STETKAR & ASSOCIATES

203 WHISPERING HILLS STREET
HOT SPRINGS, AR USA 71901
TELEPHONE: (501) 627-0500 • E-MAIL: jwstetkar@aol.com

October 11, 2020

Dr. Michael Cheok
Director, Division of Risk Analysis
Office of Nuclear Regulatory Research
Mail Stop T10-A12
U.S. Nuclear Regulatory Commission
11545 Rockville Pike
Rockville, MD 20852

Dear Dr. Cheok:

Comments on RIL-2020-02

Attached are my comments on the version of report RIL-2020-02, "Integrated Human Event Analysis System for Event and Condition Assessment (IDHEAS-ECA)", that was discussed during the September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment.

More than half of my comments pertain to the examples in Appendix C of this report. The examples are intended to show analysts how the IDHEAS-ECA methods and guidance are used in practice. Therefore, I think that it is very important for each example to clearly demonstrate the staff's expectations for how a systematic, comprehensive, and well-documented human reliability analysis should be performed.

I do not request, or expect, formal responses to any of these comments and questions. They are provided only for your consideration as you finalize this important project.

Please contact me if you have any questions, or if you need any additional information.

Very truly yours,

John W. Stetkar

John W. Stetkar

Attachment

Comments and Questions on
RIL-2020-02
Integrated Human Event Analysis System for
Event and Condition Assessment (IDHEAS-ECA)
February 2020
John W. Stetkar

JWS Note: These comments are based of the version of this report that was discussed during the September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment (ADAMS accession number ML20016A481).

1. General Comment

Reference [1] for this report is:

> J. Xing, Y. J. Chang, and J. DeJesus, "The General Methodology of an Integrated Human Event Analysis System (IDHEAS-G) - Draft Report," U.S. Nuclear Regulatory Commission, NUREG-2198 (ADAMS Accession No. ML19235A161), Aug. 2019.

The current version of NUREG-2198 (ADAMS accession number ML20238B988) was issued to support the September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment. Several of my comments identify differences between the guidance in this report and the guidance in the current version of NUREG-2198. That is an unfortunate result from publishing this version of the IDHEAS-ECA application before the final version of the general IDHEAS methodology is completed. I have tried to highlight items where the differences may have an important influence on how analysts interpret or apply the IDHEAS-ECA guidance. It is very important that the final published version of this guidance should be consistent with the final version of NUREG-2198.

2. General Question

This report indicates that the NRC staff developed a software tool to implement and document the IDHEAS-ECA analysis process.

Is that software available to the public?

If so, where can it be obtained?

3. Section 2, IDHEAS-ECA Basics, General Comment

This is a very good overview of the basic concepts, including appropriate references to the general guidance in NUREG-2198.

4. Section 3, Guidance for the IDHEAS-ECA Process, Figure 3-1, Step 8, Uncertainty Analysis; Section 3.9, Summary of IDHEAS-ECA, Relation between the IDHEAS-ECA Steps

Step 8 of the methodology includes identification, documentation, characterization, and quantification of uncertainties. That is an essential element of the integrated human reliability analysis (HRA) process. The text that summarizes Figure 3-1 acknowledges that this step is intentionally not shown in the figure. The summary of the IDHEAS-ECA steps in Section 3.9 does not mention Step 8.

JWS Note: Although Table 3-8 lists Step 8, its entries are curiously blank.

These omissions inappropriately imply that the evaluation of uncertainty is an ancillary task that is subsidiary to the other steps in the analysis, or it is only an after-thought. To emphasize the fact that uncertainty analysis should be an integral part of the methodology, I think that it is very important for the figure to explicitly show that step of the process. I also think that it is important for the summary to emphasize that identification, documentation, characterization, and quantification of uncertainties is an integral element of each step of the analysis process.

Why do Figure 3-1 and the summary in Section 3.9 not include Step 8 of the analysis process?

5. Section 3, Guidance for the IDHEAS-ECA Process, Figure 3-1, General Comment

This figure clearly shows the major steps in the analysis process and their interrelationships. It is very useful for an understanding of how the qualitative and quantitative elements of the analysis are related, and how the information that is derived from the qualitative analysis affects quantification of the two contributions to the overall human error probability (i.e., P_c and P_t).

This depiction of the overall analysis process is not unique to the IDHEAS-ECA application. In particular, I think that it would be very useful to include this figure and the summary of the steps from Section 3.9 of this report (with Step 8) in the current version or the next update to the IDHEAS-G methodology in NUREG-2198.

6. Section 3.1.1, Develop the Operational Narrative

The last paragraph in the introduction to this section notes that:

> "For the purposes of IDHEAS-ECA, the operational narrative (scenario narrative and scenario timeline) should be developed ***based on the PRA*** (i.e., the event tree where the HFE is being credited) and documented in Section A.1 of Worksheet A found in Appendix A." [emphasis added]

Section 1.2 indicates that the scope of IDHEAS-ECA includes applications that may require the development of new PRA models or modifications to existing PRA models (e.g., to evaluate a particular license amendment request, an unexpected operational event, an inspection finding, an evolving regulatory issue, etc.).

It is certainly true that the operational narrative should pertain to the event scenario that is being evaluated in the PRA. However, this discussion and several other similar statements in the report imply that the PRA model is a pre-determined and fixed input to the human reliability analysis (HRA) process. In other words, the report seems to imply that the HRA process does not involve the identification of alternative (or "deviation") scenarios that require modifications to the PRA event trees and fault trees, and definitions of new human failure events (HFEs). Those new HFEs may address the same functions that apply to the nominal (or "baseline") HFEs, but in a different scenario context. In some cases, the HRA process may also identify new HFEs that were not considered when the initial PRA models were developed. This search for alternative scenarios and the identification of new HFEs is an integral part of the HRA process and the development of a PRA model that appropriately accounts for human performance.

The introduction to Section A.1, "Operational Narrative" in Worksheet A in Appendix A notes that:

> "It may also include ***important scenario deviations*** and past operational experience review." [emphasis added]

The report does not contain any other mention of the need to identify scenarios that deviate from the conditions in the nominal PRA models or to identify and evaluate HFEs that apply in the context of those alternative scenarios. This seems to reinforce the archaic notion that "PRA analysts" develop the PRA models and "HRA analysts" simply evaluate the human error probabilities (HEPs) for the HFEs that are pre-defined by those models.

JWS Note: The discussion of "Guidance for the Identification of HFEs" in Section 3.1.2 notes that it may be necessary to define a new HFE to analyze a "real event". However, that section begins with the observation that "the PRA should already identify the HFEs that need to be analyzed". Section 3.1.2 briefly summarizes how an analyst may search for and identify HFEs when the PRA model does not evaluate the observed scenario. However, if an existing PRA contains a nominal ("baseline") HFE for the desired personnel response, the guidance does not alert analysts to conduct a critical assessment of the those models and determine whether deviations from the specific modeled scenarios may contain conditions that have an important effect on human performance and, hence, require the definition and evaluation of additional HFEs (with corresponding changes to the event trees or fault trees).

Does this report essentially assume that the PRA model is a pre-determined and fixed input to the HRA process?

In particular, is it intended that the scope of the IDHEAS-ECA analysis process includes the identification of deviation scenarios that require changes to the nominal PRA models, with corresponding new HFEs?

If so, why does the guidance in Section 3.1 not explicitly address that element of the overall HRA process?

If the IDHEAS-ECA analysis process is not intended to identify and evaluate scenarios that deviate from the nominal PRA models, how does this guidance provide confidence that the PRA model structure (i.e., the analyzed scenarios) adequately accounts for an integrated evaluation of human performance?

7. Section 3.1.1, Develop the Operational Narrative, Guidance for Developing the Scenario Narrative; Section 3.1.3, Identify the Scenario / Event Context, Guidance for Assessing the System Context

The guidance in Section 3.1.1 emphasizes a focus on only those "safety considerations" of the scenario which are directly relevant to the particular scope and success criteria of the human failure event (HFE) that is being analyzed. That information is certainly very important. However, it does not necessarily capture all of the potentially important influences on human performance, and the guidance does not explicitly prompt analysts to search for those influences and document them in the narrative. Experience from actual events has shown that these other influences can be important. Restricting the focus of the narrative to describe only the scenario "safety issues" that are depicted in the PRA models may cause the analyst to overlook these influences and perhaps develop an overly optimistic perspective about the desired personnel response.

In principle, the scenario narrative should describe everything that is happening in the plant, because that is the actual context within which personnel must respond. Of course, that ideal can rarely be achieved in a practical analysis. However, the narrative should describe all conditions that may have a potentially important effect on human performance, even if those conditions are not included explicitly in the PRA models. That description helps the analysts to identify and evaluate the states of relevant performance-influencing factors (PIFs) that account for distractions, interruptions, multi-tasking, conflicting priorities, time pressure, stress, etc. It also helps others to understand what conditions were considered by the analysts and to question the reasons for possible omissions.

For example, experience from actual events has shown that personnel may be distracted by failures or damage to non-safety systems that are important for overall plant investment protection or are perceived to affect the stability of overall plant conditions, but are not modeled explicitly in the PRA. In some scenarios that involve severe plant damage (e.g., fires, floods, seismic events, etc.), operators may also need to attend to treatment and relocation of personnel who are physically injured. These concerns introduce conflicting strategic and time priorities for decision-makers and constraints on the assignment of limited personnel resources. These types of diversions and distractions have occurred in practice, and analysts should account for them. That is why it is essential that the integrated scenario narrative must describe the entire context of the plant damage, and not focus only on systems and equipment that are modeled explicitly in the PRA, and the distinct human actions that are needed to cope with only those failures.

JWS Note: The discussion of the "Initial Condition" under "Beginning of the Scenario" appropriately notes that the narrative should identify "other ongoing activities performed at the same time of the initiating event that can have effects on the scenario". Because those activities are associated with the plant status when the initiating event occurs (e.g., testing, maintenance, specific operational evolutions, etc.), they do no not include the types of post-initiator effects that are addressed by this comment. The discussion of "Scenario Progression and End State" focuses only on the "safety consideration" for the modeled HFE. It does not address the need for the narrative to describe the overall plant conditions that are highlighted in this comment. Furthermore, these types of considerations are not mentioned in the "Guidance for Developing the Scenario Timeline" or in the "Guidance for Assessing System Context" in Section 3.1.3. For reference, the discussion of "System Context" in Section 4.2.3 of NUREG-2198 explicitly mentions these considerations.

Why does Section 3.1.1 not emphasize the need for analysts to examine the totality of what is happening in the plant and document any conditions that may have a potentially important influence on personnel response, even if those conditions are not related directly to the equipment, human actions, and "safety issues" that are modeled explicitly in the PRA and the defined HFE success criteria?

Why do the guidance and the probing questions for the system context in Section 3.1.3 not explicitly mention these considerations?

8. Section 3.1.1, Develop the Operational Narrative, Guidance for Developing the Scenario Narrative, Beginning of the Scenario, Initial Condition

The first bullet item in this section notes that the summary should identify:

> "Structures, systems, and components (SSCs) with latent failures, that are unavailable (tagged out), or have ***historically unreliable performance*** (especially the ones that would affect operator's decisions and the scenario)." [emphasis added]

I think that I understand the intent of the highlighted phrase. However, this type of "speculative" information is not typically included in most operational narratives, and the intent of this guidance might confuse some analysts.

In practice, PRA models should account explicitly for successes and failures of all SSCs that are needed to mitigate a particular evolving scenario. The human failure event (HFE) definitions should account for those scenario-specific conditions. For example, HFE1 may apply during a scenario when Component X is operating, and HFE2 may apply during a scenario when Component X is failed.

In the second scenario, the analyzed contributions to failure of Component X should include latent failures and unavailability due to maintenance. In practice, the evaluation of HFE2 should account for possible distractions or delays that are introduced by personnel actions to troubleshoot the particular reason why Component X is not available. Therefore, I do not think that it is necessary for the narrative of that scenario to explicitly address the component's historical reliability, unless that perspective might have a significant effect on the troubleshooting time or effort.

