	4.1	sg_a_control_mode	1	0	0
15:07:40.224	4.1	sg_b_control_mode	1	0	0
15:07:54.822	4.2	fw_pump_a	1	0	0
15:07:57.021	4.2	fw_pump_b	1	0	0
15:08:48.354	4	mark_procedure	1	0	0

The recorded completion of the procedures and execution of actions were coded against a procedure script to determine if participants were accurately performing the task (see Table 9). This categorized each procedure or action event as a:

1. **Correct execution**—correct action completed in the proper sequence
2. **Correct out-of-order execution**—a correct execution performed out of sequence in relation to the prescribed procedure sequence
3. **Error of commission**—executed action was not prescribed by the procedure
4. **Error of omission**—prescribed action was not executed
5. **Optional omission**—prescribed action was not executed but was not classified as an omission since other methods could satisfy the successful execution of the procedure.

The correct execution, error of commission, and error of omission categories are established human error classifications within the HRA domain. The correct out-of-order execution captures nuanced situations in which the order of completing an action may have little negative consequence on the overall task, especially within the Rancor simulation due to failsafe interlocks preventing the participant from entirely omitting key actions for a scenario. Lastly, an optional error of omission classification was included to capture operator discretionary actions. These types of actions include multiple methods to achieve the same net effect in which it is up to the participant to choose a method. This category is a limitation of the analysis and in future work will be removed in lieu of a more complicated logic set to identify that one of the available options was selected. Here it was used programmatically to prevent an overestimation of the error of omission. These could have been coded as errors of omission or commission if a more rigid interpretation of human error were adopted.

Table 10. Sum total counts of classified action events.

Action Event Type	Correct Execution	Correct Execution Out of Order	Error of Commission	Error of Omission	Optional Omission
activate_si	55	14	2	12	0
bypass_demand	62	0	30	8	0
fw_cv_a_demand	7	15	8	0	45
fw_cv_b_demand	12	3	2	0	48
fw_iv_a_demand	0	0	2	0	0
fw_pump_a	23	21	42	10	0
fw_pump_b	16	8	19	12	0
rc_pump_a	36	8	0	5	0

Action Event Type	Correct Execution	Correct Execution Out of Order	Error of Commission	Error of Omission	Optional Omission
rc_pump_a_seal_iv_demand	0	0	1	0	0
rc_pump_b	36	7	1	4	0
rc_pump_b_seal_iv_demand	0	0	1	0	0
reactivity_rx_target	0	0	14	0	0
reactor_control_mode	0	0	1	0	0
reactor_trip	21	4	2	48	0
rod_control_mode	0	0	2	0	0
sg_a_control_mode	39	15	1	19	0
sg_b_control_mode	29	19	2	21	0
turbine_trip	22	4	0	47	0
	358	118	130	186	93

After classifying the action and procedure events, the sum total for each type of action performed within the feedwater scenario was calculated to generate Table 10 and Table 11.

To the best of our knowledge, this is the first study able to capture data at this fine of a resolution level to support dependency analyses. The following section demonstrates one method to evaluate these data for dependency effects by examining cutsets, or subsequent errors observed within a single trial, of errors of commission.

Table 11. Sum total counts of classified procedure events.

Procedure Event Type	Correct Execution	Correct Execution Out of Order	Error of Commission	Error of Omission	Optional Omission
mark_procedure	439	10	0	1	0
open_procedure	117	6	0	1	0
unmark_procedure	0	0	3	0	0
	556	16	3	2	0

3.4 Loss-of-Feedwater Scenario

PRA is a systematic and quantitative method for analyzing the risk of undesired events or accidents using event and fault trees. PRA provides a framework for determining not only the frequency of accident occurrences but also identifying various accident scenarios and their contributions to the risk by using minimal cutsets, which are generated as the output. Event trees analyze accident scenarios based on several safety functions that follow an initiating event. Fault trees are used to analyze the causes of the safety function failures. Operator actions performed according to operating procedures are included in the analysis of safety functions. For example, operator actions such as controlling the bypass valve or atmospheric dump valve of a steam generator or performing feed and bleed operation for reactor cooldown are required. These operator actions play an important role in determining the availability of the safety function. They are modeled as HFEs in the PRA.

