United States Nuclear Regulatory Commission
Protecting People and the Environment

RIL2024-03
PNNL-35267

INTEGRATED HUMAN EVENT ANALYSIS SYSTEM (IDHEAS) – TIME AND TIME UNCERTAINTY

Considerations for Supplemental Guidance for NUREG-2256

March 2024

Prepared by:
G. Coles
B. Jefferson
J. Baweja
R. Prasad

Pacific Northwest National Laboratory
P.O. Box 999
Richland, WA. 99352

Y. James Chang, NRC Project Manager

Research Information Letter
Office of Nuclear Regulatory Research

Disclaimer

Legally binding regulatory requirements are stated only in laws, U.S. Nuclear Regulatory Commission (NRC) regulations, licenses, including technical specifications, or orders; not in Research Information Letters (RILs). A RIL is not regulatory guidance, although NRC's regulatory offices may consider the information in a RIL to determine whether any regulatory actions are warranted.

Abstract

As part of their evaluation, human reliability analysts must often evaluate if crews in nuclear power plants (NPPs) can complete tasks associated with a human-failure event within specified time limits. The time-required for operator response in NPP accident scenarios is normally determined by systematic and structured walk-throughs, feasibility studies, recorded times from training exercises, and interviews with experienced operators and experts. Typically, a point estimate is derived for the estimate (mean, maximum, or 95th percentile) of time-required (depending on the human reliability analysis [HRA] approach employed). Using point-estimate values can mask the risk associated with variability among crews, plant conditions and set-up, environmental conditions, and other impact factors under which these actions are executed. Point estimates for time-required and time-available have served the industry well; however, without considering uncertainty, they could lead to biased understanding about the risk.

The intention of this report is to provide supplemental guidance on a specific aspect of the Integrated Human Event Analysis System (IDHEAS) for Event and Condition Assessment (ECA) HRA method documented in NUREG-2256. The IDHEAS-ECA HRA method is based on the General Methodology of an Integrated Human Event Analysis System (IDHEAS-G) which can be found in NUREG-2198. The method is intended to be used in HRA applications within a probabilistic risk assessment for an NPP or in any safety assessments of an engineering system in which humans have a role. The specific aspect addressed in this report is determining the probability that the time-required to perform an action exceeds the time-available for that action, as determined by the plant success criteria in applicable accident scenarios. To this end, the focus of the material in this report is on the development of pertinent probability distributions.

Table of Contents

List of Figures

List of Tables

ACRONYMS AND ABBREVIATIONS

AIC	Akaike Information Criterion
ASME/ANS	American Society of Mechanical Engineers/American Nuclear Society
ATHEANA	A technique for human event analysis
ECA	Event and condition assessment
EPRI	Electric Power Research Institute
ETP	Equipment and Tools
FSAR	Final safety analysis report
HCR	Human cognitive reliability
HEP	Human error probability
HFE	Human failure event
HRA	Human reliability analysis
HIS	Human-System Interface
IDHEAS	Integrated human event analysis system
INF	Information Availability and Reliability
I&C	Instrumentation and Control
JPM	Job performance measure
LOFW	Loss of feedwater
MF	Mental Fatigue
MT	Multi-Tasking, Interruptions and Distractions
NLL	Negative log-likelihood
NOS	Noise in Workplace and Communication Pathways
NRC	U.S. Nuclear Regulatory Commission
NPP	Nuclear power plants
OECD	Organization for Economic Co-operation and Development
ORE	Operator reliability experiments
PD	Physical Demands
PG	Procedures, Guidelines, and Instructions
PIF	Performance influencing factor
PNNL	Pacific Northwest National Laboratory
PR	Resistance to Physical Movement
PRA	Probabilistic risk assessment
RICT	Risk informed completion time
RIL	Research information letter
SF	Scenario Familiarity
SGTR	Steam generator tube ruptures
SIC	System and Instrumentation and Control Transparency to Personnel
STA	Staffing
TC	Task Complexity
TE	Training

TEP	Cold/ Heat/Humidity
TOF	Team and Organization Factors
TPS	Time Pressure and Stress
VIS	Workplace Visibility
VVER	Water-water energetic reactor
WAH	Work Location Accessibility and Habitability
WP	Work Processes

1 INTRODUCTION

1.1 Background

The intent of this report is to provide supplemental guidance on a specific aspect of the Integrated Human Event Analysis System (IDHEAS) for Event and Condition Assessment (ECA) human reliability analysis (HRA) method documented in NUREG-2256 [1]. The IDHEAS-ECA HRA method is based on the General Methodology of an Integrated Human Event Analysis System (IDHEAS-G) that can be found in NUREG-2198 [2]. The method is intended to be used in HRA applications within a probabilistic risk assessment (PRA) for a nuclear power plant (NPP) or in any safety assessments of an engineering system where humans have a role. The specific aspect addressed in this report is determining the probability that the time-required to perform an action exceeds the time-available for that action, as determined by the plant success criteria in applicable accident scenarios. To this end, the focus of the material in this report is on the development of pertinent probability distributions.

The IDHEAS-ECA HRA method was developed to support risk-informed applications in which the PRA model used to support regulatory decision-making is intended to be consistent with Regulatory Guides 1.200 Revision 2 [3] and Revision 3 [5] and Regulatory Guide 1.247 [4]. IDHEAS-ECA supports PRA and safety assessment applications by analyzing human events and estimating human error probabilities (HEPs). The application scope of IDHEAS-ECA is broad because the performance-influencing factor (PIF) structure, which models the context of a human failure event (HFE), is comprehensive. The method nominally covers all the PIFs identified in light water reactor PRA standards [6] and non-light water reactor PRA standards [7], existing HRA methods for assessing nuclear facility safety, and the factors reported in the broad literature and nuclear-specific human events.

IDHEAS-ECA is cognition centered, technology neutral, and applicable to all U.S. Nuclear Regulatory Commission (NRC) HRA applications; for example, PRA, integrated safety analysis, spent fuel handling, nuclear material users, and nuclear medicine. For PRA applications, the scope includes:

- Level 1 and Level 2 PRA
- Internal and external hazards
- At-power, low power, and shutdown operations
- Conventional (analog) and digital control rooms
- Control room and field actions
- Actions with installed components and portable equipment
- Base (or baseline) PRA development
- License amendment request reviews
- Significance determination process (SDP) evaluations
- Accident Sequence Precursor (ASP) program
- Pre-initiator, at-initiator, and post-initiator human failure events (operator actions)

This report provides information intended to be supplemental to the guidance provided in NUREG-2256 [1]. However, the information provided could benefit from further investigation or consideration as identified in the report.

1.2 Purpose and Scope of the Supplement

The two main concerns for evaluating NPP operator performance are committing a critical cognitive error and not completing a task within the time limit available. These two aspects of human error form the mathematical framework for IDHEAS-G [2]. In IDHEAS-G, the HEP, P, for an HFE is defined as:

$$P = 1 - (1 - Pc)(1 - Pt)$$

In the equation, P represents the HEP, P_c represents the error probability attributed to the cognitive failure modes of all critical tasks, and P_t represents the probability that the time needed (or time-required) exceeds the time-available. This supplement focuses on characterizing P_t.

The purpose of the information presented in this report is intended to augment and clarify guidance presented in the front-end of Section 3.6 of NUREG-2256. This report aims to provide further practical guidance to supplement the largely theoretical guidance provided in NUREG-2256 about determining the probability that an operator exceeds the time-available to perform a task. Section 3.6 of NUREG-2256 addresses estimation of P_t, which is defined in Section 3.5 of the regulation as the “probability that the time-required to perform an action exceeds the time-available for that action, as determined by the success criteria.” NUREG-2256 refers to this as the convolution of time-available and time-required, and, by definition, necessitates developing probability distributions. See Section 2.3 of this report for more detail.

In IDHEAS-ECA, the time-required is defined as the time taken for the actions associated with an HFE to be completed, including time taken for information detection, diagnosis, decision-making, execution, and inter-team coordination [1]. Time-available is defined as the duration from the onset of cues indicating that an action is needed to the time beyond which the action is no longer useful in mitigating the event consequence [1]. Figure 1-1. shows the timeline from the start of an NPP event at time T_0 and its progression as the crew responds to the event. Time-required terms are defined below:

T_{delay} denotes the duration of time from the start of the event until the cue is available to be received by the crew.

T_{sw} denotes the system time window, the duration from the start of an event until crew action is no longer beneficial to mitigate the adverse effect of the event.

T_{avail} is the time-available to the crew to respond to the event and is the difference between T_{sw} and T_{delay}.

T_{cog} denotes the cognition time of the crew to detect, diagnose, and make the decision to take an action.

T_{exe} is the crew’s action execution time, including the time for travel, tool collection, personal protective equipment donning, and manipulation of equipment.

T_{reqd} is the time-required which us the sum of T_{cog} and T_{exe}, denotes the time needed by the crew to accomplish the action; it also is called the crew response time.

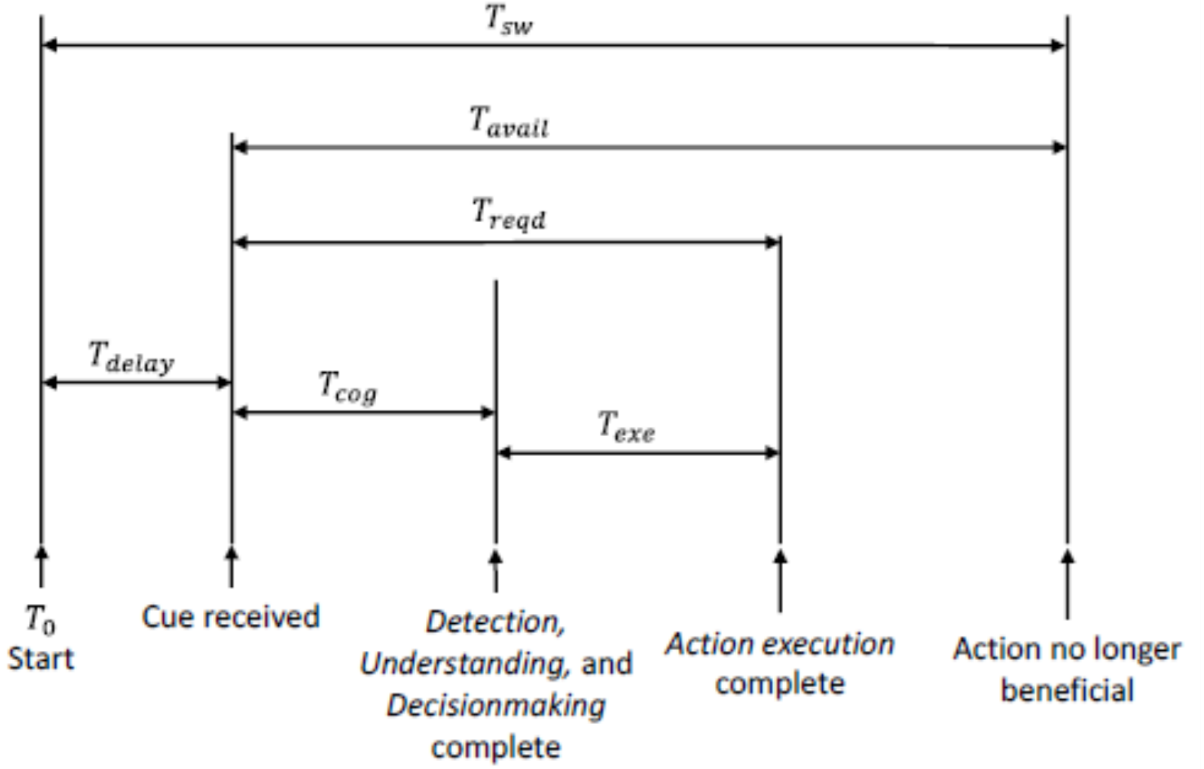
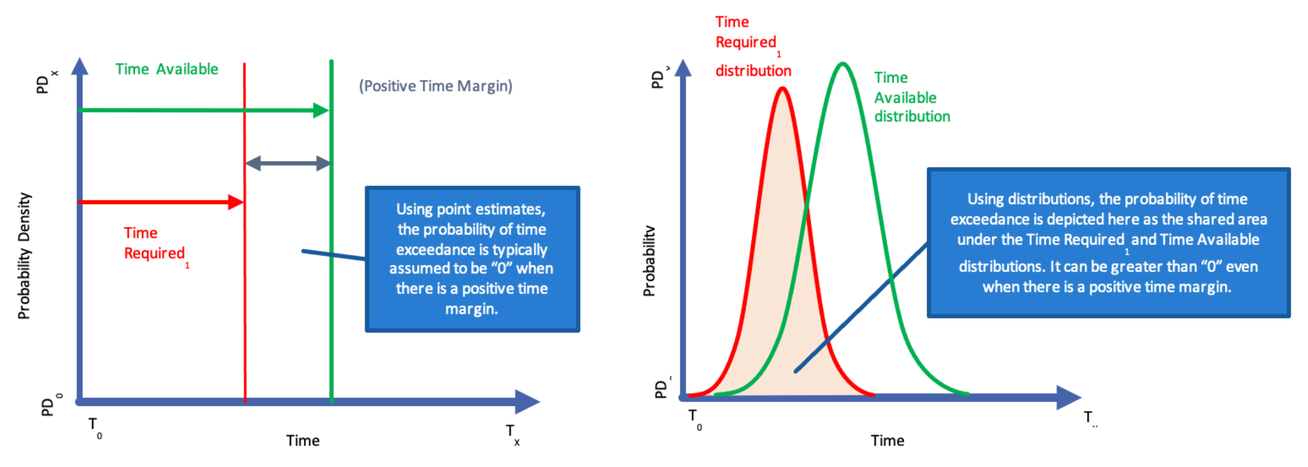


Figure 1-1. Timeline Used in Assessing an HFE (Source: NUREG-2256 [1])

Pacific Northwest National Laboratory (PNNL) undertook a task to better quantify the variability in the crew time-required (T_{reqd}), also referred to here as completion time. Colloquially, time-required represents the time taken by crews to complete tasks associated with resolving abnormal scenarios at NPPs. The focus of this report is in providing methodologies for estimating probability distributions of time-required for HFEs. The IDHEAS-G framework posits that the specific conditions of the scenario context might hinder or enhance human performance (including time-required), thus affecting the HEP. The framework uses PIFs to model the specific conditions associated with a scenario context [2]. PIFs are conceptually distinct from time uncertainty factors, in that PIFs impact cognitive processing, while time uncertainty factors explicitly impact time.

The typical approach used by HRA analysts to evaluate if crews can complete tasks associated with an HFE within time limits is based on point estimates of the time-required and time-available. HRA analysts base point estimate values on the best available information, such as times recorded from talk-throughs or walk-throughs of the applicable procedure, simulator observations of training exercises, or feasibility studies. Normally, these times are not supported by statistical data, although they may be in some cases. For example, average completion times (time-required) could potentially be developed from Job Performance Measure (JPM) testing data, if available. JPM data show the time-required to complete a given action for which operators are trained and tested. If data are compiled on how long it takes for crews to complete certain tasks, then it may be possible to determine a probability distribution and calculate mean values.

Use of point estimate values can mask performance variability among crews, the impact caused by characteristics of the NPPs at which operator actions occur, the different conditions under which they are executed, and the type of action being performed. Use of point estimates can oversimplify the impact of the environment, plant, task, and crew on completion time and obfuscate their impact on safety and risk. For example, when a mean point estimate of time-required is used and it is less than the time-available as illustrated (Figure 1-2(a)) for $TimeRequired_1$, then it is typically assumed that the probability of time exceedance is "0" because there is a positive time margin.



(a) Using Point Estimates

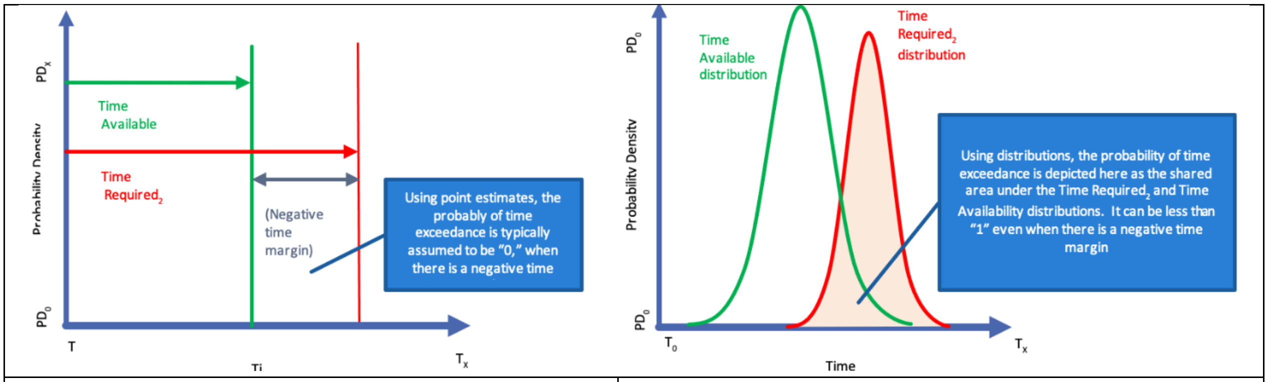
(b) Using Distributions

Figure 1-2. Two Methods to Determine Time Exceedance When $T_{reqd} \le T_{avail}$

However, if time exceedance is determined using probability distributions for time-available and time-required then the probability of time exceedance can be a non-zero (and non-trivial) value. Figure 1-2(b) also shows that when probability distributions are used for the same case (i.e., the point estimate of time-required is less than the point estimate of time-available), there is a reasonably critical probability that time-required exceeds time-available as depicted by the shared area under the $TimeRequired_1$ and *TimeAvailable* probability distribution curves in Figure 1-2(b). If only point estimates are used to determine the probability of exceeding time-available, then the probability of exceedance could be systemically underestimated.

Conversely, if a mean completion time is used for time-required, and it is more than the time-available as shown for $TimeRequired_2$ in Figure 1-3(a), then in this case it is typically assumed that the probability of time exceedance is "1" because there is negative time margin. However, this is a conservative simplification.

For cases in which the point estimate of time-required is greater than the point estimate of time-available, Figure 1-3(b) shows that, using time-required and time-available probability distribution curves, there is a probability that time-required does not exceed time-available. The probability of time non-exceedance is depicted by the shared area under the $TimeRequired_2$ and *TimeAvailable* probability distribution curves. Accordingly, the assumption made that the probability of exceedance is "1" when the time margin is negative (i.e., the time-required is greater than the time-available) can be conservative.



Using point estimates, the probably of time exceedance is typically assumed to be "0," when there is a negative time

Using distributions, the probability of time exceedance is depicted here as the shared area under the Time Required$_2$ and Time Availability distributions. It can be less than "1" even when there is a negative time margin

(a) Using Point Estimates

(b) Using Distributions

Figure 1-3. Two Methods to Determine Time Exceedance When Treqd > Tavail

The primary benefit of using a distribution over point estimates is that analysts gain a more informed and complete perspective of variance in performance and can provide a more accurate estimate of probability of exceedance of time-available. However, while distributions are a goal, guidance is needed to construct them accurately. To supplement the guidance provided in NUREG-2256, this report provides details on the different approaches to building distributions and when each approach is appropriate. Table 1-1 provides a high-level mapping of estimation method with data and resource availability.

Table 1-1. When to Use Distribution Approximation Techniques

Method	Operation Expertise Required	Data Availability
Single Point Estimate	High	Low
Five Point Estimate	High	Low
The PNNL Method	Medium	Low–Medium
Full Distribution Approximation	Low	High

The IDHEAS methodology addresses the illustrated limitations associated with using point estimates to characterize P_t by using probability distributions of time-available and time-required. The primary objective of this report is to provide supplemental information to the information provided in NUREG-2256 that 1) supports estimation of time-required and time-available for HRA and 2) provides further guidance to HRA analysts to inform these estimations.

Section 2 of this report will discuss when point estimates should be used and the usual procedure used to derive them. Also, this section will cover 5-point estimate methods. A note about time-available distributions is shared in Section 3. Section 4 discusses the ideal approximation for a probability distribution of time-required and the limitations with achieving this in practice. In Section 4.2, the PNNL method for producing a probability distribution for time-required is detailed. This method is conceptually a compromise between traditional distribution approximations and point estimates. The final section, Section 4.3, addresses adjusting distributions to account for PIFs and other factors that may impact the shape of the distribution.

Supplemental information is provided in four appendices: Appendix A – Data Selection for the Required Evaluation; Appendix B – PIF Impacts and Distribution Adjustments; Appendix C – Expert Elicitation on Methods; and Appendix D – Statistical Derivation of Time Required.

2 USE OF PROBABILITY DISTRIBUTION VERSUS POINT ESTIMATES

Although NUREG-2256 promotes the use of time-required and time-available probability distributions as an important part of the IDHEAS methodology, it is acknowledged that there are cases for which the use of point estimate values of time-required and time-available is sufficient. There are also cases for which the use of point estimate values is clearly insufficient, and cases for which it is not clear whether point estimate values are sufficient or not. There may also be cases for which an "in between" method may be appropriate. This section (1) addresses the question of when point estimates may be sufficient and when propagation of probability distributions is needed, (2) describes an approach for when a full distribution is not required and yet point estimates are not sufficient, and (3) provides general statistical information about development of probability distribution as a foundation for later discussion for cases in which probability distribution are needed.

2.1 Single Point Estimates

The typical approach used by HRA analysts to evaluate if crews can complete tasks associated with an HFE within the time-available is based on point estimates of the time-required (e.g., mean, median, or maximum time-required). HRA analysts base their point-estimate values on the best available information, such as times recorded from talk-throughs or walk-throughs of the applicable procedure, simulator observations of training exercises, or feasibility studies. As stated in Section 1.2, use of point estimate values can mask (1) the variability among crews in the time it takes to complete an action, (2) the characteristics of the NPPs where operator actions occur, and (3) the different conditions under which these actions are executed. Using point estimates, the probability of exceeding time-available is often set to a value of 1.0, if the point of time-available is less than time-required depending on the HRA approach used (i.e., methods such as Human Cognitive Reliability (HCR)/Operator Reliability Experiments (ORE) do not depend on this determination because the probability of failure is correlated with the time-required only). Accordingly, the development of probability distributions is encouraged in NUREG-2256 to determine P_t. This precision can be important in a risk-informed application if the operator error can have a meaningful impact on the risk-informed conclusions. Additionally, using point estimates may result in conservatively setting the probability of exceeding time-available (P_t) to a value of 1.0, which could skew the results from a PRA supporting a risk-informed application. In a risk-informed application, it is important to identify and focus on the true risk-significant contributors.

However, even though there are clear advantages to developing probability distributions for determining P_t (e.g., precision), there may be compelling reasons not to devote significant resources to this effort, as noted by HRA experts. To generate insights about determining P_t, an expert knowledge elicitation process was used based on a workshop and email solicitations to gather information from experts based on their experiences, lessons learned, insights gained, suppositions and mental model. Appendix C describes the expert elicitation process that was designed to produce insights on ways to determine P_t. This primarily involved asking experts detailed questions about how to determine time-required and time-available and their corresponding probability distributions. The workshop itself was devised to reduce social bias by employing principles presented in an NRC white paper titled *Practical Insights and Lessons Learned in Implementing Expert Elicitation* authored by Xing and Morrow [8]. A detailed discussion of the observations by the experts on the use of probability distributions versus point estimates is summarized in Appendix C.4.

Following are summaries of insights from the knowledge elicitation on when it may be sufficient to use point estimates of time-required and time-available to set opposed to probability distributions.

- Only a subset of HFEs modeled in a PRA are typically important to risk; therefore, modeling the uncertainty of time-available and time-required is not needed for all HFEs. For screened-in HEPs, the importance of an HEP might be gauged using importance measures (e.g., Fusel-Vesely or Risk Achievement Worth values) as determined by the PRA. However, the challenge of this approach is that importance measures cannot be determined until the PRA is complete.
- In many instances, the time-available exceeds the time-required by a significant margin, so the probability of exceeding the time-available is likely negligible. The challenge in this case is knowing how much time margin is enough to render the P_t negligible. Quantitative assessment is one way to estimate how much time margin between time-required and time-available is needed to exclude determining the probability of exceeding the time-available using probability distributions of time.
- Additionally, even though a particular HFE may have a certain level of risk significance, the HFE may not be important to the risk-informed application that the PRA supports. Also, for risk-informed programs, such as the Risk Informed Completion Time program, it may be possible address key uncertainties using measures allowed by the program such as Risk Management Actions to prevent or mitigate the risk.
- Finally, in some cases, NPPs already use the HCR/ORE HRA approach to calculate non-response probability based on time-required.

Accordingly, the expert elicitation suggested that time-required and time-available probability distributions should be used if:

1. The associated HFE is not determined using the HCR/ORE HRA method which already accounts for time related non-response.

   AND

2. The risk importance of the associated HFE is not known or is known to impact the risk results so that risk-informed decisions could be impacted.

   AND

3. It cannot be conclusively judged that the time margin between time-available and time-required renders the probability of exceeding the time-available as negligible.

   AND

4. The associated HFE might be important to the risk-informed application or risk-informed decision that is being supported by the PRA or HRA, or if the HFE is estimated incorrectly, it could skew the risk conclusions.

A final note about single-point estimates is that one may use the estimate as a parameter for a suspected distribution form. For example, if the time-required is suspected of being lognormal, the single-point estimate could serve as the median parameter for the lognormal distribution. This approach requires some thought about why the distributional form was selected and how to estimate other distribution parameters accurately. We revisit this in Section 4.2.

2.2 Five-Point Estimate

In scenarios in which a full distribution is not required and yet point estimates are not sufficient, a five-point estimate provides a useful alternative. NUREG-2256 [1] states the following:

> *With respect to the five-point estimation of probability distribution, if operational data are not adequate for confident estimation of the parameters of an assumed parametric probability distribution, or if evidence suggests that a parametric distribution is not appropriate for the situation (for example, the personnel modeled fall into two distinctive groups), HRA analysts can estimate five points of the time distribution at the 5th, 25th, 50th, 75th, and 95th percentiles. The IDHEAS-ECA software interpolates the full distribution based on the five-point estimates.*

The five-point time distribution can be estimated based on the same best available information used for single point estimates but requires a more nuanced view of the data. NUREG-2256 also offers some guidance in thinking about these estimates. Although in the context of the "range" of distributions, NUREG-2256 ask two questions:

- What percentage (or fraction) of crews would perform the action by the lower estimate of the T_{reqd} range? (i.e., the 'faster' crews)?
- What percentage (or fraction) of crews would perform the action by the higher estimate of the T_{reqd} range? (i.e., the 'not-so-fast' crews in addition to the 'faster' crews)?

A slight modification to these questions can be adopted for each percentile:

- At what time would [*percentile*] percentage (or fraction) of crews perform the action?

We use the example from NUREG-2256:

> *… assume a hypothetical scenario in which the time to detect, understand, decide, and execute the use of a portable generator was determined to be 30 minutes. Next, by considering the staff experience in Table 3-5, this 30-minute estimate could be 25 minutes if highly experienced staff perform the action or 40 minutes if less experienced staff perform the action. The parameters of the T_{reqd} distribution can be based on the 25-to-40-minute or 30-to-40-minute T_{reqd} range.*

One may then ask, "At what time would 5% of crews perform the action?" This may be 17 minutes, while the 25th percentile may be 20 minutes, and so on. This five-point method (and the single-point-estimate method) requires knowledge of plant operations, conditions, and other factors for accurate estimates. The additional benefit of the five-point estimate over the single point estimate is that a distribution can be estimated with fewer (or no) distribution shape assumptions (i.e., Gaussian vs. Weibull). NUREG-2256 highlights this:

> *With respect to the five-point estimation of a probability distribution, often the time-available for an action does not fall into a parameterized probability distribution. HRA analysts can estimate five points of the time distribution at the 5th, 25th, 50th, 75th, and 95th percentiles. The IDHEAS-ECA software interpolates the full distribution based on these estimates using the step function. In the IDHEAS-ECA software, the probability density functions of between 0 to 5th percentile is specified as a half of the probability density function of between 5th and 25th percentiles, and the probability density function between 95th and 100th percentile is specified as a half of the probability density functions of between 75th and 95th percentiles.*

This provides the second coarsest level of granularity to estimating an accurate distribution, with the coarsest being to make several assumptions regarding the distribution shape and form and using the single point estimate. The benefit of this method (over the single point method) is that one does not need to make assumptions about the distribution shape and form. The estimated percentiles provide a basis for deriving a very basic view of the distribution shape. Again, see the guidance provided above for single point estimates for when this method is appropriate.

2.3 Statistical Development

This section provides general statistical information about development of probability distributions as a foundation for later discussion for cases in which probability distributions are needed. Section 2.3 provides a general discussion of statistical development needed to support determination of the probability of time exceedance and specifically discusses selecting a distributional form.

P_t uses the time-available (T_{avail}) and time-required (T_{reqd}) to perform an action. To calculate P_t, T_{reqd} is represented by its cumulative distribution function $F_{Treqd}(t)$, and T_{avail} is represented by its probability density function $f_{Tavail}(t)$ and P_t is estimated as the convolution of the two probability distributions, that is

$$P_t = P\left(T_{reqd} \geq T_{avail}\right) = \int_0^{\infty} \left(1 - F_{T_{reqd}}(t)\right) \cdot f_{T_{avail}}(t)dt$$

IDHEAS-ECA requires the estimation and justification of the probability distributions for T_{avail} and T_{reqd}. The IDHEAS-ECA software or any general-purpose computation software can then use the formulas presented in reference [1] to calculate P_t. In the previous sections, the five-point estimate allows for a distribution to be estimated via interpolation over the five points to determine a cumulative distribution function. In general, there are parametric and non-parametric approaches to developing a distribution. Parametric approaches require an analyst to choose a distribution form (e.g., lognormal, Weibull, etc.) based on some reasonable justification and then determine the parameters of that distribution (i.e., the central tendency, the variance, or a shape parameter). Non-parametric approaches use interpolation methods that directly create a probability density function with some analyst-specified smoothness constraints. The later approach requires some familiarity with statistical model fitting and is not within the scope for this report. NUREG-2256 describes parametric approaches as follows:

> *The observed completion times for most crews are typically clustered around a central value (i.e., the median response time). However, it is often the case that the times for a small number of crews deviate substantially from that behavior. In particular, a small number of crews often need much more time to complete the desired action. There are many reasons for these deviations (i.e., not only differences in training), and they often depend on the context of the specific response scenario. The shape and the range of the distribution for T_{reqd} should account for this observed behavior. Thus, it is often appropriate to characterize the uncertainty in T_{reqd} with a skewed distribution, such gamma, Weibull, or lognormal. It is important for the shape and the range of the uncertainty distribution to account for the analysts' consideration of these "outlier" effects. The quantification results for P_t can be affected significantly by the "overlap" in the low-probability "tails" of the distributions for T_{avail} and T_{reqd}.*

When sufficient data are available for estimating a distribution, analysts should identify a distribution form based on properties of the data. As NUREG-2256 points out, the skew associated with most completion time data along with unimodality suggest a few choice distributions. A maximum likelihood estimation procedure is typical for determining distribution parameter values. The negative log-likelihood (*NLL*) is computed for a candidate density function, *f*, and candidate parameter values, *θ*.

$$NLL = -\sum_{x \ni \mathrm{Data}} \log f(x|\theta)$$

Available data are tested with a variety of parameter values until a minimum *NLL* is produced. Among the distribution candidates, the distribution providing the minimum NLL is deemed to be the best fitting distribution. This approach requires data that is representative of the tasks, environments, and scenarios that crews face in order to be an accurate representation of time-required. The accuracy of the recommendation is completely determined by the quality of data available.

Four distributions are commonly considered based on guidance about the general shape of time-required data in nuclear power plant environments (a distribution with a mass of lower values and a long tail, where most crews complete the task well-under previously specified time frames): lognormal, Weibull, Exponential, and a truncated Normal distribution. Each of these distributions have the desired long tail and a large mass closer to smaller times. There are some key differences, however. The exponential distribution has density that increases as $t \to 0$. Completion times related to human performance will not have this property because humans typically are subject to a lower bound on how fast time-required can be. However, without exact knowledge of that lower bound and no constraints on sample size, exponential distributions may technically fit the data available. The lognormal distribution classically captures the lower bound of human performance in time-required. The distribution has three parameters that impact its shape (how different the tail appears), location (where the distribution is centered), and scale (the number of values covered by the majority of the distribution). The Weibull distribution is a generalization of the exponential distribution and can also be interpreted as a distribution of failure times. As such, the Weibull distribution provides a reasonable candidate for time-required. Lastly, there are some applications where the tail of a time distribution is not as long as the named distributions, and the distribution is much closer to a normal distribution. Thus, the truncated normal distribution is also commonly considered.

3 DEVELOPING A PROBABILITY DISTRIBUTION FOR TIME AVAILABLE

Supplemental information presented in this report is intended to augment and clarify guidance presented in Section 3.6 of NUREG-2256 on developing a probability distribution for time-available. The discussion in NUREG-2256 Section 3.6 on time-available states that an estimated range may be based on (1) several thermal-hydraulic code runs, (2) first-principle engineering calculations, and/or (3) values provided by experts. HRA analysts may be required to base the point estimate of time-available on appropriately realistic thermal-hydraulic analysis or simulations in support of risk-informed applications. Section 3.1 discusses ways to estimate time-available and Section 3.2 discusses developing a corresponding probability distribution.

Although IDHEAS-ECA does not limit the probability distributions that may be used to calculate P_t, the IDHEAS-ECA software (version 1.2) offers six options to represent T_{avail}: (1) normal distribution,[1] (2) lognormal distribution, (3) gamma distribution, (4) Weibull distribution, (5) five-point estimation of a probability distribution, and (6) single-value threshold.

3.1 Ways to Estimate Time Available

Section 3.6 of NUREG-2256 titled “Guidance for Estimating the Distribution Parameters of Time Available” includes identification of and limited discussion about ways to estimate time-available. This Research Information Letter (RIL) provides supplemental information from an expert knowledge elicitation that was designed to produce insights on ways to determine the probability the time-available to perform a task is exceeded. The material presented here is described in detail in Appendix C. The elicitation focused on determining way to estimate the time-required and time-available and their corresponding probability distributions. The list below provides the approaches to identify time-available identified by the workshop and/or non-workshop experts in the order of preference. Although this list is a robust identification of ways to estimate time-available, it should be noted that to meet the requirements for risk-informed applications supported by PRA, the estimate of time-available should meet the requirements of the PRA standard (American Society of Mechanical Engineers/American Nuclear Society [ASME/ANS]) PRA standard [6] endorsed by NRC through Regulatory Guide 1.200, Revision 2 [3] and Revision 3 [5]. The PRA standard requires that time-available estimates be based “on appropriately realistic generic thermal/hydraulic analysis or simulation from similar plants (e.g., plant of similar design and operability)” for Capability Category (CC), such as:

1. Thermal-hydraulic analysis performed for the Final Safety Analysis Report (FSAR) in support of design basis success criteria or Modular Accident Analysis Program runs in support of PRA success criteria
2. Design basis information other than the FSAR
3. Calculations based on engineering first principles such as determining how long would it take the water in a tank to move from one level to another given a volumetric flow rate

---

[1] Special caution should be taken when the probability distributions of T_{avail} and T_{reqd} are assumed to be normal (Gaussian): “Since a normally distributed [random variable] can take on a value from the $(-\infty,+\infty)$ range, it has limited applications in reliability problems that involve time-to-failure estimations because time cannot take on negative values. However, for cases in which the mean μ is positive and is larger than σ [i.e., the standard deviation] by several folds, the probability that the [random variable] T takes negative values can be negligible. For cases where the probability that [random variable] T takes negative values is not negligible, the respective truncated normal distribution can be used.” [1]

4. Other engineering modeling that supports the success criteria used in a PRA such as room heat-up calculations and internal flood height calculations
5. The time-available stipulated in the procedure (if provided)
6. Plant and NRC Technical Training Center simulator runs (i.e., time-available assumed in accident simulation)
7. Use of expert elicitation with an interdisciplinary team including plant licensed operators and trainers, maintenance personnel, engineers, and experts in the PRA and HRA at the plant
8. Operator experience about timing used in training (i.e., time-available assumed in accident simulation)
9. Delay time until a cue is received by the operators that an action is required can come from indicators, annunciators, alarms, or from procedural steps and might be impeded depending on the situation

3.2 Developing a Probability Distribution for Time Available

NUREG-2256 acknowledges that running enough thermal-hydraulic code calculations to develop a probability distribution is typically not feasible. This can also be true when using engineering calculations to develop a probability distribution because of the large number of runs it takes to generate a distribution. As stated above, Appendix C of this RIL provides supplemental information from an expert knowledge elicitation that was designed to produce insights on ways to determine the probability the time-available to perform a task is exceeded. The elicitation focused on determining a way to estimate the time-required and time-available and their corresponding probability distributions. The primary insight from the expert knowledge elicitation was to acknowledge the sources modeling uncertainty to developing a probability distribution cited but pointed out that there are other important of uncertainty that must be considered:

1. Uncertainty associated with thermal-hydraulic analysis inputs
2. Uncertainty associated with available margin until equipment fails or in case of failed equipment
3. Process of using representative results to address the large number of variations in HFEs across scenarios, and
4. Uncertainty associated determining the delay time until a cue is received by the operators indicating an action is needed.
5. Uncertainty resulting from unmodeled operator intervention that results in increased or decreased available time margin (identified after the expert elicitation was performed)

It might be useful to consider these sources of uncertainty when using the guidance in NUREG-2256. However, the experts suggested that to make meaningful progress on characterizing the probability distribution of time-available, a study is needed that includes expert elicitation and addresses these sources of modeling uncertainty.

4 ESTIMATING A FULL DISTRIBUTION FOR TIME REQUIRED

4.1 Practical Recommendations for PRA Estimates

Section 3.6 0f NUREG-2256, “Guidance for Estimating the Distribution Parameters of Time Required,” describes ways to estimate a probability distribution but refers to ways to estimate time-required. This section of the RIL provides supplemental information from an expert knowledge elicitation that was designed to produce insights on ways to determine the probability the time-available to perform a task is exceeded, which is described in Appendix C of this report. The elicitation focused on determining ways to estimate the time-required and time-available and their corresponding probability distributions. The cited subsection, Section 3.6 of NUREG-2256, refers to review of operational and simulator data and interviews with operators. This is not inconsistent with CC-II requirements of the PRA standard (ASME/ANS) PRA standard [6] endorsed by NRC through Regulatory Guide 1.200, Revision 2 [3] and Revision 3 [5] that stipulate estimates be based on walk-throughs or talk-throughs of the procedures or simulator observations for significant HFEs. Moreover, as discussed in Appendix C.3, the PRA standards are moving towards requiring actual plant or simulator performance verification. The list below provides the approaches to identify time-required identified by the workshop and/or non-workshop experts that supplement the guidance in NUREG-2256, Section 3.6 on time-required:

1. For risk significant HFE events, walk-throughs or talk-throughs of the procedures with operators or trainers and use of simulator observations are the primary ways used to estimate the time-required to perform an operator action and meets the PRA ASME/ANS standard.
2. For some plants, a good source of simulator observations are the results of applicable JPM tests. This information is plant- and crew-specific, and in addition to helping estimate nominal time-required, it could possibly be used to develop probability distribution depending on extent of the records and level of detail (e.g., the probability distribution of time-required for complex tasks could be compared to simple tasks.)
3. If compliance with CC-II of the PRA ASME/ANS standard is not required, then an approach that might be used in certain cases is an analyst’s use of their own experience (or the experience of a trusted source) to establish benchmark times. These times can then be adjusted to match the HFE being evaluated, such as adjusting the time for number of steps or the additional complexity in the scenarios. This approach might be used in HRAs for which the risk results are not sensitive to uncertainty in the time-required or when the time margin between time-available and time-required is large.
4. Caution should be taken when using the “rule of thumb” that estimates time-required by assuming each procedural step takes 1 minute. This approach, even when used as an approximation, can overestimate time-required in some cases and underestimate in other cases. The experts noted that (1) there can be significantly more time uncertainty associated with detecting, understanding, and deciding on a course of action than there is with executing a physical action; (2) it is not clear how to define a single step, as a numbered procedural step can have many subparts of varying degrees of difficulty; and (3) the source of the rule-of-thumb is not clear.
5. When evaluating performance impacts for time-required, the uncertainty associated with activities that occur in the plant besides those specifically associated with safe shutdown should be considering, including routine tasks and actions needed to avoid equipment loss or damage or actions needed to avoid safety concerns not related to nuclear safety

(e.g., radiation safety or occupational safety). These uncertainties and activities can divert an operator's attention, affect a scenario, or affect the operator's understanding of a scenario. If using the second approach to Option 1 described in that NUREG-2256 using Table 3.5, the factor described above is related to the second "Uncertainty Factor" – "Plant Conditions" to determine time-required estimates. If using the PNNL approach described in Section 4.2 of this report, it is related to the PIF Multitasking, Interruptions and Distractions.