This guidance seems to imply that the narrative for the first scenario should note that Component X has been historically unreliable. Therefore, personnel may be worried about its imminent failure, despite the fact that the scenario applies for conditions when it operates successfully for the duration of the PRA mission time (or at least up until the time when HFE1 is questioned). This is certainly a consideration that may affect personnel performance during an actual event. However, in practice, I am not sure how analysts may interpret this guidance, especially in the context of the defined PRA scenario. Therefore, if I have correctly interpreted the intent of this guidance, I think that it should be expanded somewhat to better alert analysts to the types of considerations that may apply for the narrative of a scenario when the equipment of concern is operating successfully.

JWS Note: Section E.2.1.1 in Appendix E of NUREG-2198 also contains the highlighted phrase. I missed this nuance when I read that guidance.

Is my interpretation of the intent of this guidance correct?

If not, what is the intent of this guidance?

Should this discussion better explain the intent, especially for the narrative of a scenario that applies when the equipment of concern is operating successfully?

9. Section 3.1.2, Identify and Define the Human Failure Events

The first paragraph in this section notes that:

"Typically, ***pre-initiator and initiator HFEs are not explicitly modeled*** in PRA because the human error contribution is included as part of the component reliability estimates and initiating event frequencies, respectively." [emphasis added]

I disagree very strongly with this statement.

It is certainly true that some types of maintenance and calibration errors may not be discovered by post-maintenance testing. Those errors are eventually revealed by equipment failures. Their prevalence depends on the plant-specific maintenance practices and the effectiveness of the plant-specific post-maintenance inspection and testing protocols. Most PRAs do not separately quantify these causes for equipment failure, because it is too resource-intensive to extract them from the composite equipment performance records.

All PRAs explicitly identify, model, and quantify many pre-initiator human actions. Examples are errors that do not restore equipment to their normal alignments after maintenance and testing activities, miscalibration of instrumentation and signal processing logic, etc. A typical PRA model may contain dozens of these pre-initiator human failure events (HFEs). In some cases, depending on the plant-specific design, maintenance protocols, and testing frequencies, these types of errors can be important contributors to risk. It is not appropriate for "NRC-approved" guidance to imply that these pre-initiator actions are "not explicitly modeled" or that analysts have an option to overlook them.

Many PRAs also explicitly model and quantify human errors that contribute to initiating events. For example, a normally-running cooling water pump may fail, and the operators may fail to start the standby pump manually. The resulting initiating event frequency is determined by the combination of the hardware failure and the subsequent operator error. Therefore, it is also not appropriate for "NRC-approved" guidance to imply that these types of initiator actions are universally "not explicitly modeled" or that analysts have an option to overlook them.

JWS Note: Section F.1 in Appendix F of NUREG-2198 discusses the scope of pre-initiator human actions that are typically modeled in a PRA. Section F.2 briefly addresses the identification of human actions that contribute to the occurrence of an initiating event. Section K.1.1 in Appendix K of NUREG-2198 addresses possible dependencies among pre-initiator HFEs, initiator HFEs, and post-initiator HFEs. Section K.1.3 addresses dependencies between initiator HFEs and post-initiator HFEs.

Why does this section indicate that "pre-initiator and initiator HFEs are not explicitly modeled"?

10. Section 3.1.2, Identify and Define the Human Failure Events, Guidance for the Identification of HFEs

The discussion of errors of commission (EOCs) refers to NUREG-1624 and NUREG-1921. Section F.4 in Appendix F of NUREG-2198 contains the information that is discussed in this section, plus some additional background. Since IDHEAS-ECA is an application of the general IDHEAS methodology, it seems that this section should refer to the guidance in NUREG-2198.

Why does the discussion of EOCs not refer to Section F.4 in Appendix F of NUREG-2198 (i.e., rather than the cited NUREGs)?

11. Section 3.2.1, Defining HFEs, Guidance for the Definition of HFEs

The last bullet item in this section notes that the human failure event (HFE) definition should include:

> "available time to perform the HFE (whether the HFE is ***time critical***)" [emphasis added]

The concept of "time critical" actions has a distinct implication in the context of personnel training and procedures. I think that this term may also have a regulatory context, in the sense that Emergency Operating Procedures (EOPs) are supposed to contain guidance for all "time-critical" actions that are credited in a plant's licensing basis, and NRC inspectors audit the procedures and training for those specific actions. Therefore, analysts may interpret this guidance to imply that it is necessary to document the available time window only for that specific subset of HFEs. I do not think that it is appropriate, or necessary, to use this term in the context of this guidance.

JWS Note: The introduction to Section 3.6 also uses the term "time-critical HFEs".

Why does this guidance use the term "time critical" to qualify the need to document the available time window for each HFE?

12. Section 2.2, Overview of the Cognition Model for IDHEAS-ECA, Cognitive Failure Modes; Section 3.3.2, Identification of Applicable Cognitive Failure Modes

Section 2.2 lists only five cognitive failure modes (CFMs), one for each of the five macrocognitive functions. Section 3.3.2 also simply acknowledges that these are the high-level CFMs from NUREG-2198.

This construct is a very significant simplification of the extensive set of CFMs that are discussed in NUREG-2198. For reference, Section 4.3.3.2 and Appendix H of NUREG-2198 describe the following CFM structure.

High-Level CFM	Middle-Level CFMs	Detailed CFMs
Detection	5	16
Understanding	5	10
Decisionmaking	6	16
Action Execution	5	19
Interteam Coordination	7	7

I certainly understand why the IDHEAS-ECA authors want to simplify the extensive set of CFMs that are described and tabulated in NUREG-2198. However, it seems evident that this simplification may have an important effect on the level of detail at which an analyst decomposes each critical task and identifies one or more relevant CFMs for that task. It also affects how analysts may identify and evaluate the applicable performance-influencing factors (PIFs) for each CFM, and thus quantify the influence of those PIFs on the overall human error probability (HEP). It is not apparent how this simplification supports a comprehensive evaluation of human cognitive performance and identification of the most important contributions to a particular human failure event (HFE). Furthermore, as described in Section 4.5 and Appendix K of NUREG-2198, the cognitive performance evaluations are important for the identification and quantification of human dependencies. Therefore, the dependency analyses may also suffer from an excessively simplified assessment.

For example, it is not apparent why the five high-level CFMs provide an appropriate assessment of the contributions to cognitive errors, compared to use of the middle-level CFMs that are listed in Table 4-6 through Table 4-10 in NUREG-2198. Figure 3-3 through Figure 3-7 show the breakdown of each high-level CFM into a set of cognitive processors, which correspond to the middle-level CFMs in NUREG-2198. However, the guidance in Section 3.3.2 does not describe how analysts should consider these processors or account for their effects when they select the relevant high-level CFMs and evaluate the applicable PIFs. In particular, the guidance simply notes that:

> "It is recommended that HRA analysts use the processors to verify the selection of the applicable CFMs and distinguish between the CFMs of a critical task."

I think that the overview of this element of the methodology in Section 2.2 and the discussion in Section 3.3.2 should explicitly acknowledge this simplification, explain why it is appropriate for the IDHEAS-ECA application, and refer the reader to the more detailed discussion of CFMs in NUREG-2198. I also think that the guidance in Section 3.3.2 should be expanded to explicitly alert analysts to examine how each of the detailed CFMs (or, at least, the middle-level CFMs) and their associated PIFs may apply when they evaluate the high-level CFM for a particular critical task. That will at least provide some confidence that the analysts may not inappropriately overlook a potentially important contributor by focusing only on the "obvious" CFM and PIF effects at a very high level.

JWS Note: I think that it would be very useful to provide an example in Section 3.3.2 that illustrates the basic concepts of this process. In particular, if any of the processors for a CFM applies for a particular critical task, then that CFM should be evaluated for the task. Furthermore, to provide a better focus for the analyses, I think that the guidance should indicate that the relevant PIFs should be evaluated in the more detailed context of the applicable processors, rather than the more general context of the high-level CFM. For example, high-level CFMx may fail because PIFa adversely affects processor CFMxn. The final HEP quantification can still be based on the high-level CFM, but the supporting analysis better identifies and documents the contributors to that failure.

Why does the IDHEAS-ECA methodology explicitly recommend use of only these five high-level CFMs (i.e., rather than the middle-level CFMs or detailed CFMs from NUREG-2198)?

Why does Section 2.2 or Section 3.3.2 not acknowledge this simplification of the construct that is described in NUREG-2198 and explain why it is appropriate for this particular application of the general methodology?

Why does Section 3.3.2 not explain how analysts should consider the applicability of each of the processors (or, perhaps, the detailed CFMs from NUREG-2198) and their associated PIFs when they select and evaluate the high-level CFMs for a particular critical task?

13. Section 3.3.2, Identification of Applicable Cognitive Failure Modes, Figure 3-5 Cognitive Activities and Processors for Decisionmaking

The summary description of processor DM1 is:

> "Adapt the infrastructure of decisionmaking"

I do not understand what this means. In particular, I do not understand how analysts will interpret the phrase "adapt the infrastructure" to address the use of an appropriate decision-making process or mental model, and selection of the appropriate decision criteria, as noted in Table 4-8 in NUREG-2198. I think that the summary description of DM1 should be revised to better characterize the basic functional intent and context of that processor in a way that is more easily understood by prospective analysts.

How does the phrase "adapt the infrastructure of decisionmaking" adequately characterize the functional intent of processor DM1 (e.g., in the context of Table 4-8 in NUREG-2198)?

14. Section 3.3.2, Identification of Applicable Cognitive Failure Modes, Figure 3-6 Cognitive Activities and Processors for Action Execution

The summary description of processor E3 is:

> "Prepare or adapt infrastructure for action implementation"

I do not understand what this means. In particular, I do not understand how analysts will interpret the phrase "prepare or adapt the infrastructure" to address the coordination of personnel and activities that is noted in Table 4-9 in NUREG-2198. I think that the summary description of E3 should be revised to better characterize the basic functional intent and context of that processor in a way that is more easily understood by prospective analysts.

How does the phrase "prepare or adapt infrastructure for action implementation" adequately characterize the functional intent of processor E3 (e.g., in the context of Table 4-9 in NUREG-2198)?

15. Section 3.3.2, Identification of Applicable Cognitive Failure Modes

The first bullet item at the end of this section notes that:

> "Whether a CFM should be selected for a critical task depends on the nature of the task, not the PIFs."

I certainly agree with this guidance. Selection of the applicable cognitive failure modes (CFMs) depends on characteristics of the respective critical task. The scenario-specific analysis then determines how each CFM is influenced by the composite effects from the relevant performance-influencing factors (PIFs).

However, the bullet item for CFM3 notes that:

> "Similarly, CFM3 is under the assumption that personnel already detected the information and made the right *understanding* of the situation. If a ***procedure*** directs operators'

> response without uncertainty (e.g., if... then ...), then ***CFM3 is negligible***." [emphasis added]

I disagree very strongly with this guidance. It inappropriately biases analysts' critical examination and consideration of all elements of the Decsion-Making CFM. It is also contrary to the overall guidance that is highlighted from the first bullet. In particular, it implies that analyst decisions for the selection of CFM3 can be based on a premature evaluation of only the PIF for "Procedures, Guidelines, and Instructions". Furthermore, this guidance also perpetuates an archaic "procedure-centered" perspective on human performance that is not consistent with the integrated cognitive framework that is presented in NUREG-2198.

Why does the guidance imply that analysts can omit consideration of CFM3 for a particular critical task, based on a premature assessment of only the PIF for procedures?

How is this item internally consistent with the overall guidance for identification and selection of the applicable CFMs, which should consider all of the Decision-Making processors, and without regard to the scenario-specific PIFs?

How is this "procedure-centered" perspective consistent with the integrated cognitive framework that is presented in NUREG-2198?

16. Section 3.3.2, Identification of Applicable Cognitive Failure Modes

The last bullet item for CFM5 at the end of this section illustrates the notion of interteam coordination by an example of activities that are performed by personnel in the technical support center (TSC) and the emergency response center. That is certainly a relevant example. However, most analysts are not familiar with the consideration or evaluation of this cognitive failure mode (CFM). Therefore, I think that it would be useful to reinforce the context with another example that includes coordination between supervisory commands and actions that are performed in the main control room (MCR) and actions that are performed locally in the plant. Without that example, analysts may inappropriately limit their evaluation of CFM5 to address only coordination among various high-level elements of the response organization.