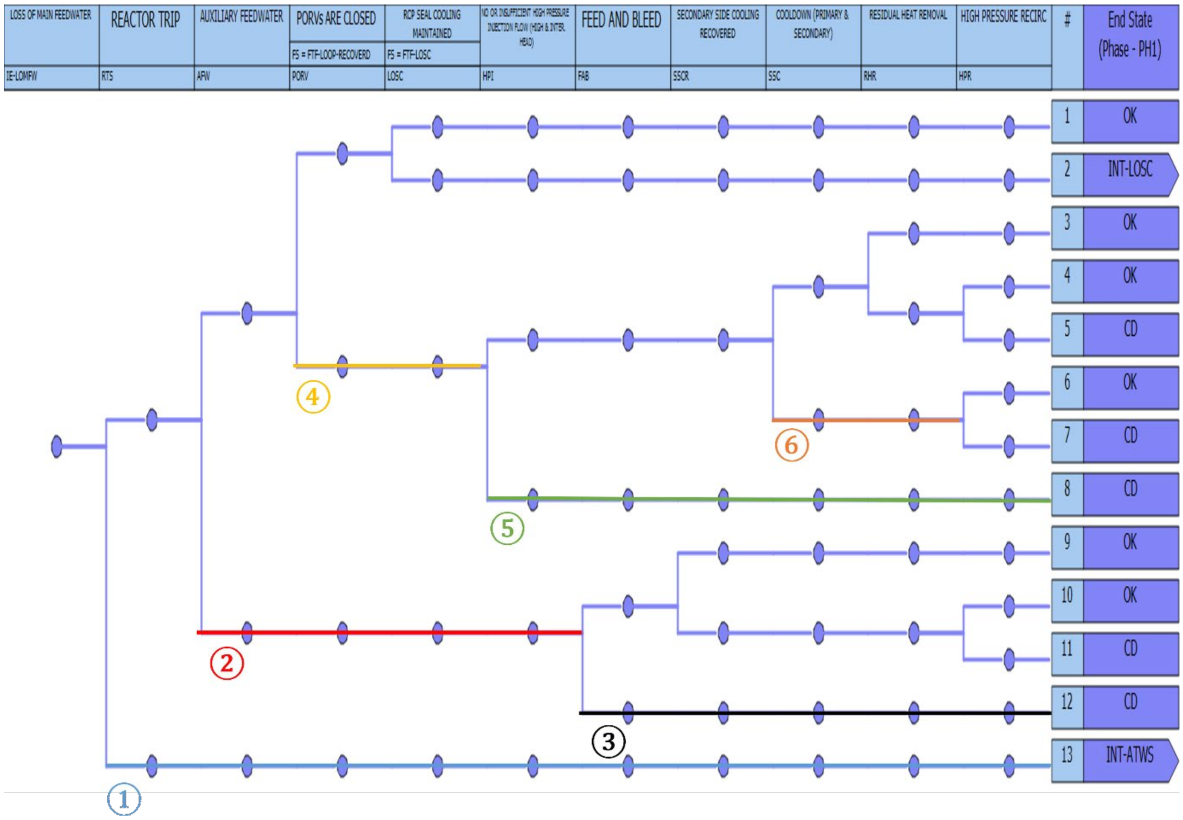

Figure 11. Generic LOFW event tree.