6. As discussed in NUREG-2256, the experts stated that when reviewing operational and simulator data the average, slowest and fastest times can be obtained, and a distribution estimated by assigning percentiles.

Supplemental information presented on developing a probability distribution in this report is intended to clarify guidance presented in Section 3.6 of NUREG-2256 for time-required distributions. The discussion in NUREG-2256, Section 3.6 identifies ways to estimate time-required and presents options for determining time-required probability distributions, including reference to an approach developed by PNNL. Section 2.3 discusses a common statistical approach to developing a probability distribution. Section 4.2 discusses the PNNL approach to develop a probability distribution for time-required. Other approaches for developing a probability distribution for time-required described in NUREG-2256 are not supplemented.

NUREG-2256 describes time requirements as follows:

> *The time-required to perform the action (T_{reqd}) should account for entire time that is needed to achieve the desired plant conditions. Estimates of T_{reqd} should not account only for the time that is needed to initiate the desired action (e.g., to open a valve, start a pump, etc.). In particular, T_{reqd} includes the subsequent time that is needed to achieve the plant conditions that determine the functional success criteria for the modeled action. For example, the success criteria may require that the operators must cool down and reduce pressure below a certain value. After the decision is made, the total execution time is the time that is needed to manipulate the relevant controls to begin the cooldown, plus the time that is needed to achieve the desired temperature and pressure, as determined by allowable cooldown rates, scenario-specific thermal-hydraulic response, etc. That time is typically much longer than the time that is needed to initiate the cooldown. It is also affected by scenario-specific limitations such as the number of available cooling water trains, pressure relief valves, etc. That total execution time determines whether the functional success criteria are achieved within the available time window, and it should be included in the estimate for T_{reqd}.*
>
> *The time-required to complete an action can be affected by many factors. Estimating the distribution of T_{reqd} should consider three key aspects: nominal contributors, uncertainty factors, and bias factors. HRA analysts should keep in mind the HFE definition (see Appendix C) when estimating the distribution parameters of T_{reqd} for which the NRC staff proposes two options.*

NUREG-2256 refers to two options for estimating statistical distributions for time-required. The first option is a two-step approach wherein analysts estimate the 5th and 95th percentiles for time-required using information gathered during walk-throughs and talk-throughs and use the information to calculate the parameters of a normal, lognormal, Weibull, or gamma distributions, NUREG-2256 Appendix C provides mathematical details. Likewise, a confidence can be provided that reflects the judged confidence that the range captures the bulk of time-required.

The second option is PNNL's proposed method referred to in NUREG-2256 and detailed in Section 4.2 of this RIL.

PNNL's proposed method is summarized as follows:

- *Given a point estimate for an operator action T_{reqd}:*
  - *Set the point estimate as the scale parameter (median) of the lognormal distribution*
  - *Use a value between 0.28 and 0.54 as the shape parameter of the lognormal distribution. More conservative values closer to 0.54.*
- *Given a conservative (i.e., 95th percentile) estimate for an operator action T_{reqd}:*
  - Set (95th percentile)/1.585 as the scale parameter (median) of the lognormal distribution
  - Use a value between 0.28 and 0.54 as the shape parameter of the lognormal distribution. More conservative values closer to 0.54

  The shape parameter range of 0.28 to 0.54 was derived from operator actions performed inside the control room of NPPs (i.e., in-control-room actions). For operator actions outside the control room of an NPP (i.e., ex-control-room actions), the range of shape parameter values, [0.28, 0.54] may be used as well. However, P_t results for ex-control-room actions may be too optimistic because the T_{reqd} variability (represented by the shape parameter) of ex-control-room actions could be greater than the T_{reqd} variability of in-control-room actions.

4.2 Developing a Probability Distribution for Time Required Using the PNNL Method

Regarding guidance for estimating the probability distribution parameters of time-required, NUREG-2256 states that PNNL proposed a method to develop the probability distribution of T_{reqd} based on a data analysis of operator action completion times time-required from Electric Power Research Institute (EPRI) NP-6937-L, "Operator Reliability Experiments Using Power Plant Simulators" [9]-[11]. Instructions in NUREG-2256, Section 3.6, state that a lognormal distribution with a 0.28 shape parameter could be applied based on a point value median or 95th percentile estimate of T_{reqd}. However, at the time NUREG-2256 was published, the PNNL method for estimating a distribution for time-required was not complete. Use of values between 0.28 and 0.54 as the shape factor (which corresponds to an error factor between 1.58 and 2.42) is just the first step in a more complete two-step overall approach that is described in this section. Note that the guidance in Section 4.2.2 as supported by Appendix D of this RIL supersedes the referenced PNNL-32384 report in NUREG-2256, Section 3.6[1].

Section 4.2.1 provides an overview on developing a time-required probability distribution using the PNNL method, called a "first-order" distribution. Section 4.2.2 explains how the "first-order" distribution for an HFE can be developed for a specific kind of PIF impact that is treated separately from the impact of other PIFs. Section 4.3 describes how to adjust of the first-order distribution using available plant data, operator interviews, and applicable literature for dominant PIFs to create the final distribution of time requested for the HFE being evaluated. That section can be applied to any distribution estimate derived without explicit consideration of PIFs.

4.2.1 Overview of Developing a Time Required Distribution Using the PNNL Method

Based on examination, review, and analysis of available completion time data (i.e., time-required by NPP operators to perform tasks), PNNL derived a two-part process that consists of developing a "first-order" distributional form from simulator data and adjusting that probability distribution to account for PIFs effects not addressed in the "first-order" distribution. Nearly all the usable data sources with enough data points to distill generalizations that support development of time-required probability distributions in HRA come from simulator experiments. PNNL's exploratory analysis of available time-required data and conclusions about how to use the data is documented in detail in Appendix A of this RIL. Appendix D provides a discussion of the data analysis that derived the "first-order" lognormal distribution and shape factor.

Most experiments that generated the applicable data examined, including the experiments that generated by far the most time-required data points, were not developed to test the impact of the 20 PIFs defined by the IDHEAS methodology. Rather, in nearly all cases, the differences in PIFs were intended to be minimized across scenarios. The crews participating in the simulator experiments were given the same tasks to perform using the same simulations, were tested under the same environmental conditions, and used the same simulated plant controls and operational procedures. All of the PIFs were therefore equivalent, with one important exception: differences amongst the crews themselves in how they performed the tasks. The uncertainty associated with these differences was acknowledged in the cited research and is referred to here as *crew-to-crew variability*. Researchers have previously noted that there are inherently slower and faster crews. This inherent variability largely makes up what PNNL terms the "first-order" distribution. The "first-order" distribution reflects the variability of factors whose impacts were not minimized or were not entirely minimized in the simulator experiments.

The second part of PNNL's two-step process is adjustment of the "first-order" distribution to address PIFs not yet accounted for when evaluating a specific operation action. When estimating the impact of PIFs, the inherent crew-to-crew variability discussed above must be considered. Although models of teamwork have been proposed to account for crew-to-crew variability, they employ many factors with complex interrelationships for which there is no data to parameterize the models. Crew-to-crew variability is therefore considered aleatory because there is not enough data to support the detailed modeling to analyze and dissect the teamwork factors contributing, resulting in randomness in crew performance that has not been reduced by increasing the analyst's knowledge of the situation and systems involved. In the PNNL method for adjusting the "first-order" distribution, the IDHEAS PIFs of Team and Organizational Factors and Work Processes were identified to primarily contribute to crew-to-crew variability. PNNL proposes to treat the PIFs of Team and Organizational Factors as aleatory and the other IDHEAS PIFs as epistemic. The adjustment of the "first-order" distribution of time-required then relies on further assessment (largely engineering judgment) to incorporate the PIF impacts. A discussion of the ways to adjust the "first-order" distribution is presented in Appendix B.

Creation of the "first-order" distribution of the HFE being evaluated involves using point value estimate(s) of time-required in combination with the distributional form and shape factor for the distribution recommended by PNNL. Adjustment of the "first-order" distribution is then performed to create the final probability distribution for time-required using available time-required data, operator interviews, and applicable research literature. The first step in PNNL's two-step approach is described in following section. Adjusting the distribution for other PIFs in described in Section 4.3. If sufficient amounts of data are available estimate distributions directly from data, that is the preferred method. However, when sufficient data are not available, the PNNL approach provides an informed option.

4.2.2 Creation of First-Order Distribution for Time Required from a Point Estimate

This section explains how to create the “first-order” distribution for time-required from a point estimate value and the recommended distribution form and shape factor for HFEs being evaluated. As described in the previous section, PNNL method consists of initially developing a “first-order” distribution of time-required to account for crew-to-crew variability. The recommended form and shape factor of the distribution was developed from existing applicable simulator data and is considered reflect aleatory uncertainty (see Appendix D). This aleatory uncertainty is judged to exist regardless of other PIFs impacts that may exist beyond Teamwork and Work Process. Barring analysts having sufficient plant data to estimate a distribution directly, analysts should assume lognormal distribution form. This distribution has three parameters: (1) location, (2) shape, and (3) scale.

$$f(x) = \frac{e^{-\left(\left(\log\left(\left(x-location\right)/scale\right)\right)/\left(\sqrt{2}\cdot shape\right)\right)^{2}}}{\sqrt{2\pi}\cdot shape\cdot x}$$

Often the location parameter is set to zero, because in many applications, values can be very small and near zero. This leads to a two-parameter lognormal distribution. PNNL initially considers the two-parameter distribution, but in verifying the parameter solutions, then considers the full three-parameter model.

To develop the “first-order” time-required distribution for the HFE being evaluated, these parameters should be determined using one of the following alternatives:

Alternative 1: The first-order normalized distribution should be combined with a point estimate of time-required. HRA analysts should set the scale parameter for a lognormal distribution to the **mean** time-required for a given HFE (i.e., the point estimate mean time-required) and set the shape parameter to a value between 0.28 (or the error factor to 1.59) and 0.54 (2.42 error factor). More conservative estimates will be closer to 0.54. Note that some sources may realize the lognormal distribution with the scale parameter outside of the logarithmic function. In that case, the logarithm of the mean time-required should be used. The location parameter is set to zero.

Alternative 2: The scale of the desired probability distribution may be matched with the **median** of observed data. However, PNNL determined that the mean of the time-required provided a reasonably good fit in the analysis of the EPRI time required data. Lastly, although the shape parameter of 0.28 (or error factor of 1.59) was determined to work best for data transformed by a normalization procedure (see Appendix D), it was also a valid value for untransformed data where, potentially, only a single point estimate may be available. Shape parameter values closer to 0.54 (error factor of 2.42) also provided reasonable fits for the untransformed and transformed data. Location parameter is set to zero.

As stated in the previous section, PNNL’s exploratory analysis of available time-required data and conclusions about how to use the data is documented in detail in Appendix D of this RIL. That appendix includes detailed discussion of the data analysis that derived the “first-order” lognormal distribution and shape factor of 0.28.

4.3 Adjustment of the First Order Distribution Based on Analyst or Operator Judgment and Context from Plant Data and Applicable Literature

The PNNL approach to creating a probability distribution for time-required involves adjusting of the "first-order" distribution based largely on HRA analyst or operator judgment about the impact of dominant PIFs on the "first-order" distribution (i.e., the dominant one to three PIFs). (Considering the PIF with the most impact or the combination of two or three PIFs with the most impact, is judged to be sufficient to adjust the "first-order" distribution given limited applicable research.) This is similar to estimating a probability distribution for time-required as described as the first approach of Option 2 in NUREG-2256, Section 3.6, based on operational and simulator data, and interviewing operators, for which entire distribution is based on PIFs impacts without first considering crew-to-crew variability. However, the analyst's judgment is more explicitly relied on to adjust the "first-order" distribution, rather than being used to develop the time-required probability distribution parameters. The cited information could be used to develop the time-required probability distribution itself (presented as the first approach of Option 1) if it is judged to provide adequate results. However, PNNL suggests based on our research that it may be difficult for an operator or analyst to make a valid estimate of the impact of crew-to-crew variability if that person's judgment is based on their own experience of working with or observing only one (or a few) crews.

In addition to review of potential plant data and/or information from interviews with operators about time-required, there is also applicable information that could be used from the general literature about the impact of PIFs on the time-required to perform tasks. PNNL has reviewed research on the quantitative impact of PIFs on the time-required to perform tasks, and this review is described in detail in Appendix C. This information can also be used to inform an analyst's estimation of the impact of PIFs on the "first-order" distribution.

Finally, HRA analysts are required by the NRC endorsed PRA standard [5] to base the point estimates of time-required for risk significant HFEs on the best available information, such as times recorded from training exercises, walk-throughs or talk-throughs, or feasibility studies. In addition, there is information from general literature about the impact of PIFs on the time-required to perform tasks. These sources of information can potentially be used for a given HFE to support estimation of the impact of the dominant PIFs on the "first-order" distribution.

In Step 1 of PNNL's method, the "first-order" time-required distribution for the HFE being evaluated is developed by using the first-order normalized distribution combined with a point estimate of time-required. An HRA analyst's judgment, supported by the data and information sources discussed above, is then made about the impact of the dominant one to three PIFs on the "first-order" distribution. The form of the distribution is assumed to stay lognormal, but the mean (or median if that is used) is impacted as well as the shape factor (or error factor for a lognormal distribution) in this modeling approach. There is not enough evidence to support otherwise. Accordingly, the HRA analyst should make a judgment about how much the mean (or median) and shape factor (error factor) increase for the HFE considering the impact of the dominant one to three PIFs. The dominant PIFs are considered to be the PIFs with the greatest impact on time-required. Using this approach, the mean (or median), is likely to increase beyond the mean (or median) of the probability distribution that accounts for "first-order" distribution (the impact of crew-to-crew variability). Likewise, the shape factors are likely to be greater than the range 0.28–0.54 for the same reason. However, it is recognized that the impact of some PIFs could decrease the mean and variance. However, for environmental, system, and personnel related PIFs the impact is not likely to be significantly better than the impact from a nominal situation. However, if a particular task is especially simple or routine, then the impact on variability from associated task related PIFs could potentially be less than nominal.

Although there are some available information sources about time-required, it is unlikely that adequate data or information exists to inform the evaluation of specific HFEs. Sources of available information discussed above include (1) an analyst's review of applicable plant data or information (such as walkthroughs or talk-throughs, feasibility studies, and JPM test data) on time-required; (2) interviews with operators, (3) an analysts' own judgment based on past experience, and/or (4) information from the general literature about the impact of PIFs on the time-required to perform a task. Nonetheless, this information at minimum can be used to provide general context and insight to estimate the increase in the central tendency and/or variance of the "first-order" distribution to estimate the adjustment needed to account for the dominant PIFs.

4.3.1 Consideration of Plant and Operational Data

Regarding use of plant data on time-required, one potential source of statistical data is JPM testing. JPM testing standards are bounding estimates on which all plant crews are trained to demonstrate they can meet time requirements. In JPM testing, plant operators are tested using the same simulator and plant indications and controls and experience the same situations and environmental conditions. Because of these controlled conditions, the PIFs associated with Systems (i.e., System and Instrumentation and Control (I&C) Transparency to Personnel, Human-System Interfaces, and Equipment and Tools) and Environment and Situation (i.e., Work Location Accessibility and Habitability, Workplace Visibility, Noise in Workplace and Communication Pathways, Cold/Heat/Humidity, and Resistance to Physical Movement) are not significantly different between tests and would not be expected to contribute much to the variability in the estimated time-required probability distribution. In addition, plant operators are tested using the same plant procedures, guidelines, and training, and are under the same staffing rules, and therefore many of the PIFs associated with Personnel (i.e., Staffing, Procedures, Guidelines, and Instruction, and Training) are not significantly different between tests and would not be expected to contribute much to variability in the estimated probability distributions. The two remaining PIFs are Teamwork and Organizational Factors and Work Processes, but their impacts are already considered to already be encompassed by the "first-order" distribution.

However, the simulator tests themselves can be different from each other, and therefore, different tests can be impacted by task related-PIFs (i.e., Information availability and reliability, Scenario familiarity, Multitasking, interruption, and distraction, Task complexity, Mental Fatigue, Time pressure and stress, and Physical demands) differently. Accordingly, some tasks may be more complex than others, require different levels concentration, impart different time pressures and stress, etc. It may be, for example, that testing data on complex tasks versus simple tasks may show more variability between for complex tasks versus simple tasks. NUREG-2198 [2] defines task complexity as "measur[ing] task demand for cognitive resources (e.g., working memory, attention, executive control). Nominal complexity refers to the level of complexity that does not overwhelm personnel." The range of the variance could be used to inform the time-required probability distribution, if, for example, the dominant PIF was task complexity. The same kind of statistical information could be used to estimate the variance of task-related PIFs such as tasks in scenarios that are familiar versus those that are not familiar.

4.3.2 Consideration of Applicable Research

Regarding use of information from the open literature on the impact of dominant PIFs on the time-required to perform human actions, the sparse nature of specific applicable data likely limits its use. Even though the data likely cannot be directly used, it can still provide at least some perspective and examples of the magnitude of impact the dominant PIFs could have context for adjusting the first-order distribution. PNNL’s evaluation of applicable literature on the impact of PIFs is described and summarized in tables presented in Appendix B, Section B.4. This information provides a way to inform or provide context to the judgment of the impact of dominant PIFs. Each table contains PIF considerations (i.e., examples of the way that the PIF might impact the probability distribution of time-required), a representative study from the general literature, and a description of how that PIF might impact the probability distribution of time-required based on the associated findings. Studies were selected based on their applicability to a NPP context (e.g., simulator studies) as well as their inclusion of task completion time information. These tables provide some information for analysts to use when considering the one to three dominant PIFs that they believe may have an impact on the distribution of time-required. After identifying the dominant PIFs, analysts and operators can use the referenced tables to help them to understand how they might adjust either the mean or the shape factor of the probability distribution of time-required.

5 References

[1] J. Xing, Y. J. Chang, and J. DeJesus Segarra, “NUREG-2256, integrated human event analysis system for event and condition assessment (IDHEAS-ECA),” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML22300A117, Oct. 2022.

[2] J. Xing, Y. J. Chang, and J. DeJesus Segarra, “NUREG-2198, the general methodology of an integrated human event analysis system (IDHEAS-G),” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML21127A272, May 2021.

[3] NRC, “Regulatory Guide 1.200, Revision 2, “An approach for determining the technical adequacy of probabilistic risk assessment results for risk-informed activities.,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML090410014, Mar. 2009.

[4] NRC, “Regulatory guide, 1.247 1.247, For Trial Use, “Acceptability of Probabilistic Risk assessment Results for Non-Light-Water Reactor Risk-Informed Activities.” U. S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML21235A008, March 2022.

[5] A. Gilbertson, “Regulatory guide 1.200, revision 3, acceptability of probabilistic risk assessment results for risk-informed activities,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML20238B871, Dec. 2020.

[6] ASME/ANS (American Society of Mechanical Engineers/American Nuclear Society), “Addenda to ASME/ANS RA-S–2008 standard for level 1/large early release frequency probabilistic risk assessment for nuclear power plant applications,” ASME/ANS (American Society of Mechanical Engineers and American Nuclear Society), New York, NY, ASME/ANS RA-Sa–2009, Feb. 2009.

[7] ASME/ANS (American Society of Mechanical Engineers/American Nuclear Society), “Probabilistic risk assessment standard for advanced non-light water reactor nuclear power plants,” ASME/ANS (American Society of Mechanical Engineers and American Nuclear Society), New York, NY, ASME/ANS RA-S-1.4–2021, Feb. 2021.

[8] J. Xing and S. Morrow, “White paper: Practical insights and lessons learned on implementing expert elicitation,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML16287A734, Oct. 2016.

[9] Electric Power Research Institute (EPRI), “Operator reliability experiments using power plant simulators, volume 1,” NUS Corporation, Gaithersburg, Maryland, EPRI NP-6937, Jul. 1990.

[10] Electric Power Research Institute (EPRI), “Operator reliability experiments using power plant simulators, volume 2,” NUS Corporation, Gaithersburg, Maryland, EPRI NP-6937, Jul. 1990.

[11] Electric Power Research Institute (EPRI), “Operator reliability experiments using power plant simulators, volume 3,” Accident Prevention Group, General Physics Corporation, and NUS Corporation, Gaithersburg, Maryland, EPRI NP-6937-L, Jan. 1991.

Appendix A
–
Data Selection for Time Required Evaluation

This appendix discusses details of the investigation into how performance influencing factors (PIF) impact time-required. Data for that investigation come from simulator experiments. Here, we discuss how these data should be considered when trying to develop insights about development of time-required probability distributions. Appendix A, Section A.1 discusses the data sources that were available to PNNL for this evaluation.

A.1 Time Required Data Sources

This section discusses the available datasets that appeared to have applicability to estimating a statistical distribution and Pacific Northwest National Laboratory's (PNNL) qualitative evaluation of those datasets. These datasets included real-world data provided in U.S. Nuclear Regulatory Commission (NRC) NUREG documents and simulation data provided by Korean Atomic Energy Research Institute (KAERI), the Nuclear Research Institute of the Czech Republic, ÚJV Rez, a.s. (UJV), the Halden Man-Machine Laboratory (HAMMLAB/Halden), and from EPRI.

A.1.1 Real-World Data from NUREG/CR-6365

NUREG/CR-6365 [1] contains a summary of real-world steam generator tube ruptures (SGTR) events and includes data on time-required, leak rates, plant information, and the reason for the rupture. In many ways, real-world data from actual events on time-required are preferable over time-required data generated from simulators, because real-world data may be free from the potential bias and uncertainty introduced by the artificial nature of a simulated experiment and environment. However, PNNL was unable to use these SGTR data because the time-required data existed at the scenario level, not at the human failure event (HFE) level. Also, the data came from a mixed set of reactor designs and, thus, the design could have contributed to the variability in time-required.

Data in Table A.1 from NUREG/CR-6365 show that the maximum leak rate for the 12 leaks ranged from 112 gpm to 760 gpm. NUREG/CR-6365 states that maximum leak rates less than 100 gpm are generally below the normal charging flow capacity of the reactor cooling systems and considered by NRC to be tube defects rather than tube ruptures. However, the NRC indicates in the IDHEAS-DATA draft report [2] that only leaks above 300 gpm are significant, which leaves only six data points in NUREG/CR-6365. This is insufficient data to inform development of time-required probability distributions.

A.1.2 Data from Korean Atomic Energy Research Institute

PNNL evaluated two KAERI reports [3], [4] that used simulator trials to evaluate crew performance for training purposes. These simulator trials were designed to be as realistic as possible to improve the crews' real-world performance. For both cases, all crews were experienced, had proper procedures and supervision, and were given similar scenarios. Dr. Jinkyun Park of KAERI met with PNNL through NRC to discuss relevant Korean datasets and to provide time-required data for individual crews that were not available in the published reports. Although the KAERI reports provide the average time-required across all crew members, PNNL could not use this data because crew-specific time-required data are needed to estimate time-required probability distributions.

Accordingly, KAERI also provided PNNL with unpublished crew-level data for a SGTR scenario associated with their published work, Analysis of Human Performance Observed Under Simulated Emergencies of Nuclear Power Plants [3]. However, the data often were incomplete at the HFE level because of limitations of their audio-visual recording system that made it difficult to identify start and stop times of key tasks. While all 12 crews successfully completed the SGTR scenario, KAERI was able to verify only the key HFE level data for three of the 12 crews. KAERI indicated that they had significant amounts of other crew-level data that could be useful for informing the goals of this report, but those data were not publicly available at the time of this analysis.

Concerning the other KAERI report, A Study on the Validity of a Task Complexity Measure for Emergency Operating Procedures of Nuclear Power Plants – Comparing Task Complexity Scores with Two Sets of Operator Response Time Data Obtained Under a Simulated SGTR [4], only the average time-required data for six crews was available from KAERI unlike the first report. So, given that crew-specific time-required data was not available for the six crews that participated in this experiment, the data could not be used to inform a time-required probability distribution.

Table A.1. Summary of Real-World SGTR Events and Their Causes (Sources Tables 12 and 13 NUREG/CR-6365)

Date	Plant	Vendor & SG Model	Max Leak Rate (gpm)	Degradation Mechanism	Rupture Size	Time Operators Recognized SGTR (min)	Time Isolate SGTR (min)	Stress and Contributing Factors
2/6/1975	Point Beach-1	WEC -loop, W-44	125	Wastage	2 adjacent ruptured bulges each about 20 mm and wide	24–28	58	Large sludge pile, ineffective leaning
9/15/1976	Surry-2	WEC 3-loop, W-51	330	PWSCC	I14.3 mm long axial crack	< 5	18	High stresses and ovalization caused by inward movement of the legs due to support plate deformation
6/25/1979	Doel-2	ACE-44	135	PWSCC	100 mm long axial crack	9	9.4	High residual stresses due to ovalization during fabrication
10/2/1979	Prairie Island	WEC 2-loop, W-51	336	Loose Parts Wear	38 mm long axial fishmouth opening	5–18.5	27	Sludge lancing equipment left in the steam generator
1/25/1982	Ginna	WEC 2-loop, W-44	760	Loose Parts Wear, Fretting	100 mm long axial fishmouth opening	< 1	15	Loose parts (baffle plate debris) left in the steam generator, wear of peripheral tubes, fretting, or inner tubes
5/16/1984	Fort Calhoun	CE 2-loop	112	ODSCC	32 mm long axial crack (small fishmouth opening)	32	40	Tube deformation caused by corrosion of the vertical batwing support bars, caustic impurities on the secondary side
7/15/1987	North Anna-I	WEC 3-loop, W-51	637	High-Cycle Fatigue	360° circumferential break	< 5	18	High-cycle vibration, denting, lack of AVB support
3/7/1989	McGuire-I	WEC 2-loop. W-D2	500	ODSCC	9.5 mm long axial crack in a 645 mm long groove, 9.5 mm wide at the maximum point	< 1	11	Long shallow groove, possibly a contaminant
2/9/1991	Mihama-2	WEC 2-loop, MHI-44	700	High-Cycle Fatigue	360° circumferential break	5	22	High-cycle vibration, lack of AVU support
3/14/1993	Palo Verde-2	CE-80	240	ODSCC	65 mm long axial fishmouth opening in a 250 mm long axial crack	37	176	Tube-to-tube crevice formation, bridging deposits, caustic secondary waler chemistry, susceptible material

A.1.3 Data from the Nuclear Research Institute of the Czech Republic

Through NRC, the Nuclear Research Institute of the Czech Republic, UJV Rez, A.S. (UJV) providedˇ PNNL with unpublished crew-level data for an loss of feedwater (LOFW) based on simulator trials of a Russian water-water energetic reactor (VVER) plant design [5]. UJV acquired these data from plant owners to evaluate the performance of their HRA methods and to provide plant owners feedback about factors that influence work performance. However, UJV provided only limited information about the simulator conditions and crew factors (e.g., experience) because these details were not provided by the plant owners. The dataset contains time-required data for 15 crews that successfully completed the tasks associated with key HFEs and the entire LOFW scenario; two crews committed errors, so those times had to be removed from the total dataset of 17 crews. The tasks for which time-required data were provided were for a Russian VVER and are unlike the tasks required for a U.S. design like a Westinghouse pressurized water reactor (PWR) even though both reactors are PWR designs. For example, the VVER feedwater collector has two halves that can be isolated in a LOFW scenario if needed. UJV indicated that isolation of the impacted side of the feedwater collector was a complex but advantageous option for a VVER but an option unavailable in U.S. designs. The central strategy in a LOFW accident for Westinghouse PWR is to perform a feed and bleed operation.

Accordingly, the UJV time-required data were 1) for a reactor design not comparable to U.S. designs and 2) based on LOFW accident response procedures not fully comparable to Westinghouse procedures. For these reasons, PNNL decided not to combine the UJV time-required data from a Russian VVER with time-required data from other (i.e., U.S.) reactor designs such as Westinghouse PWR designs used by KAERI and the HAMMLAB (Halden Man-Machine Laboratory) to increase the size of a dataset.

A.1.4 Data from the Halden Man-Machine Laboratory

PNNL evaluated the results from three Halden Man-Machine Laboratory (HAMMLAB/Halden) simulator studies [6]–[8]. PNNL through NRC also met with HAMMLAB staff to discuss the utility of their studies for informing development of a probability distribution for time-required and how that distribution can be impacted by PIFs. The main purpose of the three HAMMLAB studies was to validate different HRA methods and identify refinement possibilities by comparing crew time-required performance predicted by the HRA methods to crew performance results from the simulations.

The HAMMLAB simulator trials produced crew-level time-required data for basic and high complexity scenarios. Basic complexity scenarios were for routine events that were familiar to the crews and were well informed by the procedures. Complex scenarios were designed to test the ability of crews to respond to challenging multiple-event scenarios that were unfamiliar, difficult to properly diagnose due to one event masking the other and were not well informed by the available procedures. For example, HAMMLAB commonly simulated a “complex induced SGTR” in which response to the LOFW scenario could involve restoration of cold auxiliary feedwater that causes hot pipes to rupture and induces a SGTR. The SGTR remains hidden to crews if they keep auxiliary feedwater flowing. The simulations also involved a base case which was a more typical less complex version of the scenarios from which contrasts could be made.

PNNL determined that a significant portion of the time-required data from the HAMMLAB reports would be difficult to use to inform development of a time-required probability distribution because they are for the complex scenarios. It is expected that the time-required for operator response in complex scenarios would be longer and could have different variability than time-required for the base case scenarios. So, the available data, which is already sparse, is reduced to the remaining time-required data (i.e., time-required for operations action in the base case scenarios). PNNL judges that by itself this data does not represent enough data to inform time-required probability distributions.

However, PNNL postulated that the data might be used to estimate the impact of the Task Complexity PIF to on time probability required distributions. Table A.2 presents the time-required data for the first three HFEs associated with a base and complex case of a SGTR scenario using data presented in the International HRA Empirical Study – Phase 2 Report Resulting from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios [7]. The fourth and fifth HFEs associated with these scenarios are not included because the fourth HFE only applies to the base case, and insufficient data is presented for the fifth HFE. Table A.2 shows a comparison of the mean and standard deviation between the time-required for HFEs in the base versus complex scenarios. HFE-1A and B is “Failure to identify and isolate SGTR,” HFE-2A and B is “Failure to cool down the reactors cooling system (RCS) expeditiously” and HFE-3A and B is “Failure to depressurize cool down the RCS expeditiously.” The base case is a SGTR while the complex case is major steam line break that induces a SGTR along with failure of secondary radiation indications.

Table A.2. Mean and Variance (Standard Deviation) Time (s) Comparison of the HFEs in the Base versus Complex Scenarios as Amended [9]

	HFE-1A	HFE-1B	HFE-2A	HFE-2B	HFE-3A	HFE-3B
	Basic	Complex	Basic	Complex	Basic	Complex
	813	1681	355	430	380	441
	799	1270	490	280	521	275
	1133	1737	610	485	366	571
	994	1634	736		442	324
	862	2727	385	552	523	986
	1125	1816	281	405	272	303
	1118	1419	530	350	380	480
	719	1483	910	710	342	242
	817	1296	485	435	158	142
	1058	1928	247	445	331	228
	909	1599	370	395	318	224
	786	1199	380	325	779	316
	623	1332	375	420	534	153
	1239	1477	630	375	369	211
Sum	13045	22598	6784	5607	5715	4896
Average	931.8	1614.1	484.6	431.3	408.2	349.7
SD	190.6	386.7	184.8	108.6	148.0	220.5
SD/Average	0.21	0.24	0.38	0.25	0.36	0.63

In the case of HFE-1A and B, the time-required for operator action is about 73% greater for the complex case compared to the base case and the standard deviation (SD) is about 105% greater. This result does not seem unreasonable. However, in the case of HFE-2A and B, the time-required for operator action for the complex case is about 11% lower than for the base case and the SD is significantly lower at 41% lower. This result is, in contrast to HFE-1, not immediately intuitive because one would expect the mean required time and SD to be greater in the complex scenario. The comparison between HFE-3A and HFE-3B is different from either of the previous two HFEs. In the case of HFE-3A and B, the average time-required for operator action for the complex case is about 14% lower than for the base case and the SD is significantly higher at 49% higher. It could be that the complexity in scenarios mostly impacts HFE-1, but without an assessment by the crews who participated in the experiment it is hard to explain these results by attributing the impact to the PIF Task Complexity as a singular impact on standard deviation or mean time-required. PNNL judges that this data, alone, is also hard to use to reliably estimate the impact of the PIF Task Complexity.

Other HAMMLAB documents reports about similar experiments also document time-required data, but these experiments employed a much smaller number of crews. The U.S. HRA Empirical Study – Assessment of HRA Methods Against Operating Crew Performance on a U.S. Nuclear Power Plant Simulator [10] assessed performance using just four crews opposed to 14 crews used by the International HRA Empirical Study [6]–[8]. The experiment documented in A HAMMLAB HRA Data Collection with U.S. Operators, 2016 – HWR-1123: OCED Halden Reactor Project [11] was performed using five crew members. In all three of these studies, a primary focus was to evaluate if crews could accomplish challenging scenarios where it was difficult to identify all of the problems because of masking. In Massaiu and Holmgren [11], for the SGTR scenario, a construction explosion was postulated to cause small SGTR that was followed by a larger volume SGTR. It was difficult for crews to identify both ruptures because a single radiation alarm activated by the smaller rupture masked the presence of multiple ruptures. While all five crews successfully identified and isolated the larger rupture, only three of the five crews identified and isolated both ruptures. Again, it is not clear whether time-required data gathered from the simulator scenarios that are created to be as challenging as described above provide a good basis for development of time-required probability distributions.

A secondary goal of the HAMMLAB studies was to evaluate how overall crew performance was affected by 12 performance shaping factors (PSF), such as teamwork and stress, and was addressed in detail for two of studies, the International HRA Empirical Study [8] referred to above and the U.S., HRA Empirical Study [10]. The researchers tried to identify the reasons that some crews were slower than other crews (e.g., "... slower crews may not feel the urge to perform cooldown expeditiously," "... too much time in crew meeting or discussing plan of action," and "... four other teams with poor team dynamics performed less well").

These PSF evaluations were conducted by experts who observed crew performance and rated how each PSF affected overall crew performance (i.e., strongly negative, slightly negative, neutral, slightly positive, or strongly positive effects). The subjective PSF ratings from the HAMMLAB studies are useful for identifying which PSFs had strong or weak influence on overall crew performance. They used a rating system from −2 for strongly negative to +2 for strongly positive that exhibited more neutral scores (in-between) on crew performance such as time pressure, stress, scenario complexity, etc. PNNL combined the PSF scores from the International HRA Empirical Study report for all HFEs into one table shown in Table A.2.

Table A.3. Summary of Semi-Quantitative Impact Scores by Observer on PSFs for HFEs from International HRA Empirical Study [7]. Assessment of the impacts of 12 PSFs on crew performance for HFEs for the basic SGTR ("A") and complex SGTR ("B") scenarios.

	Overall PSF Assessment for SGTR Scenario HFEs						
PSFs	**HFE-2A**	**HFE-2B**	**HFE-3A**	**HFE-3B**	**HFE-4A**	**HFE 5B1**	**HFE 5B2**
1 – Time pressure	0	0	0	0	0	0	0
2 – Stress	0	0(-0.5)	-0.5	0(-1)	0	0	0
3 – Scenario complexity	0(1)	(-1)	0	0(-1.5)	0	-2	2
4 – Indication of conditions	0	0	0	0	0	-2	2
5 – Execution complexity	0(-1)	(-1)	-1	-1	2	2	2
6 – Training	1	1	1	1	1	1	2
7 – Experience	0	0	0	0	0	0	0
8 – Procedural guidance	-1	1	1	1	2	2	1
9 – Human Machine Interface	0	0	0	0	0	0	0
10 – Work Processes	1	1(-0.5)	0	0	0.5	-1	0
11 – Communications	1	1(-0.5)	0.5	1	0.5	0	0
12 – Team Dynamics	1(-1)	0.5(-2)	1(-2)	1(-1.5)	0.5	0	0

The evaluation of PSFs identified scenario complexity, execution complexity, and team dynamics to have the most influence over team performance. Time pressure, stress, and human machine interfaces were not rated to have a positive or negative effect on team performance. In both the International HRA Empirical Study [8] and the U.S., HRA Empirical Study [10], a similar rating system was used to evaluate the effects of PSFs on overall crew performance. The ratings were attributed to specific HFEs against PSFs as a whole and not to specific individual crews. In the cited reports, there is an evaluation table for each HFE, and observer comments are provided for many of the ratings. In some cases, observations are made about specific crews by their crew label. However, there is not enough information about variations in the impact of PIFs on individual crews to draw conclusions. Moreover, the ranking system, though quantitative values are used, are subjective.

A.1.5 Data from the Electric Power Research Institute

PNNL identified that the results from EPRI simulator trials published in the early 1990s had a goal of measuring time-required for operator actions across a significant number of crews and accident scenarios. This study was documented in EPRI NP-6937 and NP-6937L, Operator Reliability Experiments Using Power Plant Simulators [12], [13]. The study included eight experiments on six NPP designs that covered 40 accident scenarios and about 148 operator actions using up to 18 different crews per simulation. Accordingly, a large set of time-required data was generated (i.e., 1,068 crew level records) from the simulations.

The purpose of the EPRI studies was to validate and improve the HRA correlations used to calculate the probability of operator non-response to key control room actions in support of what was referred to as the operator reliability experiment. Those data primarily pertained to time-required for human interactions (HI) and associated probability distributions but also included data on the effects of important PSFs. The term HFE came into use well after the EPRI report on HIs was issued. However, there is a reasonable correlation between the tasks that are evaluated in an HFE and the HIs that were defined for the EPRI study. Data were collected at the training simulators of three pressurized water reactors (PWRs), which included two Westinghouse designs and a Babcock and Wilcox design, and three boiling water reactors (BWRs), all of which were General Electric designs. Table A.4 provide a summary of the accident scenarios used in the study. The number of required operator actions addressed for each scenario varied from a few to as many as a dozen or more. Accordingly, the EPRI study provides useful information for informing the probability distribution of time-required.

The data collected in the EPRI study included the following information that was found to be useful to investigating time-required probability distributions:

- Cumulative-time and delta-time tables for HIs for each crew (i.e., the time for an HI is the time from a cue such as an alarm to the time when the action is taken)
- Description of key HIs
- Data qualification
- Summary statistics such as sample size, mean, median, standard deviation, mean for normalized times, and standard deviation for normalized times

The study provided combined individual HI probability distributions by normalizing/centering them together along the medians. The authors described the benefits of combining the HI distributions as follows: 1) it lowers the probability estimates for areas without data, 2) it reduces statistical uncertainty, and 3) it allows the results to be more applied to a broader spectrum of plants, scenarios, crews, etc. (see page B-28 of [13]).

Though the main goal of the EPRI study was to validate and improve the HRA correlations used to calculate the probability of operator non-response, the study also included collection of survey information on the effects of important PSFs, which are conceptually similar to PIFs of IDHEAS. The effects of PSFs were assessed using multiple-choice crew debriefing questionnaires and observer multiple-choice questionnaires. In general, the observers assigned to an experiment had experience in PRA, plant operations, nuclear engineering, training, and human reliability. The “questionnaire approach” for individual crews was different than the HAMMLAB approach discussed in Appendix A, Section A.1.4 above in which ratings were assigned based on overall negative and positive crew performance impact by each PSF. In general, the observer questionnaire contained questions regarding PSFs and the crew debriefing questionnaire contained questions regarding crew information processing, diagnosis and decision making, action, plant-simulator differences, and operator experience for each HI.

Also noteworthy, Appendix D, Section D.3.1.2 of EPRI NP-6937L report specifically noted that certain crews were inherently faster or slower than other crews across different tasks based on quantitative comparisons.

A.2 Evaluation of Usefulness of Explored Time Required Datasets

A.2.1 Quantity of Applicable Data

With one exception, the primary challenge with the datasets presented in Appendix A, Section A.1 (i.e., real-world data from NUREG/CR-6365, KAERI data, UJV data, HAMMLAB data, and EPRI data) was a lack of applicable data from which to develop a general distributional form for time-required. The only exception was the EPRI study, which generated a large set of time-required data as described in Appendix A.1.5. After investigation, it was determined that much of the other data could not be used for the following reasons:

1. The lack of HFE-level data for individual crews in NUREG/CR-6365 as described in Appendix A.1.1.
2. The lack of HFE-level data for individual crews in the KAERI data although total time-required was provided at the accident sequence level as discussed in Appendix A.1.2.