Should the example for CFM5 also note that it applies for coordination of actions that are performed in the MCR and actions that are performed locally in the plant, to reinforce the need for analysts to evaluate interteam coordination for those types of integrated tasks?

17. Section 3.4, Step 4 - Assessing PIF Attributes Applicable to CFMs, Guidance for Assessing PIF Attributes

Item (1) and Item (2) in this section indicate that analysts should document their rationale for conclusions that specific performance-influencing factors (PIFs) and specific PIF attributes are not relevant to the analysis of a particular critical task. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

The discussion in Item (2) notes that:

> "If eliminating a PIF attribute is ***not obvious***, a rationale should be provided for the elimination." [emphasis added]

I am somewhat concerned about the caveat that a documented rationale is needed only if the reason for eliminating a PIF attribute is "not obvious". For example, reasons that are "obvious" to Analyst A may not be equally "obvious" to Analyst B.

Why is the "not obvious" caveat needed in Item (2)?

18. Section 3.4, Step 4 - Assessing PIF Attributes Applicable to CFMs, Guidance for Assessing PIF Attributes

Item (4) in this section notes that:

> "The effect of an attribute on HEP can vary continuously with the quantitative measure of the attribute. ***Multiple discrete scales*** are used to model those attributes, which are referred as multi-scale attributes because they have multiple scales instead of being just present versus absent. The PIF tables in Appendix B present ***several measures*** for such attributes. The IDHEAS-ECA software uses those measures as benchmark scales and allows HRA analysts to select a scale value between ***one to ten***, with one being the lower limit and ten being the upper limit of the attribute being modeled. For each scale selected, the software assigns the corresponding base HEP or PIF weight (i.e., a multiplier) based on a ***linear interpolation between the benchmarks***." [emphasis added]

After I studied the tables in Appendix B, I am very confused by this discussion. In particular, I do not understand how this scaling and interpolation process applies to the entries in those tables, or how it is performed for an actual analysis.

Almost all of the entries in the tables seem to simply address a discrete effect. For example, Tables B-1 through B-3 provide estimates for the human error probability (HEP) that is associated with a particular performance-influencing factor (PIF) attribute. Tables B-4 through B-15 provide the PIF weights that are associated with a particular PIF attribute. Each attribute evaluates a distinct element of the respective PIF. There is not a continuous relationship among the attributes or an implied scale that varies from the "best" to the "worst" attributes.

The guidance for a small number of attributes identifies multiple degrees of severity (e.g., ENV1 in Table B-4, TE1 in Table B-10, and MT1 through MT-5 in Table B-13). However, it is not apparent whether the discussion of the 1-to-10 scale and linear interpolation along that scale applies to only those few examples, or whether it is intended to somehow apply over the entire range of the attributes that are listed for a particular PIF. For example, I do not know whether the 1-to-10 scales are intended to apply over the range of SF1 through SF4 for "Understanding" in Table B-1, the range of C1 through C9 for "Detection" in Table B-3, the range of ENV1 through ENV13 for "Execution" in Table B-5, etc.

This process may be very important, because it affects how an analyst's evaluation of the scenario-specific influence on the relevant PIF attributes is transformed into a numerical effect on the HEP. However, except for this brief (and confusing) discussion, neither the overview of the methodology in the main report, nor the material in Appendix B explains how it is implemented.

JWS Note: The summary of "Calculate P_c" in Appendix D and Item 1.5 in Table D-1 also mention the 1-to-10 scales, but the appendix does not provide any more information about when, or how, they are used in practice.

How does the 1-to-10 scale that is mentioned in this section relate to the tables in Appendix B or their use in an actual analysis?

How does the IDHEAS-ECA software implement these 1-to-10 scales and the associated linear interpolation process?

What are some specific examples of this process for the use of Tables B-1 through B-3, and for the use of Tables B-4 though B-15 in Appendix B?

19. Section 3.4, Step 4 - Assessing PIF Attributes Applicable to CFMs, Guidance for Assessing PIF Attributes

The first bullet item at the end of this section notes that:

> "Some factors, e.g., staffing, can be of concern, but there is no information on the 'average' staffing level. The analysts believe that plants follow the ***minimal adequate staffing*** rules. Thus, the boundary conditions would include the ***reasonable assumption*** that the Staffing PIF has ***'No Impact'*** on the HEPs for this HRA." [emphasis added]

I disagree very strongly with this guidance. It inappropriately biases analysts' assessments of the specific performance-influencing factor (PIF) for "Staffing", and it does not support an objective scenario-specific analysis.

JWS Note: This concern applies more broadly to the implication that an assessment that any particular PIF meets nominal regulatory or administrative requirements is adequate justification for assuming that the PIF never has any effect on any scenario-specific human failure events (HFEs). However, to explain my concern, I will limit my discussion to the specific example of staffing.

For example, the minimum staffing level at every plant is expected to meet nominal regulatory and administrative requirements. This guidance seems to imply that this staffing level is always adequate to cope with any evolving scenario conditions. Thus, if analysts assume that the minimum requirements are met, the "Staffing" PIF always has "no impact" on the analysis of any HFE. This assertion is not correct in practice, and it is not appropriate for "NRC-approved" guidance. Experience from actual events (e.g., some fires) has clearly demonstrated that the demands for coordinated actions in multiple locations throughout the plant can introduce challenges that require difficult decisions about the priorities and allocation of limited personnel to accomplish the desired actions. Furthermore, PRA models often contain scenarios that include multiple coincident and sequential actions that are well beyond the demands that are anticipated by the nominal administrative staffing requirements (e.g., multiple local manual actions to compensate for failures of automatic equipment responses, multiple functional requirements that result from cascading failures, some complex procedure-directed actions that are not fully integrated, scenarios that involve "beyond design basis" events, etc.). Thus, although the nominal minimum staffing may often be adequate for many actions that are limited to the personnel in the Main Control Room (MCR), that is certainly not the case for an objective and comprehensive analysis of many scenarios that are evaluated in a PRA.

Why does this section contain "NRC-approved" guidance which implies that analysts need not evaluate how the plant staffing affects supervisory decisions about the coordination of available resources to perform multiple actions, the feasibility of multiple coincident actions, and the time that is needed to accomplish those actions during a specific scenario?

Why is this type of guidance included in the methodology (i.e., regardless of the specific PIF)?

20. Section 3.4, Step 4 - Assessing PIF Attributes Applicable to CFMs, Guidance for Assessing PIF Attributes

The second bullet item at the end of this section notes that:

> "When assessing the remaining PIFs, analysts do not need to select the attributes if those are ***already represented in the selected base PIFs***." [emphasis added]

I do not understand the intent or the basis for this guidance.

According to the general IDHEAS methodology, the 20 performance-influencing factors (PIFs) are intended to be orthogonal, and they are intended to be evaluated as mutually exclusive effects on human performance. Of course, in practice, it is very difficult to demonstrate complete independence among some specific PIF attributes. Section 6.2.3 and Appendix D of NUREG-2198 address composite effects that may occur from interactions among multiple PIFs. However, without an explicit scenario-specific quantitative evaluation of those effects, the assumption that each PIF, and its attributes, should be evaluated independently is a fundamental element of the methodology. Considering that basic assumption, I do not understand why this guidance implies that an evaluation of the three base PIFs can account for the attributes of any of the 17 modifying PIFs. In other words, it is not apparent why it is asserted that an evaluation of only the three base PIFs could be used as justification for not evaluating the attributes of one or more of the 17 modifying PIFs.

Considering the fact that the 20 PIFs are intended to be evaluated as mutually exclusive effects on human performance, why is it implied that an evaluation of the three base PIFs can account for the attributes of any of the 17 modifying PIFs?

What is the intent and basis for this guidance?

Why is it included in this methodology?

21. Section 3.4, Step 4 - Assessing PIF Attributes Applicable to CFMs, Guidance for Assessing PIF Attributes

The third bullet item at the end of this section notes that:

> "Table 2-2 through Table 2-5 describe the PIFs in their ***No-Impact state***. The description is generic. For NPP control room operation, a simple way to think about No Impact PIFs is the following context: Experienced crews perform emergency operating procedures (EOPs) in control room simulators on routinely trained scenarios without complications, such as the SGTR example in Appendix C. With this context, ***all the PIFs are No Impact*** except the Task Complexity which is specific to a critical task." [emphasis added]

I disagree very strongly with this guidance. It inappropriately biases analysts' assessments of multiple performance-influencing factors (PIFs), and it does not support an objective scenario-specific analysis.

JWS Note: Table 2-2 through Table 2-5 simply summarize the scope and intent of each PIF.

They do not "describe the PIFs in their No-Impact state". However, that misrepresentation of the tables is not directly relevant to my concern in this comment.

A preceding comment on the first bullet item at the end of this section addresses a similar concern about "NRC-approved" guidance which inappropriately implies that analysts need not objectively evaluate how the scenario-specific conditions can affect each PIF. If a comprehensive objective analysis concludes that the "no impact" attribute applies for one or more of the 20 PIFs, then those assessments are justified. However, it is not appropriate for the guidance to implicitly bias those analyses by citing the general assertions about experience, procedures, and training that are mentioned in this excerpt. It is also certainly not apparent why the assessments of experience, procedures, and training can justify the conclusion that no other PIFs are potentially important, except for "Task Complexity".

Why does this section contain "NRC-approved" guidance which implies that analysts need not perform an objective scenario-specific assessment of each PIF that may affect personnel performance and document the bases for their assessment?

Why is this type of guidance included in the methodology (i.e., regardless of the specific PIF)?

22. Section 3.5, Step 5 - Estimation of P_c - The Sum of Human Error Probabilities of Cognitive Failure Modes

The bullet item that discusses P_c notes that:

> "If operators' responses are as trained, then the time available to complete the action is sufficient."

This statement seems to inappropriately bias analysts' perspectives on the evaluation of P_t. In particular, analysts may interpret this guidance to imply that P_t is always negligible if the operators' training is adequate for the particular scenario. That is certainly not the case, and it does not appropriately characterize P_t as accounting for uncertainties in the time that is available and the time that is needed to complete an action, within the context of the operators' training, experience, etc. (The bullet item that discusses P_t provides that general concept.) This sentence is not necessary to explain the intent of P_c, or what it evaluates.

Why does the discussion contain this sentence?

Can it be simply removed from the report (i.e., rather than modified or elaborated further), without altering an analyst's basic understanding of the methodology?

23. Section 3.5.1, Estimation of P_c; Appendix B, Base Human Error Probabilities and Performance-Influencing Factor Weights

Reference [1] for this report is:

> J. Xing, Y. J. Chang, and J. DeJesus, "The General Methodology of an Integrated Human Event Analysis System (IDHEAS-G) - Draft Report," U.S. Nuclear Regulatory Commission, NUREG-2198 (ADAMS Accession No. ML19235A161), Aug. 2019.

The current version of NUREG-2198 (ADAMS accession number ML20238B988) was issued to support the September 23, 2020 meeting of the ACRS Subcommittee on Reliability and

Probabilistic Risk Assessment.

One important difference between the 2019 draft version of NUREG-2198 and the current version of that report is that the formulations of Equation 4.6 and Equation 4.7 changed. Those changes have a fundamental effect on the way that the performance-influencing factor (PIF) weights are calculated and are applied for quantification of the human error probability (HEP).

For reference, Equation 4.7 in the 2019 version of NUREG-2198 was:

$$P_{CFM} = P_{CFM_{Base}} \cdot \left(1 + \sum_{i=1}^{n} w_i\right) \cdot C \cdot \frac{1}{Re} = \frac{P_{CFM_{Base}} \cdot (1 + w_1 + w_2 + \cdots + w_n) \cdot C}{Re}$$

Equation 4.7 in the current version of NUREG-2198 is:

$$P_{CFM} = P_{CFM_{Base}} \cdot \left(1 + \sum_{i=1}^{n} (w_i - 1)\right) \cdot C \cdot \frac{1}{Re} = \frac{P_{CFM_{Base}} \cdot \big(1 + (w_1 - 1) + (w_2 - 1) + \cdots + (w_n - 1)\big) \cdot C}{Re}$$

The format of Equation 3.4 in this report is identical to Equation 4.7 in the current version of NUREG-2198. The format of Equation 3.6 is also identical to Equation 4.6 in the current version of NUREG-2198. This implies that the numerical PIF weights that are tabulated in Appendix B of this report are derived and applied according to the current formulation of the general quantification model (i.e., not according to the model in Section 4.4.3.2 of Reference [1]).