Figure 11 shows one of the event trees for the risk analysis of a generic pressurized-water reactor, which was developed by INL (Ma et al., 2019). The particular event tree is focused on a LOFW scenario, which typically requires several safety functions to mitigate the event. These safety functions include reactor trip, injection of auxiliary feedwater, control of power-operated relief valve (PORV) and others. If the LOFW occurs but the safety functions operate as intended, the reactor will be safe. For instance, there is a success scenario where the reactor is tripped, the auxiliary feedwater is successfully injected, PORVs are properly closed, and reactor coolant pump seal cooling is maintained. The operating procedures for Rancor closely follow the emergency operating procedures developed by Westinghouse Electric Company and are used as the basis of pressurized-water reactor plant procedures. It is worth noting that the safety functions and operating procedures used in Rancor are somewhat simplified compared to those in the event tree presented by INL, but most human actions are functionally equivalent.

According to simplified Rancor procedures, several operator actions exist in the branches of event tree and they are identified as follows:

1. Operator fails to manually trip reactor
2. Operator fails to start feedwater pump manually
3. Operator fails to perform feed and bleed operation
4. Operator fails to close PORVs
5. Operator fails to activate safety injection (SI) manually
6. Operator fails to cool down reactor using bypass valve.

After the LOFW occurs, the reactor should be tripped to insert all control rods. If the reactor trip system does not work automatically, the first operator action is to manually trip the reactor. In the LOFW scenario, the operator can start the feedwater pump manually, which is the second operator action. The third operator action is the feed and bleed operation that activates SI and opens PORVs to cool down the reactor. The fourth operator action is closing the PORVs. This is because when the LOFW occurs, the reactor pressure may be increased and PORVs may be opened to prevent overpressure. Thus, after the feedwater system is successfully initiated, the PORVs should be closed. If the PORVs cannot be closed, SI should be initiated to make up the inventory lost through open PORVs. For this reason, the fifth operator action is activating SI. There is another operator action to cool down the reactor by opening bypass valves. This sixth operator action would be performed after the feedwater can be successfully established to maintain the inventory of the steam generator.

There exist various combinations of operator actions in the event tree. For instance, if the feedwater system fails to operate, the reactor should be cooled down by performing the feed and bleed operation. In the scenario, the failure of the second operator action may impact the third operator action. Furthermore, if an operator is unable to close the PORVs, they should activate SI and cool down the reactor by using bypass valves. In this scenario, the failure of the fourth operator action may impact the fifth or sixth operator actions. This highlights the existence of dependencies between these operator actions in Table 12.

Table 12. Combinations of HFEs.

Combinations of HFEs		
1	② Operator fails to start feedwater pump manually	③ Operator fails to feed and bleed
2	④ Operator fails to close PORVs	⑤ Operator fails to activate SI manually
3	④ Operator fails to close PORVs	⑥ Operator fails to reactor cooldown using bypass valve

3.5 Preliminary Dependency Findings from Empirical Data

We performed several preliminary analyses of the data as reported in this section. Note that analyses are ongoing and will be published subsequently, including results from supplemental studies being conducted.

Table 13 presents the counts of errors of commission performed across all LOFW trials according to each action event type. The study featured timed blocks of trials presented in controlled semirandomized order according to the three types of CBPs, resulting in 11 total blocks of scenarios. The first digits of the trial numbers (1–11) indicate the block number and account for when the scenario was performed. Because of timing, some trials consisted of two runs of the scenario as indicated by the last two digits of the trial number and some featured three runs. In this manner, it is possible to see error sequences within each of the 11 blocks of scenarios.

Table 13. Errors of commission for action event types for the LOFW scenario.