Table A.4. Scenarios Used in EPRI Study NP-6937L

	Boiling Water Reactor (BWR) Series II
1	Turbine Thrust Bearing Failure/Anticipated Transient without Scram (ATWS)/Stuck Open Relief Valves (SORV)
2	ATWS with Reactor Core Isolation Cooling (RCIC) failure
3	Transient with narrow range level instrument function
4	Transient with narrow and wide range level instrument malfunction
5	Station Blackout
6	Delayed Station Blackout
7	Transient with Loss of a DC bus
8	Feedwater Line Break
	Pressurized Water Reactor (PWR) Series I
1	ATWS/SORV
2	Main Line Steam Break/SGTR
3	Steam Generator Tube Rupture (SGTR)/ Loss of Secondary Heat Sink
	BWR 2
1	Main Turbine Thrust Bearing failure with failure to Scram
2	60% power AWTS with Turbine Trip
3	100% power AWTS with Standby Liquid Control System (SLCS) failure
4	Station Blackout
5	Partial Station Blackout with a Loss of High Pressure Injection (HPI)
6	Blowdown Cooling
	BWR 3
1	ATWS
2	Station Blackout
	PWR 2
1	ATWS/SORV
2	Steam Line Break inside Concealment/SGTR
3	SGTR/ Loss of Secondary Heat Sink
4	
5	Spurious Open Pressurizer Spray Valve/Loss of Component Cooling Water
6	Inadvertent Safety Injection with Failure of Auxiliary Feedwater System (AFWS)
7	Large Loss of Coolant Accident (LOCA)
	BWR I Series III
1	Partial Station Blackout with a Loss of HPI
2	Main Turbine Thrust Bearing failure/ATWS/SORV
3	ATWS with failure of SLCS/SORV
4	Seismic Failure of Main Feedwater Line with failure of RCIC/Flood-induced Failure of Service water pumps
	PWR I Series II
1	Steam Line Break/SGTR
2	SGTR/Loss of Secondary Heat Sink
3	ATWS/SORV
4	Medium LOCA/Failure of High Pressure Charging
5	Loss of Component Cooling Water
	PWR 3
1	Loss of all Feedwater Resulting in HPI Cooldown
2	Switch Yard Isolated with no Offsite Alternative and Turbine Driven AFWS Inoperable
3	Small Break LOCA with Borated Water Storage Tank Supply Valves Failed Closed
4	Main Steam Line Failure Outside Reactor Building with Steam Generator Tube leak

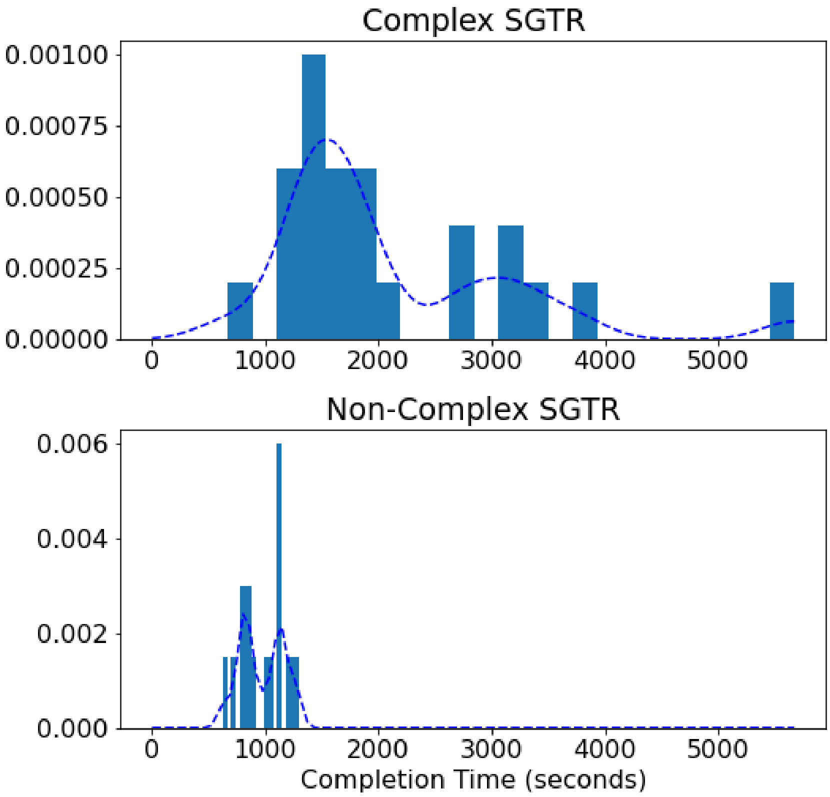


Figure A.1. Estimated Probability Density Function for Complex SGTR Scenario Completion Times (top) from the HAMMLAB Sources is Longer and More Variable than the Estimated Probability Density Function for Non-Complex SGTR Scenario Completion Times (bottom)

1. Significant plant and operational differences in the Russian designs reflected in the UJV time-required data that was provided compared to Westinghouse designs as discussed in Appendix A.1.3.
2. Reduction in usable data from the HAMMLAB studies given that half of the time-required data were for HFEs used in accident scenarios of high complexity compared to the data for HFEs used in accident scenarios of basic complexity as discussed in Appendix A.1.4. The time-required for the high complexity accident scenarios had a significantly higher variance than for the basic complexity scenarios. Moreover, the time-required for HFEs associated with the high complexity scenarios could include operator error and subsequent recovery.

Concerning the HAMMLAB time-required data, the IDHEAS-ECA methodology treats operator failure from cognitive error and subsequent recovery as a separate contributor to Human Error Probability from exceeding time-available. Accordingly, it is inappropriate to use time-required data from instances in which there was operator error and subsequent recovery. As indicated above, the HAMMLAB completion times for high complexity scenarios (the total time-required for all tasks) had a significantly higher variance and more complex distributional form than for non-complex scenarios (as shown in Figure A.1). Given this observation and the fact that the proportion of completion times from high complex scenarios represented half of the data and could unduly influence the distribution recommendation, it was judged that only the time related data from the basic accident scenarios could be used inform development of a general probability distribution for time-required (i.e., so just half of the HAMMLAB data). However, Appendix D.4 presents a discussion of the testing of the proposed “first order” distribution to the entirety of the HAMMLAB data, showing that there is a reasonable fit even with complex scenarios.

Additionally, the number of crews completing the same simulated scenarios in the HAMMLAB studies was small (i.e., 3 or 5) except for the International Empirical Study in four out of six of the studies examined in detail as shown in Table A.3. Accordingly, except for the EPRI dataset, there appears to be an insufficient quantity of applicable data to develop a distributional form for time-required with the exception of the EPRI data. When distinguishing distributional forms statistically, more data are required to definitively choose one form over another. We note that HAMMLAB data was used later to test the developed distributional form.

Table A.5. Number of Crews Used in HAMMLAB Studies

Accident Scenario	HAMMLAB Study	Number of Crews in Simulation	
		Complex	Noncomplex
Loss of Coolant Accident (LOCA)	U.S. HRA Data Collection	5	-
Loss of Feedwater (LOFW)	International Empirical Study	10	10
LOFW	U.S. Empirical Study	4	-
LOFW	U.S. HRA Data Collection	-	4
Steam Generator Tube Rupture (SGTR)	International Empirical Study	14	14
SGTR	U.S. Empirical Study	3	3
SGTR	U.S. HRA Data Collection	5	-

A.2.2 Conclusions from Data Examination Time Required Datasets

Analyses in this report for developing a general distributional form for time-required were conducted using simulator data from EPRI studies reported in EPRI NP-6937 and NP-6937L, Operator Reliability Experiments Using Power Plant Simulators [12]–[14]. EPRI data were determined to be suitable for analysis for the following four reasons.

1. The purpose of the EPRI study was to evaluate time-required probability distributions for individual HIs (operator tasks), which makes their results useful and generalizable for this report.
2. EPRI collected crew-level data, which is necessary for analyzing time-required probability distributions that reflect variable crew performance. In contrast, summary statistics (e.g., means/medians and standard deviations) do not inform the shape of distributions.
3. EPRI was the most data-rich source as it contained 1,068 records of crew-level completion times. As shown in Appendix A.1.5, EPRI data contains crew-level completion times for several crews and for multiple HIs. Although some of the HIs can be considered complex, they represent a small proportion of the scenarios.
4. The EPRI study also used a multiple-choice questionnaire to evaluate if PIFs (e.g., communication) affected time-required, which provided a useful dataset to evaluate if PIFs shifted or reshaped time-required probability distributions.

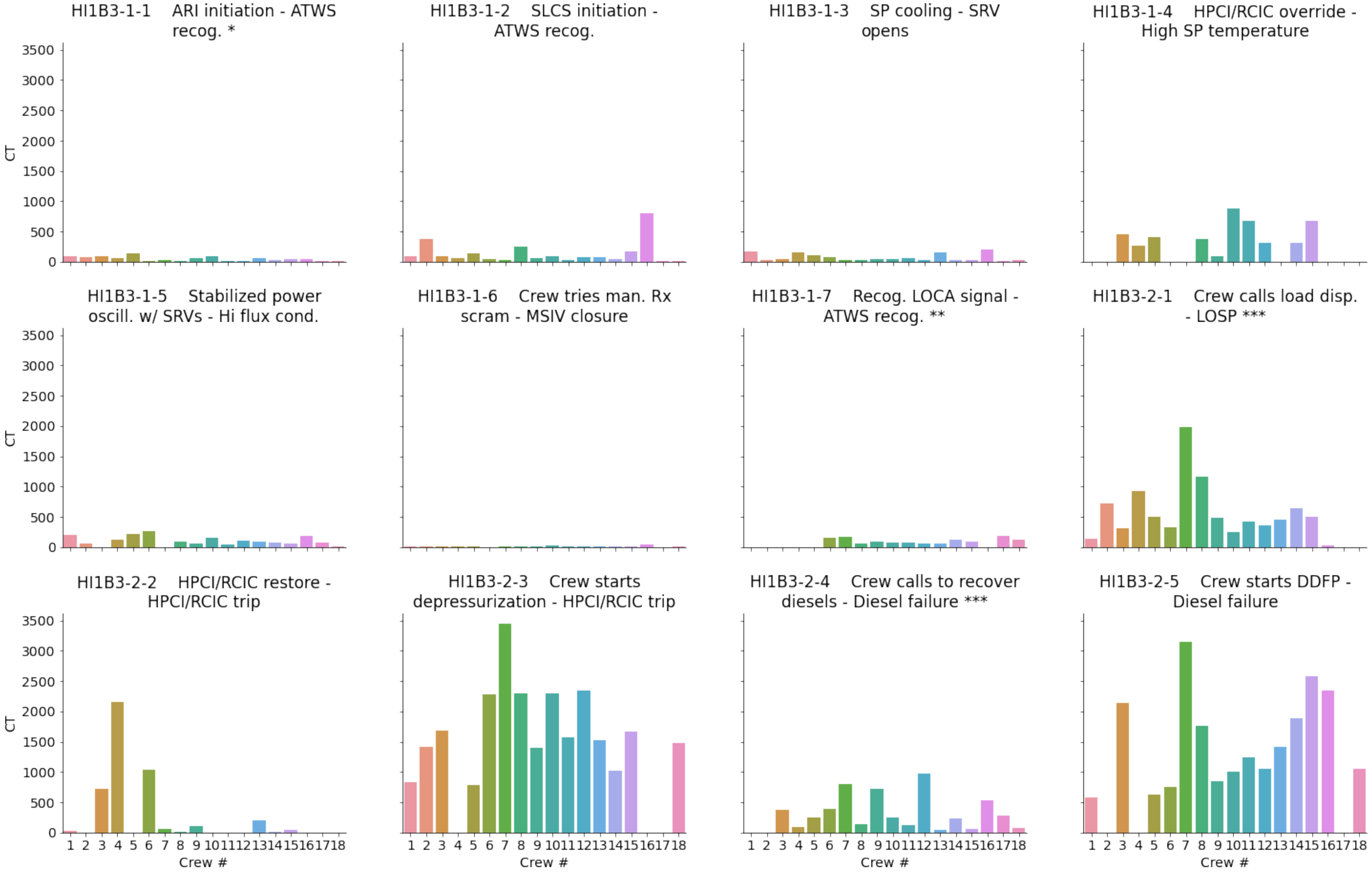


Figure A.2. Crew-to-Crew Variability in Task Completion Time from BWR-3 Reactor Design. The data source is the EPRI NP-6937L study and HI refers to human interaction

Limitations identified in the EPRI dataset are summarized below:

1. Time-required data reported in EPRI NP-6937L were compiled in the late 1980s to the early 1990s. Because of this, the data may not be entirely representative of current operations at NPPs. PNNL notes that operating procedures at NPPs have evolved since the EPRI study. However, actions required in response to accident initiators and plant controls that are used have changed minimally since the EPRI study was performed, and though dated, the study remains the best source of time-required information.
2. The 1,068 records in EPRI NP-6937L were a combination of multiple HIs, and the sample size for each individual HI was no larger than 18.
3. Across HIs, there was also a high degree of variability in time-required (see Figure A.2) that is discussed further in the next section.
4. Because EPRI data were from simulator experiments, they may not be entirely representative of real-world situations in NPP control rooms. However, very limited real-world data were reported in the sources examined for this report.
5. Since the EPRI experiments, NPPs have transitioned to "symptom-based" emergency Response Procedures (EOPs). Therefore, it is expected that improved procedural clarity have impacted time-required and may have impacted its variability. EPRI NP-6937L states:

   > The "symptom-based" EOPs provide the structure for crew diagnosis and action during accident sequences. While an improvement over earlier procedures, the EOPs used at all plants reportedly had deficiencies in logic and clarity that affected crew responses. Several examples were noted where the lack of clarity or incomplete guidance required the crews to interpret the intent of the EOPs before they could act. This lack of clarity and/or logic leads to alternative strategies such that one cannot then expect all crews to respond in a consistent manner to a given accident situation.

   The authors note that the plants and required response actions remain largely the same while the EOP instructions are improved overtime.

A.3 References

[1] J. P. Kotra, M. P. Lee, N. A. Eisenberg, and A. R. DeWispelare, "NUREG-1563, branch technical position on the use of expert elicitation in the high-level radioactive waste program.," Center for Nu clear Waste Regulatory Analyses (CNWRA), U.S. Nuclear Regulatory Commission/NMSS/DWM, Washington, DC, ADAMS Accession No. ML033500190, Nov. 1996.

[2] J. Xing, Y. J. Chang, and J. DeJesus Segarra, "Draft RIL-2021-XX, integrated human event analysis system for human reliability data (idheas-data)," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML20238B982, 2020.

[3] J. K. Park, W. D. Jung, J. W. Kim, and J. J. Ha, "Analysis of human performance observed under simulated emergencies of nuclear power plants," Jan. 2005. [Online]. Available: https: //www.osti.gov/etdeweb/servlets/purl/20658965.

[4] J. Park and W. Jung, "A study on the validity of a task complexity measure for emergency operating procedures of nuclear power plants—comparing task complexity scores with two sets of operator response time data obtained under a simulated SGTR," *Reliability Engineering & System Safety*, vol. 93, no. 4, pp. 557–566, 2008, ISSN: 0951-8320. doi: https://doi.org/10.1016/j. ress.2007.02.002. [Online]. Available:

https://www.sciencedirect.com/science/article/ pii/S0951832007000543.

[5] “Unpublished data from the dukovany npp simulator trials from 2011 as described in slides 18-29 of ml18079a165,” Nuclear Research Institute of the Czech Republic, Husinec, Provided by Dr. Radim Dolezal of the UJV Rez (Nuclear Research Institute of the Czech Republic), 2011.

[6] E. Lois, V. N. Dang, J. Forester, H. Broberg, S. Massaiu, M. Hildebrandt, P. Ø. Braarud, G. Parry, J. Julius, R. Boring, I. Männistö, and A. Bye, “NUREG/IA-0216, Vol. 1, international hra empirical study - phase 1 report. description of overall approach and pilot phase results from comparing HRA methods to simulator performance data.,” U.S. Nuclear Regulatory Commission, OECD Halden Reactor Project, Washington, DC, ADAMS Accession No. ML093380283, Dec. 2009.

[7] A. Bye, E. Lois, V. N. Dang, G. Parry, J. Forester, S. Massaiu, R. Boring, P. Ø. Braarud, H. Broberg, J. Julius, I. Männistö, and P. Nelson, “NUREG/IA-0216, Vol. 2, international HRA empirical study - phase 2 report. results from comparing HRA method predictions to simulator data from SGTR scenarios.,” U.S. Nuclear Regulatory Commission, OECD Halden Reactor Project, Washing ton, DC, ADAMS Accession No. ML11250A010, Aug. 2011.

[8] V. N. Dang, J. Forester, R. Boring, H. Broberg, S. Massaiu, J. Julius, I. Männistö, H. Liao, P. Nelson, E. Lois, and A. Bye, “NUREG-IA-0216, Vol. 3, international HRA empirical study - phase 3 report. results from comparing HRA methods predictions to hammlab simulator data on LOFW scenarios,” U.S. Nuclear Regulatory Commission, OECD Halden Reactor Project, Washington, DC, NUREG/IA-0216 Vol. 3, ADAMS Accession No. ML14358A254, Dec. 2014.

[9] P. Liu and J. Liu, “Combined effect of multiple performance shaping factors on human reliability: Multiplicative or additive?” *International Journal of Human–Computer Interaction*, vol. 36, no. 9, pp. 828–838, 2020. doi: 10.1080/10447318.2019.1695461. eprint: https://doi.org/10.1080/10447318 . 2019. 1695461. [Online]. Available: https:// doi. org/ 10. 1080 / 10447318. 2019 . 1695461.

[10] J. Forester, H. Liao, V. N. Dang, A. Bye, E. Lois, M. Presley, J. Marble, R. Nowell, H. Broberg, M. Hildenbrandt, B. Hallbert, and T. Morgan, “NUREG-2156, the U.S. HRA empirical study – assessment of HRA method predictions against operating crew performance on a U.S. nuclear power plant simulator,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML16179A124, Jun. 2016.

[11] S. Massaiu and L. Holmgren, “A hammlab HRA data collection with U.S. operators,” Institutt for energiteknikk, Halden, Norway, HWR-1123, May 2016.

[12] EPRI (Electric Power Research Institute), “Operator reliability experiments using power plant simulators, volume 2,” NUS Corporation, Gaithersburg, Maryland, EPRI NP-6937, Jul. 1990.

[13] EPRI (Electric Power Research Institute), “Operator reliability experiments using power plant simulators, volume 3,” Accident Prevention Group, General Physics Corporation, and NUS Corporation, Gaithersburg, Maryland, EPRI NP-6937-L, Jan. 1991.

[14] EPRI (Electric Power Research Institute), “Operator reliability experiments using power plant simulators, Volume 1,” NUS Corporation, Gaithersburg, Maryland, EPRI NP-6937, Jul. 1990.

Appendix B
–
PIF Impacts and Distribution Adjustments

This section discusses investigation of how the impact of performance influencing factors (PIF) on time-required data that come from simulator experiments should be considered when trying to develop insights about development of time-required probability distributions. Appendix B.1 discusses, in general, why consideration of the impact that PIFs have on time-required probability distribution is important. Appendix B.2 presents, in general, the impact of PIFs on time-required data that comes simulator experiments. Appendix B.3 discusses the notion of crew-to-crew variability and why it is important to the consideration of the impact of PIFs on time-required data that comes simulator experiments. Finally, Appendix B.4 presents a brief literature review of the impact of PIFs on time-required uncertainty.

B.1 Possible Impact of PIFs on Time Required Distribution

PIFs can have an impact time-required by providing greater cognitive challenges to operators that impact the time-required to complete a task. For the same reason, PIFs can have important effects on time-required probability distributions. Figure B.1 provides a conceptual example where PIFs negatively affect time-required by causing the probability distribution to shift to the right (from the blue line to black line). This shift could result in increase in the mean or median of time-required and/or increase in the variance of time-required. The distribution shift illustrated could be a concern for HRA analysts because a portion of the curve could now be outside the time limit (depicted as the red vertical line), which means there is a probability that a task associated with an HFE task may not be completed in the time-available. Therefore, it is important to understand which PIFs are likely to affect the time-required of tasks associated with HFEs and the magnitude of their effects on the time-required probability distributions.

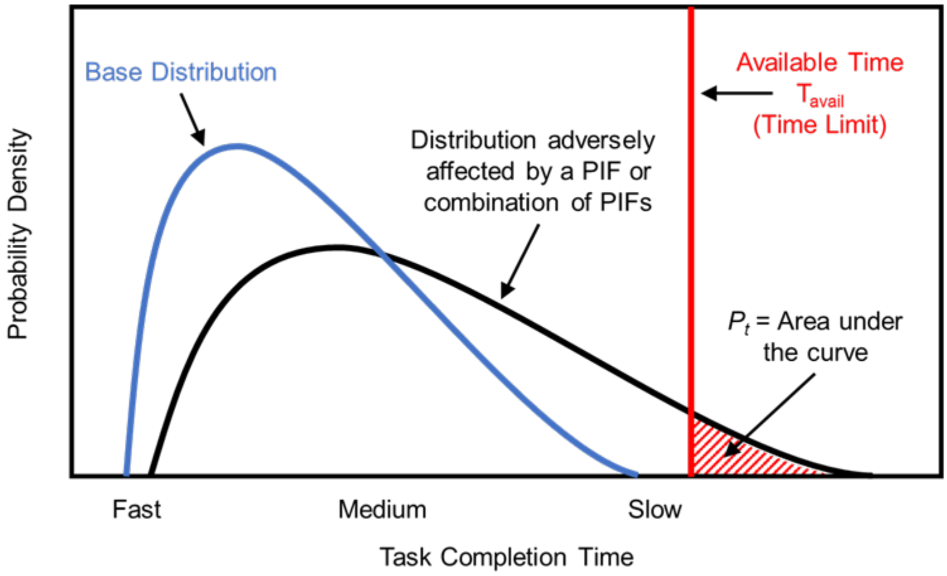


Figure B.1. Illustration of a Base Distribution of Time Required and Adverse Effect of PIFs on the Time Required Probability Distribution. Conceptually, the base distribution includes Completion Times that are not affected (either adversely or favorably) by PIFs individually or in combination.

To investigate the impact of PIFs on time-required data that comes simulator experiments, it was useful to organize the PIFs groups. NUREG-2198, The General Methodology of An Integrated Human Event System (IDHEAS-G) [1], already uses a PIF organizational structure that groups the 20 PIFs[2] into four context types—environment and situation-related, system related, personnel related, and task related—as shown in Table B.1 below. It is assumed in Figure B.1 that the base case is an optimal state and that the context associated with PIFs can only be worse than the base case.

Organization by these context types provided a way to discuss general insights above time-required data that comes from simulator experiments. The following sections discuss plant, task, and personnel related PIFs associated with the four context types identified above. Descriptions of the experiments performed by the HAMMLAB as discussed in Appendix A.1.4 and by EPRI as discussed in Appendix A.1.5 provided enough information to understand how the experiments were conducted and to get a general sense of the physical environment and equipment. The experiments explored are similar to simulator training tests. The EPRI experiments that were used to develop insights on time related probability distributions were performed using NPP simulators. With some exceptions, PIFs appear to have been largely controlled in the experiments. The following sections discuss how plant, task, and personnel factor related PIFs impacted the variability in time-required observed in the simulator experiments cited above.

Table B.1. Performance Influencing Factors in IDHEAS-G [1]

	Category			
	Environment and Situation	**System**	**Personnel**	**Task**
Performance Influencing Factors	**WAH.** Work Location Accessibility and Habitability **VIS.** Workplace Visibility **NOS.** Noise in Workplace and Communication Pathways **TEP.** Cold/ Heat/Humidity **PR.** Resistance to Physical Movement	**SIC.** System and Instrumentation and Control (I&C) Transparency to Personnel **HSI.** Human-System Interface **ETP.** Equipment and Tools	**STA.** Staffing **PG.** Procedures, Guidelines, and Instructions **TE.** Training **TOF.** Team and Organization Factors **WP.** Work Processes	**INF.** Information Availability and Reliability **SF.** Scenario Familiarity **MT.** Multi-Tasking, Interruptions and Distractions **TC.** Task Complexity **MF.** Mental Fatigue **TPS**. Time Pressure and Stress **PD.** Physical Demands

B.1.1 Plant-Related Factors

NUREG-2198 [1] on the general IDHEAS methodology identifies two context types that are considered to be encompassed by plant factors (i.e., environment- and situation-related PIFs and system-related PIFs). Environment- and situation-related PIFs are hazards, such as steam, fire, toxic gas, seismic events, or flooding, that can introduce environmental conditions that impede personnel performance. Harsh or abnormal environmental conditions can impact information detection, decision-making, motor activities, and team collaboration. The PIFs grouped into these two context types are shown in Table B.2 along with dispositions about how these PIFs are judged to impact the variability in the time-required from the examined experimental data. In general, no harsh or abnormal environmental conditions were introduced into the experiments. Furthermore, all crews were subject to the same environment and situation conditions that were controlled in the experiment to be

---

[2] For a list of PIFs and PIF attributes used in IDHEAS, see PNNL's Task 1 report [2]

normal (i.e., such as the environmental conditions found in an NPP control room). The PIFs related to environment and situation are more applicable to ex-control room actions. None of the experiments simulated those kinds of actions.

System-related PIFs are conditions associated with using I&C, HSIs to plant systems, and portable equipment and tools. An acknowledged drawback of using time-required data generated from NPP simulators is that there are some differences between operating a simulator and an actual NPP. However, all crews used the same simulator I&C and HSIs. The PIFs regarding equipment and tools are associated with the use of portable equipment and tools; therefore, they apply only to ex-control room actions, which–again–were not simulated in the data examined.

Based on examination of the EPRI experiments discussed in Appendix A.1.5, it is difficult to conclude that any of the PIFs listed in Table B.2 (or any combination of the PIFs listed) had a significant effect on the variation of time-required because plant factors are controlled to be as nominal as possible. Their impact on time-required variability tends to be minimized in the research results because all crews were exposed to the same plant factors throughout the experiment.

Table B.2. Impact of PIFs Associated with Plant Factors

Type of Context	PIFs	PIF Impact on Variability in Time Required
Environment and Situation	Work Location Accessibility and Habitability‡	Work location accessibility and habitability in the simulator are designed to be conducive to operations similar to a NPP control room. Moreover, all crews were subject to the same work location, accessibility, and habitability conditions. Accordingly, there is no expectation that this PIF would impact the variability of time-required between tasks observed in the simulator experiments.
	Workplace Visibility‡	Workplace visibility of the simulator is designed to be conducive to operations similar to an NPP control room. Moreover, all crews were subject to the same workplace visibility conditions. Accordingly, there is no expectation that this PIF would impact the variability of time-required between tasks observed in the simulator experiments.
	Noise in Workplace and Communication Pathways‡	Noise levels and communication pathways in the simulator are designed to be conducive to operations similar to an NPP control room. Moreover, all crews were subject to the same noise and communication pathway conditions. Accordingly, there is no expectation that this PIF would impact the variability of time-required between tasks observed in the simulator experiments.
	Cold/ Heat/ Humidity‡	The temperature and humidity environment in the simulator is designed to be conducive to operations similar to an NPP control room. Moreover, all crews were subject to the same temperature and humidity conditions. Accordingly, there is no expectation that this PIF would impact the variability of time-required between tasks observed in the simulator experiments.

Type of Context	PIFs	PIF Impact on Variability in Time Required
System	Resistance to Physical Movement‡	This PIF pertains to resistance to physical movement by forces such as wind, rain, and flooding. No ex-control room actions were simulated where resistance to physical movement might be encountered.
	System and instrumentation and Control (I&C) Transparency to Personnel‡	This PIF pertains to plant systems and I&C designs, which should operate and respond in a predictable way to the operators in various operating conditions. The transparency of the operation of plant systems and I&C is the same for all crews for a given experiment because all crews used the same simulators. Also, the simulators used in the experiments were the same as or similar to the operator's home plant control rooms for which they were trained and licensed. Additionally, there is no or minimal difference in the transparency of the operation of plant systems and I&C between tasks. Therefore, there is no expectation that this PIF would impact the variability of time-required between tasks observed in the simulator experiments.
	Human-System Interface (HSI) ‡	HSI pertains to the indications (e.g., displays, indicators, labels) and controls used by personnel to execute actions on systems similar to an NPP control room. However, within an experiment, the impact of HSI was the same for all crews because all crews used the same simulator. In the HAMMLAB experiments, there may have been differences between the simulator design and an operator's home plant, but the participating crew's home plants all used Westinghouse pressurized water reactor (PWR) designs. The EPRI NP-6937 experiments were performed using the operators' home simulator. No ex-control actions were simulated were simulated where the HSI may be different from the control room. Accordingly, there is no or minimal difference in the HIS between tasks. Therefore, there is no expectation that this PIF would impact the variability of time-required between tasks observed in the simulator experiments.
	Equipment and Tools‡	This PIF pertains to the use of portable equipment and tools which only applies to ex-control room action actions that are out in the plant or outdoors. However, none of the experiments simulated those kinds of actions.

‡ There was no or little impact from this PIF on the variability of time-required between crews or tasks observed in the simulator experiments

B.1.2 Task-Related Factors

NUREG-2198 on the general IDHEAS methodology identifies a third context type referred to as Task Factors in this report. The PIFs grouped into this context type are shown in Table B.3 along with assessments about how these PIFs are judged to impact the variability of time-required from the examined experimental data. In general, all crews were subject to the same scenarios and performed the same tasks within a given experiment and, therefore, were subject to the same task-related PIFs. However, staff or crews may vary in both their susceptibility to a particular PIF as well as their ability to manage team performance while under the influence of a given PIF.

Given the dispositions in Table B.3, it appears the impact of task-related PIFs on the variability of time-required across crews was mitigated in these experiments because all crews for a given experiment performed the same tasks. Therefore, for the examined experiments, it was concluded that the impact of task factors on time-required variability across crews are minimized in the experiments. However, crews and crew members may be affected differently by a given task-related PIFs. Accordingly, a task- related PIF could impact the variability of the time-required across crews differently for different kinds of tasks. For example, the time-required variability associated with complex or unfamiliar scenarios may be higher than for basic or familiar scenarios. However, this variability due to the impact of task-related PIFs cannot be determined by examining the time-required differences between crews because all crews perform the same tasks. The EPRI NP-6937 experiments and the other examined experiments were not set up to examine these kinds of impacts on time-required variability. An exception to that, are the HAMMLAB experiments that examined the difference for at least one PIF by using both complex and basic scenarios. However, as explained in Appendix A.1.4, there is in insufficient HAMMLAB data to inform the general development of probability distributions for time-required due to the disproportionately high number of complex scenarios (see Appendix A 2.1).

Table B.3. Impact of PIFs Associated with Task-Related Factors

Type of Context	PIF Impact on Variability in Time Required	
Task	Information Availability and Reliability‡	In certain accident scenarios, plant information from instrumentation, observations, or other operating staff may be incomplete, unreliable, untimely, or even incorrect or misleading. Information availability and reliability were designed to be conducive to operations similar to an NPP control room. However, within an experiment all crews were subject to the same scenarios and information availability and reliability conditions. Additionally, there is no or minimal difference in information availability and reliability between tasks. Therefore, there is no expectation that this PIF would impact the variability of time-required between scenarios observed in the simulator experiments.
	Scenario Familiarity§	Unfamiliar scenarios can pose challenges to crews in understanding the situation and making decisions which produces uncertainty. However, within an experiment all crews were subject to the same scenarios and all crews had, in general an equivalent level of training. However, there could be a difference in scenario familiarity between tasks, which could contribute to variability when comparing the time-required for low versus high scenario familiarity. For example, the personal experience of staff within a given accident scenario might impact time-required.
	Multi-Tasking, Interruptions and Distraction§	Tasks require multiple cognitive functions, such as detecting cues or parameters, assessing information, and mentally programming sequences of actions, personnel must frequently switch between these tasks during multitasking. Switching between tasks can make errors more likely. However, within an experiment all crews were subject to the same scenarios and tasks. However, there could be a difference in the need for multi-tasking or managing interruptions and distractions between tasks, which could contribute to variability. For example, different personal may have different learned or inherent abilities to perform under the influence of this PIF.

Type of Context	PIF Impact on Variability in Time Required	
	Task Complexity§	Task complexity (also referred to as “cognitive complexity”) creates demand for cognitive resources (e.g., working memory, attention, and executive control) that can overwhelm the personnel. However, within an experiment all crews were subject to the same scenarios and tasks, which has the effect of reducing variability. However, certain staff or crews may have an inherent ability or inability to manage task complexity, which contributes to time-required variability which has the effect of increasing variability between crews. However, there could be a difference in task complexity which could contribute to variability. For example, different personal may have different learned or inherent abilities to perform under the influence of this PIF.
	Mental Fatigue§	Mental fatigue can result from tasks that take extended time, non-routine tasks, and cognitively demanding tasks leading to loss of vigilance, difficulty in maintaining attention, reduced working memory capacity, and use of shortcuts. However, within an experiment all crews were subject to the same scenarios and tasks. However, there could be a difference in mental fatigue between different kinds of tasks, which could contribute to variability such as those that require long periods of mental concentration or that are tiring. Also, tasks were performed at the end of a shift could contribute to mental fatigue. Different personal may have different learned or inherent abilities to perform under the influence of this PIF.
	Time Pressure and Stress§	Time pressure refers to the perceived sense of time urgency to complete a task that can create psychological pressure affecting personnel performance. However, within an experiment all crews were subject to the same scenarios and tasks and, therefore, the same time pressures. That said, staff or crews may be impacted differently by time pressure, which contributes to time-required variability. However, there could be a difference in time pressure between different tasks which could contribute to variability. For example, different personal may have different learned or inherent abilities to perform under the influence of this PIF.
	Physical Demands‡	Physical demands indicate that a task requires extraordinary physical effort, such as twisting, reaching, dexterity, or strong force. This PIF is primarily applicable to ex-control room action actions that are out in the plant or outdoors. None the experiments simulated those kinds of actions.

§This PIF appeared to have little/no impact on variability associated with crews but could have an impact on the variability of time-required between tasks observed in the simulator experiments.

‡There was no or little impact from this PIF on the variability of time-required between crews or tasks observed in the simulator experiments

B.1.3 Personnel-Related Factors

NUREG-2198 on the general IDHEAS methodology identifies a fourth type of context, which is referred to as Personnel Factors in this report. The PIFs grouped into this type of context are shown in Table B.4 along with dispositions about how these PIFs are judged to impact the variability in time-required in the examined experimental data.

Given the dispositions in Table B.4, it appears that the some of the personnel-related PIFs have little or no impact on the variability in time-required in the examined experiments. It was concluded that impacts associated with staffing, procedures, guidelines, instructions, and training are controlled to be as nominal as possible in the experiments and that their impact on time-required variability tends to be minimized in the research results because there is no significant difference in crews for these factors. However, two of the PIFs were identified as having some potential impact on the variability of time-required (i.e., Team and Organization factors and Work Processes, which are identified as TOFs and WP in Table B.1 above). It is hypothesized that crew-level or team-level variability produce practices and biases that affect teamwork and impact time-required differently for different crews. The variability in teamwork is likely the product of several factors. The effect of these factors might be mitigated by training, but training is unlikely to eliminate all variability at the crew and team levels. Similarly, the variability in work processes is likely to be primarily influenced by the practices and culture of the crew's home plant but could also be influenced by individual crew-level and team-level characteristics.

Table B.4. Impact of PIFs Associated with Personnel-Related Factors

Type of Context	PIFs	PIF Impact on Variability in Time Required
Personnel	Staffing‡	Staffing refers to having adequate, qualified personnel to perform the required tasks. The experiments examined were staffed by licensed operators with the prerequisite training and experience representative of the nuclear power industry in general. The make-up of the crews was representative of NPP control room crews and was usually simplified by the exclusion of equipment operators to perform local actions ordered by the control room. Within an experiment, all crews were stated essentially the same way using licensed operators with experience commensurate with the nuclear industry. This was true for all tests within an experiment as well as for a single task.
	Procedures, Guidelines, and Instructions‡	This PIF refers to the availability and usefulness of operating procedures, guidance, instructions (including protocols). Within an experiment (in some cases across experiments) all crews used the same set of procedures and guidance. In general, this was true for all tests within an experiment, as well as for a single task. In the HAMMLAB experiments, there may have been differences between the simulator design and an operator's home plant, but the participating crew's home plants all used pressurized water reactor (PWR) procedures developed from the Emergency Response Guides by the Westinghouse Owner's Group and all the crews received the same simulator-specific training before the experiments. The EPRI NP-6937 experiments were performed using the operators' home simulator. None of the HAMMLAB or EPRI findings identified staffing as having a significant negative impact on operator error or time-required.

Type of Context	PIFs	PIF Impact on Variability in Time Required
	Training§	This PIF refers to the required training that operators receive to perform their tasks. Within an experiment, all crews were licensed operators at their home plant with the prerequisite training and experience representative of the nuclear power industry in general. Within an experiment, all the training across crews is judged to be essentially the same. None of the study's findings (i.e., EPRI and HAMMLAB studies as discussed earlier in this section) identified training as having a significant negative impact on operator error or time-required.
	Teamwork and Organization Factors¶	Teamwork factors include influences that impact team communication, coordination and cooperation and include planning, communicating, executing actions across individuals, teams and organizations. Depending on their individual crew member and group characteristics, different crews could have practices and biases that affect teamwork factors and impact time-required. Some of the study's findings identifieded team dynamics and team communication as having an impact on operator error or time-required.
	Work Processes¶	This PIF refers to work processes and conduct of operation and includes work controls and authorization and operating practices such as verification of task performance and attention to procedural guidance. Depending on the practices and culture of their home plants, different crews could have different work processes. Some of the study's findings identifieded work processes as having an impact on operator error or time-required.

§This PIF appeared to have little/no impact on variability associated with crews but could have an impact on the variability of time-required between tasks observed in the simulator experiments.

‡There was no or little impact from this PIF on the variability of time-required between crews or tasks observed in the simulator experiments

¶This PIF was judged to have an impact on the variability of time-required between crews observed in the simulator experiments.

B.2 Review of PIF Impacts in Available Data

While discussed in more detail in Appendix D, it is possible to generate a first-order normalized distribution from the EPRI-6937 time-required data by aggregating data across all HIs in the study. The first-order normalized distribution, which only accounts for crew-to-crew variability, can then be used in combination with other PIF information to create first-order time-required distribution for that HFE. This section describes the investigation into the possibility of using survey data from two studies that gathered PSF information (i.e., EPRI NP-6937L and the HAMMLAB studies) to adjust the first-order normalized distribution to account for the impact of PIFs and achieve the first-order time-required distribution. In addition, this section includes a discussion on the potential impact of some PIFs according to a literature review.

B.2.1 Review of PIF-Related Survey Data from EPRI NP-3937 Experiments

During the EPRI NP-6937L study [2], qualitative information about the impact of PSFs on time-required was collected from crew debriefings that used 19 multiple choice questions as shown in Table B.5. For example, Question 1.2 asked if they had difficulty using procedures (yes or no), Question 4.3 asked what type of supervisor they had (e.g., authoritative, or participative that is open to crew feedback), and Question 4.5 asked if the quality of communications was good, average, or poor. There is partial overage by the questions asked concerning PSF impacts over the 20 PIFs used in IDHEAS-ECAs that seemed worth investigating.

Boxplots were used to evaluate whether there is correlation of the time-required probability distributions to the responses to the PSF multiple choice questions. This was done across all nine of the HIs in the BWR-3 experiment, which was selected because it had the most comprehensive PSF questionnaire. Prior to producing these boxplots, the time-required data table was joined with the PSF questionnaire data table for all 18 crews that participated in the BWR-3 experiment. Time-required boxplots were plotted on a logarithmic scale so that the time-required probability distributions could be compared for slower and faster completion times (i.e., time-required to complete) across all nine HIs for each question.

Boxplots were used (see Figures B.2 and B.3) to visually evaluate whether the crew responses to Questions 3.3 (i.e., did the crews use non-procedural job aids) and 2.3 (i.e., if simultaneous communication occurred) were correlated to the time-required probability distributions. Figure B.2 shows that the time-required probability distributions were slower for crews that used non-procedural job aids in four of the HIs (B3-1-1 through B3-1-5), but the opposite was true for two different HIs (B3-1-4 and B3-2-3). Figure B.3 shows no clear correlation between different frequencies of simultaneous communication (i.e., none, rarely, frequently, seldomly, occasionally) and time-required probability distributions.

It is concluded that there is not enough evidence to support a correlation between use of non-procedural job aids or frequency of simultaneous communications and time-required probability distributions. This same conclusion extends to the other 17 multiple choice PSF questions that showed similar results.