I think that is the case. For example, suppose that only one attribute for only one of the 17 modifying PIFs is relevant for a particular critical task with a "base" HEPb of 1E-02. Suppose also that the applicable table in Appendix B indicates that the attribute has a weight of 1.

- According to the current version of Equation 4.7, the modified HEP would be:
  HEPm = HEPb * [1 + (w – 1)] = HEPb * [1 + (1 – 1)] = 1E-02

This seems to be consistent with the general guidance in Appendix B that a weight of 1 applies if the attribute has "no impact" on the "base" HEP.

- According to the 2019 draft version of Equation 4.7, the modified HEP would be:
  HEPm = HEPb * [1 + w] = HEPb * [1 + 1] = 2E-02

It is not apparent why an attribute that has "no impact" would double the "base" HEP.

Thus, based on the formulation of Equation 3.4 in this report and this simple example, it seems evident that the PIF weights that are tabulated in Appendix B are derived according to the current version of Equation 4.6 in NUREG-2198, and they are applied according to the current version of Equation 4.7.

To avoid the possibility that readers may become very confused and misuse the guidance or the numerical values in this report, I think that it is very important for Section 3.5 and Appendix B to explicitly document the fact that the formulas which were used to derive and apply the PIF weights are not the formulas from Reference [1].

Are the weights that are listed in the tables in Appendix B of this report actually derived according to the current formulation of Equation 4.6 in NUREG-2198, and are the weights applied according to the current formulation of Equation 4.7 (i.e., according to Equation 3.6 and Equation 3.4 in this report)?

Does the IDHEAS-ECA software also use those equations?

If so, why does this report not explicitly document those differences from the quantification model that is described in Reference [1]?

If the current versions of Equation 4.6 and Equation 4.7 were not used consistently to derive the numerical values in Appendix B (or if a hybrid of the 2019 and current formulations was used), why does the "NRC-approved" guidance for IDHEAS-ECA support a methodology where the quantitative relationship between the PIF weights and HEPs is different from the general quantification model that is described in the current version of Section 4.4.3.2 in NUREG-2198?

24. Section 3.5.1, Estimation of P_c

The discussion of Equation 3.5 notes that P_{SF}, P_{INF}, and P_{TC} account for the "base" human error probabilities (HEPs) that are attributed to the performance-influencing factors (PIFs) for "Scenario Familiarity", "Information Availability and Reliability", and "Task Complexity". The HEPs that correspond to various attributes of each PIF are listed in Table B-1 through Table B-3 in Appendix B.

The methodology does not describe how analysts should account for scenario-specific conditions that adversely affect multiple attributes for a particular PIF. For example, consider the PIF for "Scenario Familiarity" in Table B-1. Suppose that an analyst concludes that the scenario-specific conditions affect attributes SF1 and SF2 and SF3. In other words, suppose that the scenario involves unpredictable dynamics, involves tasks or strategies that are not consistent with normal practice, and requires that the operators must notice ancillary or supplemental indications.

Suppose also that the critical task for a particular human failure event (HFE) involves the cognitive failure mode (CFM) for "Detection" (D). According to Table B-1, the following "base" HEPs apply for that CFM under these scenario conditions.

- SF1 6.6E-04
- SF2 5.0E-03
- SF3 1.2E-02

The simple sum of these HEPs is 1.766E-02.

A calculation formula of $HEP_{SF} = 1 - (1 - HEP_{SF1})*(1 - HEP_{SF2})*(1 - HEP_{SF3})$ would provide a result of 1.759E-02.

This comment also applies to the treatment of scenario conditions that affect multiple attributes for each of the 17 modifying PIFs (i.e., the composite effects on multiple attributes for a single PIF in Table B-4 through Table B-14). For example, the scenario environmental conditions may involve elevated temperatures (ENV3) and poor lighting (ENV4) and loud noise (ENV7). Based on the values that are listed in Table B-4, it seems that the composite PIF weight for the "Execution" CFM (E) might be 1.5 + 2.0 + 1.1 = 4.6. However, if the applied combination formula is $w_{ENV} = 1 + (w_{ENV3} - 1) + (w_{ENV4} - 1) + (w_{ENV7} - 1)$, the composite weight would be 2.6.

In this example, is the "base" HEP for the "Detection" CFM the sum of these HEPs (i.e., approximately 1.76E-02)?

If not, what is the "base" HEP that applies for these composite effects?

If the PIF attribute effects are not added, what is the technical basis for the quantification model that is actually used?

Is the same calculation process used to derive a composite weight for a particular PIF (i.e., if the scenario conditions affect multiple attributes for a single PIF in Table B-4 through Table B-14)?

In particular, in this example for the "Environmental" PIF, is the composite weight for the "Execution" CFM equal to 4.6, 2.6, or some other value?

Why does the guidance in this section not describe how the quantification process accounts for scenario-specific conditions that adversely affect multiple attributes for a particular PIF?

25. Section 3.5.1, Estimation of P_c, Guidance for Estimating P_c

This section notes that:

> "Note that the current version of IDHEAS-ECA and the software only provide the ***mean values*** of the base HEPs and PIF weights without giving the information of the main body and range of the distribution of those values. ***If an HRA requires*** the inclusion of the HEP distribution, the analysts need to ***make their own judgment*** of the distribution of Pc." [emphasis added]

This is a fundamental technical deficiency in the IDHEAS-ECA methodology. This report should not be published in a final form until guidance and available data are provided to support quantification of the uncertainties in the "base" human error probabilities (HEPs), the performance-influencing factor (PIF) weights, and the overall distribution for P_c.

It is not appropriate for "NRC-approved" guidance to imply that a quantitative evaluation of uncertainty is not a fundamental element of the human reliability analysis (HRA) process. That implication is contrary to the state-of-practice, applied standards, and NRC policy for the performance of risk assessments and the use of PRA for risk-informed decision-making. In short, quantification of uncertainty is not an option that is left to the judgment of an analyst. It is a fundamental part of every analysis.

It is also not appropriate for "NRC-approved" guidance to imply that analysts should use their judgment to "back-fit" an uncertainty distribution to the overall results for P_c. The uncertainty distribution for P_c should be derived from a process that consistently combines the uncertainties in the "base" HEPs with uncertainties in the PIF weights that modify those HEPs. Thus, the overall uncertainty in P_c should be the result from a systematic evaluation of the uncertainties in each element of the analysis, using readily-available tools to combine the constituent uncertainty distributions. The overall uncertainty is not an afterthought that is "back-fit" according to the notions of each analyst.

I obtained a draft copy of RIL-2021-XX, "Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA)", that was released publicly to support a September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment. I hope that the analyses in that report identify and quantify the uncertainties for each parameter that is evaluated (i.e., the "base" HEPs and the PIF weights that are summarized in Appendix B of this report). If the IDHEAS-DATA report does not contain that information, that is a fundamental technical flaw in the way that the data were developed and represented for use in a modern PRA. I will comment separately on that report.

JWS Note: Item 2) in Section 4.3 indicates that "it is desirable to develop the probabilistic distribution of the base HEPs and PIF weights". This is not a "desirable" improvement. It is necessary to correct an important technical deficiency in the methodology.

Why does the IDHEAS-ECA methodology contain "NRC-approved" guidance which implies that a quantitative evaluation of uncertainty is not a fundamental element of the HRA process?

Are measures of the uncertainty in each "base" HEP and each PIF weight available from the supporting data analyses?

If so, why are those uncertainties not documented in Appendix B?

If measures of the uncertainties are not currently available, will the NRC staff postpone final publication of the IDHEAS-ECA report (and the IDHEAS-DATA report) until the uncertainties are documented?

If it is intended to publish the IDHEAS-ECA report without clear guidance to quantify the uncertainties and without supporting estimates of the underlying data uncertainties, how is that decision consistent with the state-of-practice, applied standards, and NRC policy for the performance of risk assessments and the use of PRA for risk-informed decision-making?

26. Section 3.5.1, Estimation of P_c, Guidance for Crediting Recovery Effect in P_c, General Comment

The first paragraph in this section clarifies the authors' distinction between a "recovery action" and "recovery of an HFE". In the context of the IDHEAS-ECA methodology, the analysis of a "recovery action" requires the definition and evaluation of a new human failure event (HFE). The "recovery of an HFE" is an integral part of the analysis of an HFE, and it affects the human error probability (HEP) for that HFE (i.e., through recovery factor R_e in Equation 3.4). This clarifies the intent and scope of each type of analysis.

This section provides a very good discussion of the considerations that are needed to assess the feasibility and evaluate the likelihood for "recovery of an HFE".

JWS Note: This guidance is not unique to the IDHEAS-ECA application. I think that it would be very useful to also include this discussion in NUREG-2198, to expand and clarify the general notions of "recovery" that are currently addressed in that report.

27. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required

Section 3.6 notes that the IDHEAS-ECA software accepts the following types of uncertainty distributions: Normal, Gamma, Weibull, and Five-Point Discrete Probability. The software also accepts a single point-estimate value for T_{avail}.

I am happy to see that the time uncertainty can be represented by a five-point discrete probability distribution. In many cases, the most informative representation of the uncertainty in a parameter value may be provided by a discrete probability distribution that does not have a defined analytical form. Experience has also shown that discrete probability distributions are often the best and most efficient format to represent the uncertainty from an expert elicitation process.

It is noteworthy that the catalog of available functions does not include a lognormal distribution. Lognormal distributions are used extensively in PRAs. Hence, they may be much more familiar to many analysts than a gamma distribution or a Weibull distribution. Of course, in practice, it may occasionally be necessary to truncate one or both of the lognormal distribution "tails" to account for physical constraints or analyst judgments that affect the maximum range of the evaluated times. However, that practice may also be necessary for the other analytical forms.

Why does the IDHEAS-ECA software not accept a lognormal probability distribution?

28. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required; Appendix A, Worksheet E. Time Uncertainty Analysis of the HFEs

The introduction to Worksheet E notes that:

> "The distribution of the time can be estimated as a single number (for T_{avail} only), the ***mean and standard deviation*** (SD) by assuming a normal distribution, or a five-point estimation of probability distribution (at 5^{th}, 25^{th}, 50^{th}, 75^{th}, and 95^{th} percentile) or Gamma distribution or Weibull distribution." [emphasis added]

The discussion of the normal probability distribution in Section 3.6 notes that it can be represented by a mean value and a standard deviation. Of course, the mean value of a normal distribution is equal to the median value. Thus, the mean value represents the "best estimate" for the central tendency of the underlying uncertainty. That is not the case for the gamma distribution or the Weibull distribution. As noted in Section 3.6, those distributions are typically specified by parameters that measure the shape and the scale of the distribution. Unless analysts have extensive experience with the use of those analytical forms, those parameters are not easily related to measures of the "best estimate" and the range of times in a practical analysis.

Experience has shown that an analyst's "best estimate" for a parameter value is often more closely associated with the median value of the uncertainty distribution, rather than the mean value. That is especially true for situations where the uncertainty is rather large and is represented by a distribution that is not linearly symmetric about the median, as is often the case in many PRA analyses. In this context, it seems more appropriate for the general guidance to indicate that the uncertainty distribution should be characterized by a median value and a standard deviation, or perhaps by a median value and estimates of an upper-bound value and a lower-bound value. That is especially important for distributions that are derived from expert judgment, because experts tend to focus on the central tendency of an estimate (i.e., the median), rather than the expected value that accounts for their uncertainty (i.e., the mean). However, this observation also applies to estimates that are derived from engineering calculations that use only "nominal" or "best estimate" input values.

Furthermore, many analysts may not be familiar with techniques to estimate a standard deviation that accurately represents the full range of their uncertainty. Thus, alternative measures are often more useful to specify the range of the distribution (e.g., 5^{th} and 95^{th} percentiles, error factor, etc.).