Action Event Type	Trials																									Errors	
	101	102	201	202	301	302	303	401	402	501	502	601	602	701	702	703	801	802	803	901	902	1001	1002	1101	1102	#	Rate
Activate_SI														C		C	C	C	C							5	2.78E-01
bypass_demand	C	C	C	C				C	C			C		C		C	C	C			C			C	2C	14	7.78E-01
fw_cv_a_demand			C		C																			C		3	1.67E-01
fw_cv_b_demand				C																						1	5.56E-02
fw_pump_a					2C				C	2C	C	C	C							C	C	C	C	C		11	6.11E-01
fw_pump_b					C					C														C		3	1.67E-01
RC_Pump_A														C	C	C										3	1.67E-01
RC_Pump_B										C																1	5.56E-02
reactivity_rx_target					C	C	C								C		C							C	2C	7	3.89E-01
reactor_control_mode																	C									1	5.56E-02
Reactor_Trip													C				C		C							3	1.67E-01
sg_a_control_mode								C																		1	5.56E-02
sg_b_control_mode																										0	0.00E+00
Turbine_Trip																			C							1	5.56E-02

The average error rate was 2.14E-1 across all action event types. This error rate is considerably higher than would be expected from typical HRAs for licensed reactor operators. It must be understood that this analysis is conducted at the step and trial level, which is a much finer grained level of analysis than an HFE typically used in HRAs. These step-level errors may be seen as unsafe actions that do not necessarily lead to the overall task failure. In other words, the step-level errors are not all consequential errors that would rise to fault the overall HFE, and they are observed to be at a higher frequency than HFE HEPs.

In addition to the fine level of analysis, the participant pool of student interns varies significantly from licensed reactor operators in that they have considerably less experience and training. According to the SPAR-H HRA method, for example, such inexperienced participants would be estimated to have error rates 10 times higher than skilled operators. Applying SPAR-H reasoning, the average error rate would be approximately 2.14E-2 if reactor operators had performed the same tasks.

In terms of dependency, Figure 12 shows the relationship between initial errors of commission and subsequent errors. If there was an error of commission in Trial 1, the likelihood of error in Trial 2 was four times higher than if there was no error in Trial 1. For Trial 3, the error was six times more likely if there had been a preceding error in Trial 1. While these results are preliminary, they do provide empirical evidence of dependency existence, specifically that an error is more likely once an initial error has been committed.

THERP as the origin for dependency analysis uses an anchor approach, whereby it yokes the overall HEP to particular values rather than transforming them in a systematic relational way. Dependence essentially serves to anchor the Conditional HEP to a higher HEP than the Basic HEP. For low dependence, this anchor is around 1/20 (5E-2); medium dependence, 1/7 (1.42E-1); high dependence, 1/2 (5E-1); complete dependence, 1. For zero dependence, the Conditional HEP remains unchanged from the Basic HEP. Thus, it is not possible to benchmark the result from this study against the predicted dependency values in THERP. Further research is necessary to validate these numbers, but a dependency multiplier, such as that suggested in these data, may prove a more credible approach to dependency calculation than is historically offered by THERP.