A second approach was also performed to evaluate if the multiple choice PSFs impacted time-required probability distributions. This was done by showing the results of all 19 multiple choice questions for select HIs that were paired by the EPRI authors. For example, Figure B.4 shows the paired results for HIB3-1-1 (Initiation of Alternate Rod Injection) and HIB3-1-2 (Initiation of Standby Liquid Control System) in an anticipated transient without "SCRAM"[3] scenario. The plots appear to indicate that there is a correlation between time-required probability and PSFs to some but not all responses. For example, the question called "Interface 3.3 Use Non-Procedural Job Aids" shows that "Yes" responses appear to be correlated with slower completion times than "No" responses.

Similarly, the plot labeled "Teams Communicate 4.5 Quality of Communications" in Figure B.4 shows that "Good" responses have an average completion time that is slower (higher) than "Average" responses, which is counter-intuitive because good communications are expected to lead to better results (i.e., shorter time-required). However, there may be scenario-based considerations that can account for good communication leading to longer time-required. To model this as a universal effect may lead to unintended and erroneous predictions of time-required. However, most plots do not reflect a discernible or statistically relevant correlation between answers to the questionnaire and time-required.

---

[3] SCRAM is the term used in the nuclear field for rapid emergency shutdown of a nuclear reactor.

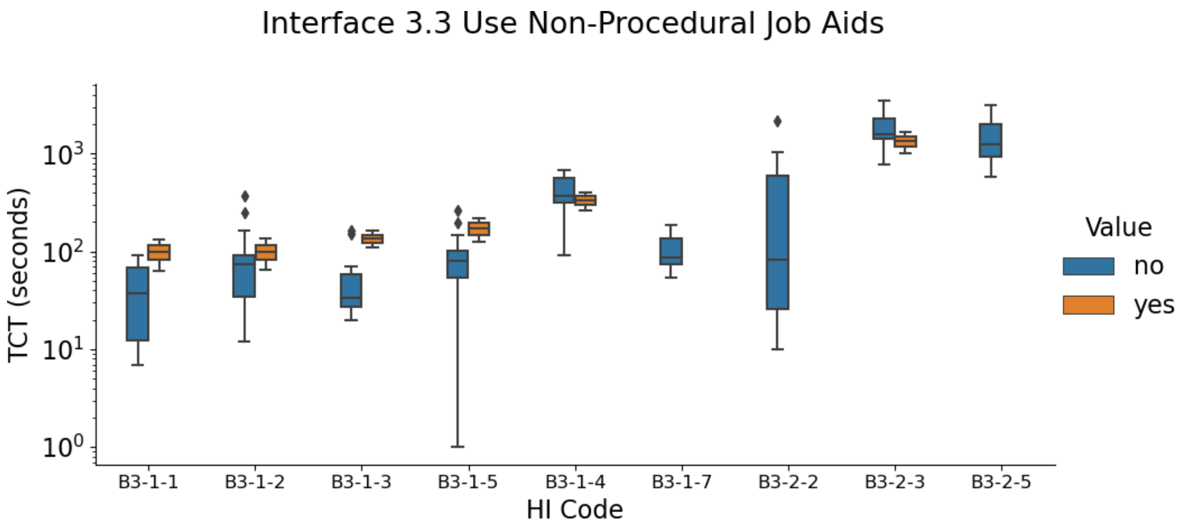


Figure B.2. **Boxplots of Time Required Probability Distributions (log scale) Across All Nine HIs in the EPRI BWR-3 Experiment for Question 3.3 Regarding Use of Non-Procedural Job Aids. TCT: Task Completion Time.**

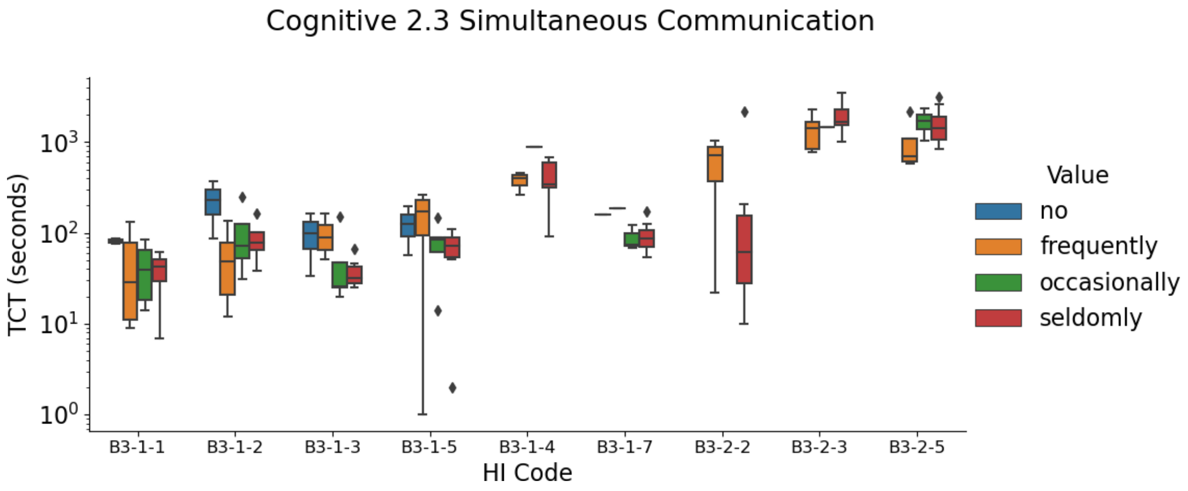


Figure B.3. **Boxplots of Time Required Probability Distributions (log scale) Across All Nine HIs in the EPRI BWR-3 Experiment for Question 2.3 Regarding Simultaneous Communications. TCT: Task Completion Time.**

Table B.5. **Observer Questions Reported in EPRI NP-6937, Volume 3, Section C.4, Table C.4-11 [2]**

ID	Topics and Questions
1	Procedures
1.1	Were procedures used? (Yes or No)
1.2	Did the crew exhibit any difficulty in following procedures? (Yes or No)
1.3	Were any modifications/deviation in procedures made to respond? (Yes or No)
1.4	Did procedures allow/require interpretation (e.g., priorities of actions, timing of actions, equipment selection)? (Yes or No)
2.1	Which best describes crew response? (a). Crew response required simultaneous actions by multiple crew members much of the time with coordination by supervisor, (b). Crew response required simultaneous actions by multiple crew members a small part of the time, and (c). Crew response did not require multiple crew members to act simultaneously.
2.2	Did the crew consider more than one possible response (including modifications, deviations, interpretations of procedures)? (Yes or No)
2.3	How often did simultaneous communication occur? (Frequently, Occasionally, or Seldom)
2.4	Was the amount of information gathered by the crew: (Low? Medium, or High?)

ID	Topics and Questions
3	Human-Machine Interface
3.1	Did the crew exhibit difficulty in responding because of the human-machine interface? (Yes or No)
3.2	Were there features of the control room interface that slowed down the crew response? (Yes or No)
3.3	Did the crew use any job aids other than procedures that facilitated its response? (Yes or No)
4.1	How often did crew members other than the supervisor contribute ideas regarding what was going on and what the crew should do? (Frequently, Occasionally, or Seldom)
4.2	How many members of the crew contributed ideas (a). All, (b). More than one, less than all, (c). One)
4.3	Which best describes the crew’s decision making? (a). Authoritative - The supervisor exhibited strong authority, directing and coordinating all activities, and seldom, if ever, requesting opinions of other crew members., (b). Participative - The supervisor directed and coordinated crew activities, obtaining opinions from other crew members, allowing crew members to make appropriate independent actions, and (c). Consensus - The supervisor coordinated crew activities based on a consensus of opinion of crew members, (d) Diffused - Decision making responsibility was diffused throughout the crew
4.4	How often did the crew discuss what was going on and where the response was going? (Frequently, Occasionally, or Seldom)
4.5	The quality of communications among crew members was (a). Good, (b). Average, (c). Poor
5.1	Was the onset of the problem sudden? (Yes or No)
5.2	Did the crew have to respond immediately and quickly? (Yes or No)
5.3	Were equipment/systems the crew would normally use in responding non-functional? (Yes or No)

Figure B.5 displays similar information to that shown in Figure B.4 but for another pair of HIs addressed in the BWR–3 experiment (i.e., HIB3-1-3 (Initiation of suppression pool cooling) and HIB3-1-5 (Stabilization of the power oscillation with the Safety Relief Valves) in an anticipated transient without an anticipated transient without a SCRAM scenario. The same general trend that is observed above in the plot presented in Figure B.2 also can be observed in the plots presented in Figure B.3. In addition, a few other results that seem non-intuitive merit discussion. For Question 5.1 (“Was the onset of the problem sudden?”), the plot in Figure B.5 shows that the answer “Yes” appears to be correlated with a significantly faster completion time (time-required) than the answer “No.” This result also may be viewed as counter-intuitive because the problems that occur suddenly might be expected to take more time to resolve than those that appear slowly and thus allowing better recognition of the situation. However, there may be scenario-based considerations that can account for the fact that problems that arise suddenly led to faster time-required in this experiment. Therefore, modeling this as a universal effect also may lead to unintended and erroneous predictions of time-required.

For Question 4.2 — “How many members of the crew contributed ideas?: all, more than one but less than all, or one?” — the plot in Figure B.5 (fourth row, fourth plot from the left) indicates that the answer “More than one but less than all” led to the longest time-required, the answer “One” led by a significant margin to the shortest time-required, and the answer “All” led to a wide range in completion times. There may be crew- or task-specific consideration related to this experiment that accounts for this result, which cannot be generalized. Though some of the results presented in the plots seem intuitive (e.g., the results for Question 1.2 on difficulty following the procedure and Question 5.3 on non-functioning equipment or systems), other results were as non-informative as the results to Question 4.2.

Concerning researcher efforts to distill insights about intra-plant PSFs, the EPRI NP-6937L report states that “no overall strong correlations were found,” and because “the indications were inconclusive, the datasets for other plants were not subjected to complete analysis.” For inter-plant PSFs, the report states that they may affect the median response time but not the response variability; however, the report also notes that further work is required as detailed examination was only performed for one HI.

Therefore, given the conclusions about the PSF questionnaire data from the EPRI researchers and the investigation described above using Figure B.2 through Figure B.5, it appears that data from the observer questionnaire on PSFs cannot be used (at least by itself) for refining or adjusting the first-order time-required probability distributions to account for the impact of PIFs even though the EPRI questionnaire addressed several factors pertinent to the PIFs defined by the IDHEAS methodology.

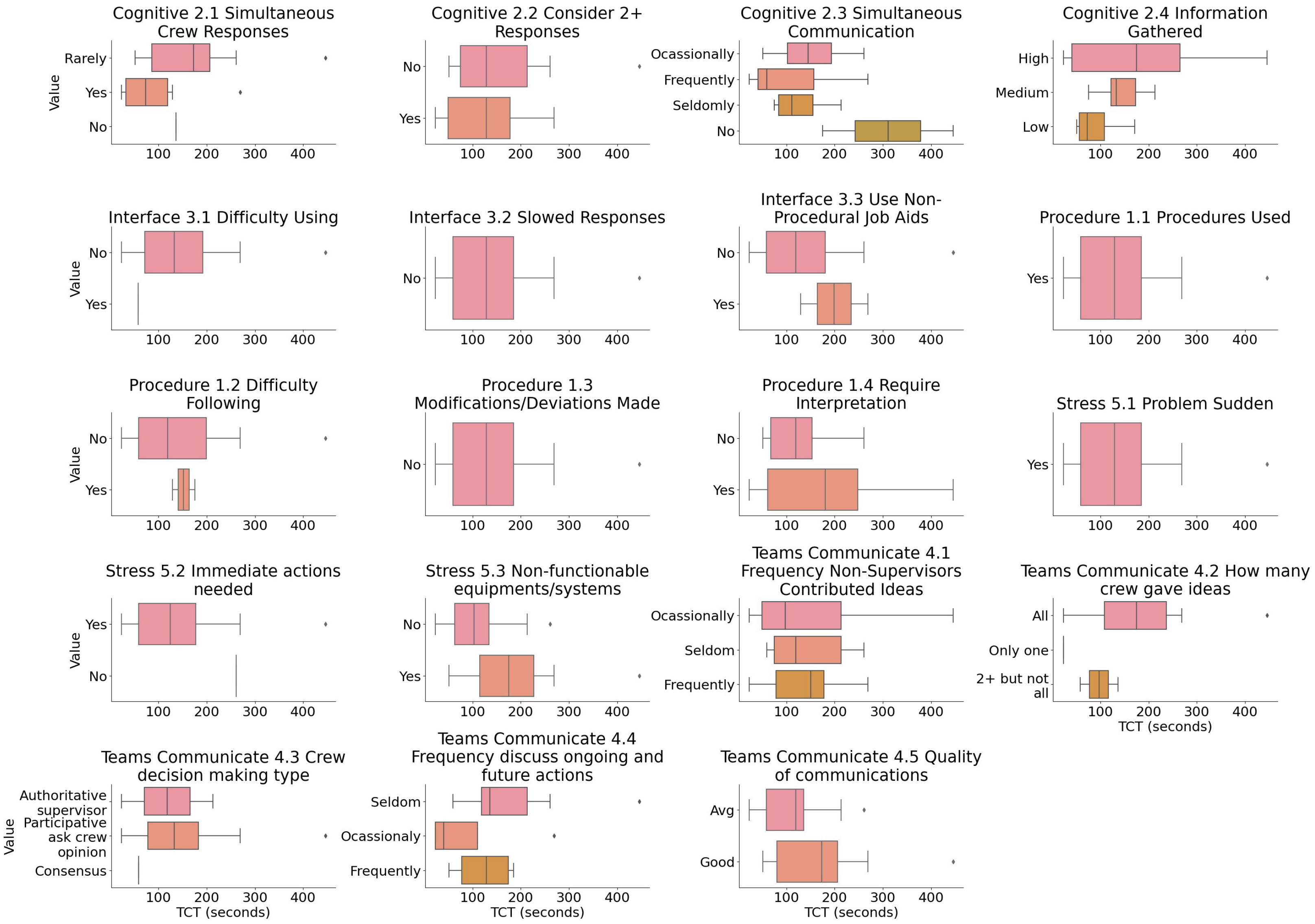


Figure B.4. Hyman Factors Evaluations for BWR-3 His B3-3-1 and B3-1-2

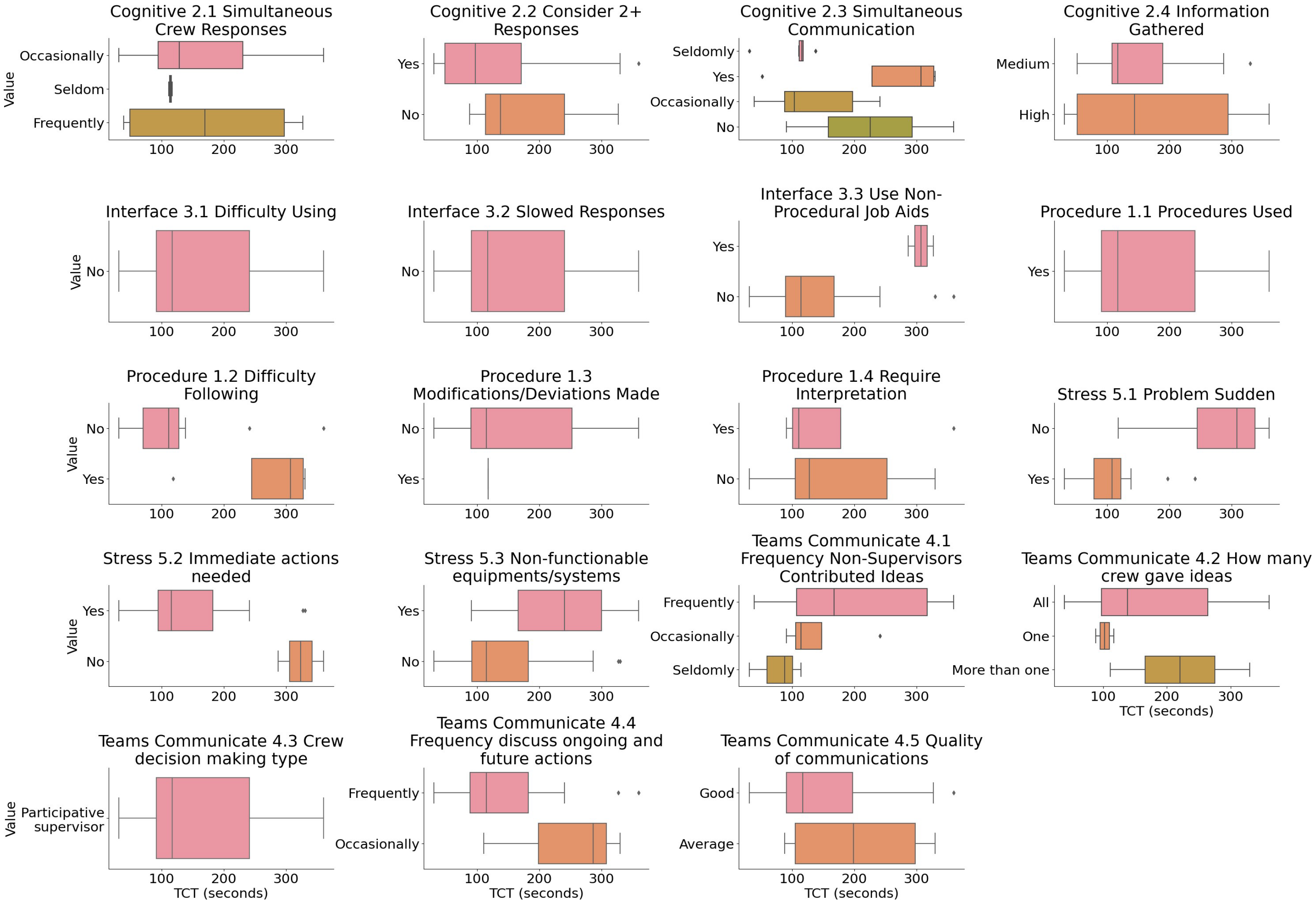


Figure B.5. Human Factors Evaluations for BWR-3 His B3-1-3 and B3-1-5

B.2.2 Review of PIF-Related Survey Data from HAMMLAB Experiments

In addition to the EPRI study questionnaire results on PSFs, PSF information collected during HAMMLAB experiments, discussed in Appendix A.1.4, was considered. The objective of the HAMMLAB experiments was to develop an empirically based understanding of the performance, strengths, and weaknesses of different HRA methods by comparing HRA predication to actual crew performance on the HAMMLAB simulators. Although very limited in the number of scenarios and operator actions tested, these experiments did result in some time-required data and qualitative and semi-quantitative information collected by observers about the impact of PSFs. The observers provided positive and negative scores about the general impact of PSFs on operator performance perceived by the observer. The scores ranged from -1.5 to +2.0 and other than being cited as a quantitative indication of the impact on the PIF they are not further defined in the HAMMLAB reports. The scores do not differentiate the impact on performance from time exceedance and cognitive errors. Table A.3 shows the consolidated semi-quantitative results recorded by the observers for the International HRA Empirical Study [3-5]. There is partial overage by the PSF impact addressed over the 20 PIFs used in IDHEAS-ECAs that seemed worth investigating.

In general, the results of the observer scores as displayed in Table A.3 show that a few PSFs had minimal to no impact on operator performance (i.e., time pressure, experience, and human-machine interface). A few other PSFs had consistently (or nearly consistently) positive impact (e.g., training and procedural guidance). Several PIFs were judged to have both a negative and positive (or neutral impact) for several of the same HFEs. Several PSFs were judged to have negative and positive impacts on different HFEs. Given that the semi-quantitative information was provided on operator performance in general, it is difficult to separate out the contribution to the impact to time-required. Moreover, the impacts appear to be consistently minimal (mostly “0s”), non- informative (i.e., training was scored as having a positive impact of “1” for every HFE), or mixed (i.e., having the opposite impacts for different crews on the same and different HFEs).

Therefore, as in the assessment of the EPRI NP-6937L observer scores, it appears that the PSF data from the HAMMLAB studies cannot be used (at least by itself) for refining or adjusting the first-order time-required probability distributions to account for the impact of PIFs. Observers attempted to capture information about the impact of PSF on operator performance, but this effort was focused primarily on operator error opposed to completion time and only provided partial coverage of the 20 PIFs in IDHEAS-ECA. Also, the impacts were characterized qualitatively (or semi-quantitatively at best) and produced non-informative or mixed results in many cases.

B.2.3 Use of Available Literature on PIF Impact to Inform Adjustment of First-Order Distribution

This appendix presents PNNL’s evaluation of applicable literature on the impact of PIFs on the prob- ability distribution of time-required. Studies were selected based on the literature review in Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA) [6], and priority was given to studies that include NPP actions (e.g., simulator studies). Each table presents some considerations associated with each PIF that provide the analyst with some guidance about how that PIF might impact time-required and presents an example of relevant research is presented to illustrate the possible impact of that PIF on time-required.

Studies presented in Appendix B were selected based on the literature review in Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA) [8]. These studies were used as a basis for this effort because they represent a reasonably comprehensive set of work that explores human performance as influenced by PIFs. Studies were selected that, in addition to including information on rate of error, also included information about time to perform an action, including reaction time, response time, or task completion time. Finally, studies that focus on NPP actions (e.g., simulator studies) were prioritized for inclusion. Each table presents some considerations associated with each PIF that provide the analyst with some guidance about how that PIF might impact time-required. In addition, an example of relevant research is presented to illustrate the possible impact of that PIF on time-required. For each study included, the tables present a summary of the task completed by study participants, the findings, and the potential implications of those findings for time-required for task completion.

Most of the PIFs examined have a derogatory impact on task completion time—that is, task completion time is increased under the influence of the PIF. For example, task complexity has a well-documented effect of increasing the time-required to complete an action [7, 8]. However, as the results show, the impact of the PIF on overall human performance may be more complex. For example, there is a well-known trade-off between speed and accuracy; a meta-analysis showed that time pressure or stress moderately increased reaction time or speed, but it also decreased accuracy [9]. In another example, interruptions resulted in decreased completion time, but increased errors [10]. The results in the appendix, therefore, attempt to provide sufficient context to describe the overall impact of the PIF observed in the study. Nonetheless, the findings should be used only to inform analyst judgment rather than to serve as a ready-made solution for estimating the impact of the PIF on time-required for task completion.

There are a number of additional considerations when exploring the literature for applicable research for adjusting the time-required to perform human actions. First, many studies in the human performance literature that examine the impact of factors on response time or reaction time (rather than task completion time). As a result, the impact of the PIF is often measured in milliseconds between different conditions (e.g., milliseconds faster or slower based on a specific PIF). It is not clear how such differences might generalize to the longer-term activities that occur within NPPs; certainly, the size of the impact may not scale in a linear fashion. To address this concern, when possible, the review of relevant literature in Appendix A prioritized operational and simulator data in the nuclear domain, meta-analytic studies of the human performance literature, or studies deemed to be as similar as possible to the nuclear context. As a result, the number of studies is limited due to a lack of available literature on certain PIFs. Nonetheless, analysts can review these studies to inform their judgments of the time-required based on the presence of the PIF; the size and strength of the impact should, however, be made based on the analyst’s experience and expertise.

Additionally, most studies explore the mean-level impacts of PIFs on the time-required to complete an action; information about the impact of the PIFs on the uncertainty of that estimate was much less widely available. Whenever possible, the tables in Appendix B.2 include any figures available for a specific study to illustrate the impact of that PIF on the probability distribution. However, many studies lack such figures and present only limited information about the probability distribution of time-required with and without the PIF’s impact. As a result, most of the tables in Appendix B.2 present information on the mean-level impact of the PIF only, and do not include information about the impact of the PIF on the shape or spread of the time-required probability distribution. Appendix B.2 does, however, provide an example of how past data might be leveraged for understanding the impact of PIFs on the probability distribution of time-required using the example of task complexity.

Ultimately, as already mentioned, these papers help to provide some general information about how a specific PIF might impact the probability distribution of time-required. However, that information is limited in nature and should be used only to inform analyst judgments as they leverage their expertise to estimate the impact of PIFs on the probability distribution of time-required.

While most studies include only the mean-level impacts of PIFs on the time-required to complete an action, they do not discuss changes in variability. To help address that gap, the impact of one PIF on the probability distribution of time-required was explored in five datasets described in greater detail in Appendix A.

- U.S. HRA Data Collection: A HAMMLAB HRA Data Collection with U.S. Operators, 2016, HWR- 1123, Organization for Economic Co-operation and Development (OECD) Halden Reactor Project [11].
- HAMMLAB International Empirical Study: Results from comparing HRA methods predictions to HAMMLAB simulator on SGTR scenarios in Phase 2 Report [3] and LOFW Scenarios in Phase 3 Report [4]
- UJV Data (Nuclear Research Institute of the Czech Republic): UJV provided PNNL with unpublished time-required data for 17 crews that participated in LOFW simulator trials that were held in 2011 [12]. Two of the crews committed errors so that information had to be removed from the dataset. The simulators were of a Russian designed water-water energetic reactor (VVER) PWR reactor and the LOFW scenario required that they isolate coolant from the ruptured half of the VVER feedwater collector and transfer flow over to the “healthy” unruptured half.
- HAMMLAB U.S. HRA Empirical Study: The U.S. HRA Empirical Study – Assessment of HRA Method Predictions against Operating Crew Performance on a U.S. Nuclear Power Plant Simulator. NUREG-2156 [13].
- Analysis of Human Performance Observed Under Simulated Emergencies of Nuclear Power Plants, 2005 – KAERI/TR-2895/2005, Korea Atomic Energy Research Institute [8].

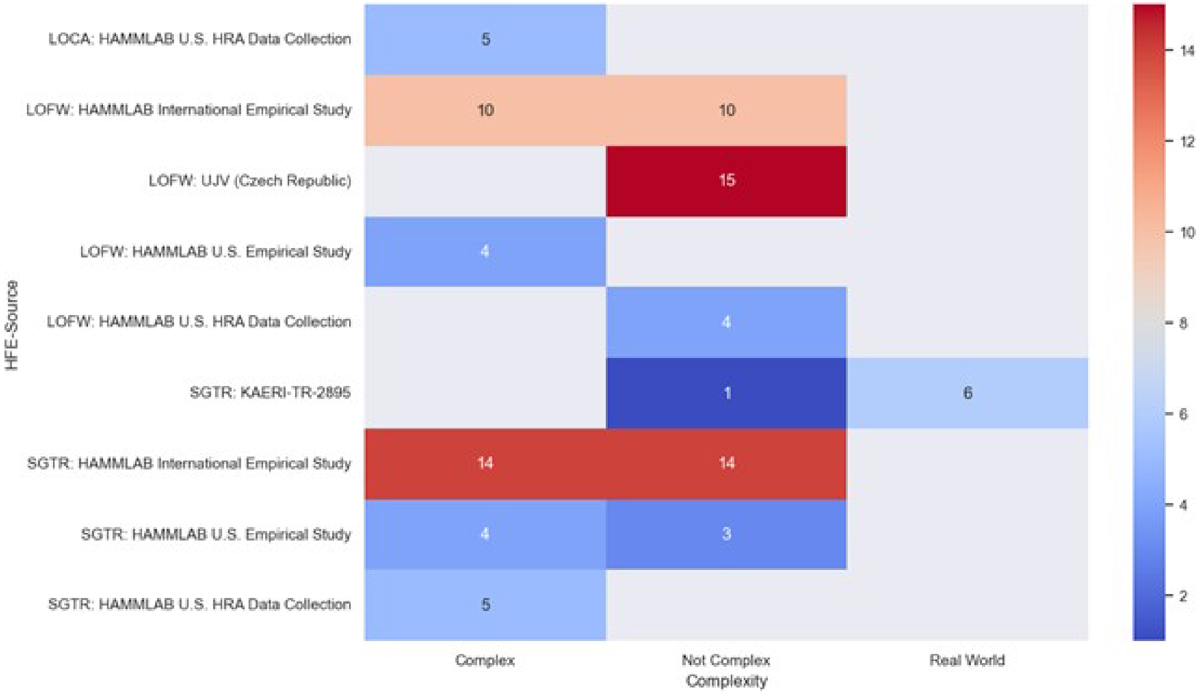


Figure B.6. Number of Records for Data Sources and HFE

Figure B.6 shows the number of records available for each data source for different HFEs as well as the level of complexity associated with the HFE (i.e., complex, not complex, or real-world, meaning that the data was operational rather than simulator data).

To examine the impact of task complexity on the probability distribution of time-required, the means and standard deviations for task completion time were examined for each data source and scenario by task complexity. Figure B.7 shows the mean task completion time in seconds; Figure B.8 shows the standard deviation of task completion time. As the figures show, time-required was generally lower when the task was less complex. There is some indication that the variability may also be lower for less complex tasks; however, given the variability across the different data sources and tasks, the impact on the distribution is less learn. Figure B.9 shows histograms of the task completion time for by complexity for each of the data sources.

This analysis provides one example of how existing operational and simulator data in the nuclear domain could be examined to inform on the impacts of PIFs on the time-required probability distribution. Such analyses can provide analysts with information about how to adjust the time-required probability distribution based on the impact of PIFs. However, as the results show, the impact of PIFs on the probability distribution is not uniform and additional data are required to understand the changes that might occur to the shape of the probability distribution given the presence of PIFs.

B.2.4 Conclusions about the Impact of PIFs on the Variability in Time Required from Simulator Data

Plant, task, and personnel factor related PIFs have differing impact on the variability in time-required observed in the simulator experiments cited above. With some exceptions, it appears that PIFs were largely controlled in the experiments that generated the examined time-required data. In simulations plant factors are typically designed to be conducive to operations like a NPP control room. Also, all crews are tested on the same scenario and in-control-room tasks, using the same simulator controls and procedures, under the same environmental conditions. Tasks for which the operating system, or equipment and environment could be different from an NPP control room are not tested in NPP simulators. Therefore, all PIFs except two (i.e., teamwork and work processes) have no or little impact on the variability of time-required between crews in the observed datasets. This variability is assumed to be inherent to time-required distributions. The variability due to these PIFs in simulator experiments is referred to as crew-to-crew variability.

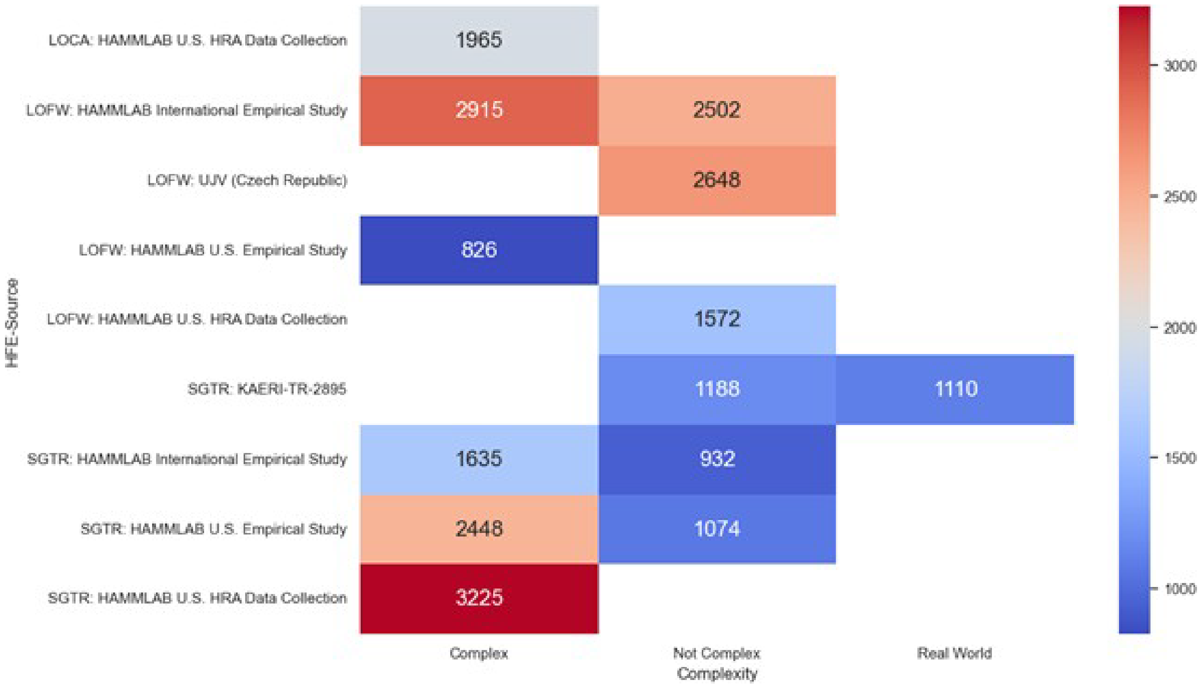


Figure B.7. Mean Task Completion Times (in second)

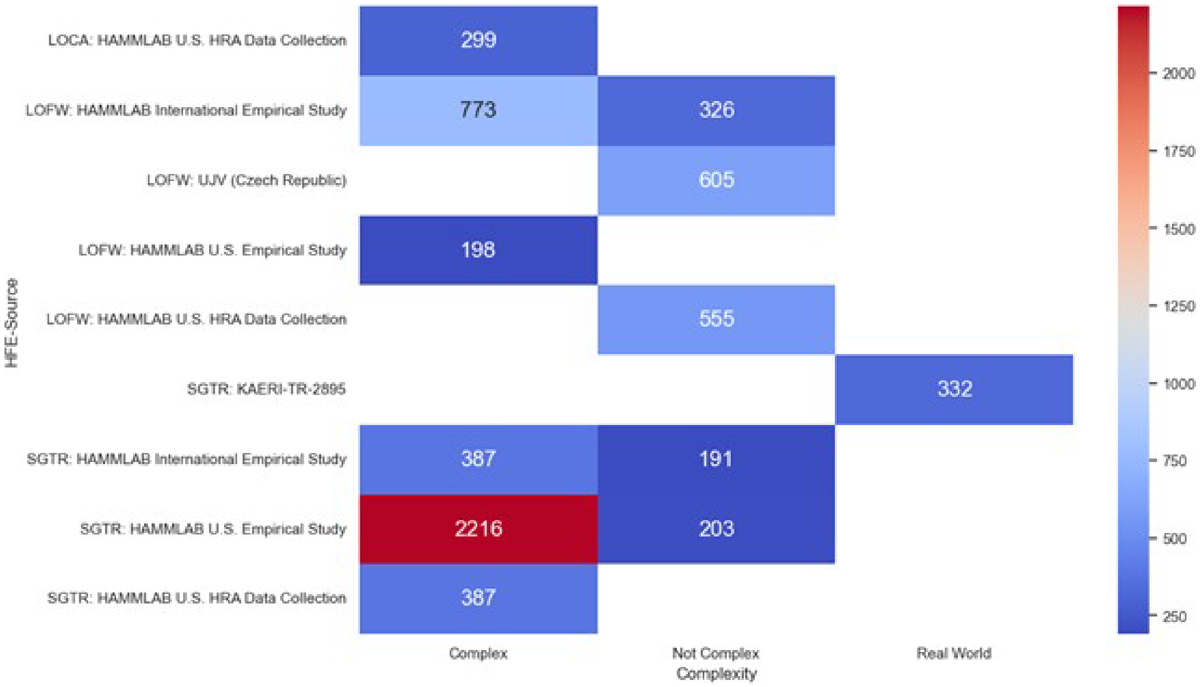


Figure B.8. Standard Deviation of Task Completion Times (in second)

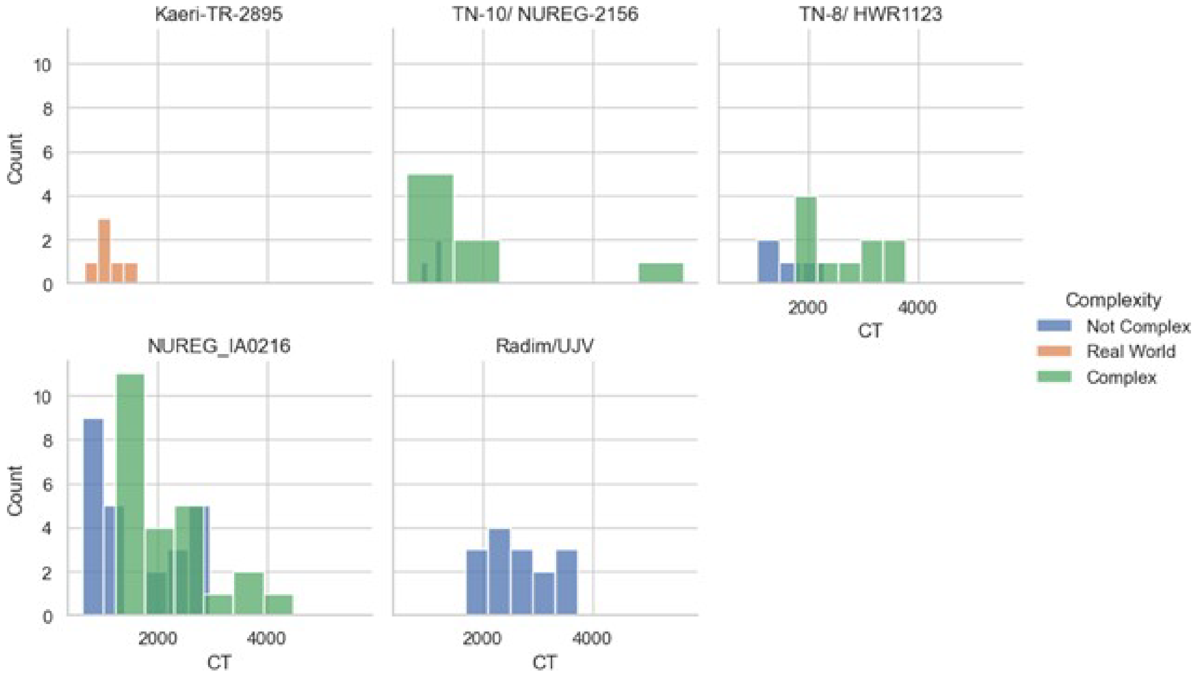


Figure B.9. Histograms of Time Required by Task Complexity and Data Source

It is also observed, based on the discussions in the sections above about PIF impact, that many PIFs do not change between tasks and scenarios addressed in an experiment. This includes PIFs associated with system, and environmental and situation related factors as well as PIFs associated with certain personnel or task related PIFs. For example, the environment and habitability PIF does not change between tasks or scenarios given that no ex-control room actions were included in the experiments where the location, accessibility and habitability could be different. Likewise, within an experiment staffing qualification, training, and kinds of procedures used did not change between tasks of the same experiment. The exceptions to PIFs that do not change between tasks are many of the task related PIFs. Task-related PIFs can change between tasks, particularly tasks associated with different scenarios. These include PIFs such as scenario familiarity and task complexity. However, the impact on variability from these PIFs cannot be determined by examining the time-required differences between crews, because all crews perform the same tasks. Moreover, with very limited exception, time-required data from experiments like EPRI NP-6937 were not set up to examine the range of task-related PIFs that could potentially impact completion time (i.e., scenario familiarity, multiple-tasking and distractions, task complexity, mental fatigue, time pressure and stress).

Additionally, it appears that neither the EPRI NP-6937L or HAMMLAB survey data on the impact of PSFs can be easily used to make consistent quantitative assessments of time-required probability distributions (be used to adjust the “first-order” distribution to account for PIFs not already accounted for). The information gathered in both sets of studies was primarily qualitative or semi-quantitative at best. For the EPRI NP-6937L survey data on the impact of PSFs, there was only a limited relationship between the survey questions and 20 PIFs used in IDHEAS-ECA and no consistently strong relationships between PIF effects and time-required were identified by either the original researchers or by PNNL. Again, for the HAMMLAB survey data on the impact of PIFs produced non-informative or mixed results in many cases and were not focused specifically on time-required but rather general operator performance. In general, the impact of PIFs on time-required variability is believed to be significantly minimized because in simulator studies many factors are designed to be conducive to operations and the crews were tested on the same tasks using the same simulators and in the same environment.

Given the observations above, it appears that the variability in time-required data from simulator studies such as EPRI NP-6937 experiments is largely due to crew-to-crew variability. The existence of crew-to-crew variability has been previously noted by researchers in the past as mentioned earlier in this report. For example, Appendix D.3.1.2 of EPRI NP-6937 specifically acknowledges and shows quantitatively that certain crews are inherently faster or slower than other crews across different tasks. Therefore, the next section discusses crew-to-crew variability and how to model it within a time-required distribution framework.

B.3 Modeling Crew-to-Crew Variability

Based on the qualitative assessment of impact of the 20 IDHEAS PIFs on the time-required based on simulator experiments as explained above, two personnel-related PIFs—Teamwork and Organizational Factors, and Work Processes—were identified as having the primary impact on time-required variability between crews. Other PIFs could also impact time-required in actual NPP tasks but are not reflected in the simulator data on time-required variability and are addressed later separably in this report. Accordingly, this section discusses research on teamwork (and by extension works processes by teams), how teamwork might be modeled, and its role in the development of time-required probability distributions.