JWS Note: After carefully reading Section 3.6 and the introduction to Worksheet E a few times in the context of this comment, I understand that the guidance explicitly indicates that the mean value and standard deviation input applies only when the normal distribution is used. However, the intent of this comment is that the guidance and the IDHEAS-ECA software should facilitate analyst specification of the uncertainty distribution in a format that is intuitive and easy to understand, such as the "best estimate" (i.e., median) time and estimates for the upper- and lower-bound times. The software can then derive the entire distribution, based on the general shape that is desired. In other words, this should be a practical exercise, and it should not require that human reliability analysts must be fully familiar with subtle mathematical features of each distribution form. In that sense, it may be better to use a lognormal form to represent a skewed uncertainty distribution, because it is familiar to most PRA analysts, and it is easily specified by intuitive parameter values.

Why does the guidance indicate that an analyst should specify the mean value for each type of uncertainty distribution, rather than the median value?

Why does the software require that the range of a distribution can be specified only by its standard deviation, rather than offering other more intuitive options to characterize that input?

Are there alternative ways to specify a gamma distribution or a Weibull distribution that are more intuitive and that relate to the types of information that are typically available from a practical analysis (i.e., a "best estimate" time and estimates of the upper and lower bounds of that time)?

29. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Available

This section notes that:

> "It represents the time lapse from time zero to the time that a selected key parameter would exceed its ***safety threshold*** without human intervention." [emphasis added]

This characterization inappropriately implies that T_{avail} is determined by "traditional" safety analyses, with their inherent assumptions and licensing-based success criteria (e.g., analyses that are documented in the plant's FSAR or other licensing reports). It does not convey the intent or use of T_{avail} in the context of a PRA model.

Section 5.3 of NUREG-2198 provides the following definition for the system time window (T_{sw}):

> "T_{sw} is defined as the system time window and is the time from the start of the event until the action is no longer beneficial (typically when irreversible damage occurs, such as core or component damage). It is typically derived from thermal-hydraulic data for the representative PRA scenario and, for HRA quantification, is considered to be a fixed input. The system time window represents the maximum amount of time available for the action."

This definition more accurately characterizes the functional intent and application of T_{avail}. In particular, T_{avail} ends when the particular action (or human failure event) that is being analyzed is no longer effective, or the scenario evolution alters the scope or the need for the modeled actions. In many PRA models, those scenario conditions are not related to the time when a "key parameter" exceeds its "safety threshold". For example, the end of T_{avail} for personnel actions to restore a source of feedwater is typically determined by the secondary heat removal conditions that require initiation of feed and bleed cooling. The end of T_{avail} for personnel actions to restore offsite power is determined by the time when the non-safety batteries are depleted. The end of T_{avail} for personnel actions to cool down and depressurize the reactor is determined by the conditions at which initiation of low pressure makeup is no longer effective to prevent fuel damage (or, perhaps, prevent containment failure in a Level 2 PRA model).

Why does this characterization of T_{avail} focus narrowly on the time when a "key parameter" exceeds a "safety threshold"?

30. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Available

This section notes that:

> "It represents the time lapse ***from time zero*** to the time that a selected key parameter would exceed its safety threshold without human intervention." [emphasis added]

The characterization of T_{avail} in this report does not seem to distinguish between the system time window (T_{SW}) and the time that is available to perform the desired action (T_{avail}). That distinction is discussed in Section 5.3 of NUREG-2198, and it is shown in Figure 5-2 of that report. In particular, the IDHEAS-ECA guidance seems to assume that T_{delay} is always negligible. That assumption can lead to an optimistic assessment of T_{avail} (i.e., an inappropriately long time window), with an associated optimistic influence on the evaluated human error probability (HEP). Of course, the amount of optimism depends on what conditions determine the definition of time zero for the scenario, and the amount of time that elapses between that time and the first cue for the desired action.

This section of the report does not contain guidance for determining when time zero occurs. The "Guidance for Developing the Scenario Timeline" in Section 3.1.1 explicitly notes that time zero for a predictive (i.e., PRA) analysis corresponds to the time when the initiating event occurs. For a retrospective analysis, it is the actual time when the event occurred.

Thus, it seems very likely that analysts will assume that time zero always corresponds to the time when the initiating event occurs. If that is the case, then T_{delay} can be very long, especially for actions that are evaluated at later stages of the scenario progression, actions that are evaluated to prevent or mitigate containment failure, severe accident management actions, etc. However, even if T_{delay} is relatively short, it can affect the analyses of actions that are needed early in the scenario, with small functional time windows to achieve the desired plant conditions.

Why does the guidance not alert analysts to the need to determine T_{delay} for the specific scenario and the modeled action?

Why does the guidance not alert analysts to account for T_{delay} when they estimate T_{avail} for the modeled action?

31. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Available

This section notes that:

> "The analytic approach starts by reviewing the preliminary risk analysis results to identify the ***dominant risk contributors***. The calculations can help analysts identify ***areas where uncertainty analysis is needed*** and where more sophisticated analyses should be performed to better define the success criteria. This phased approach makes ***uncertainty analysis feasible***." [emphasis added]

I disagree very strongly with this guidance and its implications. It seems to tell analysts that uncertainty analysis is needed only for the "dominant risk contributors", which are identified from analyses that do not account for uncertainty. Uncertainty analysis is always "feasible", even if it involves only a subjective quantitative assessment of the uncertainties, based on expert judgment.

The NRC State-of-the-Art Reactor Consequence Analysis (SOARCA) project has clearly shown the importance of evaluating and quantifying uncertainties as an integral part of the risk assessment process (i.e., as an integral element of the parameter estimations). In particular, that project has shown that the uncertainties may determine which specific scenarios are important contributors to overall risk. A retrospective assessment of uncertainty for only the "dominant contributors" from a "point estimate" calculation may not identify the correct risk contributors, and it is not consistent with the general guidance for performing a risk assessment.

It is certainly reasonable to indicate that analysts may use a variety of methods to evaluate the uncertainty in T_{avail}, without performing numerous resource-intensive thermal-hydraulic simulations. However, it is not appropriate for "NRC-approved" guidance to imply that uncertainties should be evaluated as an after-thought or as an add-on that is not an integral part of the analysis process.

JWS Note: Section 5.2 in the current version of NUREG-2198 was revised to remove the emphasis on "dominant risk contributors". It now notes that a variety of methods may be used to evaluate uncertainties in T_{avail}, with more detailed analyses performed as needed.

JWS Note: This comment also applies to Section 3.8. The first paragraph in that section appropriately notes that "The assessment of uncertainty on HEPs is a required part of the PRA". However, the second paragraph notes that "Assessment of the uncertainty in the HEPs should be performed (***at least for the significant HEPs***)…". Item 2) in the guidance for Step 8 also indicates that uncertainty distributions are developed only for the "significant HEPs".

Why does "NRC-approved" guidance imply that uncertainty analysis for T_{avail} is needed only for the "dominant risk contributors", which are identified from analyses that do not account for uncertainty?

Why does "NRC-approved" guidance imply that an assessment of uncertainty may be needed for only the "significant HEPs", which are identified from analyses that do not account for uncertainty?

32. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Required

This section notes that:

> "P_t is the probability that personnel ***could not complete*** the required human action within the available time." [emphasis added]

This is an accurate characterization of P_t. The highlighted phrase may alert analysts to the fact that the time required to perform the action (T_{reqd}) should account for entire time that is needed to achieve the desired plant conditions.

In practice, I have seen many estimates of T_{reqd} that have accounted only for the time that is needed to initiate the desired action (e.g., to open a valve, start a pump, etc.). In particular, the analyses did not include the subsequent time that is needed to achieve the plant conditions which 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}. A simple example may be useful to clearly reinforce this notion and provide confidence that analysts do not inadvertently overlook this contribution.

Why does this guidance not explicitly alert analysts to account for the total amount of time that is needed to achieve the desired plant conditions?

33. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Required

This section notes that:

> "Human actions in HRA are ***assumed being performed as trained***. The distribution of the time-required to complete a ***trained action*** can be caused by many factors." [emphasis added]

I do not understand why it is necessary to explicitly emphasize training in the context of this discussion. The evaluation of P_c examines the effectiveness of training in the scenario-specific context of the modeled human failure event (HFE). For some actions, that training might not be very comprehensive. However, despite the lack of training, the analysts have concluded that the action is feasible, and it is included in the PRA model. Therefore, it is presumed that personnel will perform the action, compensating as necessary for their lack of training. Estimates of the time that is needed to perform the action, including the uncertainty in that time, should account for the scenario context, without assuming that the operators are well-trained for the particular response or that any other performance-influencing factors (PIFs) are necessarily optimal.

JWS Note: The discussion of the process for estimating T_{reqd} (T_n) in Section 5.3.1 of NUREG-2198 does not contain this qualification or mention the effects from any other PIFs. It simply notes that "Estimation of T_n is based on the baseline scenario and its context identified in the scenario analysis of an HRA".

JWS Note: This comment also applies to the first bullet item in the "Guidance on Selecting a Time Distribution". That guidance implies that differences in training may be the primary consideration for not using a normal distribution to characterize the uncertainty in T_{reqd}.

Why is it necessary to explicitly emphasize training in the context of this discussion?

34. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Required

This section notes that:

> "Estimating the distribution of T_{reqd} should consider three key aspects: nominal contributors, uncertainty factors, and ***bias factors***." [emphasis added]

Table 3-4 summarizes the typical factors that should be considered in the estimation of T_{reqd} (i.e., the "nominal contributors"). Table 3-5 summarizes sources of uncertainty. The guidance does not address the consideration of "bias factors". Section 5.3.1 in NUREG-2198 discusses three specific sources of bias that may affect the estimation of T_{reqd}.

Why does this guidance not address the consideration of bias factors?

35. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Required

The first bullet item in this section notes that:

> "Average crew response time should be obtained, as well as an estimate of the time by which the slowest operating crews would be expected to complete the actions."

I think that the guidance should emphasize the importance of developing a complete uncertainty distribution, and not focus only on the central tendency and the "upper tail" of that distribution. Therefore, I think that the guidance should indicate that an average crew response time should be obtained, as well as estimates of the times by which the fastest and slowest operating crews would be expected to complete the actions.

JWS Note: The guidance in Section 5.3.1 of NUREG-2198 indicates that estimates of the times for both the fastest and the slowest crews should be obtained.

Why does this guidance not indicate that the initial uncertainty distribution should also estimate the fastest crew response time?

36. Table 3-4, Typical Factors Contributing to T_{reqd}

A time uncertainty analysis is also needed for integrated human actions that require coordination of multiple teams of personnel. Although the timelines and inter-relationships among various teams may be rather complex, methods are available to display and account for combinations of series, parallel, and functionally dependent activities. Therefore, it is not appropriate to imply that an evaluation of the uncertainty in parameter T_{reqd} is not needed, or that the evaluation need not consider the effects from the Interteam Coordination macrocognitive function. In fact, the time needed for effective team coordination may be the most important source of uncertainty in those analyses.

JWS Note: Table 5-1 in NUREG-2198 includes factors that contribute to the Interteam Coordination macrocognitive function.

Why does Table 3-4 not include any factors that contribute to the time required for the Interteam Coordination macrocognitive function?

37. Table 3-5, Uncertainty Factors that Modify the Distribution of T_{reqd}

The considerations for "Plant Condition" are:

> "Simultaneous multiple events that demand the same set of resources. Multiunit events (e.g., an external hazard impacts multiple units in the same site)."

Because of the spacing and the indent in the table, I do not know if these are separate entries, or whether consideration of multi-unit events is intended as a specific example of "simultaneous multiple events". However, clarification of that intent is peripheral to my primary concern in this comment.

None of the considerations in this table seem to explicitly address scenario-specific sources of distractions, interruptions, possibly conflicting priorities, stress, etc. that may divert supervisors' and operators' attention away from the desired course of action for the defined human failure event (HFE). In practice, many of those sources of uncertainty for T_{reqd} may arise from conditions that are not modeled explicitly in the PRA, and are documented only in the scenario narratives.