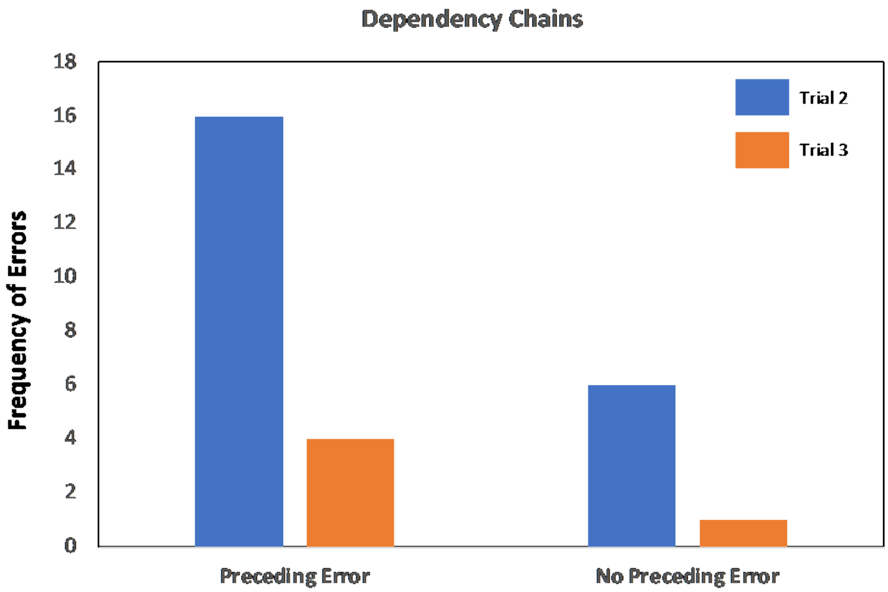

Figure 12. Frequency for errors of commission with and without preceding errors.

A further analysis at the procedure step level is shown in Table 14, which shows a nonparametric correlation analysis between procedure steps that resulted in errors of commission, errors of omission, and the success or failure of the immediately preceding step. No correlations were significant. As is expected, there is no meaningful relationship (correlation coefficient [*r*]=-0.104) between errors of commission and omission, supporting the idea that these are distinct types of errors. There is a slight negative relationship between success on the previous task and an immediately following error of commission (*r*=-0.126) or error of omission (*r*=-0.054). At the procedure step level, success or error is not predictive of subsequent step performance. While Figure 12 offers a longer event horizon across multiple procedure steps and linkage between initial and later errors, Table 14 implies that error priming may not be immediate on subsequent actions from one step to the immediate next step.

Table 14. Nonparametric correlations of errors of commission and omission vs. success on previous task.

	Error of Commission	**Error of Omission**	**Previous Task**
Error of commission	1		
Error of omission	-0.1036546	1	
Previous task	-0.1255596	-0.0539929	1

3.6 Discussion

This section describes an initial use of human performance data from a simulator study to investigate the phenomenon of HRA dependency. There are numerous limitations in this study, from the use of students as operators to the complexities of extracting successful vs. failure outcomes for each procedure step. Moreover, the detailed level of data proves somewhat incongruous with the HFE level of analysis. Still, the process of conducting this first-of-a-kind empirical dependency analysis is fruitful. It demonstrates a process to use CBPs to extract success and failure data in a manner that does not require an analysts' subjective categorization of performance. It also provides some preliminary data on the interaction of erroneous task execution in series. If dependency attempts to capture "error begets error," the data suggest that, overall, an initial error leads to increased error rates downstream, but initial errors may not be immediate triggers of subsequent errors.

As noted, this is a first-of-a-kind study. The researchers have already planned three additional studies to gather empirical evidence for dependency:

- A new study using Rancor with specific error seeding to test the effects of inducing errors for dependency using student operators
- The same induced error study with licensed reactor operators at a nuclear power plant
- The use of a nonnuclear task to test error chains over simple repetitive tasks.

The two Rancor studies planned will help compare the error proneness of novice vs. experienced operators, where we expect that errors will be more likely with inexperienced operators. The studies will also establish key differences in error performance across experience levels to determine if students are a suitable surrogate for reactor operators. HRA and dependency studies require considerable data collection due to generally low probability error events under investigation. The ability to conduct empirical dependency studies using more readily available student samples greatly facilitates the data collection process. However, such studies are only meaningful if the results generalize to the target population of reactor operators. It is therefore necessary to benchmark between the two operator groups.

The third study overcomes the challenge of simulators and scenarios and affords the opportunity to collect organic errors. Again, due to the low probability of errors, it becomes necessary to collect large samples or force the errors—a process that may confound any pure effects of dependency due to possible indirect dependency by co-occurent PSFs. Using a simplified task that can be completed quickly across hundreds of trials allows for a fundamental understanding of dependency as a psychological phenomenon. Establishing the first principles of dependency allows subsequent studies to investigate and model additional nuances of dependency to validate mathematical models of the effects on HEPs.