B.3.1 General Literature on Teamwork

The general literature on teamwork has identified several factors that impact team performance (see [12] for a review). Several models of teamwork have been proposed to show the complexity of the relationship between these factors. For example, Salas et al. [15] proposed a model that suggests how five core components of teamwork and three supporting coordinating mechanisms may interact with each other to affect teamwork and lead to its variability across teams (see Figure B.10).

Although standardized training may reduce some of this variability, a significant amount of variability across teams will likely remain. The teamwork factors addressed in this model include core components (*Team Leadership, Mutual Performance Monitoring, Backup Behavior, Adaptability*, and *Team Orientation*) and coordinating mechanisms (*Closed Loop Communication, Shared Mental Model*, and *Mutual Trus*t). The following parts discuss teamwork factors presented in Figure B.10 that illustrate the complexity of the model and the potential difficulty in applying it using the time-required data that is available.

Team Leadership. In Salas et al. [15], team leadership was identified as a core component of team- work. Team leadership influences team effectiveness by setting the standards of performance that include both mutual performance monitoring and backup behavior. This proposition is expressed in the model as P1.

Mutual Performance Monitoring. One of the core components affected by team leadership—mutual performance monitoring—does not affect team effectiveness directly. This factor impacts performance indirectly by helping the team identify when and where backup behavior is needed. Team members who monitor each other’s performance recognize when an intervention is needed and act by providing backup to maintain (or improve) team performance. P2 in the model proposes this indirect relationship between mutual performance monitoring and team effectiveness.

Shared Mental Models. For mutual performance monitoring and backup behavior to be effective, the team must have shared mental models. Shared mental models are important for team members to be able to determine if another team member is struggling with their task or experiencing high workload. If team members do not share a model of each other's tasks and responsibilities, they cannot identify if their team members' task performance begins to slip.

Mutual Trust. Another coordinating mechanism that is necessary for effective mutual performance monitoring is mutual trust. Each team member must trust that their team is monitoring everyone's performance for the benefit of the team and not to single out or expose individual failures. The proposed role these two coordinating mechanisms play in mutual performance monitoring and backup behavior is expressed as P3 and P6 in the model, respectively.

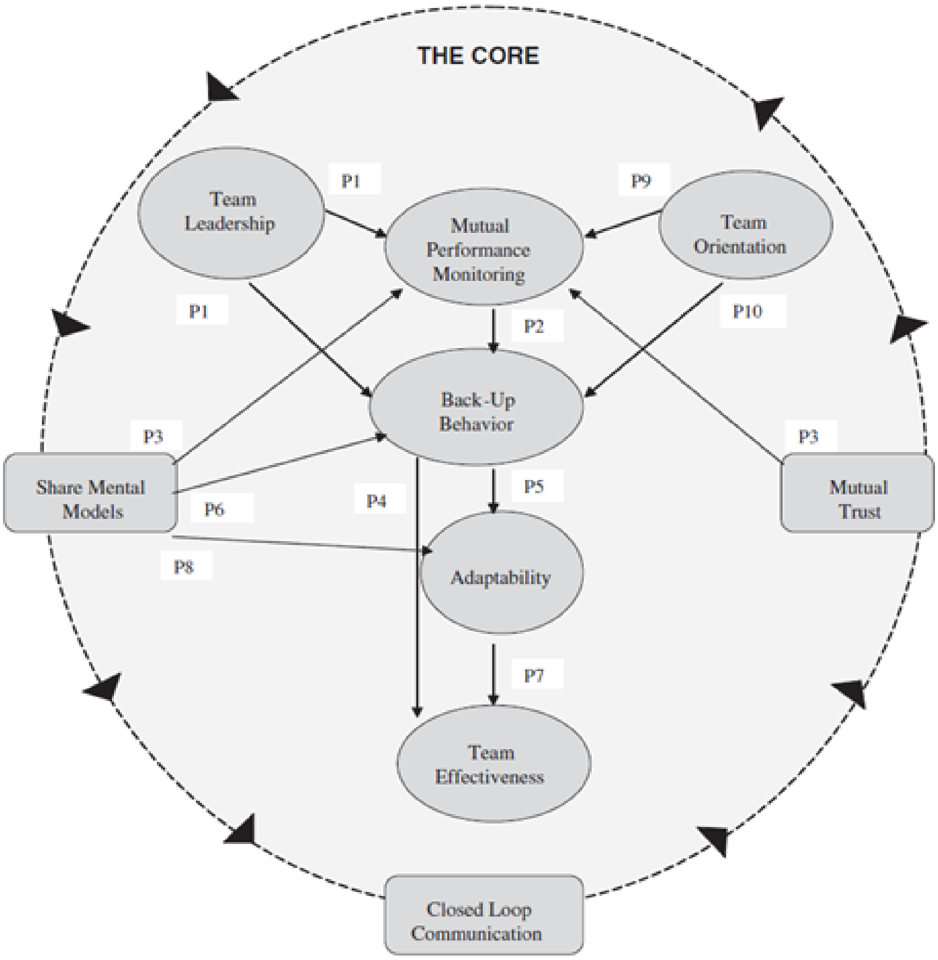


Figure B.10. Model from Salas et al. [15] Illustrating High-Level Relationships Among Core Components and Supporting Coordinating Mechanisms that Influence Teams

Team Orientation. Team orientation also influences the effectiveness of mutual performance monitoring and backup behavior. Team orientation is based on the attitudes of the individual team members. Teams that are high in team orientation have members who prefer to work with each other and improve individual performance with coordination and evaluation. According to the Salas et al. [15] model, teams high in team orientation are more willing to engage in mutual performance monitoring and more accepting of backup behavior from fellow teammates. These propositions are expressed in P9 and P10 of the model respectively.

Backup Behavior. Unlike mutual performance monitoring, backup behavior directly affects team effectiveness. This behavior can improve team effectiveness by making sure all necessary team tasks are performed. Backup behavior involves providing constructive feedback to other team members to improve performance, supporting team members with their task responsibilities, and completing another team member's tasks when necessary. Backup behavior often is needed in high workload situations when the demands of the task outweigh a team member's ability. This direct relationship between backup behavior and team effectiveness is expressed as P4 in the model. In addition, the model suggests that part of backup behavior's benefits may be in allowing greater team adaptation to workload demands. This relationship is identified in P5.

Adaptation. The model reported by Salas et al. [15] suggests that adaptation plays a direct role in team effectiveness. Adaptability can be characterized as a multi-step process that includes recognizing the need for change (e.g., noticing a decline in performance), diagnosing the problem, identifying the correct adjustments, and executing those changes. The model proposes that a team's ability to adapt will directly impact its effectiveness (P7). Successful adaptations require all the relevant team members to adjust their behavior in concert. In P8 of the model, Salas et al. [15] suggest that the team can benefit from shared mental models to execute this coordinated activity.

B.3.2 Effects of Teamwork on Task Completion Time

As mentioned above, not only do the listed factors differ across teams, they also likely interact with other task-related factors such as Task Complexity, Scenario Familiarity, Mental Fatigue, Time Pressure, and Stress to cause differences in time-required across crews. For example, the PIF, Task Complexity, may interact with some subset of the teamwork components, such as Mutual Performance Monitoring (the tendency of team members to monitor each other's performance) and Backup Behavior (assisting other team members with their tasks when needed), that vary across teams. One might hypothesize that strong Mutual Performance Monitoring and Backup Behavior in teams may act as a buffer against performance degradation due to Task Complexity. In less complex tasks these teamwork components do not play a significant role in impacting time-required. However, when faced with complex tasks, strong Mutual Performance Monitoring and Backup Behavior are needed to keep the team functioning efficiently.

One useful way to distinguish Teamwork from other PIFs is to categorize Teamwork as aleatory uncertainty while other PIFs can be considered epistemic uncertainty. Aleatory uncertainty in crew behavior leads to variability including differences in time-required across crews and within a particular crew across different scenarios. For the purposes of this project, Teamwork (and to some extent Work Processes because it is integral to Teamwork) is treated as aleatory uncertainty. Even though there may be conceptual models that explain Teamwork [13], they are too complex to apply and require specific information about crews and individual crew member characteristics that are very difficult for an HRA analyst to obtain. In contrast, the other PIFs in the IDHEAS framework can be considered to reflect epistemic uncertainty. By controlling and/or measuring the other PIFs in the IDHEAS model, epistemic uncertainty from these sources of can be reduced.

HRA methodologies recognize that PIFs can interact (i.e., they are not independent of each other; see [16, 17]). Considering the interaction between Teamwork and task related PIFs adds additional complexity. Certain crews or crew members may be impacted differently by a given task related PIF and/or vary in their ability to manage a given task related PIF as described above. Accordingly, the examined experiments revealed combinations of some of the task related PIFs and Teamwork appear to impact time-required variability. These combinations should also be treated as aleatory uncertainty because the researchers do not have the model to explain the aleatory component (i.e., Teamwork) of these interactions.

B.3.3 Modeling Crew-to-Crew Variability Time Required Distributions

One useful way to address crew-to-crew variability is to distinguish Teamwork from other PIFs and to categorize its uncertainty as aleatory uncertainty while other PIFs can be considered epistemic uncertainty. Aleatory uncertainty due to its nature (in this case the complexity of further modeling) is not reduced as more data is gathered. It, therefore, represents a contribution that can be accounted for separately from other uncertainty. Aleatory uncertainty in crew behavior leads to variability including differences in time-required across crews and within a particular crew across different scenarios. For the purposes of this research, Teamwork (and Work Processes because it is integral to Teamwork) is treated as aleatory uncertainty.

Given Appendix B.1, the variability in time-required data from simulator studies such as EPRI NP-6937 experiments is [2, 18, 19] largely due to crew-to-crew variability, it seems reasonable to use the variability in time-required found in the simulator studies to model just this contribution to time-required. After this first contribution to a time-required probability distribution (or so-called “first order normalized distribution” hereafter) is developed from the EPRI data, it can be adjusted to account for other PIFs that may impact the probability distribution. Development of the “first-order” distribution is subject of Appendix D.

B.4 Review of Literature on PIF Impacts

The tables in Sections B.4.1–B.4.15 present considerations associated with 15 of the PIFs where there was relevant available literature for the potential impact of the PIF on time-required or time-required uncertainty.

The literature review emphasized work that was most relevant to the NPP context, prioritizing studies (for example) in NPP control rooms or scenarios. For each of the 15 PIFs, considerations and sample evidence are presented. For each study reviewed, there is a description of the task, associated findings, and the PIF effect on time uncertainty or time-required to complete the task. Note that many of the studies do not discuss variability in time-required, but the mean impact; this is because many of the studies do not report differences in variability (or standard deviation) based on the PIF. Nonetheless, information about the mean impact may be helpful. These studies are presented as example to help guide the analysts as they develop their time-required distribution; analysts should use their best judgment when determining the impact of each PIF on time-required uncertainty.

B.4.1 Work Location, Accessibility, and Habitability

Possible Considerations

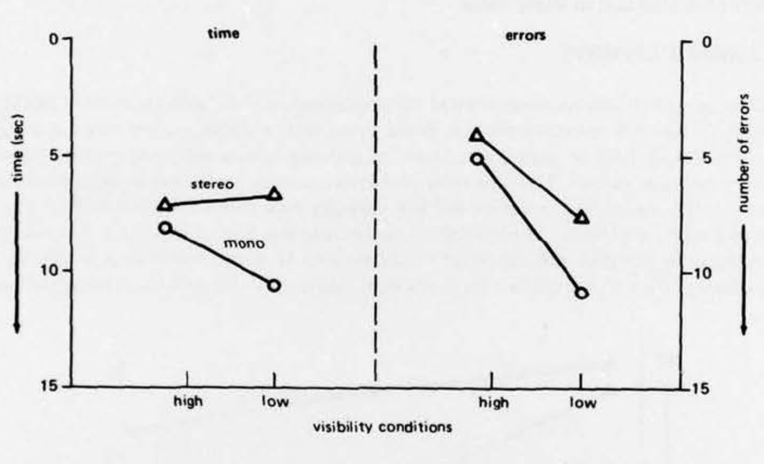
- Accessibility (travel paths, security barriers, and sustained habituation of worksite) is limited because of physical threats to life in the environment (e.g., traffic or weather impeding vehicle movement)
- Habitability is reduced; personnel cannot stay long at the worksite, or they experience degraded conditions for work, challenges to living conditions (e.g., isolation, confinement, microgravity), or environmental hazards like radiation or earthquake aftershocks.
- The worksite is flooded or underwater.
- The surface of systems, structures, or objects to be worked on cannot be reached or touched (e.g., because the surface is too hot to touch, or the object is too high to reach).
- Different paths to work site
- Hurdles to access work site (e.g., security system denies access)
- Radiation dose limit to work in a high-radiation environment may limit personnel access.

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Researchers used traffic data to estimate the impact of precipitation (rain and snow) and temperature on travel time in the Greater London Area.	Precipitation increases total travel time on average. Specifically, total travel time increased when snow or rain were present, with greater decreases associated with heavier precipitation. Temperature has nearly negligible effects on travel times.	Precipitation can increase travel time when driving.	[20]

B.4.2 Visibility

Possible Considerations
• Low levels of lighting that cause delay in detection • Low visibility of work environment (e.g., smoke, rain, fog) • Glare or strong reflection of the object to be detected or recognized

Sample Evidence			
Task	**Findings**	**PIF Effect on Time Uncertainty**	**Reference**
Participants were asked to position a manipulator under high and low visibility conditions using stereo or mono TV displays.	Stereo TV displays generally reduced response time but this advantage was eliminated in conditions of low visibility. In conditions of low visibility, participants were able to position the manipulator more quickly using mono versus stereo TVs.  Figure 5. Operator performance on peg-in-hole task.	Low visibility led to increased response time in general, but this effect was mitigated when using simpler (i.e., mono vs. stereo) displays.	[21]

B.4.3 Noise

Possible Considerations

- Loud noise levels that cause distraction or disruption in crew communications
- Continuous mixture of noisy sounds
- Intermittent non-speech noise
- Speech noise
- Intermittent mixture of speed/noise

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Authors conducted a meta-analysis of 242 studies examining the impact of noise on human performance, including both accuracy and response speed. Authors also examined the moderating impact of task type.	Noise has a very small but nonzero negative effect on response speed. Studies showed that noise decreased response speed in general.	Noise may reduce response speed (i.e., increase task completion time).	[9]

B.4.4 Cold/Heat/Humidity

Possible Considerations

- Exposure to temperature and humidity may degrade performance, requiring additional time for task completion
- Longer exposure times (somewhat related to times required) can result in greater degradation in performance
- For long-duration tasks, rest-work schedules may need to be considered (rest periods will add to task completion times)
- Perception and psychomotor functions may be affected more than cognitive functions

Sample Evidence

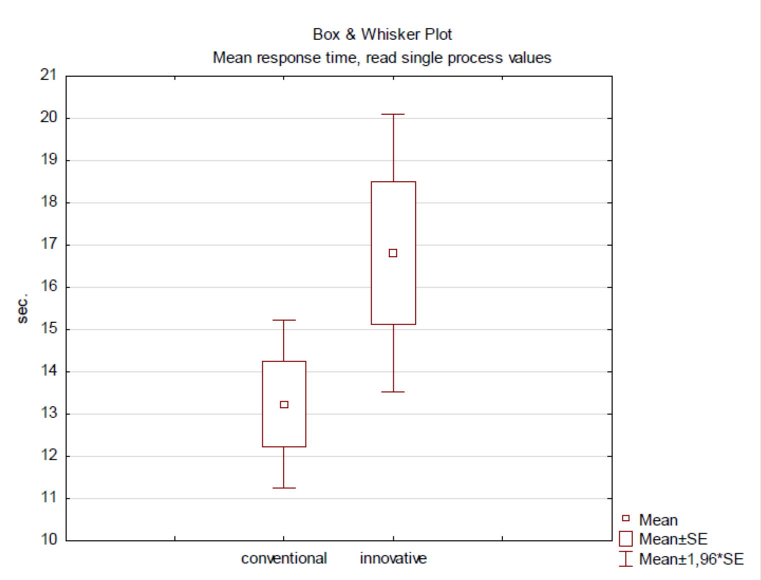
Task	Findings	PIF Effect on Time Uncertainty	Reference
Authors conducted a meta-analysis of 57 studies exploring the impact of temperature on human performance by task type.	Results indicates that heat and cold can, for some tasks, decrease reaction time. Specifically, heat led to faster response time for psychomotor tasks and cold led to faster reaction times for cognitive tasks. Heat, however, reduced accuracy, whereas cold did not; in addition, the results were dependent on the length of extremity of exposure.	Cold may lead to faster response time for cognitive tasks, and heat may lead to faster response time for psychomotor tasks.	[22]

B.4.5 Resistance to Physical Movement

Possible Considerations			
• Delay in personnel and equipment movement because of external hazards • Wearing heavy protective clothes (i.e., personal protection equipment) or gloves • Resistance to personnel movement, limited available space, or postural instability • Whole-body vibration			
Sample Evidence			
Task	**Findings**	**PIF Effect on Time Uncertainty**	**Reference**
Seven members of the Missouri National Guard's Civil Support Team completed a dexterity test (moving small objects various distances) and a tracer task (where they had to trace a visual stimulus without line-of-sight, using only physical sensory feedback). There were several levels of difficulty tested. In addition, participants completed the tasks in and outside of a Level A suit, a bulky piece of personal protective equipment that impedes movement and dexterity. The goal was to assess the impact of the suit on task performance, including the time spent in the suit.	Results showed that wearing the suit increased the time to complete tasks substantially for the dexterity task and to a lesser extent for the tracing task. Participants also made more errors when wearing the suits than without.	Impeding physical movement is likely to increase task completion time.	[23]

B.4.6 Human System Interface

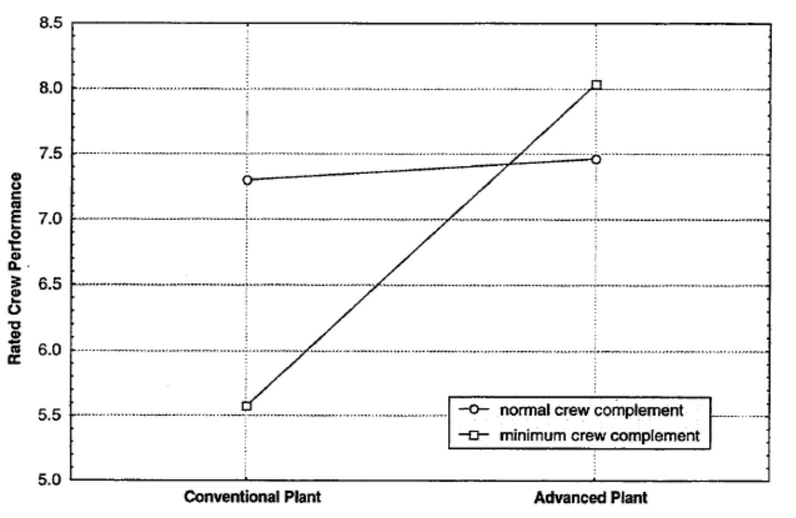
Possible Considerations
• Related information for a task is spatially distributed, disorganized, or cannot be accessed simultaneously • Unintuitive or unconventional indicators • Inconsistent interpretation of displays or insistent representation of information, measurement units, symbols, or tables • System contains ergonomic deficits, such as difficult to maneuver controls, low readability or saliency of labels and controls, confusing labels, inadequate indication of states of controls, or unintuitive

Sample Evidence			
Task	**Findings**	**PIF Effect on Time Uncertainty**	**Reference**
Nine Swedish operators monitored two different types of displays in a dynamic or static scenario in a simulator. Innovative displays contained pictorial graphical displays of information; convention displays included numerical information only. Operators answered questions about everyday tasks performed in the control room and accuracy and response time were measured.	The effects of display type on response time were task dependent, but in general, the innovative displays had no effect on response time. However, there were a large number of experimental displays explored that may not have been optimized; the authors note that charts and graphics need to be balanced with information density.  	The quality of the HSI may impact task completion time, but effects are dependent upon the task under completion and the density of the information display.	[24]

B.4.7 Staffing

Possible Considerations
• Staff adequacy (e.g., whether concurrent activities would reduce the staff available for the action or whether tasks can be performed concurrently with more than adequate staff) • Crew-to-crew variability in time-required to perform the same actions; different crews may take different procedure paths, which leads to variability in the time-required • Key personnel are missing, unavailable, or delayed in arrival; staff pulled away to perform other duties • Lack of certain knowledge, skills, or abilities needed for key personnel in unusual events (e.g., key decision maker's knowledge and ability are inadequate to make the decision; lack of required qualifications or experience)

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Eight crews of operators performed five scenarios with minimal or normal staffing levels in a conventional or advanced plant design. Authors compared the performance of the minimal vs. normal crew staffing levels and the plant design on scenario performance.	Minimum sized crews had similar performance to normal sized crews in terms of performance ratings. However, minimum sized crews experienced higher workload and completed fewer tasks, shedding tasks that were not perceived as critical.  **Figure 5.3 Plot of rated crew performance by crew size and plant type**	Minimum sized crews experienced higher workload and completed fewer tasks than normal sized crews, but overall performance was unaffected. Minimal staffing levels may lead to increased performance time, especially for non-critical tasks.	[25]

B.4.8 Procedures, Guidance, and Instruction

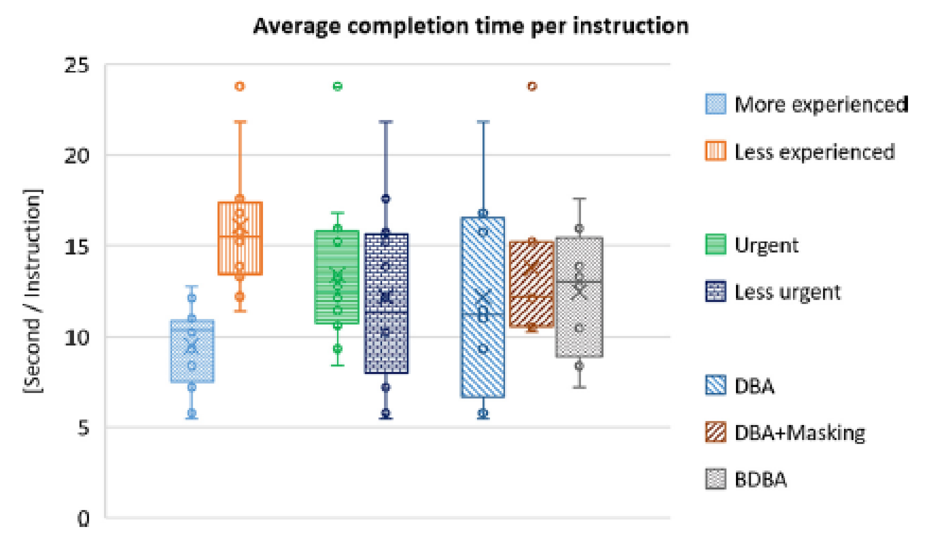
Possible Considerations

- Procedure design is inadequate or difficult to use
- Graphics, symbols not intuitive
- Complicated logic or mental calculation required (e.g., unit conversion)
- Inconsistency between procedures and displays
- Fold-out page not salient or not used
- Poor standardization in terminology
- Difficult layout, lack of placeholders
- Procedure lacks details, e.g., lack of verification for key parameters for detection or execution, lack of guidance for confirmatory data
- Procedure is ambiguous or confusing (e.g., wrong or incomplete descriptions, conflict between literal meaning and intention)
- Procedure is available but does not match the situation (e.g., needs deviation or adaptation)
- Procedure is not applicable or not available; procedure is misleading

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Three operators completed a scenario on a simulator and experienced five different failures of a computerized procedure system (CPS). Operators also completed a benchmarking scenario.	Operators completed the same scenario with CPS faster than using paper procedures. However, failures of the CPS were frequently unidentified.	Computerized procedures may improve task completion time, but also have the potential to lead to increased error if not properly monitored.	[26]
The authors proposed a measure of variability in procedure progression and explored the values in different scenarios using an existing simulator dataset (OPERA; Park and Jung 2007).	Higher variability in procedure progression was associated with longer performance times in the simulator. A high value of their variability metric suggests that there is a lack of clear ways of executing the procedure or that there are more efficient ways of doing so.	High variability in completion of procedures leads to increased task completion times and to more variable task completion time. This variability in completion of procedures may be indicative of unclear procedures or opportunistic attempts to progress through the procedure more efficiently.	[27]

B.4.9 Training

Possible Considerations			
• Recency of training • Inadequate amount or quality of training • Lack of or poor administrative control on training			
Sample Evidence			
Task	**Findings**	**PIF Effect on Time Uncertainty**	**Reference**
Four crews of three (12 operators) completed six scenarios that varied in time urgency and task complexity. In addition, crews were divided into more and less experienced crew groups. The intent of the study was to examine the effect of three PIFs on performance as measured by completion time, error rate, and other outcomes.	On average, the more experienced group spent less time to complete instructions compared to the less experienced group. Time urgency and task complexity had no effect on completion time.  Fig. 7. Comparison of average completion time per instruction for the PSFs.	Operator experience (i.e., training) significantly decreases task completion time.	[28]

B.4.10 Work Processes

Possible Considerations

- Lack of self-verification or cross-verification, peer-checking, independent checking or advising, or close supervision
- Poor attendance to task goals, roles, or responsibilities
- Poor shift handovers
- Poor work prioritization, planning, or scheduling

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Participants completed "micro-tasks," which were designed to be similar to a procedure step or part of a step. Three studies were conducted with operators (16, 9, and 10 operators, respectively). In the first study, operators completed tasks on analog panels vs. digital HSI. In the second, they completed tasks on a conventional vs. advanced display. In the third, the researchers compared the performance of individuals vs. teams independently checking each other's work using large screen overview displays or workstation displays.	Operators performed faster with analog vs. digital displays. There was no difference in completion time for conventional vs. innovative displays. Teams were faster than individuals. Large screen overview displays were also faster than workstation displays.	Teams performed tasks faster than individuals, even when cross-checking was performed.	[11]

B.4.11 Multi-Tasking, Interruptions and Distractions

Possible Considerations
• Distraction by other on-going activities that demand attention • Interruption taking away from the main task • Concurrent visual detection and other tasks • Concurrent auditory detection and other tasks • Concurrent diagnosis and other tasks • Concurrently making two or more simple decisions/plans • Concurrently making intermingled complex decisions/plans • Concurrently executing action sequence and performing another attention/working memory task • Concurrently executing intermingled or inter-dependent action plans • Concurrently communicating or coordinating multiple distributed individuals or teams

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Fifty-eight student nurses either did or did not experience an extended interruption (an alarm) during a simulated medication administration task. Participants performed a standard medication administration task, and errors were measured as well as time taken to dispense medication.	Interruption actually decreased the time taken to dispense medication, although the difference was not significant.	People who are interrupted during tasks may compensate by speeding up work, resulting in shorter class completion time, but also committed more errors than individuals who were not interrupted.	[10]

B.4.12 Task Complexity

Possible Considerations			
• Requires staff to track the states of multiple systems or to monitor many parameters • Staff must memorize many pieces of information detected or detect many types or categories of information • Criteria are not straightforward; information of interest involves complicated mental computation • Detection demands high attention, split attention, sustained attention over a period of time, or intermittent attention • Cues for detection are not obvious – e.g., detection is not directly cued by alarms or instructions and personnel need to actively search for the information • Working memory overload; need to decipher numerous messages (indications, alarms, spoken messages) • Very long procedures, voluminous documents with checkoff provisions; multiple procedures needed, action sequences are parallel and intermingled			
Sample Evidence			
Task	**Findings**	**PIF Effect on Time Uncertainty**	**Reference**
Participants were individuals undergoing the senior reactor operator (SRO) qualifying examination in Korea. They were asked to complete several tasks within the simulator, and the authors measured the association between step completion time and step complexity.	Examinees for the SRO qualifying examination completed a series of simulator exercises during their exam to mimic stressful conditions. The completion time for each step of the procedure was compared to a measure of step complexity which considered the amount of information the operator has to process, the logical complexity of each step, and the number of activities required. Results showed that step complexity was strongly correlated to step completion time.	Task complexity increases task completion time in stressful conditions.	[14]
Authors used simulator data to examine the relations between task complexity and completion time under emergency operating conditions.	Authors explore the association between a proposed measure of task complexity and performance time for operators in Korean nuclear power plant simulators in emergency scenarios. Results showed that there was a significant and strong relationship between task complexity and performance time.	Task complexity increases task completion time under emergency conditions.	[7]
Authors used simulator data to examine the relations between task complexity and completion time under ordinary and emergency operating conditions.	Authors explore the association between a proposed measure of step complexity and performance time of operators under ordinary and emergency conditions. Results showed that operators took more time to complete steps with a higher step complexity score.	Task complexity increases task completion time.	[29]
Students completed a production management task that was either simple or complex and either were interrupted or not during task completion.	Interruptions caused an increase in decision time for complex tasks, but a decrease in decision times for simple tasks.	Interruptions may impact task completion time, but the direction of the task is dependent upon task complexity.	[30]

B.4.13 Time Pressure and Stress

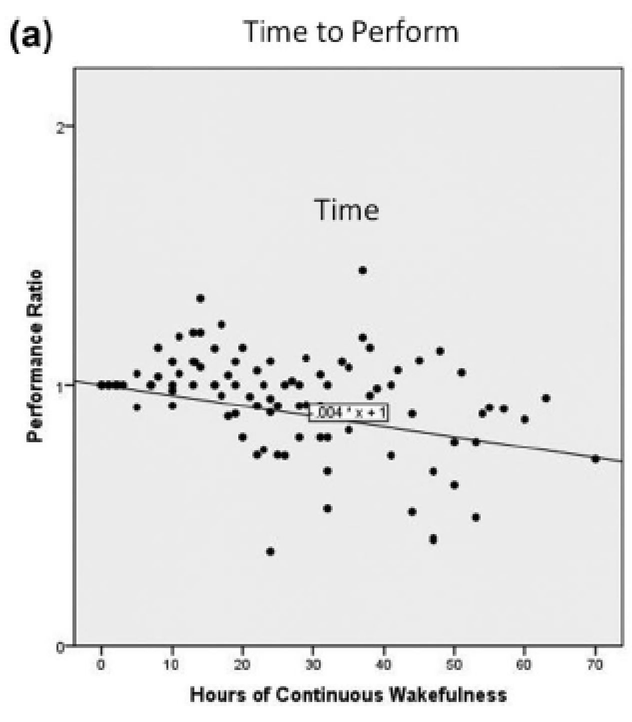
Possible Considerations

- Time pressure may cause operators to spend less time completing tasks
- Receiving instructions to complete tasks as quickly as possible, deadlines, or stimulus presentation rate
- Emotional stress (e.g., anxiety, frustration)
- Cumulative physical stress (e.g., long hours exposure to ambient noise, disturbed dark and light rhythms, air pollution, disruption of normal work-sleep cycles)

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
Authors conducted a meta-analysis of 125 papers that examined the impact of time pressure on accuracy and performance, and categorized the tasks completed in the studies as perceptual, cognitive, or psychomotor.	In general, time pressure had a moderate effect on reaction time or speed. When examining by task type, performance was faster for perceptual and cognitive tasks, but slower for psychomotor tasks. However, due to the small number of studies examining psychomotor tasks (n = 5), this result should be interpreted with significant caution.	There is a moderate effect of time pressure on task completion time or response time, whereby individuals under time pressure responded faster than individuals under no time pressure. Notably, there is a tradeoff with accuracy, where time pressure resulted in responses that were faster but less accurate.	[9]

B.4.14 Mental Fatigue

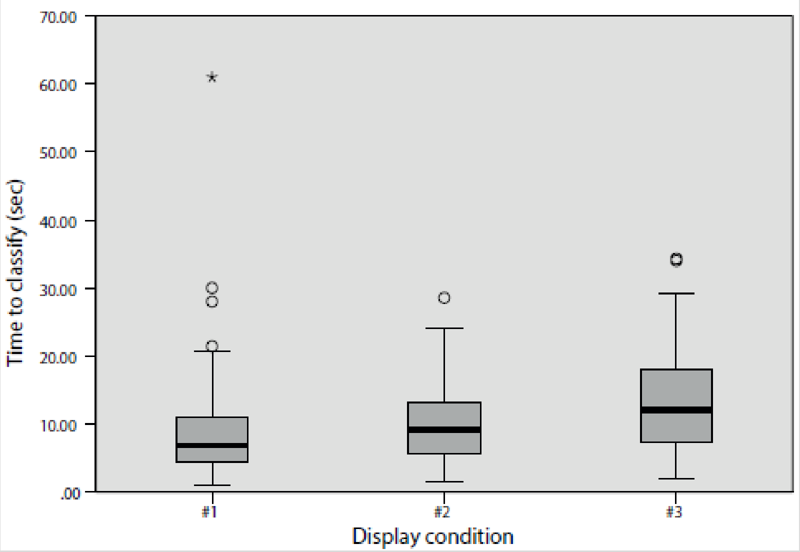
<table>
<tr><th colspan="4">Possible Considerations</th></tr>
<tr><td colspan="4">• Long working hours, non-routing shifts, persistent high mental load may slow down task performance
• Time of day or duration of being on shift may increase task completion time</td></tr>
<tr><th colspan="4">Sample Evidence</th></tr>
<tr><th>Task</th><th>Findings</th><th>PIF Effect on Time Uncertainty</th><th>Reference</th></tr>
<tr><td>Authors conducted a meta-analysis of 28 papers that examined the impact of sleep deprivation (generally 24 - 72 hours of continuous wakefulness) on performance. They explored the time to perform as well as the accuracy.</td><td>There was a negative association between hours of continuous wakefulness and time to perform (i.e., task completion time). Longer continuous wakefulness was associated with longer performance times.
 </td><td>Sleep deprivation may negatively impact task completion time.</td><td>[31]</td></tr>
</table>

B.4.15 System and I&C Transparency to Personnel

Possible Considerations

- System behavior is complex to understand or not transparent to personnel
- Feedback about system state, action, or intention is not provided
- Inappropriate system functional allocation between human and automation
- Overreliance on automation, staff not alerted to actions needed
- System failure modes are not transparent, consistent, or obvious to personnel
- System failures are coupled or interdependent
- I&C logic is not transparent or clear; I&C failure modes are not transparent to personnel

Sample Evidence

Task	Findings	PIF Effect on Time Uncertainty	Reference
6 experienced fighter pilots were presented with three different display conditions in computer-based flight scenarios. Participants were asked to identify an object (a threat classification task). One display was intended to mimic current display showing a detected object; one used a class suggestion; the third provided a class suggestion with an accompanying explanation.	Time increased with the amount of information presented, with the current displays requiring the least time and the more complex and visually involved displays requiring additional time to classify the object.  	Additional information regarding the system may require additional time for personnel to process.	[32]

B.5 References

1. Xing, J., Y.J. Chang, and J. DeJesus, *The General Methodology of an Integrated Human Event Analysis System (IDHEAS-G)*, 2021, NUREG-2198 U.S. Nuclear Regulatory Commission, ADAMS Accession No. ML21127A272.
2. Spurgin, A.J., P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, and G.W. Hannaman, *Operator Reliability Experiments Using Power Plant Simulators Volume 3: Appendixes*, 1990, EPRI NP-6937, Vol 3, Electric Power Research Institute, https://www.epri.com/research/products/NP-6937-V1.
3. Bye, A., E. Lois, V.N. Dang, G. Parry, J. Forester, S. Massaiu, R. Boring, P.Ø. Braarud, H. Broberg, J. Julius, I. Männistö, and P. Nelson, *International HRA Empirical Study - Phase 2 Report: Results from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios*, 2011, NUREG/IA-0216, Vol. 2 (ADAMS Accession No. ML11250A010), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1125/ML11250A010.pdf.
4. Dang, V.N., J. Forester, R. Boring, H. Broberg, S. Massaiu, J. Julius, I. Männistö, H. Liao, P. Nelson, E. Lois, and A. Bye, *International HRA Empirical Study - Phase 3 Report: Results from Comparing HRA Methods Predictions to HAMMLAB Simulator Data on LOFW Scenarios*, 2014, NUREG/IA-0216, Vol. 3 (ADAMS Accession No. ML14358A254), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1435/ML14358A254.pdf.
5. Lois, E., V.N. Dang, J. Forester, H. Broberg, S. Massaiu, M. Hildebrandt, P.Ø. Braarud, G. Parry, J. Julius, R. Boring, I. Männistö, and A. Bye, *International HRA Empirical Study - Phase 1 Report: Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Performance Data*, 2009, NUREG/IA-0216, Vol. 1 (ADAMS Accession No. ML093380283), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0933/ML093380283.pdf.
6. Xing, J., Y.J. Chang, and J. DeJesus, *Draft - Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA)*, 2020, RIL-2021-XX U.S. Nuclear Regulatory Commission, ADAMS Accession No. ML20238B982.
7. Park, J. and W. Jung, *A study on the validity of a task complexity measure for emergency operating procedures of nuclear power plants - comparing task complexity scores with two sets of operator response time data obtained under a simulated SGTR.* Reliability Engineering and System Safety, 2008. **93**(4): p. 557-566.
8. Park, J., W. Jung, J. Kim, and J. Ha, *Analysis of human performance observed under simulated emergencies of nuclear power plants*, 2005, KAERI/TR-2895/2005, Korea Atomic Energy Research Institute.
9. Szalma, J., P. Hancock, and S. Quinn, *A Meta-Analysis of the Effect of Time Pressure on Human Performance*. 2008.
10. Nicholas, C.A., *The Effect of Interruptions on Primary Task Performance in Safety-Critical Environments*, in *Industrial Engineering & Operations Research*. 2016, University of Massachusetts - Amherst. https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1798&context=dissertations_2.
11. Massaiu, S. and L. Holmgren, *A HAMMLAB HRA Data Collection with U.S. Operators*, 2016, HWR-1123, OECD Halden Reactor Project.
12. UJV Rez a.s., *Unpublished data from the dukovany npp simulator trials from 2011 as described in slides 18-29 of ml18079a165*. 2011, Nuclear Research Institute of the Czech Republic (UJV Rez a.s.).
13. Forester, J., H. Liao, V.N. Dang, A. Bye, E. Lois, M. Presley, J. Marble, R. Nowell, H. Broberg, M. Hildebrandt, B. Halbert, and T. Morgan, *The U.S. HRA Empirical Study – Assessment of HRA Method Predictions against Operating Crew Performance on a*

U.S. Nuclear Power Plant Simulator, 2016, NUREG-2156 (ADAMS Accession No. ML16179A124), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1617/ML16179A124.pdf.

14. Park, J., J. Kim, and W. Jung, *Comparing the complexity of procedural steps with the operators' performance observed under stressful conditions.* Reliability Engineering & System Safety, 2004. **83**(1): p. 79-91. https://www.sciencedirect.com/science/article/pii/S0951832003001959.
15. Salas, E., D.E. Sims, and C.S. Burke, *Is there a “big five” in teamwork?* Small Group Research, 2005. **36**(5): p. 555-559.
16. Acevedo, M.R., J. Ake, J. Bommer, K.J. Coppersmith, M. Juckett, C.G. Munson, and J. Stamatakos, *Updated implementstion guidelines for SSHAC hazard studies*, 2018, NUREG-2213, Nuclear Regulatory Commission, ADAMS Accession No. ML18282A082.
17. Hollnagel, E., *Cognitive Reliability and Error Analysis Method (CREAM)*. 1998, Elsevier.0-08-0428487.
18. Spurgin, A.J., P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, and G.W. Hannaman, *Operator Reliability Experiments Using Power Plant Simulators Volume 1: Executive Summary*, 1990, EPRI NP-6937, Vol 1, Electric Power Research Institute, https://www.epri.com/research/products/NP-6937-V1.
19. Spurgin, A.J., P. Moieni, C.D. Gaddy, G.W. Parry, D.D. Orvis, J.P. Spurgin, V. Joksimovich, D.P. Gaver, and G.W. Hannaman, *Operator Reliability Experiments Using Power Plant Simulators Volume 2: Technical Report*, 1990, EPRI NP-6937, Vol 2, Electric Power Research Institute, https://www.epri.com/research/products/NP-6937-V1.
20. Tsapakis, I., T. Cheng, and A. Bolbol, *Impact of weather conditions on macroscopic urban travel times.* Journal of Transport Geography, 2013. **28**: p. 204-211. http://www.sciencedirect.com/science/article/pii/S0966692312002694.
21. Pepper, R.L. and R.E. Cole, *Display System Variables Affecting Operator Performance in Undersea Vehicles and Work Systems*, 1978, Naval Ocean Systems Center San Diego CA, https://apps.dtic.mil/dtic/tr/fulltext/u2/a056791.pdf.
22. Hancock, P., J. Ross, and J. Szalma, *A Meta-Analysis of Performance Response Under Thermal Stressors.* Human factors, 2007. **49**: p. 851-77.
23. Simon, Y.L., *The effects of personal protective equipment Level A suits on human task performance*, in *Engineering Management*. 2010, Missouri University of Science and Technology. https://scholarsmine.mst.edu/masters_theses/4994.
24. Rø Eitrheim, M.H., A. Fernandes, and H. Svengren, *Evaluation of design features in the HAMBO operator displays*, 2017, HWR-1212, OECD Halden Reactor Project.
25. Halbert, B., A. Sebok, and D. Morisseau, *A Study of Control Room Staffing Levels for Advanced Reactors*, 2000, NUREG/IA-0137, U.S. Nuclear Regulatory Commission.
26. Taylor, C., M. Hildebrandt, R. McDonald, and N. Hughes, *Operator Response to Failures of a Computerised Procedure System: Results from a Training Simulator Study*, 2017, HWR-1198, OECD Halden Reactor Project.
27. Kim, Y., J. Park, and W. Jung, *Measuring variability of procedure progression in proceduralized scenarios.* Annals of Nuclear Energy, 2012. **49**: p. 41-47. http://www.sciencedirect.com/science/article/pii/S0306454912002228.
28. Park, J., D. Lee, W. Jung, and J. Kim, *An experimental investigation on relationship between PSFs and operator performances in the digital main control room.* Annals of Nuclear Energy, 2017. **101**: p. 58-68. https://linkinghub.elsevier.com/retrieve/pii/S0306454916307381.
29. Park, J., K. Jeong, and W. Jung, *Identifying cognitive complexity factors affecting the complexity of procedural steps in emergency operating procedures of a nuclear power plant.* Reliability Engineering & System Safety, 2005. **89**(2): p. 121-136. https://linkinghub.elsevier.com/retrieve/pii/S0951832004001899.