JWS Note: The considerations for "Plant Condition" in Table 5-2 in NUREG-2198 alert analysts to account for plant-wide conditions that may introduce distractions, interruptions, competing demands, or conflicting priorities.

Should the considerations in this table explicitly prompt analysts to account for plant-wide conditions that may distract supervisors' and operators' attention or introduce competing demands and delays, despite the availability of adequate personnel and equipment to perform the PRA-modeled action?

38. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance for Estimating the Distribution of Time Required, Guidance on Selecting a Time Distribution

The discussion of the normal distribution notes that:

> "Normal distribution is ***often used*** to represent the uncertainties in the time that it takes humans to perform an action." [emphasis added]

The discussion of the five-point discrete probability distribution notes that:

> "If operational data are not adequate for confident estimation of the mean and standard deviation of the ***assumed normal distribution***, or if evidence suggests that normal distribution is not appropriate for the situation…." [emphasis added]

The first bullet item in the "Guidance on Selecting a Time Distribution" notes that:

> "For example, the time it takes for individuals or crews to perform a well-trained task in the same way as trained ***mostly likely would fit to a normal distribution*** because the variability mainly comes from individual differences. However, if the individuals fall into two categories: well trained and less experienced with little training, then the performance time will not fit to a single normal distribution." [emphasis added]

This guidance inappropriately biases analysts to use a normal distribution to characterize the uncertainty in T_{reqd}. In fact, experience has shown that it is very often not appropriate to represent this uncertainty with a normal distribution. In practice, 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 uncertainty 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 as a gamma, a Weibull, or perhaps a lognormal distribution.

It is important for the shape and the range of the uncertainty distribution to account for the analysts' consideration of these "outlier" effects. In practice, 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}.

Why does this "NRC-approved" guidance emphasize the use of a normal distribution to characterize the uncertainty in T_{reqd}?

Why does this guidance not alert analysts to the types of considerations that are mentioned in this comment when they select the shape and the range of the distribution for a particular scenario-specific response?

39. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance on Selecting a Time Distribution

The second bullet item in this section notes that:

> "For example, Gamma distribution should be a better choice than Normal distribution in modeling debris removal times where heterogeneous factors are involved."

In practice, except for extremely simple responses like pushing a button on a main control panel according to an explicit procedural requirement, it seems that the vast majority of human actions that are evaluated in a PRA are influenced by "heterogeneous factors". That seems especially true for analyst considerations of the variety of factors that may affect the uncertainty in the amount of time that is needed to complete an action. Thus, it is not apparent why this discussion uses the rather unusual example of removing debris to illustrate a situation where a gamma distribution might provide an appropriate shape for the uncertainty in T_{reqd}.

Why does this discussion imply that "heterogeneous factors" affect personnel response for only a limited number of behaviors?

What is the technical basis for this implied limitation?

Why is the rather unusual example of removing debris used to illustrate a situation where a gamma distribution might provide an appropriate shape for the uncertainty in T_{reqd}?

40. Section 3.6, Step 6 - Estimation of P_t - The Convolution of the Distributions of Time Available and Time Required, Guidance on Selecting a Time Distribution

The third bullet item in this section notes that:

> "Weibull distribution is the most widely used for modeling reliability data."

I disagree with this assertion. In my experience, lognormal distributions have been used historically to characterize uncertainties for the vast preponderance of equipment reliability data that are used in PRAs. NUREG/CR-6928 uses beta distributions for demand failure probabilities (e.g., failure to start) and gamma distributions for time-related failure rates (e.g., failure to run). The authors of that report explain why they preferred to use those distributions, rather than the lognormal.

What is the technical basis for this assertion?

41. Section 3.8, Step 8: Analyze HRA Uncertainties and Perform Sensitivity Analysis

Item 3) in this section notes that:

> "For the uncertainty that results from ***whether a PIF attribute should be included*** in the HEP calculation, IDHEAS-ECA recommends using sensitivity analysis or bounding analysis for the PIF attributes that are ***important contributors to the HEPs*** of the HFEs which have ***significant impact*** on the risk (at PRA level)." [emphasis added]

I am a bit confused by this guidance. It seems to address only the performance of a sensitivity analysis to examine the effects from a particular performance-influencing factor (PIF) attribute that is included in the human error probability (HEP) calculation, when there is uncertainty about whether that attribute should have been omitted. For example, the HEP can be re-quantified without the effects from the questionable attribute, and the PRA model can be re-quantified to determine the corresponding change (decrease) in risk.

The guidance does not seem to address the performance of a sensitivity analysis to examine the effects from a particular PIF attribute that is omitted from the HEP calculation, when there is uncertainty about whether that attribute should have been included. In that situation, it is not appropriate to focus only on the human failure events (HFEs) that have a "significant impact on the risk". In particular, omission of the PIF attribute may inappropriately suppress the importance of specific HFEs which would have been significant if the attribute had been included in the analysis. Thus, to examine the sensitivity to an omitted PIF attribute, it seems that it is necessary to add the attribute, re-quantify the HEPs for all affected HFEs (i.e., regardless of their nominal significance), and then re-quantify the PRA model to determine the corresponding change (increase) in risk.

How does the guidance in Item 3) address the performance of sensitivity analyses to examine the effects from a particular PIF attribute that is omitted from the HEP calculation, when there is uncertainty about whether that attribute should have been included?

42. Table 3-7, Summary of IDHEAS-ECA Worksheets, Worksheet E, Editorial Comment

I do not think that it is necessary to explicitly list the types of distributions in this summary. However, if the authors prefer to retain the list, it should be complete. The summary omits the gamma distribution and the Weibull distribution.

43. Section 3.9, Summary of IDHEAS-ECA, IDHEAS-ECA Summary

This section notes that:

> "Alternatively, analysts can manually calculate ***the HEP*** using the base HEPs and PIF weights in Appendix B." [emphasis added]

Analysts can use the "base" human error probabilities (HEPs) and performance-influencing factor (PIF) weights from Appendix B to calculate the cognitive error contribution to the overall HEP (i.e., P_c). They must use either the IDHEAS-ECA software or another method to convolute the time uncertainty distributions to calculate P_t.

Why does this summary imply that the overall HEP can be calculated from the "base" HEPs and PIF weights in Appendix B?

44. Section 4.1, From IDHEAS-G to IDHEAS-ECA, General Comment

This is a very good summary.

45. Section 4.2, Integration of Human Error Data for IDHEAS-ECA; Section 4.3, Future Development and Improvement

Section 4.2 notes that the IDHEAS-G methodology contains three types of tables: HEP tables, PIF weight tables, and PIF interaction tables. The IDHEAS-ECA application adopts the HEP tables and PIF weight tables. However, it does not contain the PIF interaction tables.

Section 6.2.3 and Appendix D of NUREG-2198 address composite effects that may occur from interactions among multiple performance-influencing factors (PIFs). Appendix D contains some examples which illustrate that the recommended linear sum of PIF weights may not always provide a good estimate for how overall human performance is affected by possibly inter-related PIFs. That is very useful information for analysts, even if it is not treated explicitly in the human error probability (HEP) quantification process. For example, analysts might adjust their evaluations of some PIF attributes to account for scenario-specific conditions that introduce potential interactions.

Section 4.2 does not discuss why the IDHEAS-ECA application does not include the PIF interaction tables. Section 4.3 does not identify this element of the data generalization process as a need for future development and improvements.

Why do Section 4.2 and Section 4.3 not discuss why the IDHEAS-ECA application omits the PIF interaction tables, or similar considerations of the effects from non-linear combinations of the PIF weights?

46. Section 4.2, Integration of Human Error Data for IDHEAS-ECA

The summary of strategy 1) in this section notes that:

> "For the multiple data points that have about the same level of applicability and certainty, the NRC staff used the ***median*** of the data points as the base HEP or PIF weight." [emphasis added]

This summary seems to indicate that some of the numerical human error probability (HEP) values that are listed in Table B-1, Table B-2, and Table B-3 of Appendix B are median values from the distributions of available data. Median values of the HEP estimates may have also been used to derive the performance-influencing factor (PIF) weights that are summarized in Table B-4 through Table B-15. The guidance in this report (e.g., Section 3.5.1) explicitly indicates that all of the numerical values that are tabulated in Appendix B are mean values, and they are used as such in the HEP quantification process.

Fundamental PRA guidance emphasizes that the mean value of the underlying uncertainty distribution must always be used for "point estimate" calculations (e.g., for initial quantification of the event tree and fault tree models). Use of median values is appropriate only if the uncertainties in the underlying data are characterized by a normal probability distribution, for which the median value and the mean value are identical. For example, if the uncertainty in the data is characterized by lognormal probability distribution, use of the median value will provide an estimate that is inappropriately lower than the mean value.

JWS Note: This comment also applies to the summary of strategy 2) in this section, where it is noted that: "When there were only a few data points or a variety of CFMs and PIFs involved in the data points, the NRC staff combined the data points to estimate the range and then used the ***middle of the range*** as the base HEPs or PIF weights". [emphasis added]

Why did the staff select the median value of the available data to represent the base HEP or PIF weight, rather than the mean value?

Are some of the numerical HEP values in Table B-1, Table B-2, and Table B-3 actually median values from the distributions of available data?

Are some of the numerical PIF weights in Table B-4 through Table B-15 derived from median values of the available HEP estimates?

47. Section 4.3, Future Development and Improvement

Item 1) in this section discusses the current status of the IDHEAS-DATA effort and the importance of adding new information and data to the database. It is noted that:

> "Even if there are multiple data points for a base HEP or PIF weight, judgment and reasoning are still needed in generalizing and integrating the human error data because of uncertainties and complications in the data sources. The data sources as well as the process and considerations in generating the base HEPs and PIF weights should be documented. The NRC staff will develop such documentation aside from this method report."

Preceding comments address the lack of estimates for the uncertainties in the "base" human error probabilities (HEPs) and performance-influencing factor (PIF) weights that are tabulated in Appendix B of this report. Item 2) in this section also addresses that issue.

It seems that the process of synthesizing and updating the IDHEAS-DATA database should involve the use of Bayesian techniques. For example, the underlying prior uncertainty distribution for each HEP and each PIF weight could be based on the experience and judgment of the subject matter experts. Of course, the uncertainties in those estimates may be rather large, but the distributions might have some reasonable bounds and shapes. In other words, the experts' knowledge may not necessarily be well-represented by something like a purely non-informative Jeffreys prior distribution.

As more relevant data are compiled, the effects from those data will systematically update the prior estimates and will (usually) reduce the associated uncertainties. This process is used by the NRC to develop estimates for equipment reliability data (e.g., as in NUREG/CR-6928), and it is accepted and used throughout the risk assessment community. Thus, the use of Bayesian techniques would provide a systematic, well-accepted method for updating the databases as more information becomes available, and it would provide consistent, technically-justified estimates for the uncertainty in each parameter value.

JWS Note: I obtained a draft copy of RIL-2021-XX, "Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA)", that was released publicly to support a September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment. I will comment separately on that report. When I prepared these comments on the IDHEAS-ECA application, I had not yet read the IDHEAS-DATA report. Therefore, I do not know the extent to which the current estimates in Appendix B of this report may benefit from the use of Bayesian methods. However, a quick word search of the IDHEAS-DATA report indicates that "Bayes" is used only once, as a brief observation in the introduction to Section 2.

Will the NRC staff use Bayesian techniques to synthesize and update the IDHEAS-DATA estimates for the "base" HEPs and the PIF weights?

If not, what is the reason for the staff's reluctance to use those methods?

48. Section 4.3, Future Development and Improvement

Item 4) in this section discusses the evaluation of "recovery" within the context of quantifying a human error probability (HEP) for a particular human failure event (HFE). The numerical effects from that evaluation are represented by recovery factor R_e in Equation 3.4.