4. SUMMARY

This project investigated an opportunity to enhance PRA modeling and overall PRA efficiency, including improved quantification speed, by enhancements made to the vital PRA task, HRA dependency analysis. Section 2 described deficiencies that exist in the current approach to perform HRA dependency analyses and introduced a novel method that is more efficient and allows to consider scenario context. The proposed approach may be used in PRAs modeled in both CAFTA and SAPHIRE, the main PRA modeling tools used in the U.S.

In addition to improvements proposed for PRA model efficiency, the long-overlooked need to improve the basis of HRA dependency analysis is explored as discussed in Section 3. An empirical study is proposed that holds a promise to collect evidence to directly support better estimations of dependencies between operator actions. This is an important research as it has never been attempted before, yet is very valuable to provide better insights on human performance of multiple actions under demanding conditions.

5. REFERENCES

ASME. (2013). *Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications*, ASME/ANS RA-Sb-2013.

Blackman, H.S., & Boring, R.L. (2017). Assessing dependency in SPAR-H: Some practical considerations. *Advances in Intelligent Systems and Computing, 589,* 13-19. https://doi.org/10.1007/978-3-319-60645-3_2.

Boring, R.L. (2015). A dynamic approach to modeling dependence between human failure events. *Proceedings of the 2015 European Safety and Reliability (ESREL) Conference*, 2845-2851. https://doi.org/10.1201/b19094-373.

Boring, R., Park, J., & Mortenson, T. (2021). Formative vs. summative dependence in human reliability analysis. *Lecture Notes in Networks and Systems, 262,* 62-69. https://doi.org/10.1007/978-3-030-80288-2_8.

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,* 956-963.

Boring, R. L., Ulrich, T. A., & Mortenson, T. J. (2019). Level-of-automation considerations for advanced reactor control rooms. *Proc. 11th Nuclear Plant Instrumentation, Control, and Human-Machine Interface Technologies,* 9-14.

Boring, R., Ulrich, T., Park, J., Ahn, J., & Heo, Y. (2022). The HUNTER dynamic human reliability analysis tool: Overview of the software framework for modeling digital human twins. *Proceedings of the 16th Probabilistic Safety Analysis and Management Conference,* Paper RO173.

Braarud, P. Ø., Eitrheim, M. H. R., & Fernandes, A. (2015). SCORE–an integrated performance measure for control room validation. *Proc. 9th International Topical Meeting on Nuclear Plant Instrumentation, Control, and Human Machine Interface Technologies,* 2217-2228.

EPRI. (2021). *Fire Human Reliability Analysis Guidelines*, EPRI 1023001, NUREG-1921.

Gertman, D., Blackman, H., Marble, J., Byers, J., Smith, C., and O'Reilly, P. (2005). *The SPAR-H Human Reliability Analysis Method, NUREG/CR- 6883.* Washington, DC: U.S. Nuclear Regulatory Commission.

Hwang, S. L., Yau, Y. J., Lin, Y. T., Chen, J. H., Huang, T. H., Yenn, T. C., & Hsu, C. C. (2008). Predicting work performance in nuclear power plants. *Safety science*, *46*(7), 1115-1124. https://doi.org/10.1016/j.ssci.2007.06.005.

IEEE. (2022). *IEEE Guide for Human Factors Applications of Computerized Operating Procedure Systems (COPS) at Nuclear Power Generating Stations and Other Nuclear Facilities, IEEE-1786-2022.* New York: IEEE. https://doi.org/10.1109/IEEESTD.2022.9793353.

Lew, R., Ulrich, T. A. and Boring, R. L. (2021). *Rancor Hybrid Energy System Microworld*. SAGE Publications CA: Los Angeles, CA, 64(1), 1760–1764. https://doi.org/10.1177/1071181320641426.

Ma, Z., Davis, C., Parisi, C., Dailey, R., Wang, J., Zhang, S., Zhang, H., & Corradini, M. (2019). *Evaluation of the Benefits of ATF, FLEX, and Passive Cooling System for an Enhanced Resilient PWR Model, INL/EXT-19-56215.* Idaho National Laboratory, Idaho Falls, Idaho. https://doi.org/10.2172/1777262.