30. Speier, C., I. Vessey, and J.S. Valacich, *The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance.* Decision Sciences, 2003. **34**(4): p. 771-797. https://doi.org/10.1111/j.1540-5414.2003.02292.x.
31. Wickens, C., S. Hutchins, L. Laux, and A. Sebok, *The Impact of Sleep Disruption on Complex Cognitive Tasks: A Meta-Analysis.* Human Factors The Journal of the Human Factors and Ergonomics Society, 2015. **57**. https://www.researchgate.net/profile/Christopher_Wickens/publication/272622494_The_Impact_of_Sleep_Disruption_on_Complex_Cognitive_Tasks_A_Meta-Analysis/links/54ec8fbb0cf28f3e6534401f/The-Impact-of-Sleep-Disruption-on-Complex-Cognitive-Tasks-A-Meta-Analysis.pdf.
32. Helldin, T., *Transparency for Future Semi-Automated Systems : Effects of transparency on operator performance, workload and trust*, in *Örebro Studies in Technology*. 2014, Örebro Universitet: Örebro. p. 266. http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-10164.

Appendix C
–
Expert Elicitation on Methods

This appendix describes an expert knowledge elicitation that was designed to produce insights on ways to determine probability that the time-available to perform a task is exceeded. The focus was on determining ways to estimate the time-required and time-available and their corresponding probability distributions. The appendix sub-sections (1) describe the how the expert knowledge elicitation process was set up and conducted, (2) summarize the results and conclusions of the exercising the process, and (3) provide insights to supplement the guidance in Section 3.6 of NUREG-2256 Integrated Human Event Analysis System for Event and Condition Assessment (IDHEAS-ECA) [1].

Appendices C.1 and C.2 provides an overview of the expert knowledge elicitation process including explanation of the purpose, discussion of the selection of experts, and description of how the email solicitation and elicitation workshop was performed.

Appendices C.3 and C.4 provide a summary of the results of the expert knowledge elicitation, which are primarily the responses from the experts to specific concerns associated with developing estimates of time-required and availability and corresponding probability distributions.

Appendices C.5 and C.6 provide insights gained from the expert knowledge elicitation process associated with determining the probability of exceeding time-available using probability distributions of time-required and time-available. It also offers suggestions about how to use the insights and what future research is needed to address certain elements of the approach.

C.1 Overview of the Expert Knowledge Elicitation Process

There are different purposes for performing an expert elicitation. The purpose of this elicitation was to gather and evaluate the experience and knowledge of a larger technical community on a technical issue. There are many possible purposes of elicitation, including identifying available technical evidence relevant to a technical issue and disseminating and sharing common databases by experts. These elicitation efforts are often a formal and time-consuming process but can generate important information on a topic. Though nominally an expert elicitation, this effort was limited in scope and resources available (e.g., the experts donated their time for a 2-hour workshop and time to review and the think about the information requests). However, as discussed below, certain expert elicitation principles were used to assure quality and transparency of results and to reduce social biases. The expert knowledge elicitation process consisted of the following elements:

1. Developing the information requests from experts needed to address basic questions about development of point estimates and probability distributions for time-available and time-required to inform the IDHEAS-ECA guidance
2. Identifying and recruiting experts
3. Disseminating the expert knowledge elicitation workshop information requests and preliminary instructions by email
4. Following up with each workshop expert after sending the email messages
5. Holding the expert knowledge elicitation workshop
6. Sending to each expert a record of the responses drafted by the workshop facilitator and observers and giving them a chance to refine their responses

7. Finalizing the workshop responses and distilling insights
8. Solicitation of feedback from two PRA and HRA leaders outside the workshop
9. Consolidating the final insights and recommendations

As noted above, in addition to the expert elicitation workshop steps described above, the process included engagement with two PRA and HRA experts and industry leaders for their feedback on estimating and characterizing time-available and time-required.

Appendix C.1 describes development of the expert knowledge elicitation information requests, Appendix C.2 describes selecting knowledge elicitation experts, and Appendix C.3 describes preparing for, holding, and documenting the expert knowledge workshop including a discussion of how expert elicitation principles were used consistent with NRC guidance. Appendix C.4 discusses solicitation of feedback from two PRA and HRA experts and industry leaders for their feedback on estimating and characterizing time-available and time-required.

The following information requests were developed to gather knowledge that would inform IDHEAS guidance to HRA analysts about how to estimate exceedance of time-available. As stated above, IDHEAS methodology uses a time uncertainty model to estimate the Human Error Probability (HEP) contribution of the exceedance of time-available determined by the probability distributions of time-available and time-required for a human action. This research focused on how to develop a probability distribution for time-required, but not for time-available. Accordingly, information requests were developed for the experts to address basic questions about development of point estimates and probability distributions for time-available and time-required. Therefore, as part of the process, the following information requests were then sent ahead of the workshop to elicitation workshop experts and non-workshop experts by email:

1. Concerning the point estimates of time-available that are currently performed at nuclear power plants (NPPs):
    a Describe the methods used to estimate the time-available that are used in developing HEPs. Include description the different applicable situations of available resources that impact the method used that IDHEAS should address.
    b For each method described above in response to part 1.a, describe the factors that could impact the uncertainty of the estimate (e.g., Was the estimate based on a plant specific or a generic evaluation? Was the estimate performed using a conservative assumption or a bounding approach? Was the estimate based on the timing for a different but similar accident scenario?)
    c Explain whether any attempt is made to characterize the uncertainty of time-available (e.g., establish a lower and upper bound) and under what circumstances.
2. Concerning development of a probability distribution for time-available:
    a Describe any methods, used currently or in the past, by NPPs to develop a probability distribution for time-available.
    b Describe methods that could be used by NPPs to develop a probability distribution for time-available based on the approaches identified in part 2.a above.
    c Discuss the pros and cons of each method identified in part 2.b above.
    d Propose a best-practice approach for developing probability distributions for time-available based on your experience and knowledge.

3. Concerning the point estimates of time-required that are currently performed at NPPs:
    a Describe the methods used to estimate the time-required that are used in developing HEPs. Include description of the different applicable situations of available resources that impact the method used that IDHEAS should address.
    b For each method described above in response to part 3.a, describe the factors that could impact the uncertainty of the estimate (e.g., Was the estimate based on a plant specific or a generic evaluation? Was the estimate performed using a conservative assumption or a bounding approach? Was the estimate based on the timing for a different but similar accident scenario?)
    c Explain whether any attempt is made to characterize the uncertainty of time-required (e.g., establish a lower and upper bound) and under what circumstances.

C.2 Selecting Experts

To identify a set of experts with knowledge representative of the larger technical community, NRC selected participants with the following expertise:

1. Current knowledge about how HRA is performed for NPP PRAs.
2. NRC staff who use HRA and risk concepts to support risk informed decisions on NPP issues. This includes application of the Significance Determination Process or Accident Sequence Precursor analyses.
3. PRA and HRA industry leaders with a significant history and perspective about HRA approaches and its applications. This includes HRA for formally reviewed and approved PRA models per the American Society of Mechanical Engineers/American Nuclear Society (ASME/ANS) PRA standard [2], [3] and Regulatory Guide 1.200, Revision 2 [4] and Revision 3 [5].

The HRA industry experts are well known consultants to licensees of NPPs and have been instructors, and researchers for EPRI. They are especially current on the HRA approaches used in support of risk-informed applications to the NRC such as those under NRC,[6] “Risk-Informed Categorization of the Treatment of Structures, Systems and Components for Nuclear Power Reactor,” and licensee amendment requests to revise technical specifications to adopt risk-informed completion times through Technical Specification Task Force-505, Revision 2 [7]. The NRC HRA experts have recent industry experience at NPPs as Senior Reactor Operators (SROs) or academic credentials specifically related to risk and reliability. The PRA and HRA industry leaders are well known and respected across industry, the NRC, and internationally. There were eight experts in total. The first six experts agreed to participate in an expert knowledge elicitation workshop. The final two experts responded to the information requests presented in Appendix C.4.5 in email messages at their desertion in ways that added value to the process.

C.3 Expert Elicitation Workshop – Preparing, Holding, and Documenting

In June 2022, the NRC sent to industry and NRC HRA experts identified above and invited them to a virtual workshop held in July 2022. PNNL followed up by emailing the information requests that had been developed for the workshop and a brief description of the structure of the virtual workshop. The experts were told that:

> *The workshop will be organized by soliciting responses to three basic questions and information requests associated with those concerns.*
>
> *Each participant will be given a set number of minutes to provide their response and each participant will provide their response to the requests before we proceed to the next request. We anticipate that up to six experts will participate*

in the workshop. Participants can choose not to answer certain requests if they feel they do not have the experience or expertise.

PNNL will document the responses from each expert.

PNNL will send the documented responses to the participating experts after the workshop for verification. This step is important given that we intend to use the insights to enhance the IDHEAS guidance on determining time-available and time-required.

The experts were told that the purpose of the workshop was not to come to consensus about a method (or methods) to determine time-available and time-required but rather to gain an understanding of the current practice and what might be done.) to enhance the IDHEAS guidance. The PNNL staff member who facilitated the workshop followed up with each expert by telephone and email to address any potential confusion or concern about the requests or our expectations. Before the workshop, experts received the agenda for the workshop and some notes about the elicitation process itself. In July 2022, the workshop was performed as presented in Table C.1. The workshop lasted about 2.5 hours and was performed consistent with the proposed agenda.

Table C.1. Expert Knowledge Workshop Schedule

Information Request	Duration	Started at
Introductions and Briefing	10 minutes	1:00 pm EDT
Information Request 1 Responses (4 minutes per participant)	25 minutes	1:10 pm EDT
Follow-up and Observer Comments	5 Minutes	1:35 pm EDT
Information Request 2 Responses (4 minutes per participant)	25 minutes	1:45 pm EDT
Follow-up and Observer Comments	5 minutes	2:10 pm EDT
Break	10 minutes	2:20 pm EDT
Information Request 3 Responses (5 minutes per participant)	25 minutes	2:30 pm EDT
Follow-up and Observer Comments	5 minutes	2:55 pm EDT
Change of opinion and follow-up discussion	20 minutes	3:00 pm EDT
Closeout	10 minutes	3:20 pm EDT
Meeting Complete		3:30 pm EDT

This workshop knowledge elicitation was structured to solicit all information on relevant topics and also provide each participant an equal opportunity to share their knowledge and to reduce social biases. The 30 minutes period following the scheduled elicitation response time was used for follow-up questions by the facilitator, the observers, and responses and comments by experts themselves.

Use of Expert Elicitation Guidance. The workshop itself was designed to reduce social bias by employing principles presented in an NRC white paper, “Practical Insights and Lessons Learned in Implementing Expert Elicitation” [8]. The white paper was developed based on past use of expert judgment in PRA and important applicable NRC guidance documents, NUREG-1563, NUREG/CR- 6372, NUREG-2117, and NUREG-2213 [9]–[12]. At the beginning of the elicitation, the facilitator provided a briefing on primary basic principles of expert elicitation principles as discussed in NRC white paper which are discussed in the following sections.

Representation of the Technical Community. The purpose of expert elicitation was not to create new knowledge, but rather to obtain the center, body, and range of the views of the technical community on the state of knowledge about an issue. The experts were asked to represent (to best of their ability given the circumstances) the community’s views and practices concerning exceedance of time-available.

Independent Intellectual Ownership. The expert panel members were told they are not representing their employer or organization on the panel but are providing their own expertise. Each expert was asked to maintain independence from the other experts in the team to avoid (or mitigate) a group-think bias risk. The objective of the workshop was not necessarily to come to consensus positions.

Avoidance of Conflicts of Interest. The expert panel were told that they should be representative of the larger technical community to obtain a range of knowledge and interpretations about the technical issue. The experts in this panel were industry consultants and instructors in HRA and NRC staff who worked at NPPs or have related academic credentials.

Interaction and Integration. Each expert was given time to express their knowledge as well as an opportunity to augment or amend what they said after hearing the contributions from other participants.

Additionally, after the workshop the experts were given a summary of their responses, which they were asked to confirm or amend.

Structured Process. An expert elicitation should employ a structured process to facilitate interaction, integration, and to reduce biases in the outcomes, and therefore, the structure reflected in the agenda was used.

Transparency. To assure that the results are used appropriately, the information was generated and documented in a transparent way. This transparency includes description of the process, the results obtained, and the caveats and limitations of the process. Transparency helps to demonstrate the stability and integrity of the results.

Approximately one week after the elicitation, PNNL sent the experts a record their responses based on PNNL observer and facilitator notes. Experts were asked the experts to confirm, correct, add to, subtract from, clarify, or augment the responses so that it reflected the messages they intended to present.

PNNL also solicited responses to the workshop information requests by email from two PRA and HRA industry leaders with a significant history and perspective about HRA approaches and their applications. The two experts were suggested by NRC to PNNL as having considerable experience in providing guidance to NRC in the PRA and HRA domains. These experts were asked to provide their insights by either responding to the same information requests as the workshop experts or offering general advice about the topics addressed. PNNL received email response from one non-workshop expert in August 2022, and from the second expert in September 2022. These responses are summarized in Appendix C.4.5 and serve to supplement the results of the expert knowledge elicitation.

C.4 Summary of Results of the Expert Knowledge Elicitation

Results from the of the expert knowledge elicitation process come from the following sources:

- The workshop convened in July 2022 using NRC selected experts with current knowledge about how HRA is performed for NPP PRAs, and NRC staff who use HRA and risk concepts to support risk informed decisions on NPP issues not based on a formal PRA.
- Separate results from two PRA and HRA industry leaders with a significant history and perspective about HRA approaches and its applications. Those results are reported separately in Appendix C.4.5.

C.4.1 Expert Knowledge Elicitation Workshop Results

Detailed documentation of the July 2022 workshop results is provided in PNNL-33857 [13]. This section summarizes the response from the six experts by distilling the workshop into key takeaways, presenting common themes imparted by the experts, and highlighting responses that provide insights to inform the IDHEAS HRA process. The summary is provided here on a per-request basis, with one exception that is described below.

One of the industry HRA experts presented a general observation in the first round that was subsequently referred to by several experts as important context for the responses. The expert stated that the general approach to time assessment of exceedance of time-available starts by using point estimates of “time-available” and “time-required.” The HRA analyst then looks for one of the three potential outcomes described below:

- Case 1: Time-required exceeds time-available. In this case the action is not feasible.
- Case 2: Time-required is much less than the time-available. In this case, the time-related failure mode is negligible.
- Case 3: Time-required is close to the time-available.

The expert noted that precise assessment such as a refined point estimate or development of probability distributions is only of potential interest to Case 3 outcomes. However, it may not be readily apparent how much time margin is needed to determine that the probability of time-available being exceeded is negligible. The probability of time-available being exceeded is dependent on how much the tails of the time-required and time-available probability distributions overlap. The need for a probability distribution of time-available also depends on the risk significance of the HFE. Even a small overlap of the probability distributions (indicating a low probability of time-required exceeding time-available) could be important for a risk significant HFE. That said, the risk significance of a given HFE may not be fully known until the PRA is completed or a quantitative assessment is performed. Accordingly, propagating probability distributions for Case 2 HFEs could also be important.

C.4.2 Estimation of Time Available

Information Request 1: Concerning the point estimates of time-available that are currently performed at NPPs.

1.1 Describe the methods used to estimate the time-available that are used in developing HEPs. Include description of the different applicable situations of available resources that impact the method used that IDHEAS should address.

1.2 For each method described above in response to part 1.1, describe the factors that could impact the uncertainty of the estimate (e.g., Was the estimate based on a plant-specific or a generic evaluation? Was the estimate performed using a conservative assumption or a bounding approach? Was the estimate based on the timing for a different but similar accident scenario?)

1.3 Explain whether any attempt is made to characterize the uncertainty of time-available (e.g., establish a lower and upper bound) and under what circumstances.

All experts agreed that if plant-specific thermal-hydraulic analyses for the conditions of the scenario being assessed was available, they were the preferred source for estimating time-available. However, they also stated that other sources also are currently used depending on the situation. The list of the methods or sources used to estimate the time-available cited in the workshop were:

- Thermal-hydraulic analysis performed for the Final Safety Analysis Report (FSAR) in support of design basis success criteria or Modular Accident Analysis Program in support of PRA success criteria
- Design basis information other than the FSAR
- Calculations based on engineering first-principles, such as determining how long would it take the water in a tank to go from one level to another given a volumetric flow rate
- Other engineering modeling that supports the success criteria used in a PRA, such as room heat-up calculations and internal flood height calculations
- Plant training simulator runs and NRC Technical Training Center simulator runs
- The time-available stipulated in the procedure (if provided)
- Operator experience about timing used in training
- Delay time until a cue is received by the operators that an action is required can come from indicators, annunciators, alarms, or from procedural steps and might be impeded depending on the situation

The industry HRA experts explained that contributors to modeling uncertainty include model completeness and model specificity (e.g., plant-specific or generic). They could also include uncertainty associated with physical plant equipment performance and administrative factors such as procedures. For example, it is difficult to know what the actual time-available margin may be in the case of failed equipment or components. An example was cited where an emergency diesel generator ran for 2 hours without service water. They mentioned that there are many inputs and assumptions made in thermal-hydraulic analysis, and it is hard to know which are key factors for any given situation. One of the industry HRA experts stated that one underappreciated contribution to the uncertainty is the time of the cue or the time delay and how that cue might be defined in the plant procedures. The cue can come from multiple sources such as instrumentation, an alarm, or a procedure-directed step.

The NRC experts did not offer further discussion on specific factors that could affect the uncertainty of the time-available estimates but stated that in many cases the results of thermal-hydraulic analysis and simulator runs tend to be conservative or are used in a conservative manner. The NRC experts explained that they may not address the sources of uncertainty directly but rather address them by examining the assumptions made and their impact on the timing. For example, the time-available to response to a large break loss of cooling accident might conservatively be applied as the time-available to respond to a smaller loss of coolant accident even though the water inventory is being lost at a slower rate, because the time-available for the larger break is known and time-available to respond to a smaller break is not known. The NRC experts also stated that sources of modeling uncertainty are normally addressed in the context of the risk-informed, decision-making process that may provide other options for addressing the sources of uncertainty. For example, instead of calculating the uncertainty associated with an HEP modeled in the PRA, the NPP may choose to use the PRA to show that the risk increases due to setting the HEP to a clearly conservative value is negligible.

Regarding whether any attempt is made to characterize the uncertainty of time-available, the industry HRA experts explained that sensitivity cases are performed on the point estimates of time-available. The results of the sensitivity cases are used to determine the impact across different success criteria and provide a way to bound uncertainty. If it is determined that there are significant differences in the estimation of time-available, a new sequence might be developed so that both the shorter and longer times are modeled in the PRA. The industry HRA experts stated that this primarily just applies to Case 3.

Concerning whether any attempt is made to characterize the uncertainty of time-available, the NRC experts reiterated that the uncertainty is not addressed directly but rather addressed by examining the assumptions made about the accident scenario and the impacts of those assumptions on timing and ultimately on risk. The NRC experts also stated that the sources of HRA modeling uncertainty, in general, may be identified as key source of uncertainty for the risk-informed decision that the HRA is supporting.

C.4.3 Development of Probability Distributions for Time Available

Information Request 2: Concerning development of a probability distribution for time-available.

2.1 Describe any methods, used currently or in the past, by NPPs to develop a probability distribution for time-available.

2.2 Describe methods that could be used by NPPs to develop a probability distribution for time-available based on the approaches identified in part 1.1 above.

2.3 Discuss the pros and cons of each method identified in part 2.1 above.

2.4 Propose a best-practice approach for developing probability distributions for time-available based on your experience and knowledge.

The experts stated that there are no existing methods for developing a probability distribution for time-available. One NRC expert who is an author of the IDHEAS-ECA method stated that a probability distribution can be developed by estimating a range of values and assigning percentiles along that range for risk significant HFEs (IDHEAS-ECA is documented in NUREG-2256 [1]). The advantage of developing a probability distribution for time-available is that it is technically rigorous and characterizes the uncertainty in time-available, but it may be perceived as “mathematically daunting,” may be hard to justify, and may create an opportunity to “play with the numbers.” The other NRC experts stated that they do not develop probability distributions for time-available and expressed concern that the time and resources required to develop probability distributions may not be worth the benefit. They reiterated that this may be particularly true when the estimates are conservative. They also stated that if developing probability distributions becomes an NRC expectation, then sufficient guidance must be provided to the HRA analysts.

The industry HRA experts stated that development of a probability distribution for time-available is currently not done and is not required for risk-informed applications subject to the guidance in NRC Regulatory Guide 1.200 and the ASME/ANS PRA standard [2]–[5]. They also stated that trying to perform a sensitivity study using the thermal-hydraulic codes to estimate the uncertainty of time-available would be difficult given the number of plant parameters needed as inputs, such as number of days at full power, decay times, temperatures and pressures. Concerning best practice, one industry HRA expert suggested that guidance might be developed based a future study using subject matter experts to select cases, determine minimum and maximum times, identify factors that would affect break points in plant parameter curves to generate insights (e.g., “rules of thumb”) that inform best practices.

C.4.4 Methods Used to Estimate Time Required

Information Request 3: Concerning the point estimates of time-required that are currently performed at NPPs.

3.1 Describe the methods used to estimate the time-required that are used in developing HEPs. Include description of the different applicable situations of available resources that impact the method used that IDHEAS should address.

3.2 For each method described above in response to part 3.1, describe the factors that could impact the uncertainty of the estimate (e.g., Was the estimate based on a plant specific or a generic evaluation? Was the estimate performed using a conservative assumption or a bounding approach? Was the estimate based on the timing for a different but similar accident scenario?)

3.3 Explain whether any attempt is made to characterize the uncertainty of time-required (e.g., establish a lower and upper bound) and under what circumstances.

Regarding the point estimates of time-required to perform operator actions, one of the industry HRA experts stated that they follow ASME/ANS PRA standard requirements in Supporting Requirement HR-G5 that stipulates the estimate be based on walk-throughs or talk-throughs of the procedures or simulator observations for risk significant HFEs [2], [3]. The expert stated that the ASME/ANS PRA standard allows the use of engineering judgment for HFEs that are not risk significant, but it is not typically known in advance of completion of the PRA whether a given HFE is risk significant. The expert stated that a best practice approach is using estimates based on walk-throughs or talk-throughs of procedures or simulator observation from the beginning even though the PRA may ultimately show that the HFE is not risk significant. The industry HRA experts also stated that another source of time-required information is the results of Job Performance Measure testing. Plant crews are trained to demonstrate that they can meet Job Performance Measure times that are considered “bounding” times. One industry HRA expert stated that, during talk-throughs of time-required with operators and trainers, typical times and a range of times are discussed. This is especially true for time-critical or time-sensitive actions and helps characterize the uncertainty related to the timing.

The NRC expert representing the IDHEAS-ECA method (NUREG-2256), stated that in reviewing the operational and simulator data that the average, slowest and fastest times should be obtained. Based on this, information percentiles can be assigned, and a distribution assumed. If it is not feasible to estimate a range, then applying a default distribution to the point estimate may be useful. This default distribution reflects the aleatory uncertainty which can be adjusted for certain PIF effects. The NRC expert also referred to a rule-of-thumb cited in NUREG-2256 that suggests one minute per procedural step is needed to perform a task. The NRC expert, who could not trace the origin of this guidance, asked the group for feedback.

Several challenges were raised regarding the “one minute per procedural step” rule of thumb. First, it could be conservative for some steps and not others. Second, it is not clear what “a step” meant because a numbered procedural step can have many subparts of varying degrees of difficulty. Finally, there is significantly more timing uncertainty associated with detecting, understanding, and deciding on a course of action than there is with executing the action itself.

Regarding the point estimates of time-required, the NRC HRA experts stated that they get time-required information from plant or NRC Technical Training Center simulator runs and other plant-specific sources such as Job Performance Measure (JPMs). However, they also stated that they use their own experience as former SROs or the experience of other former SROs who worked at a plant of similar design. They use their own experience as a frame of reference and then adjust it according to the number of steps in the procedure or other factors to account for additional complexities in the accident scenario. The experts explained that if the time-required was close to time-available, they would seek out other resources to get a better time estimate. In response to one of the observers, the NRC experts clarified that they assumed that their experience was average as it seemed to be consistent with the experience of other crews.

Regarding development of uncertainty associated with estimating time-required, the NRC experts, with the exception of the NRC staff representing IDHEAS-ECA, expressed the concern as stated above that the time and resources required to develop a distribution may not be worth the benefits. Moreover, uncertainty can be addressed in different ways, such as employing conservative estimates or by exploring how making different assumptions about the scenario can impact the risk-informed decision to which the operator action pertains. For example, it could be that conservatism associated with assumptions made in modeling an accident scenario in the PRA (in which an HFE is one of the failures in the scenario) is far greater than the possible uncertainty associated with the HFE. Accordingly, refined uncertainty modeling of the HFE in the scenario may not be beneficial.

The industry HRA experts clarified that that, except for Case 3, development of a distribution for time-required is not recommended due to cost-benefit constraints.

C.4.5 Email Feedback from Two PRA/HRA Industry Leaders

Separate from the results summarized above from the workshop, email responses to the information requests were received from two PRA and HRA experts and leaders in this domain. Both experts have significant history with and perspective about the development of HRA approaches and their applications. These experts were asked to provide their insights by either responding to the same information requests as the workshop experts or offering general advice about the topics addressed. Both stated their hands-on HRA experience as not current enough to reply specifically to the information requests and chose to provide general comments on evaluation of time availability associated with the requests that were solicited. Their comments are documented in unprocessed form in provided in PNNL-33857 by Coles et al. [13] which provides a record the email exchange. This section summarizes important insights from these non-workshop experts based on their experience in evaluating exceedance of time-available. This summary focuses on insights that help support the conclusions of this study based on what they add to insights from the workshop results.

This first non-workshop expert stated that there is a concern in the industry about the need to perform “plant-specific thermal-hydraulic analyses” for large numbers of accident scenarios to support HRAs. The expert stated that he thought running the scenarios using the plant-specific simulator is the most effective and efficient way to derive these estimates.

This expert also stated that there are two primary sources that contribute to time-available uncertainty: (1) uncertainty in the nominal values of the parameters that are used as inputs in the thermal-hydraulic models and (2) variability in the nominal scenario that is being analyzed. The expert stated that, in his experience, the effects from the first are generally rather small compared to the effects from the second. The expert stated that PRA analysts define an HFE in the context of a nominal event scenario, which is typically used to inform a large number of individual event sequences (cutsets) that are judged to be “appropriately similar” to the nominal scenario. There are, of course, variations in the actual progression of each specific detailed event sequence. In principle, if those variations are large, the analysts should define distinct HFEs to account for their effects on human performance, but this decision is based on judgment. In practice, the uncertainty associated with time-available for the analyzed HFE is strongly affected by how the analyst decides to group the individual event sequences that are characterized by the nominal scenario. Simulator runs can be adjusted to account for the range of parameter values in the grouped sequences. Alternatively, thermal-hydraulic runs may be able to provide estimates of the upper- and lower-bound times, based on the analyst’s understanding of the range of conditions in the grouped sequences.

Concerning time-required, the first non-workshop expert stated that the only practical way to develop realistic plant-specific estimates is through a combination of simulator observations, physical walk-throughs (especially for local actions), and table-top elicitation sessions with the actual plant operators. The expert stated that the estimated uncertainty should account for 1) uncertainty in the estimates from a particular operating crew and 2) variability in the estimates from multiple crews (e.g., a merged distribution with equal weights assigned to each crew's individual probability distributions). The expert stated that this uncertainty can be rather large, especially if different crews have varying understandings of very "unusual" scenarios. However, he stated the belief that this is a real effect, and it would be observed by actual crew-to-crew variability in practice.

Summary of key takeaways from the first non-workshop expert with a focus on those that add to insights gained from workshop knowledge elicitation experts are:

- Running accident scenarios from the PRA using the plant-specific simulator is an effective and efficient way to determine time-available.
- In many HRAs, the time-available is estimated for a nominal scenario which is then used to adjust or bound time-available for variations for similar scenarios. This approach contributes to the uncertainty which the expert believes is greater than the uncertainty associated with the thermal- hydraulic analysis results.
- The impact of crew-to-crew variability on the time-required probability distribution is real and can be large. This statement is an independent confirmation of conclusions developed and reported in Appendix B of this report.

The second non-workshop expert expressed a preference for estimating time-available using expert elicitation with an interdisciplinary team, including plant licensed operators and trainers, maintenance personnel, engineers, and experts in the PRA and HRA for the plant. The expert explained that the probability distributions should be developed for time-available and time-required before the point estimates are determined. The expert stated that substantial research shows that starting with an estimate of the point value (e.g., the mean) creates a strong bias, and therefore it is better to start by establishing the minimum and maximum values.

The second non-workshop expert stated that in the 1980s and 1990s the EPRI supported simulator experiments to "develop time-required and time-available probability distributions." The expert pointed out that the EPRI work failed to account for differences between operating a simulator and the real plant where concerns of damage to equipment, the economics and potential damage to communities by loss of electricity supply, and potential radiological risk to workers and the public are involved. The study also did not consider the impacts of items not modeled in the PRA on accident progression and operators' attention.

The expert stated that the NRC's a technique for human event analysis (ATHEANA) presented in NUREG-1624, Revision 1, Technical Basis and Implementation Guidelines for ATHEANA [14], [15], discusses the many factors to be considered when evaluating timing. The expert cited the following six factors from the ATHEANA methodology:

1. Specific cause of the initiating event
2. Initial condition of the plant
3. Whether the PRA models equipment as failed completely or degraded
4. Possible status of equipment not modeled in the PRA
5. How operator training and expectations and procedures match the actual sequence of events

6. How the time of day, time during the shift, crew practices and tendencies, informal rules, and communication practices can affect performance shaping factors (PSFs)

The conditions that these factors create are generally covered by PIFs defined in IDHEAS-General method presented in NUREG-2198, The General Methodology of An Integrated Human Event Analysis System (IDHEAS-G) [16]. The expert also mentioned that the plant details and operation not modeled in the PRA can change the accident sequences, create mismatches between what actually happens in an accident scenario and what the procedures anticipate, and can divert an operator's attention or cause confusion.

Summary of key takeaways from the second non-workshop expert with a focus on those that add to insights gained from knowledge elicitation experts are:

- The expert stated that experience and research show that when time-available and time-required estimates are elicited starting with an estimate of the point value, this creates a strong bias in favor of the point value (e.g., the mean). The expert stated that it is better to start by establishing the minimum and maximum values and then determine the point estimate value.
- Simulator data fails to account for differences between operating the simulator and the real plant such as the following which can contribute to the uncertainty of time-available and time-required: 1) damage to equipment, 2) the economics and potential damage to communities by loss of electricity supply, 3) potential radiological risk to workers and the public, and 4) system and situations not modeled in the PRA that can impact the accident progression and/or impact the operators' attention. The expert stated that the ATHEANA methodology addresses these factors.

C.5 Key Insights Gained from the Expert Knowledge Elicitation

This section summarizes the key insights and conclusions of the expert knowledge elicitation and solicitation of feedback from two industry leaders on the information requests regarding time-available and time-required.

C.5.1 Estimation of Time Available

Regarding estimation of time-available, the workshop experts agreed that plant-specific thermal-hydraulic analysis for the conditions of the scenario being assessed was the preferred basis, but that in many cases other bases had to be used. The ASME/ANS PRA standard requires time-available estimates be based "on appropriately realistic generic thermal-hydraulic analysis or simulation from similar plants (e.g., plant of similar design and operability)" for CC-II [2], [3]. Therefore, some suggestions made by workshop experts and non-workshop leaders should not be used in cases for which conformance to CC-II of the PRA standard is required. The list below provides the approaches cited by the workshop and/or non-workshop experts in the order of preference based on the authors' understanding of the elected information.

1. Thermal-hydraulic analysis performed for the FSAR in support of design basis success criteria or Modular Accident Analysis Program runs in support of PRA success criteria.
2. Design basis information other than the FSAR.
3. Calculations based on engineering first-principles such as determining how long would it take the water in a tank to move from one level to another given a volumetric flow rate.
4. Other engineering modeling that supports the success criteria used in a PRA such as room heat-up calculations and internal flood height calculations.
5. The time-available stipulated in the procedure (if provided).

6. Plant and NRC Technical Training Center simulator runs.
7. Use of expert elicitation with an interdisciplinary team including plant licensed operators and trainers, maintenance personnel, engineers, and experts in the PRA and HRA at the plant.
8. Operator experience about timing used in training.
9. Delay time until a cue is received by the operators that an action is required can come from indicators, annunciators, alarms, or from procedural steps and might be impeded depending on the situation.

PNNL Conclusion 1: Although the information presented in the list above is not inconsistent with guidance provided in NUREG-2256, Chapter 3.6 [1], on determining time-available, the list above provides more detail and options.

C.5.2 Development of Probability Distribution for Time Available

Concerning development of a probability distribution for time-available, one of the NRC HRA experts stated the advantage of developing a probability distribution for time-available is that it is technically rigorous and characterizes the uncertainty. However, other NRC HRA experts stated that development of a distribution may not always be worth the effort particularly when estimates are made conservatively. The industry HRA experts stated characterization of the uncertainty associated with time-available is not required by the PRA standard. Moreover, they stated that in many cases sensitivity studies are performed on the assumed point estimate values to assess the impact of the uncertainty on the risk results or ultimately on the conclusions of the risk-informed application supported by the PRA and HRA. This can include the uncertainty associated with time related failure of operator actions. The industry HRA experts also explained that the point estimate of time-required may be significantly less than the time-available or higher than the time-available, and that in those cases characterization of the distribution of time-available is not beneficial. Accordingly, there may not be a compelling reason to characterize the time distribution depending on the point estimate difference between time-available and time-required and the results of applicable sensitivity studies.

The industry HRA experts suggested that trying to perform a sensitivity study using the thermal- hydraulic codes to estimate the uncertainty of time-available would be difficult given the number of plant parameters needed as inputs. However, they proposed a study be performed that addresses this concern using subject matter experts. In addition, the non-workshop expert stated that there was an additional uncertainty contribution he considered more important than the uncertainty associated with thermal- hydraulic code inputs. The expert stated that, typically, because of the resources needed to perform thermal-hydraulic code runs, time-available is estimated for a nominal case (i.e., nominal scenario) which is then used to adjust or bound time-available for variations from the nominal case (i.e., similar scenarios). Accordingly, this source of uncertainty along with thermal-hydraulic codes input uncertainty would need to be addressed (e.g., identify break-points in plant parameter curves). The industry HRA experts also stated that there can significant uncertainty in the delay time until a cue is received and detected by the crew and then diagnosed and acted upon. This uncertainty contribution associated with this delay time also impacts the determination of time-available.

In summary, concerning development of a probability distribution for time-available, the following are key finding from the workshop and non-workshop experts:

1. Development of a probability distribution for time-available is not currently performed and is not required by the PRA ASME/ANS standard even though a technically rigorous characterization of the uncertainty could be beneficial.

2. It may not be beneficial to characterize the uncertainty of time-available for several reasons, including cases in which there is significant difference between the point estimates of time-available and time-required or low sensitivity of the overall PRA risk results to the failure of the operator action.
3. To make meaningful progress on characterizing the probability distribution of time-available, it appears that a study is needed that includes expert elicitation to address 1) the uncertainty associated with thermal-hydraulic analysis inputs, 2) the uncertainty associated with available margin until equipment fails or in case of failed equipment, 3) the process of using representative results to address the large number of variations in HFEs across scenarios and 4) the uncertainty associated with the time delay until a cue is received by the operators indicating an action is needed.

PNNL Conclusion 2: There are several compelling reasons why modeling the uncertainty associated with time-available and time-required for an HFE may not be needed or be beneficial that are not described in NUREG-2256 [1]. Only a subset of HFEs modeled in a PRA are typically important to risk; therefore, modeling the uncertainty of time-available and required Is not needed for many HFEs. The need for determining the contribution of time exceedance to an HEP might be gauged using importance measures (e.g., Fusel-Vesely or Risk Achievement Worth values) as determined by the PRA. However, the challenge of this approach is that importance measures cannot be determined until the PRA is complete. Also, in many instances, the time-available exceeds the time-required by a significant margin, and therefore, the probability of exceeding the time-available is negligible. The challenge in this case is knowing how much time margin is enough to render the probability of exceeding the time-available negligible. Additionally, even though a particular HFE may have a level of risk significance, it 1) may not be important to the risk- informed application that the PRA is supports such as a Risk Informed Completion Time (RICT) program[4] or 2) the uncertainty of an HEP may be addressed within the context of the risk informed program. For example, if the failure of an operator action is important to a particular RICT plant configuration then the program may allow Risk Management Actions to prevent or mitigate the risk during the RICT. Finally, in some cases, NPPs already use Human Cognitive Reliability/Operator Reliability Experiments HRA for modeling non-response based on time-available versus time-required.

PNNL suggests that reluctance by industry to devote resources to determine the probability of exceeding time-available might be reduced by acknowledging the possibilities cited above or by providing guidance about when determining the probability of exceeding time-available is necessary.

PNNL Conclusion 3: Experts indicated that much more work is needed to reliably develop distributions for time-available. They suggested that a new study is needed on developing a probability distribution for time-available using industry experts and an expert elicitation process. The experts identified at least four sources of uncertainty that would need to be addressed in such a study, namely:

1. Uncertainty associated with thermal-hydraulic analysis inputs
2. Uncertainty associated with available margin until equipment fails or in case of failed equipment
3. Process of using representative results to address the large number of variations in HFEs across scenarios

---

[4] NRC endorses the use of RICTs in accordance with Technical Specification Task Force (TSTF)-505, Revision 2, “Provide Risk Informed Extended Completion Times – RITSTF Initiative 4b” [17]. This program provides a way to increase the allowed outage time of NPP equipment past their Technical Specification limits using risk information from PRA.

4. Uncertainty associated determining the delay time until a cue is received by the operators indicating an action is needed

PNNL suggests that given the complexity of assessing multiple sources of uncertainty on determining the distribution of time-available, a separate study is warranted using industry experts including thermal-hydraulic code experts, HRA analysts, and reactor operations experts if progress in this area is desired.

C.6 Methods Used to Estimate Time Required

Concerning methods used to estimate the time-required to perform an operator action, one of industry HRA experts stated that the PRA ASME/ANS standard requires these estimates for risk significant HFEs to be based on (1) walk-throughs or talk-throughs of the procedures, or (2) simulator observations to meet CC-II. The expert stated that even though engineering judgment is allowed for non-risk significant HFEs, it is difficult to know the significance level of an HFE in advance of completing the PRA. Therefore, ranges of time-available are discussed as part of talk-throughs and walk-throughs with operators and trainers for all HFEs. The industry experts stated that another source of time-required information for some plants are the results of JPM tests. This information is plant and crew specific and could possibly be used to develop a probability distribution depending on extent of the records and level of detail.

One of the NRC HRA experts stated that in reviewing operational and simulator data the average, slowest, and fastest times should be obtained, and a distribution estimated by assigning percentiles. This is supported by another expert who stated that it is better to start by establishing the minimum and maximum time values and then determining the point estimate value. The expert further stated, however, that if it is not feasible to estimate a range, then the approach could be employed using the approach presented in Section 4.2.2 of this RIL as supported Appendix D.