The "Guidance for Crediting Recovery Effect in P_c" in Section 3.5.1 emphasizes the scenario context of each HFE and the need for analysts to carefully account for that context when they assess the feasibility and evaluate the likelihood for recovery. The discussion in Item 4) of this section indicates that the numerical value that is assigned to factor R_e is "left open to the analysts' judgment". I disagree with that assertion. In particular, the analyses that are performed to evaluate potential recovery are the same as those needed to evaluate the initial (un-recovered) HEP. The numerical combination of recovery factor R_e and the initial HEP answers the question: "What HEP applies after considering the effects from possible recovery in the context of this scenario?" Thus, the assigned value of recovery factor R_e is not simply "left open to the analysts' judgment". It is derived from a systematic assessment of human performance in the context of the scenario-specific HFE. It is not, and should not be, just a number that is applied generically to reduce the HEP. Furthermore, the dependency methodology that is summarized in Appendix K of the current version of NUREG-2198 provides additional guidance for these assessments, because the evaluation of recovery factor R_e is functionally very similar to a dependency analysis.

I certainly agree with the conclusion that "it is premature to provide numeric recovery factors for a method that is intended for a broad range of HRA applications". However, I do not think that the guidance in this report, if it is applied as intended, merits the implication that the recovery factors are developed and applied arbitrarily. In fact, it is not apparent how operating experience or simulator testing data can be used to derive nominal recovery factors that appropriately account for the scenario-specific influences on human performance, consistently link those influences to estimates of the un-recovered and recovered HEPs, and account for how an analyst may structure the models and evaluations of those HEPs in a particular PRA. In summary, although there may be a desire to tabulate nominal numerical values for factor R_e, I think that derivation and technical justification of those factors would be extremely challenging.

As a final thought on this topic, it is not apparent why the concept of a separate recovery factor is needed, or is technically justified, in the IDHEAS methodology. For example, depending on the data that are used to derive the "base" HEPs and the performance-influencing factor (PIF) weights, it seems conceivable that an HEP which applies when all PIFs are set to their "no impact" state may already account for recovery. Thus, the analyst's scenario-specific evaluations of the relevant PIF attributes and their associated weights may effectively account for the numerical effects from most sources of recovery (e.g., procedural reminders, self-checking, supervisory oversight, recognition of unexpected system responses, etc.). A separate evaluation of factor R_e may inappropriately "double-account" for those effects.

Why is it asserted that the numerical value that is assigned to factor R_e is simply "left open to the analysts' judgment"?

49. Section 4.3, Future Development and Improvement

Item 5) in this section acknowledges that the IDHEAS-ECA application does not include guidance for the analysis of dependencies. This is an important omission, and it is a fundamental technical deficiency in the methodology.

Section 1.4 and Figure 1-2 in NUREG-2198 indicate that Stage 4 of the human reliability analysis process involves "Integrative Analysis", which includes the evaluation of dependencies. Appendix K in the current version of NUREG-2198 describes a proposed methodology to identify, evaluate, and quantify the effects from human dependence. I like the conceptual construct of that model. I understand why it is difficult for NUREG-2198 to provide examples of how the methodology is applied in practice. I think that the IDHEAS-ECA application should adopt that methodology, and the report should include some clear examples of how it is used in practice.

If the authors do not adopt the methodology that is described in Appendix K of NUREG-2198, it is essential that the final published version of this report should at least contain guidance for the analysis of dependencies that is consistent with the current state-of-practice (e.g., the process that is summarized in Section 6.2 and Figure 6-1 of NUREG-1921).

Will the final IDHEAS-ECA report include the dependency methodology that is summarized in Appendix K of the current version of NUREG-2198, with examples that show how the methodology is used in practice?

If not, why not?

If that methodology will not be adopted, what specific dependency analysis methodology will be used for the IDHEAS-ECA application?

50. Section 4.3, Future Development and Improvement

Item 6) in this section discusses testing and validation of the IDHEAS-ECA methodology. It is noted that:

> "In 2019, the NRC held a workshop in which six HRA analysts used the IDHEAS-ECA software to calculate HEPs of the HFEs in implementing FLEX strategies. The analysts were not required to fill out the IDHEAS-ECA Worksheets and they directly started from the software. Thus, the analysts essentially performed Steps 4, 5, 6, and 7 of the IDHEAS-ECA process without performing Steps 1, 2, and 3."

Steps 1, 2, and 3 of the methodology involve the qualitative analyses of the scenario context and timing, definition of the human failure events (HFEs) and their associated critical tasks, and identification of the cognitive failure modes (CFMs) that apply for each critical task. The scenario context also determines which performance-influencing factors (PIFs) apply for each HFE, and it affects how analysts evaluate specific PIF attributes for each CFM. Thus, Steps 1, 2, and 3 are the fundamental elements of the methodology.

The experience from practical analyses and benchmark studies has shown that comprehensive and systematic qualitative analyses are essential for realism and fidelity in the human reliability analysis (HRA) process. Benchmark studies have also shown that differences in the qualitative analyses are an important source of analyst-to-analyst variability when any HRA methodology is used. In fact, deficiencies in most contemporary guidance for the performance of those qualitative analyses were one of the primary motivations for development of the general IDHEAS methodology.

An exercise that requires analysts to perform only Steps 4, 5, 6, and 7 is neither verification nor validation of the IDHEAS-ECA methodology. It simply confirms that people can use the tables in Appendix B and the IDHEAS-ECA software to calculate a numerical value for a human error probability (HEP). That exercise may provide useful feedback on whether the software interface is user-friendly. However, the tables and the software simply assist the analysts to compute a number after the fundamental elements of the analysis are completed. The calculator is ***not*** the methodology. Thus, the cited exercise seems to have very marginal relevance for a critical examination of whether the IDHEAS-ECA methodology and guidance have achieved the goals of (1) supporting a realistic scenario-specific evaluation of human performance and (2) reducing analyst-to-analyst variability when the entire HRA process is implemented.

Why is an exercise that simply uses the IDHEAS-ECA software to calculate a numerical value for an HEP characterized as "testing and validating" the methods and guidance in this report?

Has the NRC staff conducted any exercises that require several experienced independent analysts (i.e., people who were not directly associated with the input, authorship, or reviews of this report or the IDHEAS-G methodology) to use the entire methodology (i.e., all 8 steps) to define and evaluate one or more HFEs for a particular event scenario?

If not, when are those comprehensive tests planned?

In particular, will they be conducted before the final version of this report is issued (e.g., as a NUREG)?

51. Appendix A, Worksheet A. Scenario Analysis, Section A.4. Initial Assessment of PIFs

I like the fact that Section A.4 indicates that the analysts should explicitly list the performance-influencing factors (PIFs) that are judged to be unimportant for the particular human failure event (HFE). That reinforces the expectation that they should think about the potential effects from each PIF when they analyze the scenario context. However, the guidance in Section 3.1.3 of the main report does not instruct the analysts to document why each identified PIF is not important, and the guidance for Section A.4 in the worksheet simply instructs the analysts to list the PIFs.

The first item in the "Guidance for Assessing PIF Attributes" in Section 3.4 of the main report notes that:

> "Based on the scenario context, many PIFs may not be relevant; therefore, they are not selected. If a PIF is not selected, a ***rationale should be given*** for why it is not relevant." [emphasis added]

I certainly agree with this guidance. However, it applies to Step 4 of the analysis process and the documentation in Section C.3 of Worksheet C. In other words, it applies only to the PIFs that are retained for further analysis from Section A.4 of Worksheet A. That intent is explicitly confirmed by the summary description in Section C.3 of Worksheet C.

I think that it is important for analysts to document the reasons for their decisions to omit specific PIFs during the initial evaluations in Worksheet A, even if the justifications are relatively brief. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

Why do the guidance in Section 3.1.3 of the main report and the guidance for Section A.4 in Worksheet A not indicate that the analysts should document the reasons why each omitted PIF is not important for the HFE?

52. Appendix A, Worksheet C. Analysis of Critical Tasks in an HFE, Section C.1. Analysis of Cognitive Activities and Identification of Applicable CFMs

Section C.1 indicates that the analysts should simply list the cognitive failure modes (CFMs) that apply for the particular critical task. The guidance in Section 3.3.2 of the main report does not instruct the analysts to document why the other possible CFMs are not important, and Section C.1 in the worksheet does not include that documentation.

I think that it is important for analysts to document the reasons for their decisions to omit specific CFMs, even if the justifications are relatively brief. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

Why do the guidance in Section 3.3.2 of the main report and the guidance for Section C.1 in the worksheet not indicate that the analysts should document the reasons why each omitted CFM is not important for the critical task?

53. Appendix B, Base Human Error Probabilities and Performance-Influencing Factor Weights, General Comments

I checked for consistency between the scope and definitions of the tabulated performance-influencing factor (PIF) attributes in the tables in this appendix, and the attributes that are listed in Table 3-1 through Table 3-20 in NUREG-2198. Except for the consolidations that are noted for Table B-4 and Table B-14, and specific details that are noted in the next comment, there is generally good agreement between the two reports.

I intentionally did not comment on details of specific PIF weights or the corresponding human error probabilities (HEPs) in these tables. I also did not comment on the recommended minimum HEPs of 1E-04 for Detection or Action Execution, and 1E-03 for Understanding, Decsionmaking, or Interteam Coordination. I obtained a draft copy of RIL-2021-XX, "Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA)", that was released publicly to support a September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment. The IDHEAS-DATA report evidently contains the information and analyses that justify these correlations and the associated HEPs. I will comment separately on that report.

54. Appendix B, Base Human Error Probabilities and Performance-Influencing Factor Weights

I also checked for consistency of the identified relationships between specific performance-influencing factor (PIF) attributes and the high-level cognitive failure modes (CFMs) in the tables in this appendix, compared to the links that are listed in Table B-1 through Table B-20 in Appendix B of NUREG-2198. I have the following observations from that comparison.

The bold table numbers correspond to the tables in Appendix B of this report. The following abbreviations apply: D = Detection, U = Understanding, DM = Decsionmaking, E = Action Execution, T = Interteam Coordination.

Table B-1, Scenario Familiarity

- The attributes that are listed in this table do not seem to include or correspond to the attributes for familiarity with system failure modes and worksite familiarity that are listed in Table B-15 in NUREG-2198.

- SF1, SF2, and SF3 are linked to D, U, DM, and E. Table B-15 in NUREG-2198 does not associate any Scenario Familiarity attribute with D.

- SF1, SF2, and SF3 are linked to E. Table B-15 in NUREG-2198 associates only worksite familiarity with E.

- SF4 is linked to U and DM. Table B-15 in NUREG-2198 associates a bias or preference for wrong strategies only with DM.

Table B-2, Information Availability and Reliability

- At a high level, the attributes that are listed in this table pertain almost entirely to completeness (Inf1) and reliability (Inf2) of the available information. However, the detailed summaries seem to encompass the 11 more specific attributes that are listed in Table B-14 in NUREG-2198.

- Inf1 and Inf2 are linked to only U and DM. Three of the attributes listed in Table B-14 in NUREG-2198 are linked to D, three are linked to E, and ten are linked to T.

Table B-3, Task Complexity

- I do not know why the entry for C38 explicitly lists the action to "transport fuel assemblies with fuel machines".

- C41, C42, C43, and C44 are linked to T. Table B-17 in NUREG-2198 does not associate any Task Complexity attribute with T.

Table B-4, Environmental PIFs

- ENV7 is the only listed attribute for noise. It is linked to D, U, DM, E, and T. It is not apparent how the ENV7 conditions are related to the four detailed attributes that are listed in Table B-3 in NUREG-2198. Table B-3 in NUREG-2198 links the noise attributes to only D and E.

- ENV8, ENV9, ENV10, ENV11, ENV12, and ENV13 are apparently intended to include the attributes in Table B-1 for "Workplace Accessibility and Habitability" and Table B-5 for "Resistance to Physical Movement" in NUREG-2198. It is not apparent that the six listed attributes include all of the conditions that are listed in Table B-1 in NUREG-2198. In particular, it is not apparent how they address considerations such as the need for keys, other environmental factors that affect workplace habitability (e.g., radiation), and other local effects where the work must be performed (e.g., hot surfaces, flooding). Table B-4 links these six environmental attributes to only E. Table B-1 in NUREG-2198 links the attributes for "Workplace Accessibility and Habitability" to D and E.

Table B-5, System and I&C Transparency

- SIC1 and SIC2 are linked to D, U, DM, and E. Table B-6 in NUREG-2198 does not associate these attributes with DM.

- SIC3 and SIC4 are linked to only U. Table B-6 in NUREG-2198 associates these attributes with D, U, and E.