Miller, A., S. Hess, & Smith, C. (2020). *R&D Roadmap to Enhance Industry Legacy Probabilistic Risk Assessment Methods and Tools. INL/EXT-20-59202.* Idaho Falls: Idaho National Laboratory.

Mortenson, T., & Boring, R. (2021). Is dependency in human reliability analysis a real phenomenon? Refining the dependency concept through research. *Lecture Notes in Networks and Systems, 262,* 55-61. https://doi.org/10.1007/978-3-030-80288-2_7.

Paglioni, V. P., Gorth, K., M. (2021). Defining dependency in HRA. *ESREL 2021*.

Parasuraman, R., & Riley, V. (1997). Humans and automation: Use, misuse, disuse, abuse. *Human Factors*, *39*(2), 230-253. https://doi.org/10.1518/001872097778543886.

Park, J., Boring, R.L., Ulrich, T.A., Lew, R., Lee, S., Park, B., & Kim, J. (2022). A framework to collect human reliability analysis data for nuclear power plants using a simplified simulator and student operators, *Reliability Engineering & System Safety, 221,* Article 108326. https://doi.org/10.1016/j.ress.2022.108326.

Park, J., Yang, T., Boring, R.L., Ulrich, T.A., & Kim, J. (2023). Analysis of human performance differences between students and operators when using the Rancor microworld simulator. *Annals of Nuclear Energy, 180,* Article 109502. https://doi.org/10.1016/j.anucene.2022.109502.

Swain, A.D., Guttmann, H.E. (1983). *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications, NUREG/CR-1278, Final Report.* Washington, DC: U.S. Nuclear Regulatory Commission. https://doi.org/10.2172/5752058.

Ulrich, T.A., Boring, R.L., & Lew, R. (2019). On the use of microworlds for an error seeding method to support human error analysis. *Proceedings of IEEE Resilience Week 2019*, 242-246. https://doi.org/10.1109/RWS47064.2019.8971969.

Ulrich, T. A., Lew, R., Hancock, S. G., Westover, T. L., Medema, H. D., Boring, R. L.,... & Minard, N. (2021). *Dynamic Human-in-the-Loop Simulated Nuclear Power Plant Thermal Dispatch System Demonstration and Evaluation Study, INL/EXT-21-64329.* Idaho National Laboratory, Idaho Falls, Idaho. https://www.osti.gov/servlets/purl/1866373.

Ulrich, T.A., Lew, R.T., Werner, S., & Boring, R.L. (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 Annual Meeting, 61*, 398 - 402. https://doi.org/10.1177/1541931213601582.

U.S. NRC. (1975). *The Reactor Safety Study.* Washington, DC: U.S. Nuclear Regulatory Commission. https://doi.org/10.2172/7134133.

U.S. NRC (1995). *Evaluation of the computerized procedures manual II (COPMA II).* Washington, DC: U.S. Nuclear Regulatory Commission. https://doi.org/10.2172/149987.

U.S. NRC. (2005). *Good Practices for Implementing Human Reliability Analysis (HRA)*, *NUREG-1792.* Washington, DC: U.S. Nuclear Regulatory Commission.

Wilke, P. K., Gmelch, W. H., & Lovrich Jr, N. P. (1985). Stress and productivity: Evidence of the inverted U function. *Public Productivity Review*, 342-356. https://doi.org/10.2307/3379944.

Vedros, K., Boring, R., Hess, S., Knudsen, J., Lawrence, S., Mandelli, D., Miller, A., Aoki, N., Park, J., Prescott, S. and Smith, C. (2021). *Enhancement of Industry Legacy Probabilistic Risk Assessment Methods and Tools. INL/EXT-21-64448.* Idaho Falls: Idaho National Laboratory. https://doi.org/10.2172/1822882.