Some experts stated that they use their own experience to establish benchmark values, and then make adjustments to these values to match the HFE being evaluated such as adjusting the time for number of steps or the additional complexity in the scenarios. These experts stated, however, that other ways of making the estimate are used if the difference between time-available and time-required is small. Note that this approach is not acceptable for PRAs, which must meet the PRA ASME/ANS standard.

One of the non-workshop experts stated that simulator data fail to account for differences between simulator experiments and actions in a real plant such as the following which can contribute to the uncertainty of time-available and time-required: (1) extent of damage to equipment, (2) the economics and potential damage to communities by loss of electricity supply, (3) potential radiological risk to workers and the public, and (4) system and situations not modeled in the PRA that can impact the accident progression and/or impact the operators’ attention. The expert stated that the ATHEANA methodology addresses these factors. These factors can divert an operator’s attention or cause confusion. These factors appear to be related to the PIF identified in NUREG-2198 [16] as Multitasking, Interruptions, and Distractions. Multitasking refers to performing concurrent and commingled tasks. Interruptions and distractions refer to activities that interfere with the operator’s performance of the primary task. The expert makes the point that the factors associated with Multitasking, Interruptions, and Distractions may not be exclusively associated with safe shutdown but are important for other reasons such as those cited above. These factors are usually not present in a simulation.

In summary, concerning methods and sources used to estimate the time-required to perform an operator action:

1. For risk significant HFE events, walk-throughs or talk-throughs of the procedures with operators or trainers and use of simulator observations are the primary way to estimate the time-required to perform an operator action and meets the PRA ASME/ANS standard.
2. For some plants, a good sources of simulator observations are the results of applicable JPM tests. This information is plant and crew specific and can possibly be used to develop probability distribution depending on extent of the records and level of detail.
3. If compliance with CC-II of the PRA ASME/ANS standard is not required, then an approach that might be used in certain cases is an analyst's use own their experience (or the experience of a trusted source) to establish benchmark times. These times can then be adjusted to match the HFE being evaluated such as adjusting the time for number of steps or the additional complexity in the scenarios. This approach might be used for which the risk results are not sensitive to uncertainty in the time-required or when the time margin between time-available and time-required is large.
4. Caution should be taken when using the "rule of thumb" that that estimates time-required by assuming each procedural step takes one minute. This approach, even when used as a rough estimate, can be conservative in some cases and non-conservative in other cases. The experts noted that: (1) there can be significantly more time uncertainty associated with detecting, understanding, and deciding on a course of action than there is with executing a physical action, (2) it is not clear what a "a step" means as a numbered procedural step can have many subparts of varying degrees of difficulty, and (3) the source of the rule-of-thumb is not clear.
5. When reviewing operational and simulator data the average, slowest and fastest times should be obtained, and a distribution estimated by assigning percentiles.
6. When considering performance impacts for time-required, the uncertainty associated with activities that occur in the plant besides those specifically associated with safe shutdown should be factored in. This includes considerations such as attending to concerns that are not modelled in the PRA such as routine tasks and actions needed to avoid equipment loss or damage or to avoid safety concerns not related to nuclear safety (e.g., radiation safety or occupational safety). These uncertainties and activities can divert an operator's attention, affect a scenario, or impact the operator's understanding of a scenario.

PNNL Conclusion 4: The insights listed above concerning the methods and sources used to estimate the time-required can be used supplement the guidance in NUREG-2256. The expert knowledge elicitation did not address development of probability distributions for time-required to the extent as the other topics, because that topic is substantially covered by guidance that already exist in in NUREG-2256 and Section 4 of this report.

C.7 References

[1] J. Xing, Y. J. Chang, and J. DeJesus Segarra, "NUREG-2256, integrated human event analysis system for event and condition assessment (IDHEAS-ECA)," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML22300A117, Oct. 2022.

[2] ASME/ANS (American Society of Mechanical Engineers/American Nuclear Society), "Addenda to ASME/ANS RA-S–2008 standard for level 1/large early release frequency probabilistic risk assessment for nuclear power plant applications," ASME/ANS (Ameri- can Society of Mechanical Engineers and American Nuclear Society), New York, NY, ASME/ANS RA-Sa–2009, Feb. 2009.

[3] ASME/ANS (American Society of Mechanical Engineers/American Nuclear Society), "Probabilistic risk assessment standard for advanced non-light water reactor nuclear power plants," New York, NY, ASME/ANS RA-S-1.4–2021, Feb. 2021.

[4] NRC (U.S., Nuclear Regulatory Commission), NRC (U.S., Nuclear Regulatory Commission), NRC (U.S., Nuclear Regulatory Commission), "Regulatory guide 1.200, revision 2, an approach for determining the technical adequacy of probabilistic risk assessment results for risk-informed activities.," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML090410014, Mar. 2009.

[5] NRC (U.S., Nuclear Regulatory Commission), "Regulatory guide 1.200, revision 3, acceptability of probabilistic risk assessment results for risk-informed activities," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML20238B871, Dec. 2020.

[6] U.S. Code of Federal Regulations (CFR), *10 CFR 50.69*, Jan. 1956. [Online]. Available: https://www.ecfr.gov/current/title- 10/part-50/section-50.69.

[7] V. G. Cusumano, "Final revised model safety evaluation of traveler tstf-505, revision 2, provide risk- informed extended completion times - RITSTF initiative 4b," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML18267A259, Nov. 2018.

[8] J. Xing and S. Morrow, "White paper: Practical insights and lessons learned on implementing expert elicitation," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML16287A734, Oct. 2016.

[9] J. P. Kotra, M. P. Lee, N. A. Eisenberg, and A. R. DeWispelare, "NUREG-1563, branch technical position on the use of expert elicitation in the high-level radioactive waste program.," Center for Nu- clear Waste Regulatory Analyses (CNWRA), U.S. Nuclear Regulatory Commission/NMSS/DWM, Washington, DC, ADAMS Accession No. ML033500190, Nov. 1996.

[10] NRC (U.S., Nuclear Regulatory Commission), "NUREG/CR-6372, Volume 1, recommendations for probabilistic seismic hazard analysis: Guidance on uncertainty and use of experts: Main report," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML080090003, Apr. 1997.

[11] NRC (U.S., Nuclear Regulatory Commission), "NUREG-2117, Rev. 1, practical implementation guidelines for SSHAC level 3 and 4 hazard studies," U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML12118A445, Apr. 2012.

[12] NRC (U.S., Nuclear Regulatory Commission), “NUREG-2213, updated implementation guidelines for sshac hazard studies,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML18282A082, Oct. 2018.

[13] G. Coles, R. Prasad, J. Baweja, and B. Jefferson, “Expert knowledge elicitation process on methods to determine the probability of time exceeded,” Pacific Northwest National Laboratory, Richland, WA, Limited Distribution. PNNL-33857, Jan. 2023.

[14] NRC, “NUREG-1624 Rev, 1, technical basis & implementation guidelines for a technique for human event analysis (ATHEANA),” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML003719212, May 2000.

[15] NRC (U.S., Nuclear Regulatory Commission), “NUREG-1624 Rev, 1, technical basis & implementation guidelines for a technique for human event analysis (ATHEANA), appendices a - g,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML003719239, May 2000.

[16] J. Xing, Y. J. Chang, and J. DeJesus Segarra, “NUREG-2198, the general methodology of an integrated human event analysis system (IDHEAS-G),” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML21127A272, May 2021.

[17] M. R. Acevedo, J. Ake, J. Bommer, K. Coppersmith, M. Juckett, C. G. Munson, and J. Stamatakos, “NUREG-2213, updated implementstion guidelines for SSHAC hazard studies,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML18282A082, Oct. 2018.

Appendix D
–
Statistical Derivation of Time Required Distribution

To begin modeling time-required distributions, it was first necessary to acquire data that represented real or realistic times required to address a wide variety of tasks and plants. Data that PNNL acquired included: [1]–[3]

- HAMMLAB U.S. HRA Empirical Study: The U.S. HRA Empirical Study – Assessment of HRA Method Predictions against Operating Crew Performance on a U.S. Nuclear Power Plant Simulator. NUREG-2156 [4]
- US HRA Data Collection: A HAMMLAB HRA Data Collection with U.S. Operators, 2016, HWR- 1123, OECD Halden Reactor Project [5]
- HAMMLAB International Empirical Study: Results from comparing HRA methods predictions to HAMMLAB simulator on SGTR scenarios in Phase 2 Report [1] and LOFW Scenarios in Phase 3 Report [2]
- UJV Data (Nuclear Research Institute of the Czech Republic): UJV provided PNNL with unpublished time-required data for 17 crews that participated in LOFW simulator trials that were held in 2011 [6]. Two crews committed errors so that information had to be removed from the dataset. The simulators were of a Russian designed water-water energetic reactor (VVER) PWR reactor and the LOFW scenario required that they isolate coolant from the rupture half of the VVER feedwater collector and transfer it over to the “healthy” unruptured half.
- Analysis of Human Performance Observed Under Simulated Emergencies of Nuclear Power Plants, 2005 – KAERI/TR-2895/2005, Korea Atomic Energy Research Institute [7]
- EPRI data was considered due to:
  - General lack of availability of publicly available data (1) from consistent control room environments and tasks, (2) from consistent simulator scenarios (training vs. experimental HRA scenarios), and (3) with granularity of completion time to allow for analysis of different tasks and events within a scenario.
  - Quantity and realism (1,068 records that are not duplicates or missing values)
  - Crew by crew specificity (max sample size of 18 for each human interaction event
  - The accompanying questionnaires that allow for some mapping of events to PIFs
- Akaike Information Criterion was used to determine distribution fits across a number of normalizing transformations and families of distributions.
- After mathematical fitting, visualizations of the distributions were also produced to provide a visual check that the distribution has properties that analysts expect to see from experience
- The 0.28 shape parameter for the Lognormal distribution is found by minimizing the Akaike Information Criterion (AIC) for all EPRI data across 4 distribution families and 3 possible transformations. Another analysis supports values closer to 0.54 but have more uncertainty.

D.1 Data Fitting Procedure

A maximum likelihood estimation procedure is implemented for determining distribution parameter values. The negative log-likelihood (NLL),

$$NLL = \sum_{x \varepsilon \mathrm{Data}} \log f(x|\theta)$$

is estimated for density function, f, and candidate parameter values, θ. An adjustment to the NLL, AIC discounts the number of free parameters for each distribution and is used for as our recommendation criterion. Parameter values are explored in a local region until a local minimum is obtained. In addition to the AIC, a visualization of the distribution is inspected for supplementary evidence that the distribution family aligns with the data provided.

Four distributions were considered based on guidance about the general shape of completion time data in nuclear power plant environments (long tail with most crews completing the task well-under previously specified timeframes): Lognormal, Weibull, Exponential, and a truncated Normal distribution. Each of these distributions have the desired long tail and a large mass closer to smaller times. There are some key differences, however. The exponential distribution has density that increases as $x \rightarrow 0$. Completion times related to human performance won't have this property because typically, humans are subject to a lower bound on how fast completion times can be. However, without exact knowledge of that lower bound and no constraints on sample size, exponential distributions may fit the data available. The lognormal distribution classically captures the lower bound of human performance in completion times. The distribution has three parameters that impact its shape (how different the tail appears), location (where the distribution is centered), and scale (the number of values covered by the majority of the distribution). The Weibull distribution is a generalization of the exponential distribution and can also be interpreted as a distribution of failure times. As such, the Weibull distribution provides a reasonable candidate for completion times. Lastly, there are some applications where the tail of a completion time distribution is not as long, and the distribution is much closer to a normal distribution. Thus, the truncated normal distribution is also considered.

Provided data were fit using the aforementioned AIC procedure for each of the four distributions. The distribution providing the minimum AIC value was deemed to be the best fitting distribution and was the distribution recommended by PNNL for the first-order distribution for the IDHEAS model. This recommendation criteria is an objective approach to modeling data distributions. This approach also requires data that is representative of the tasks, environments, and scenarios that crews face in order to be an accurate representation of completion times. The accuracy of the recommendation is completely determined by the quality of data provided to PNNL.

D.2 Data Normalization

The EPRI data we were provided contained eight broad level experiments on six nuclear power plant designs. These experiments spanned 36 scenarios and 148 different operator actions (human interactions, HI) that were analyzed on a crew-by-crew level basis where data were not missing or duplicated. Appendix A contains a summary count of the number of records for each HI for the entire, unfiltered dataset. Across the HI's the average completion times of crews within a HI range from just a few seconds (minimum 1.96 – PWR-3/HI1P3-2-1) to over a half hour (maximum 37 minutes – PWR-2/HI1P2-7-5B). The range of these values can be seen in Figure 1. The entire spread of completion times provided per HI is shown in Appendix B. Given that PNNL was tasked with providing a distribution recommendation on the task level, we could not simply aggregate the raw data (with this much spread and scenario difference) and provide a robust recommendation. Thus, we applied three normalization techniques to attempt to coerce the data onto comparable terms for a robust fit.

The three normalization techniques considered are standard mathematical techniques for getting data on the same scale without compromising some distributions qualities like skew. Linear transformations do not change the shape of the distribution but may impact the appearance by shifting the data to left (dividing by a large number) or centering the data (division by the mean or median). Additionally, since we aggregate the transformed data to derive a distribution, we must manage how the data is centered along with the scaling values (minimum and maximum values). Candidate transformations included

- Scale by maximum: Divide the HI completion time (CT) data by the corresponding HI's maximum CT. This transformation scales data from each HI to a bounded interval but does not center all CTs in an equitable way across HIs.

$$Transformed\ CT_{HI} = \frac{Raw\ CT_{HI}}{\max\,(CT_{HI})}$$

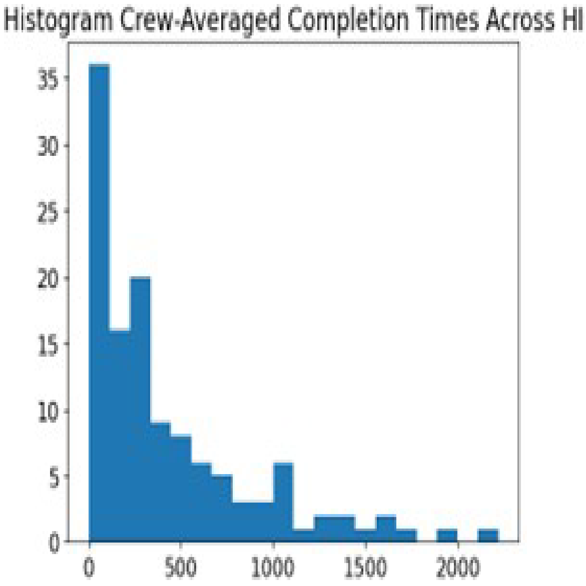


Figure D.1. **Range of Crew-Averaged Completion Times**

- Scale by median: Divide the CT data by the corresponding HI's median CT. This transformation scales the data and centers around the median (a less biased center for skewed distributions) via a single transformation but may not account for relative variances in an equitable way across HIs.

$$Transformed\ CT_{HI} = \frac{Raw\ CT_{HI}}{median(CT_{HI})}$$

- Linear transformation: Apply a linear model that first divides the HI completion time data by the corresponding HI's maximum task completion time and then subtract the mean of these ratios. This transformation scales the data to the [0, 1] interval and centers it.

$$Transformed\ CT_{HI} = \frac{Raw\ CT_{HI}}{\max(CT_{HI})} - mean\left(\frac{RawCT_{HI}}{\max\,(CT_{HI})}\right) + \text{Generalized Intercept}$$

The generalized intercept is derived by linearly regressing completion times scaled by the HI maximum to the HI categories themselves. This intercept acts to center the data across HIs.

For each normalization technique, data provided to PNNL are first normalized and then fit independently to each of the four distributions discussed in the Data Fitting Procedure subsection. Across the 12 (3 transformations x 4 distributions) distribution fits, the distribution with the smallest AIC is considered as the recommended distribution family. The data are fit at two scales. First, each EPRI series (e.g., BWR-1, PWR-3, etc.) is normalized and all series' normalized data are aggregated together prior to fitting. This provided the most informed result as it combines the data from all experiments in the EPRI dataset. Second, each series is normalized and fit independently for a series-level distribution fit. This aims to provide additional assurance that the recommended distribution family is appropriate at this level of applying the IDHEAS model. A third scale (the HI-level) was applied to the series with the most data, BWR-3. Finally, for the resultant distribution family from the above investigation, parameters are extracted. Since analysts typically provide a single point estimate for time-available, PNNL tested the generalizability of these extracted parameters so that analysts can continue to provide a single (or two) point estimate(s) and still get reasonable distribution results.

D.3 Results

Table D.1 shows the distribution along with its AIC for each series (including where all series data are aggregated) and each transformation. After transforming the data by the linear model described in the previous subsection, AIC values are lowest for the lognormal distribution for most series and for the aggregated data. In most cases, the distribution family found by fitting all series' data is the same distribution family for the independent series. PWR-1 Series 1 is the only non-conforming dataset. Given the comparatively better fit of the lognormal distribution for the linear transformation on most series and the aggregated data, the lognormal distribution is our initial candidate for modeling the first-order distribution for time-available in the IDHEAS model. The histogram of the transformed and aggregated data can be seen below.

Table D.1. Best Fits for Each Transformation x Series

Series	Linear Model	Scale by Max	Scale by Median	Unnormalized
All Data	LogNorm (-155.29)	Weibull (319.44)	LogNorm (2142.09)	Weibull (16068.83)
BWR-1 Srs 2	LogNorm (9.54)	Weibull (19.56)	LogNorm (216.95)	Weibull (1482.82)
BWR-1 Srs 3	LogNorm (-9.19)	Weibull (20.24)	LogNorm (242.96)	Weibull (1724.4)
BWR-2	LogNorm (-8.25)	Weibull (49.39)	LogNorm (253.85)	Weibull (1634.86)
BWR-3	LogNorm (-31.86)	Weibull (-1.09)	LogNorm (465.82)	Weibull (2497.42)
PWR 1-Srs 3	LogNorm (-83.76)	Weibull (28.48)	LogNorm (395.92)	Weibull (3564.61)
PWR-1 Srs 1	Normal (-13.44)	Weibull (-6.1)	LogNorm (74.99)	Exp. (1086.61)
PWR-2	LogNorm (4.29)	Weibull (49.06)	LogNorm (264.65)	Exp. (2544.67)
PWR-3	Weibull (-21.98)	Weibull (0.13)	Weibull (142.5)	Weibull (1357.28)

The remainder of this subsection will discuss the derivation of the shape parameter, 0.28 (corresponding to an error factor of 1.58) along with providing evidence for general applicability and some visualizations that compare this shape parameter with HI-level lognormal fits.

****Note:** One may wonder about the Weibull distribution given that that Weibull distribution does appear to fit the unnormalized data. Recall, the unnormalized data varies significantly in range with HI being the main driver of the variance within a series. Without accounting for the impact of individual HI’s there is an artificial tail in the distribution that corresponds to the few HI’s that have completion times well over a few minutes. To use the Weibull distribution based on this analysis would not be appropriate since the IDHEAS model is designed to be applied for a single event at a time and not a host of different events co-occurring. We only include this data for completeness.

D.4 Shape Parameter 0.28

Shape parameters were extracted for all instances of a lognormal fit (see Table D.2). The shape parameter 0.28 is corresponds to the linear transformation of all series’ data in EPRI. As the transformation and distribution that provides the lowest AIC and considering that all EPRI data were used for this fit, this shape parameter is a candidate starting point for the first-order distribution.

Table D.2. Shape Parameters for Lognormal Fits

Series	Linear Model	Scale by Max	Scale by Median	Unnormalized
All Data	0.28	Weibull	0.51	Weibull
BWR-1 Series 2	0.29	Weibull	0.58	Weibull
BWR-1 Series 3	0.39	Weibull	0.55	Weibull
BWR-2	0.38	Weibull	0.61	Weibull
BWR-3	0.45	Weibull	0.78	Weibull
PWR 1-Series 3	0.33	Weibull	0.35	Weibull
PWR-1 Series 1	Normal	Weibull	0.15	Exponential
PWR-2	0.14	Weibull	0.39	Exponential
PWR-3	Weibull	Weibull	Weibull	Weibull

Next, we test the robustness of the 0.28 shape parameter (equivalently, 1.58 error factor) against each HI in the EPRI dataset. Since the fitting procedure above only considered transformed data, it is prudent to test the 0.28 shape parameter against untransformed data. For each HI, we fit a lognormal distribution with location parameter 0, scale parameter equal to the mean completion time for that HI and vary the shape parameter. Shape parameters are either

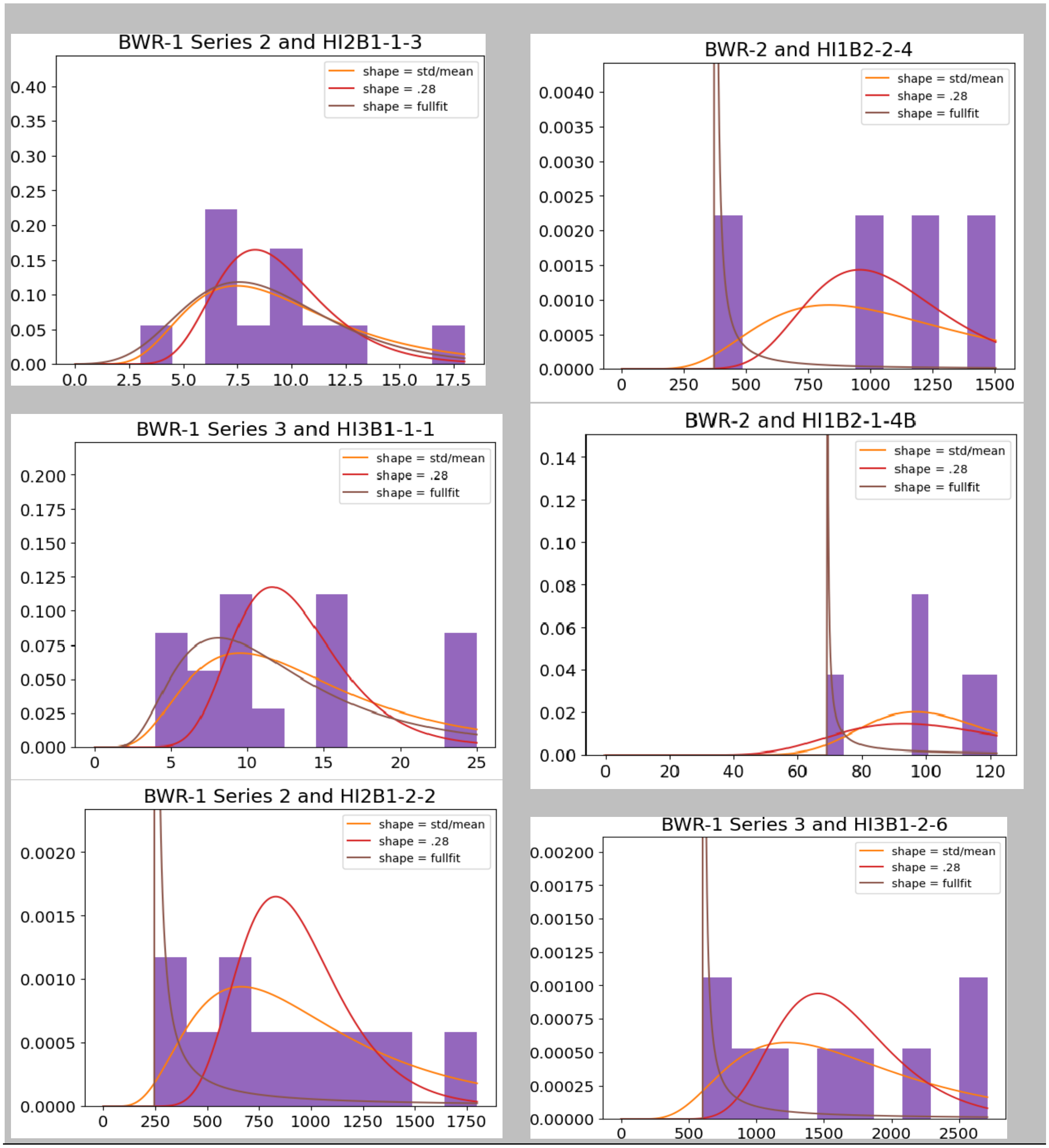
- Standard Deviation/mean – The standard deviation of the set of completion times for the given HI divided by the mean completion time
- 0.28

OR

- Full-fit shape parameter – The shape parameter derived from applying the AIC fitting procedure for the lognormal distribution to the completion time data for the given HI

To appropriately compare the full-fit shape parameter, it was necessary in this case to also adjust the scale parameter to the scale derived from the fitting procedure.

Measuring NLL, the shape parameter standard deviation/mean provided best fits for 112 of the 123 HIs in this analysis. The 0.28 shape parameter best fit 8 HIs and full fitting best fit 3. The fact that the full fit model did not have the lowest AIC could be an indicator that there are too few data to provide accurate fits at the HI level. Nevertheless, we visualize the fits for representative HIs in Figure D.2. Qualitatively, the standard deviation/mean shape parameter and 0.28 are very comparable. The average standard deviation/mean value across events is 0.537 (or 2.419 error factor). Since there are so few data for each HI (18 or fewer), the NLL may not provide an accurate measure of parameter fit. PNNL concludes that the 0.28 shape parameter has technical merit and higher values (towards 0.537) may fall within a reasonable range for untransformed data. Since there are so few data for each HI (18 or fewer data points), the NLL or AIC may not provide an appropriate measure of parameter fit. Thus, the higher values for shape parameter estimates have a degree of uncertainty.


BWR-1 Series 2 and HI2B1-1-3

BWR-2 and HI1B2-2-4

BWR-1 Series 3 and HI3B1-1-1

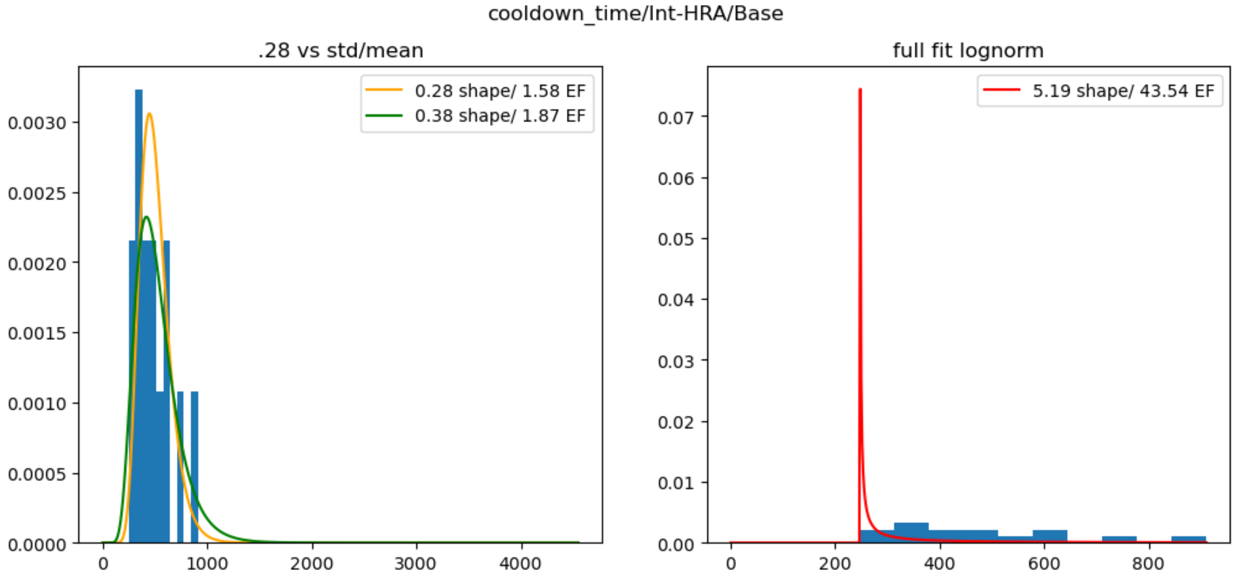
BWR-2 and HI1B2-1-4B

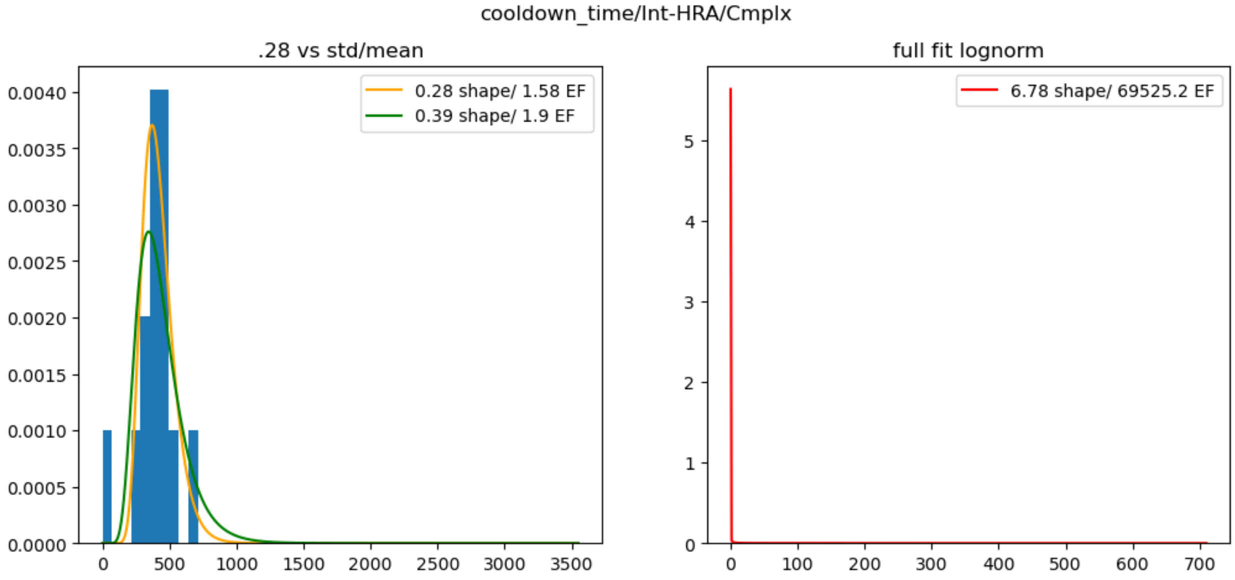
BWR-1 Series 2 and HI2B1-2-2

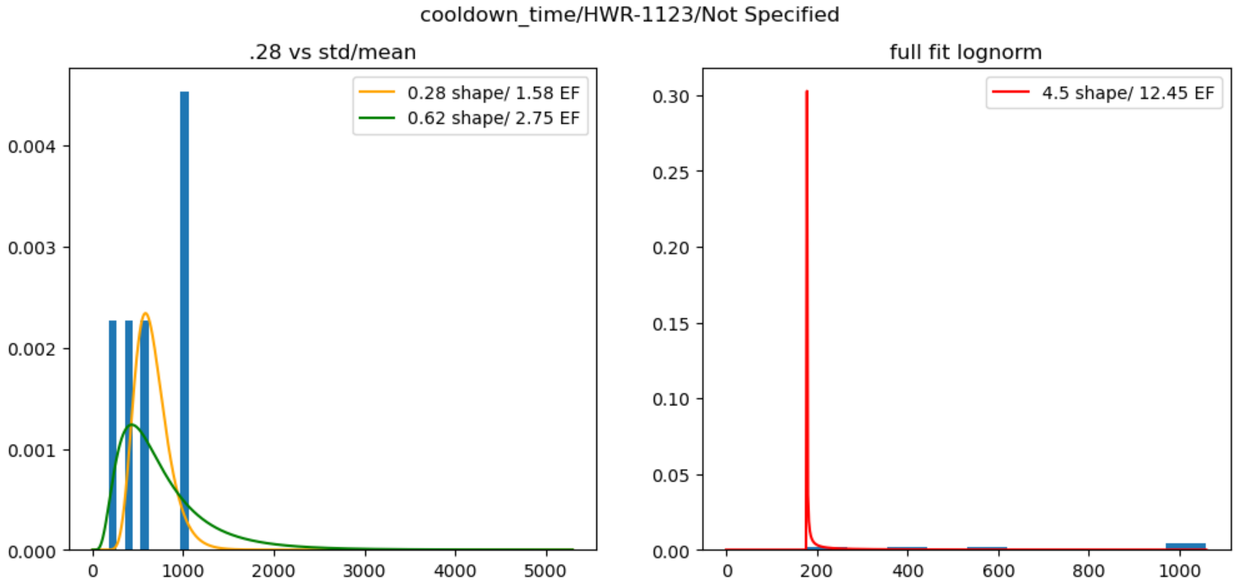
BWR-1 Series 3 and HI3B1-2-6

Figure D.2. Distribution Layover to Histograms Using Different Shape Parameters

Similar testing was conducted with HAMMLAB data from HWR-1123, NUREG/IA-0216, Vol. 2, and HWR-981. In the following figures, PNNL applied the above fitting procedure with varying shape factor to events of the Halden studies. Both complex scenarios and base scenarios are fit to test the robustness of the fits. Error factors are also recorded. In many cases there are too few data to attain an accurate full-fitting distribution. The data contained a single crew's data for the isolation time (starting from the alarm onset) in the complex scenario for the international study, thus a distribution was not estimated. However, for the remaining events, 0.28 provides a reasonable approximation of the available data. NLL values are comparable for the standard deviation/mean shape values and the 0.28 shape values.



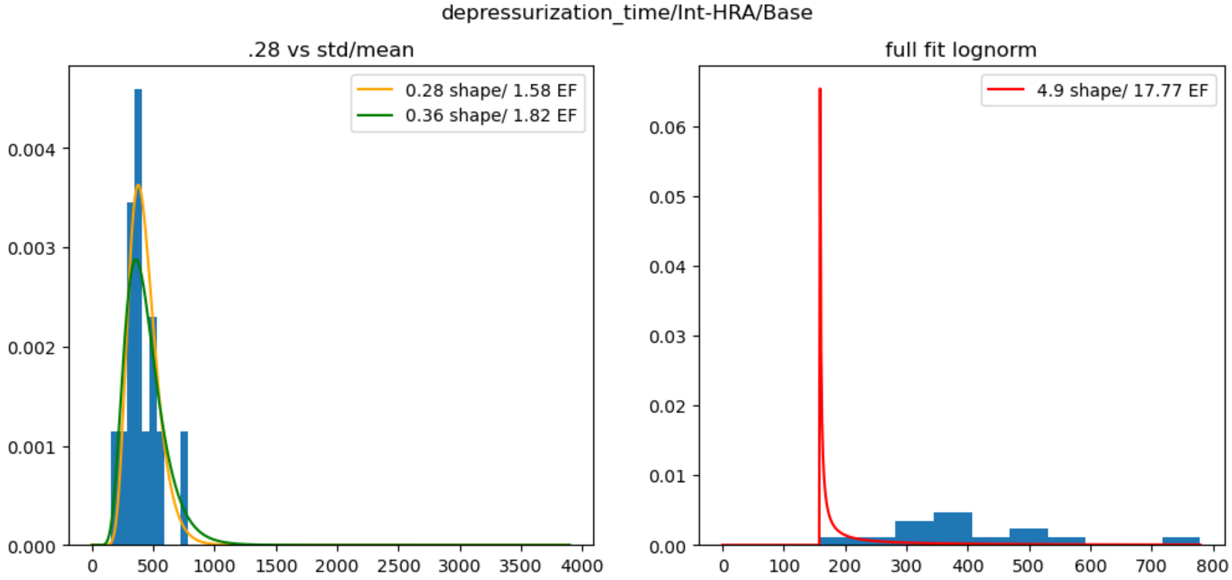




cooldown_time/HWR-1123/Not Specified

.28 vs std/mean

full fit lognorm

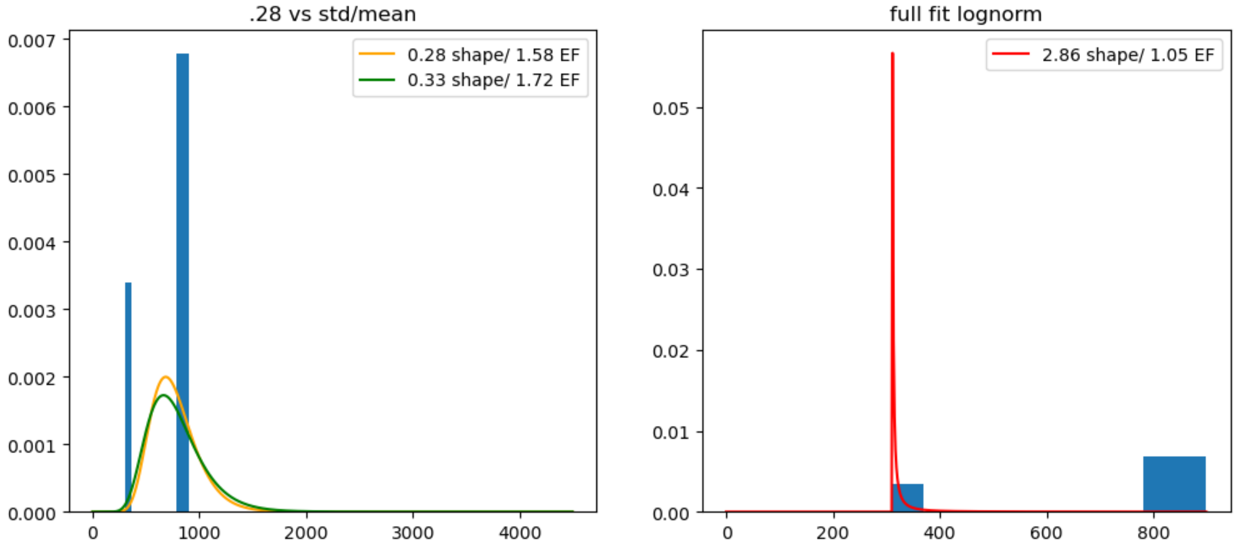


depressurization_time/Int-HRA/Base

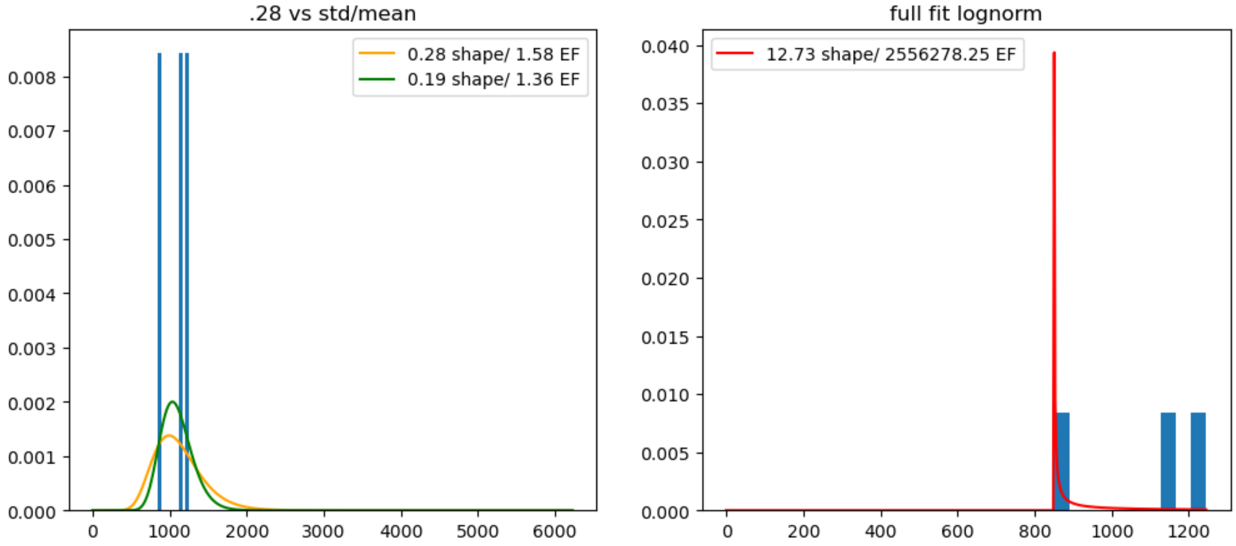
.28 vs std/mean

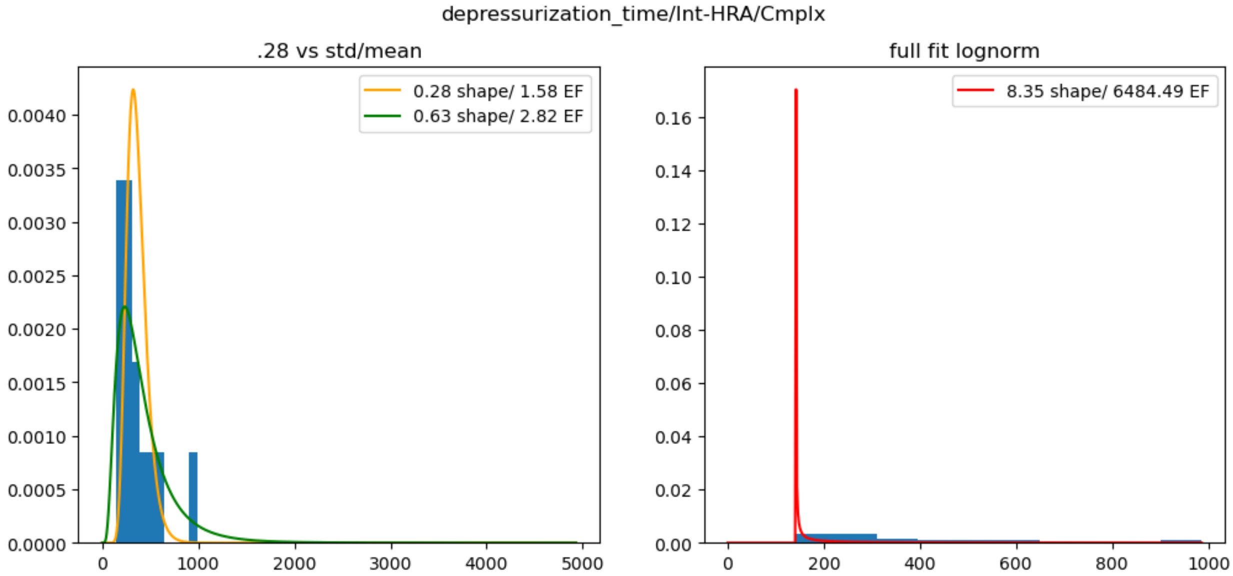
full fit lognorm

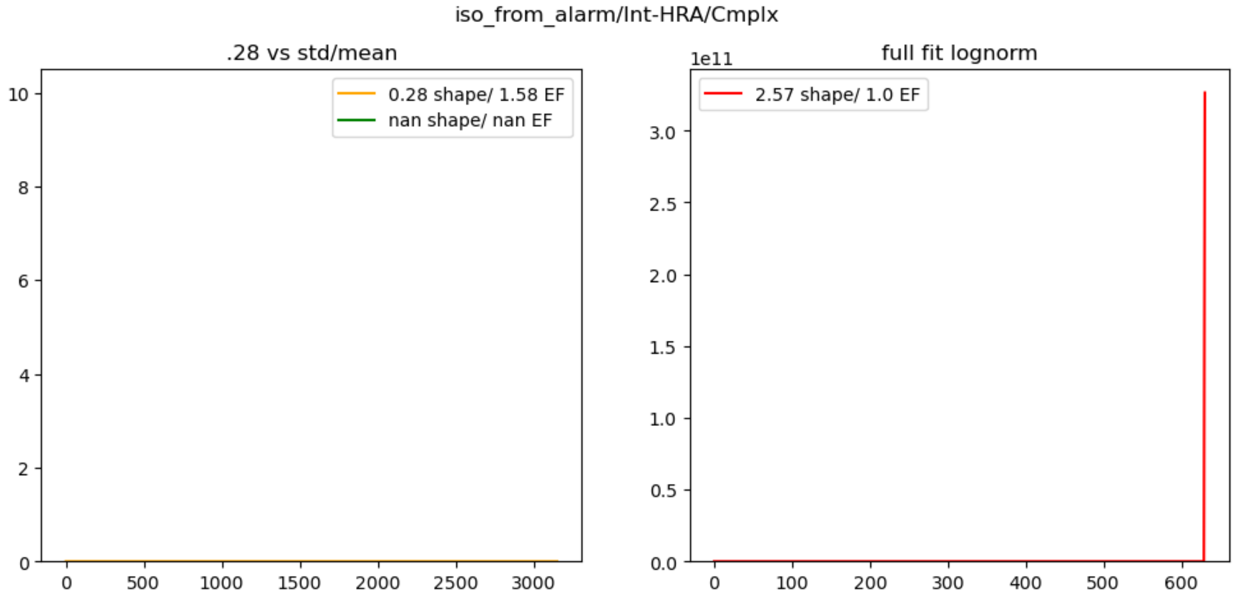
depressurization_time/HWR-1123/Not Specified



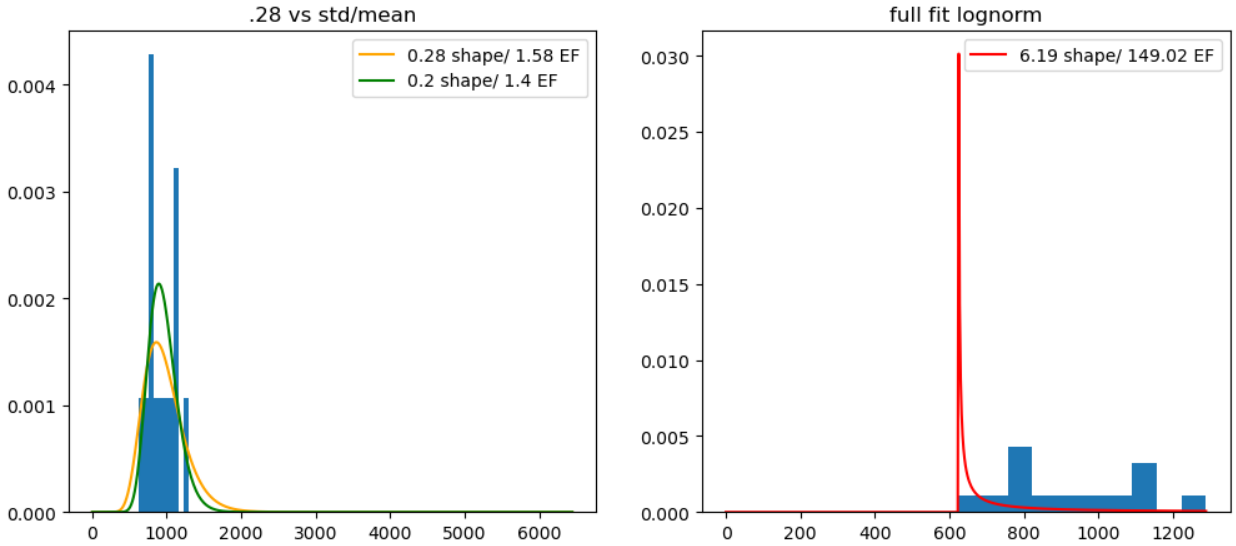
iso_from_alarm/HWR-981/Not Specified



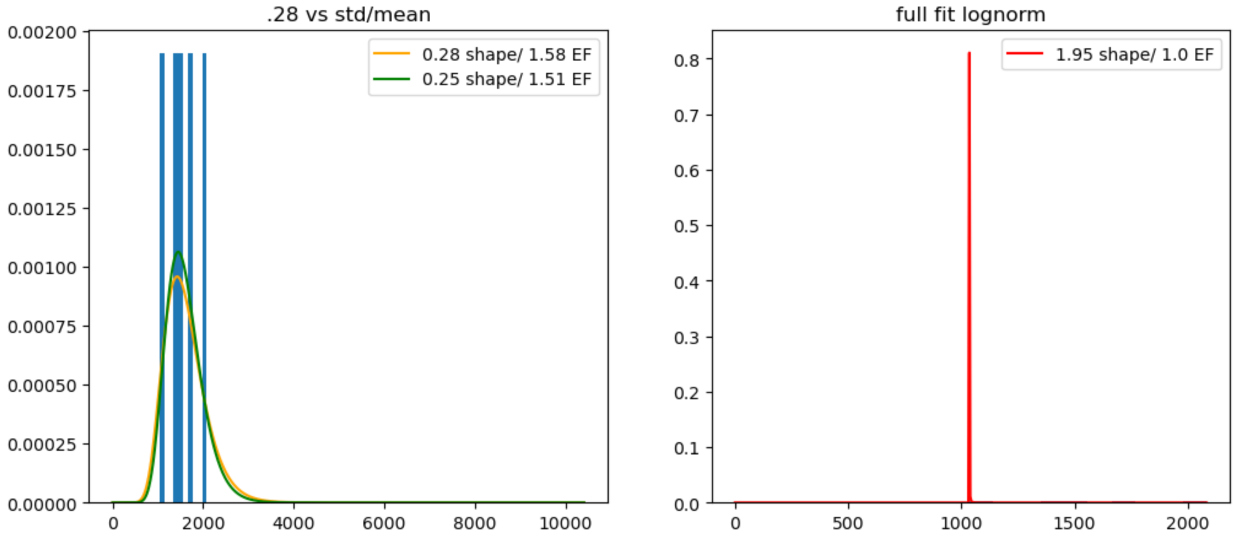

depressurization_time/Int-HRA/Cmplx
.28 vs std/mean
0.28 shape/ 1.58 EF
0.63 shape/ 2.82 EF
0.0040
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0
1000
2000
3000
4000
5000
full fit lognorm
8.35 shape/ 6484.49 EF
0.16
0.14
0.12
0.10
0.08
0.06
0.04
0.02
0.00
0
200
400
600
800
1000

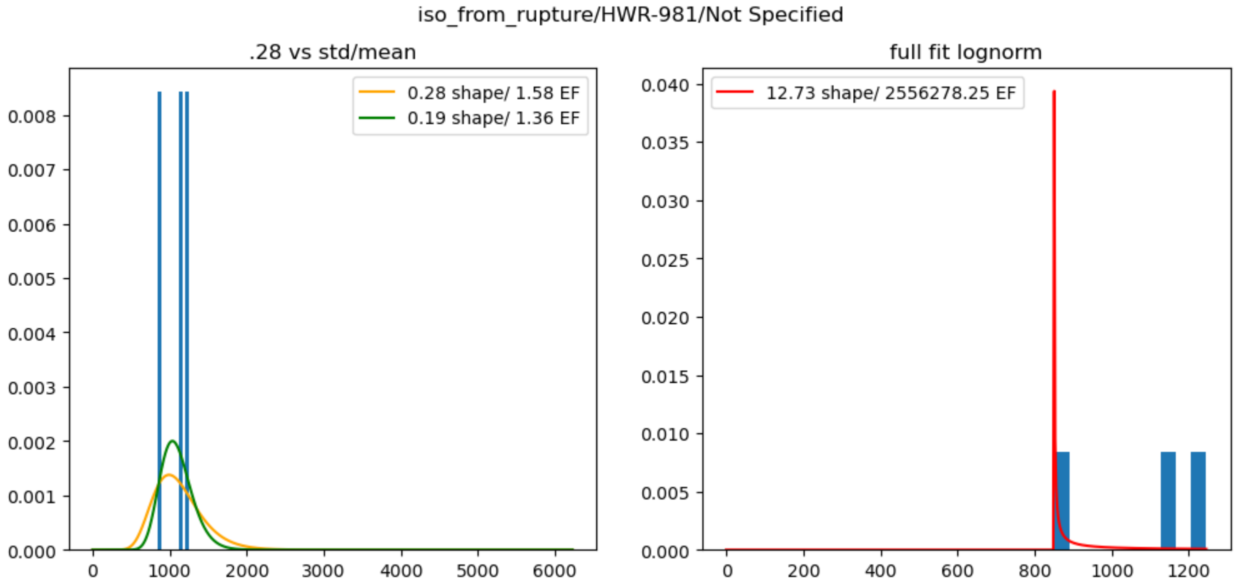

iso_from_alarm/Int-HRA/Cmplx
.28 vs std/mean
0.28 shape/ 1.58 EF
nan shape/ nan EF
10
8
6
4
2
0
0
500
1000
1500
2000
2500
3000
full fit lognorm
1e11
2.57 shape/ 1.0 EF
3.0
2.5
2.0
1.5
1.0
0.5
0.0
0
100
200
300
400
500
600

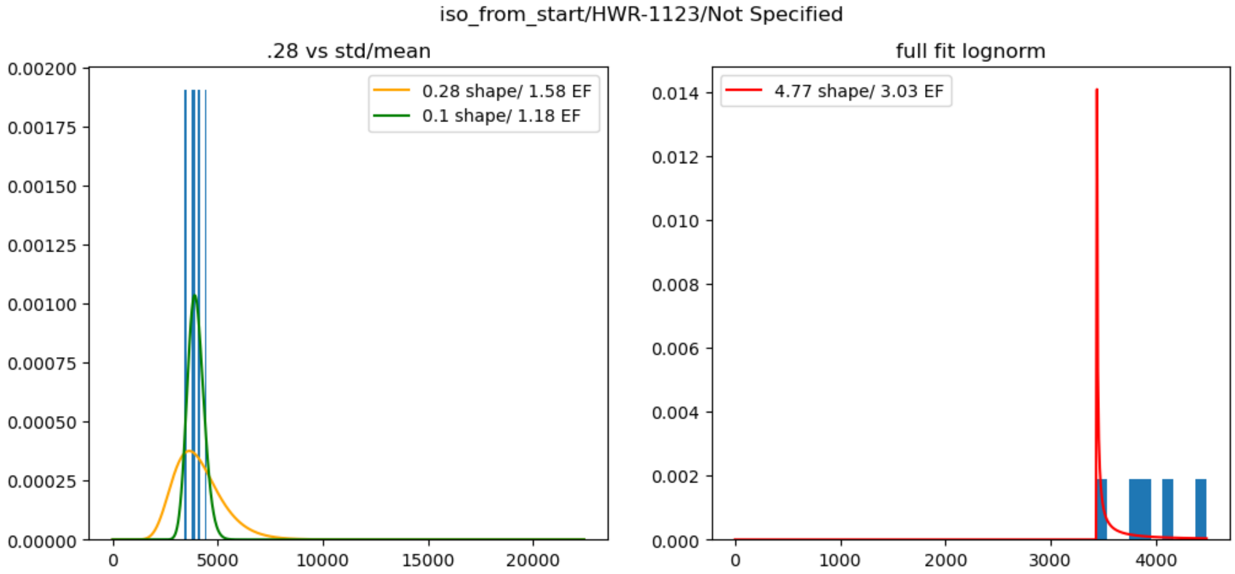
iso_from_alarm/Int-HRA/Base

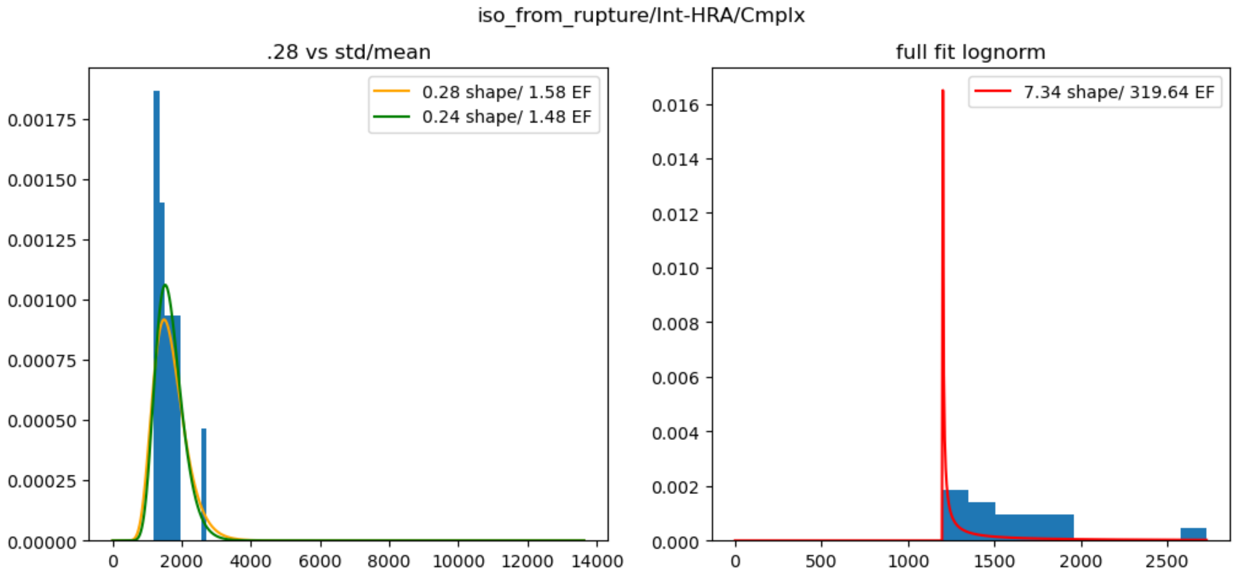


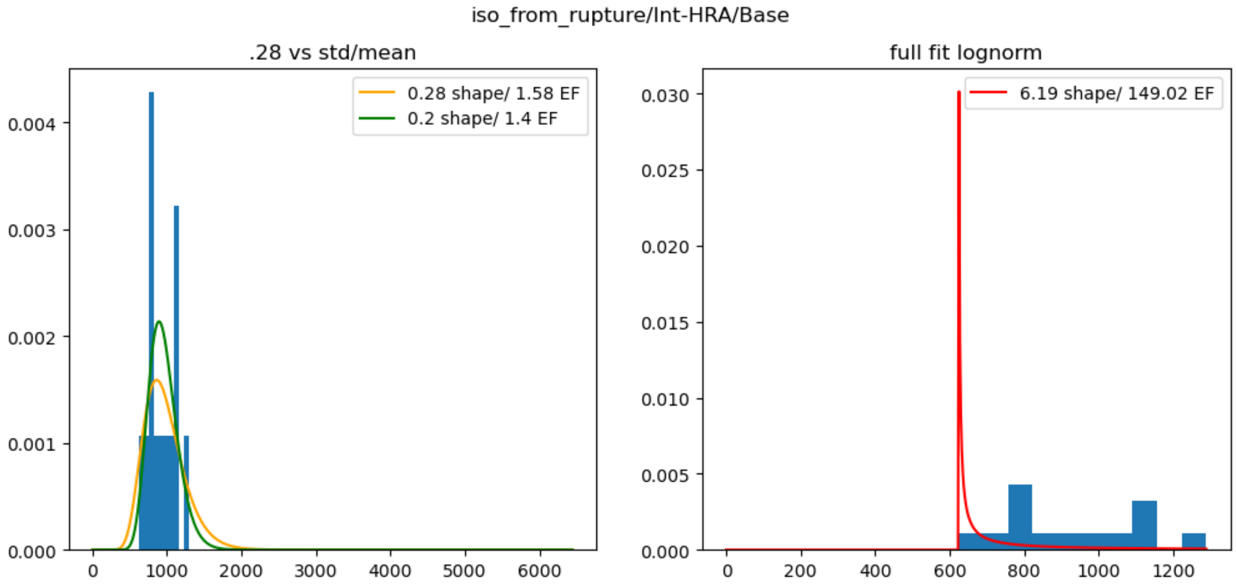
iso_from_rupture/HWR-1123/Not Specified

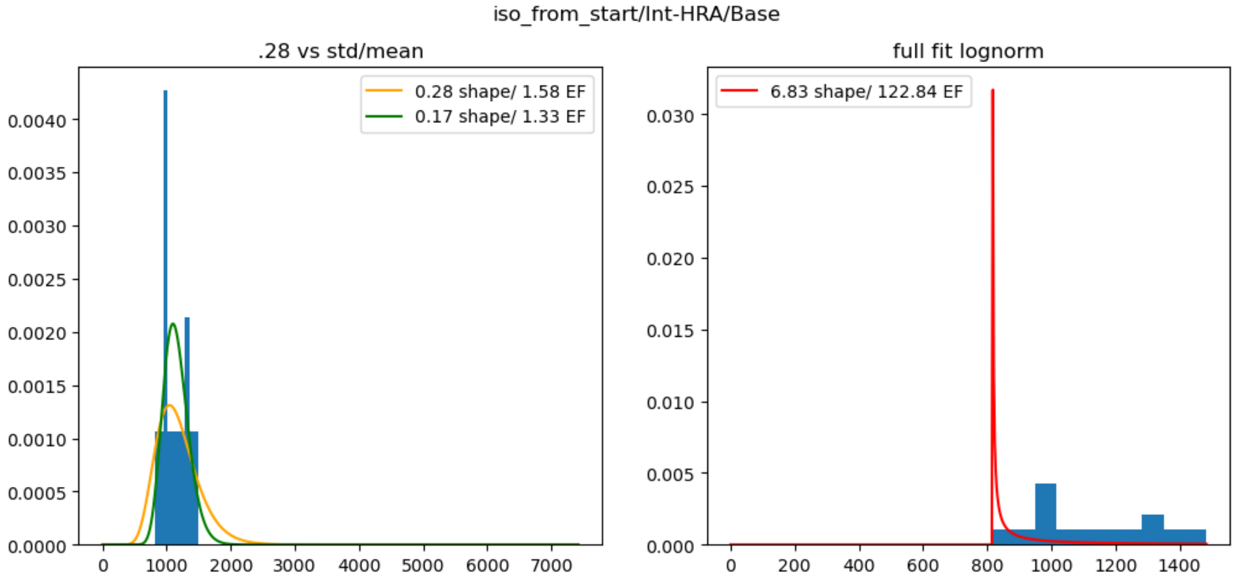


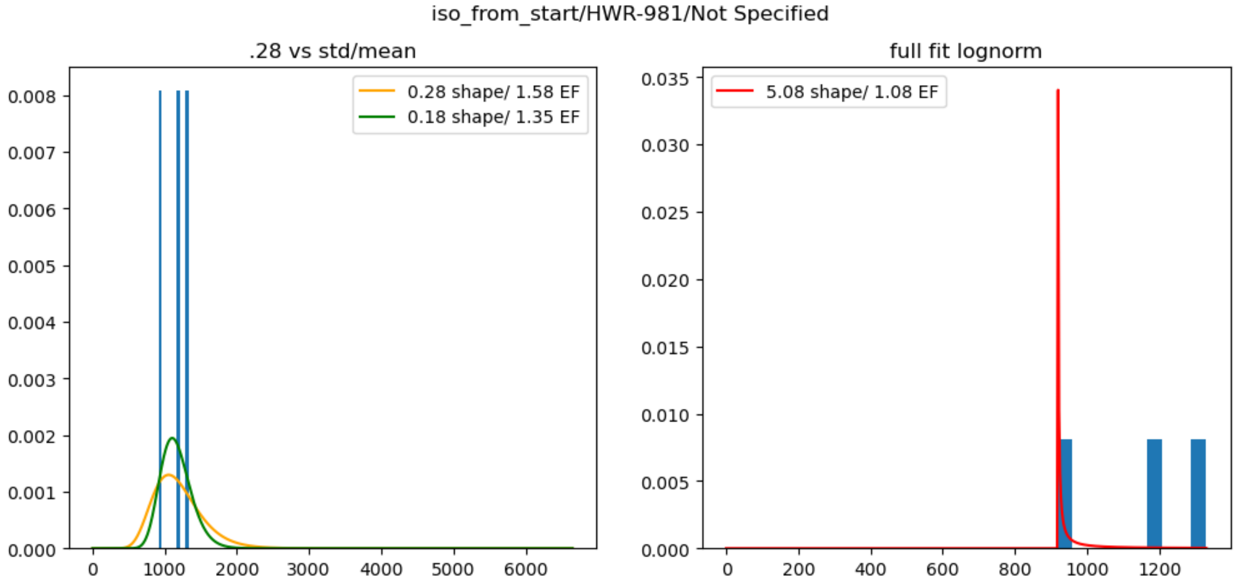

iso_from_rupture/HWR-981/Not Specified
.28 vs std/mean
0.28 shape/ 1.58 EF
0.19 shape/ 1.36 EF
0.000
0.001
0.002
0.003
0.004
0.005
0.006
0.007
0.008
0
1000
2000
3000
4000
5000
6000
full fit lognorm
12.73 shape/ 2556278.25 EF
0.000
0.005
0.010
0.015
0.020
0.025
0.030
0.035
0.040
0
200
400
600
800
1000
1200

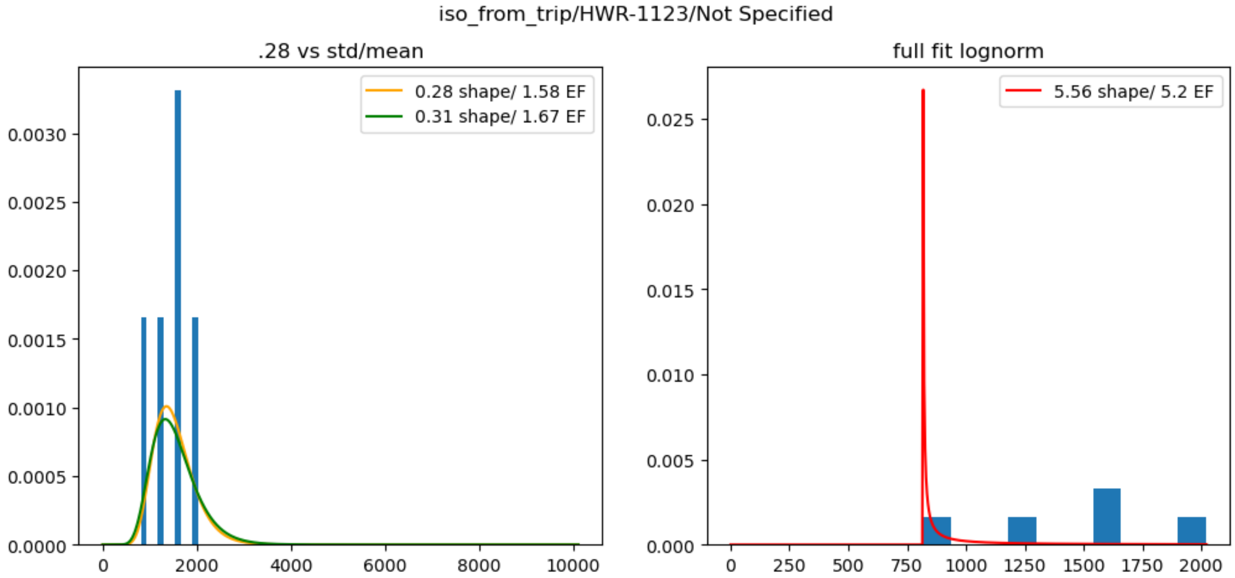

iso_from_start/HWR-1123/Not Specified
.28 vs std/mean
0.28 shape/ 1.58 EF
0.1 shape/ 1.18 EF
0.00000
0.00025
0.00050
0.00075
0.00100
0.00125
0.00150
0.00175
0.00200
0
5000
10000
15000
20000
full fit lognorm
4.77 shape/ 3.03 EF
0.000
0.002
0.004
0.006
0.008
0.010
0.012
0.014
0
1000
2000
3000
4000

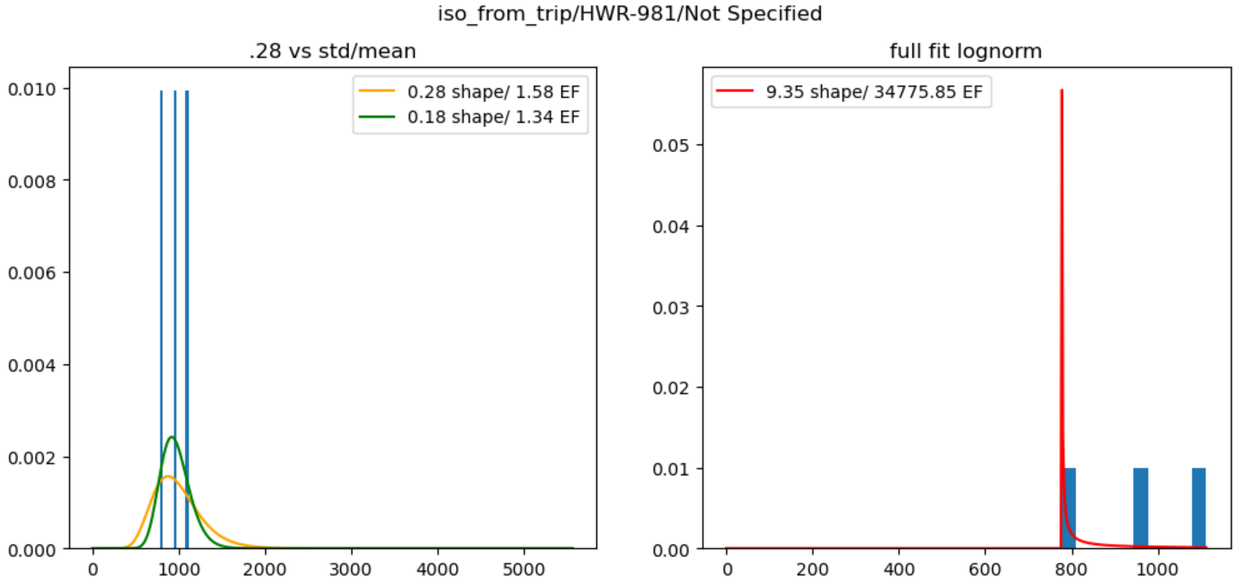

iso_from_rupture/Int-HRA/Cmplx
.28 vs std/mean
0.28 shape/ 1.58 EF
0.24 shape/ 1.48 EF
0.00175
0.00150
0.00125
0.00100
0.00075
0.00050
0.00025
0.00000
0
2000
4000
6000
8000
10000
12000
14000
full fit lognorm
7.34 shape/ 319.64 EF
0.016
0.014
0.012
0.010
0.008
0.006
0.004
0.002
0.000
0
500
1000
1500
2000
2500


iso_from_rupture/Int-HRA/Base
.28 vs std/mean
0.28 shape/ 1.58 EF
0.2 shape/ 1.4 EF
0.004
0.003
0.002
0.001
0.000
0
1000
2000
3000
4000
5000
6000
full fit lognorm
6.19 shape/ 149.02 EF
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
200
400
600
800
1000
1200

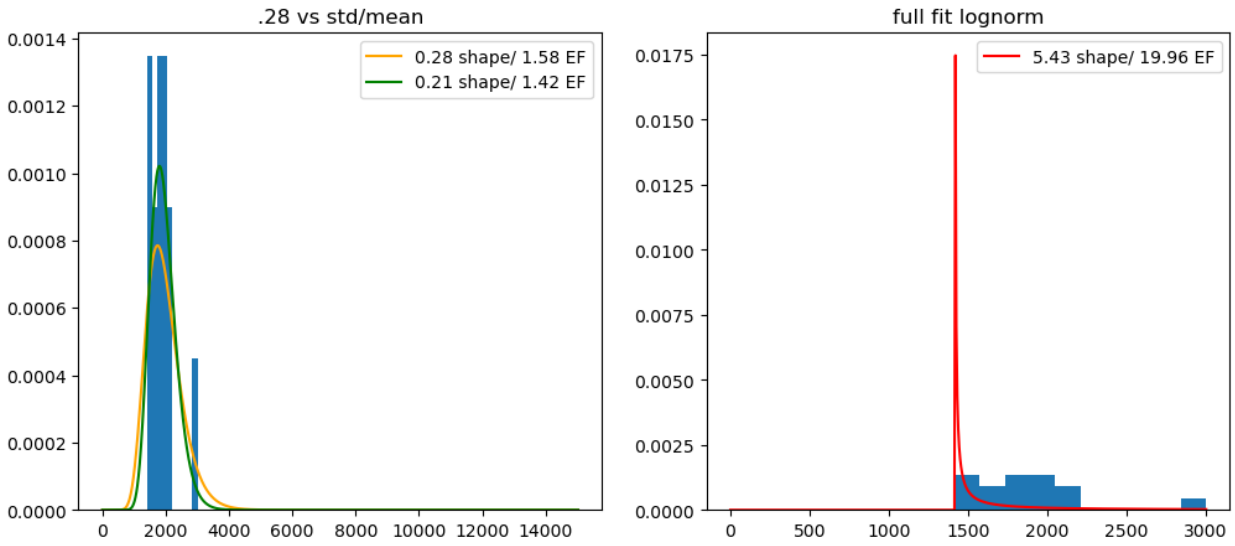

iso_from_start/Int-HRA/Base
.28 vs std/mean
0.28 shape/ 1.58 EF
0.17 shape/ 1.33 EF
0.0040
0.0035
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0
1000
2000
3000
4000
5000
6000
7000
full fit lognorm
6.83 shape/ 122.84 EF
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
200
400
600
800
1000
1200
1400


iso_from_start/HWR-981/Not Specified
.28 vs std/mean
0.28 shape/ 1.58 EF
0.18 shape/ 1.35 EF
0.008
0.007
0.006
0.005
0.004
0.003
0.002
0.001
0.000
0
1000
2000
3000
4000
5000
6000
full fit lognorm
5.08 shape/ 1.08 EF
0.035
0.030
0.025
0.020
0.015
0.010
0.005
0.000
0
200
400
600
800
1000
1200

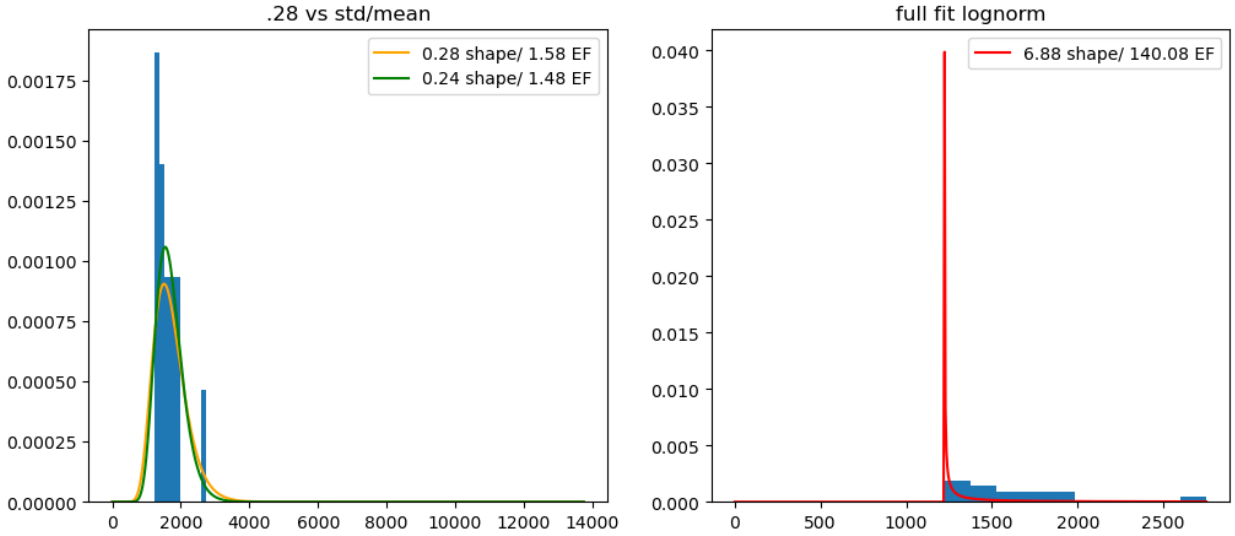

iso_from_trip/HWR-1123/Not Specified
.28 vs std/mean
0.28 shape/ 1.58 EF
0.31 shape/ 1.67 EF
0.0030
0.0025
0.0020
0.0015
0.0010
0.0005
0.0000
0
2000
4000
6000
8000
10000
full fit lognorm
5.56 shape/ 5.2 EF
0.025
0.020
0.015
0.010
0.005
0.000
0
250
500
750
1000
1250
1500
1750
2000

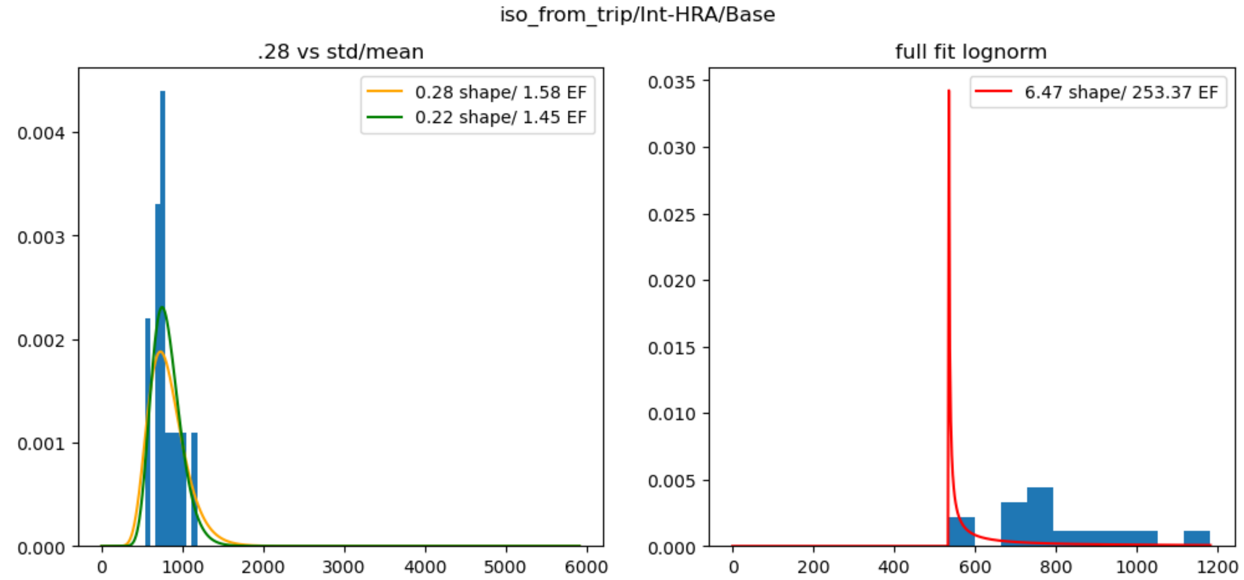

iso_from_trip/HWR-981/Not Specified
.28 vs std/mean
0.28 shape/ 1.58 EF
0.18 shape/ 1.34 EF
0.010
0.008
0.006
0.004
0.002
0.000
0
1000
2000
3000
4000
5000
full fit lognorm
9.35 shape/ 34775.85 EF
0.05
0.04
0.03
0.02
0.01
0.00
0
200
400
600
800
1000

iso_from_start/Int-HRA/Cmplx



iso_from_trip/Int-HRA/Cmplx





D.5 References

[1] A. Bye, E. Lois, V. N. Dang, G. Parry, J. Forester, S. Massaiu, R. Boring, P. Ø. Braarud, H. Broberg, J. Julius, I. Männistö, and P. Nelson, “NUREG/IA-0216, Vol. 2, international hra empir- ical study - phase 2 report. results from comparing hra method predictions to simulator data from sgtr scenarios.,” U.S. Nuclear Regulatory Commission, OECD Halden Reactor Project, Washing- ton, DC, ADAMS Accession No. ML11250A010, Aug. 2011.

[2] V. N. Dang, J. Forester, R. Boring, H. Broberg, S. Massaiu, J. Julius, I. Männistö, H. Liao, P. Nelson, E. Lois, and A. Bye, “NUREG-IA-0216, Vol. 3, international hra empirical study - phase 3 report. results from comparing hra methods predictions to hammlab simulator data on lofw scenarios,” U.S. Nuclear Regulatory Commission, OECD Halden Reactor Project, Washington, DC, NUREG/IA-0216 Vol. 3, ADAMS Accession No. ML14358A254, Dec. 2014.

[3] E. Lois, V. N. Dang, J. Forester, H. Broberg, S. Massaiu, M. Hildebrandt, P. Ø. Braarud, G. Parry, J. Julius, R. Boring, I. Männistö, and A. Bye, “NUREG/IA-0216, Vol. 1, international hra empirical study - phase 1 report. description of overall approach and pilot phase results from comparing hra methods to simulator performance data.,” U.S. Nuclear Regulatory Commission, OECD Halden Reactor Project, Washington, DC, ADAMS Accession No. ML093380283, Dec. 2009.

[4] J. Forester, H. Liao, V. N. Dang, A. Bye, E. Lois, M. Presley, J. Marble, R. Nowell, H. Broberg, M. Hildenbrandt, B. Hallbert, and T. Morgan, “NUREG-2156, the u.s. hra empirical study – assessment of hra method predictions against operating crew performance on a u.s. nuclear power plant simulator,” U.S. Nuclear Regulatory Commission, Washington, DC, ADAMS Accession No. ML16179A124, Jun. 2016.

[5] S. Massaiu and L. Holmgren, “A hammlab hra data collection with u.s. operators,” Institutt for energiteknikk, Halden, Norway, HWR-1123, May 2016.

[6] “Unpublished data from the dukovany npp simulator trials from 2011 as described in slides 18-29 of ml18079a165,” Nuclear Research Institute of the Czech Republic, Husinec, Provided by Dr. Radim Dolezal of the UJV Rez (Nuclear Research Institute of the Czech Republic), 2011.

[7] J. K. Park, W. D. Jung, J. W. Kim, and J. J. Ha, “Analysis of human performance observed under simulated emergencies of nuclear power plants,” Jan. 2005. [Online]. Available: https ://www.osti.gov/etdeweb/servlets/purl/20658965.