Table B-6, Human-System Interface

- HSI3 is linked to D and U. Table B-7 in NUREG-2198 associates this attribute with only D.

- HSI7 is linked to only U. It is not apparent how this attribute is related to any of the attributes that are listed in Table B-7 in NUREG-2198. Furthermore, none of the attributes in Table B-7 in NUREG-2198 are associated with U.

- It is not apparent how attributes HSI8 and HSI9 are related to any of the attributes that are listed in Table B-7 in NUREG-2198.

Table B-8, Staffing

- STA3 and STA4 are linked to D, U, DM, E, and T. Table B-9 in NUREG-2198 does not associate these attributes with T (i.e., they are not listed as "Overarching").

- STA5 is linked to D, U, DM, E, and T. Table B-9 in NUREG-2198 associates this attribute with only DM.

Table B-9, Procedures, Guidance, and Instructions

- PG5 is linked to D, U, DM, E, and T. Table B-10 in NUREG-2198 does not associate this attribute with T (i.e., it is not listed as "Overarching").

- It is not apparent how attributes PG6 and PG7 are related to any of the attributes that are listed in Table B-10 in NUREG-2198. PG6 is linked to D, E, and T. PG7 is linked to U, DM, and T. None of the attributes in Table B-10 in NUREG-2198 have either of these specific sets of linkages.

- The attributes that are listed in this table do not seem to include or correspond to the attributes for availability of PGI and misleading PGI that are listed in Table B-10 in NUREG-2198 (i.e., no PGI for skill-based tasks, no PGI available, PGI are available but are misleading).

Table B-10, Training

- TE3 is linked to D, U, DM, and E. Table B-11 in NUREG-2198 associates this attribute with only U and DM.

- It is not apparent how attributes TE4, TE5, TE6, and TE11 are related to any of the attributes that are listed in Table B-11 in NUREG-2198.

- TE7 is linked to only D. Table B-11 in NUREG-2198 associates this attribute with U and DM.

- TE8 is linked to only D. Table B-11 in NUREG-2198 associates this attribute with U.

- TE9 is linked to U and DM. Table B-11 in NUREG-2198 associates this attribute with only U.

- The attributes that are listed in this table do not seem to include or correspond to the attributes for complete lack of training that are listed in Table B-11 in NUREG-2198 (e.g., no training on the scenario-specific procedures, guidelines, or actions).

Table B-11, Teamwork and Organizational Factors

- TF2 seems to include several of the attributes that are listed in Table B-12 in NUREG-2198. However, the summary descriptions of the elements of TF2 do not correspond very well with the descriptions of the attributes in Table B-12 in NUREG-2198. Therefore, it is difficult to understand how TF2 relates to specific attributes in NUREG-2198.

- TF3 is linked to only T. Table B-12 in NUREG-2198 associates this attribute with D, U, DM, E, and T.

- It is not apparent how attribute TF5 is related to any of the attributes that are listed in Table B-12 in NUREG-2198.

- The attributes that are listed in this table do not seem to include or correspond to the attributes for organizational practices (or "safety culture") that are listed in Table B-12 in NUREG-2198 (i.e., job briefing, safety issue monitoring, reporting, and corrective actions).

Table B-12, Work Processes

- WP4, WP5, and WP6 seem to correspond to the attributes for organizational practices (or "safety culture") that are listed for "Team and Organization Factors" in Table B-12 in NUREG-2198 (i.e., job briefing, safety issue monitoring, reporting, and corrective actions). Use of the term "instrumentation" in the descriptions of these attributes is also very confusing.

- It is not apparent how attribute WP7 is related to any of the attributes that are listed in Table B-13 in NUREG-2198. This attribute also seems more relevant to the organizational practices (or "safety culture") that are listed in Table B-12 in NUREG-2198.

- The attributes that are listed in this table do not seem to include or correspond to the attributes for attention to task goals and reviewing the status of event progression that are listed in Table B-13 in NUREG-2198.

Table B-13, Multitasking, Interruption, and Distraction

- MT1 and MT2 are linked to D, U, DM, E, and T. Table B-16 in NUREG-2198 associates these attributes with only D, U, and E.

Table B-14, Mental Fatigue and Time Pressure and Stress

- MF2 is linked to D, U, DM, E, and T. Table B-19 in NUREG-2198 does not associate this attribute with T.

- It is not apparent how attributes MF3, MF6, and MF7 are related to any of the attributes that are listed in Table B-18 or Table B-19 in NUREG-2198.

Table B-15, Physical Demands

- The attributes that are listed in this table do not seem to include or correspond to the attributes for physical resistance to movement and concern for personnel safety that are listed in Table B-20 in NUREG-2198.

JWS Note: The links in NUREG-2198 are derived from the authors' reviews and interpretation of the supporting literature for human cognitive performance. When I read NUREG-2198, I thought that the relationships in Appendix B of that report were intended primarily for a forensic assessment of an analysis and to identify possible ways that the evaluated human performance could be improved. Based on Appendix B in this report, I now know that those relationships are an integral element of the human error probability (HEP) quantification process. Therefore, it is important that the IDHEAS-ECA application identifies and evaluates these links consistently with the general IDHEAS methodology. The following examples illustrate why these relationships are important.

- A critical task for human failure event HFE1 involves the CFM for Understanding (U). The scenario conditions for HFE1 adversely affect Attribute N for PIF X. The table for PIF X incorrectly omits a link between Attribute N and U. Thus, the HEP for HFE1 will not correctly account for the scenario effects on PIF X.

- A critical task for human failure event HFE2 involves the CFM for Decisionmaking (DM). The scenario conditions for HFE2 adversely affect Attribute M for PIF Y. The table for PIF Y incorrectly includes a link between Attribute M and DM. Thus, the HEP for HFE2 will be increased inappropriately by the scenario effects on PIF Y.

55. Appendix C, Examples, General Comment

I intentionally did not comment on any elements of these examples that involve conclusions which are based on analyst judgment. Although I might disagree with some specific conclusions, my judgment is neither more relevant nor necessarily better justified technically than the analysts'. Therefore, my comments address situations that involve analyst judgment only when I could not understand the rationale for a particular decision, or when the judgment does not seem to be consistent with the IDHEAS-ECA guidance.

56. Appendix C, Section C.1, Example 1: A Simple Steam Generator Tube Rupture (SGTR) Event, General Comment

The description of this scenario and the functional success criteria for human failure event HFE1 are based on the information that is provided for Scenario 3 in NUREG-2156. I did not study the analyses of that scenario in NUREG-2156 as background for these comments. I occasionally referred to that report when I had questions about information that seems to be lacking from the scenario description, timeline, functional basis for the available time window, time required to perform the action, etc. I tried to avoid comments on issues that are related primarily to the lack of source information from NUREG-2156.

57. Appendix C, Section C.1, Example 1, Worksheet A, Section A.1, Boundary Conditions

The list of boundary conditions includes the following items:

- "This is a standard scenario that crews are trained on frequently"
- "Adequate staffing"
- "Procedures are available and are well trained on"

These items seem to involve assumptions or conclusions about the specific performance-influencing factors (PIFs) for "Training", "Staffing", and "Procedures, Guidance, and Instructions". In particular, they seem to imply that the analysts have already concluded that the "no impact" attribute applies for these PIFs for any human failure event (HFE) that may be evaluated during this scenario.

The IDHEAS-ECA guidance indicates that the narrative should objectively describe the evolution of the scenario from an operational perspective. I thought that the scenario "boundary conditions" in that narrative are supposed to clarify important information about physical conditions in the plant and specific assumptions or constraints that may be introduced by how the scenario is modeled in the PRA. I do not think that it is appropriate for the scenario "boundary conditions" to include analyst assumptions about the relevance of specific PIFs or the status of specific PIF attributes, without an actual evaluation of those PIFs, including justification for the analysts' conclusions. Those PIF evaluations are an integral part of the human reliability analysis (HRA) process for each HFE. They are not "boundary conditions" that are intrinsic to the physical scenario progression or the PRA logic model structure for the scenario.

For example, it might be reasonable to eventually conclude that the "no impact" attribute applies for these particular PIFs for every cognitive failure mode (CFM), every critical task, and every HFE that is evaluated in this scenario. However, that conclusion should be justified by a systematic evaluation of each HFE, critical task, CFM, and PIF according to the guidance in this report. It should not be a "boundary condition" for the scenario narrative.

Why do the boundary conditions for this scenario include premature assumptions about the effects from specific PIFs for training, staffing, and procedures?

58. Appendix C, Section C.1, Example 1, Worksheet A, Section A.3, Environment and Situation Context

The second item in this list is:

- "No additional complications beside the initiating event"

This item does not seem consistent with the considerations that are discussed in the "Guidance for Assessing the Environment and Situation Context" in Section 3.1.3 of the main report. In particular, this item seems to address the System Context.

Considering the guidance in Section 3.1.3, why is this item relevant to the Environment and Situation Context?

59. Appendix C, Section C.1, Example 1, Worksheet A, Section A.3, Personnel Context

The first item in this list is:

- "The human-machine interface in the control room is ***well designed***" [emphasis added]

The last item in the list is:

- "The Symptom based EOPs are in place and ***well designed*** for the scenario under evaluation; operators ***recognize*** the event and ***are familiar*** with applicable procedures and actions." [emphasis added]

These items seem to involve assumptions or conclusions about the specific performance-influencing factors (PIFs) for "Human-System Interface" and "Procedures, Guidance, and Instructions". In particular, they seem to imply that the analysts have already concluded that the "no impact" attribute applies for these PIFs for any human failure event (HFE) that may be evaluated during this scenario.

A preceding comment addresses my concerns about premature assumptions or conclusions about the relevance of specific PIFs or the status of specific PIF attributes, without an actual evaluation of those PIFs for each HFE, including justification for the analysts' conclusions.

Furthermore, these items do not seem consistent with the considerations that are discussed in the "Guidance for Assessing the Personnel Context" in Section 3.1.3 of the main report. In particular, these items seem to address the Task Context.

JWS Note: This comment also applies to similar assessments of specific PIFs under "System Context" (e.g., a "well-designed" human-machine interface) and "Personnel Context" in Section A.3 for Example 2. It also applies to similar PIF assessments for "System Context" and "Personnel Context" (e.g., for procedures, training, and interteam coordination) in Section A.3 for Example 3.

Why do these items include premature assumptions about the effects from specific PIFs for the human-system interface and procedures?

Considering the guidance in Section 3.1.3, why are these items relevant to the Personnel Context?

60. Appendix C, Section C.1, Example 1, Worksheet A, Section A.3, Task Context

This entry lists the following items:

- "The scenario is a ***well-practiced*** classic event, covered in training, and practiced in simulator requalification exercises, such that crews are ***expected to know*** the alarm pattern of an SGTR event." [emphasis added]
- "Crew responses are ***clearly specified*** in EOP steps and ***frequently practiced*** in training scenarios" [emphasis added]

- "Control panel indications needed for diagnosis are ***simple and easily found***." [emphasis added]
- "Parameters and trends are ***easily available*** and no calculations or trend annotation or memorization is needed" [emphasis added]
- "No simultaneous event occurs, thus the crews do not need to perform parallel multiple tasks and distractions and interruptions of the crews are ***expected to be nominal***." [emphasis added]

For some brevity in these comments, I will not discuss each highlighted item. They further illustrate my concerns about premature assumptions or conclusions about the relevance of specific performance-influencing factor (PIFs) or the status of specific PIF attributes, without an actual evaluation of those PIFs, including justification for the analysts' conclusions.

Furthermore, these items do not seem consistent with the considerations that are discussed in the "Guidance for Assessing the Task Context" in Section 3.1.3 of the main report. In particular, they do not document the types of objective factual information that is needed for an analyst to evaluate the relevant PIFs for each critical task in the context of this scenario.

JWS Note: This comment also applies to similar assessments of specific PIFs under "Task Context" in Section A.3 for Example 2 and in Section A.3 for Example 3.

Why do these items include premature assumptions about the effects from specific PIFs for scenario familiarity, information availability, task complexity, training, procedures, human-system interface, and multi-tasking?

Why do these items not document the types of objective factual information that is summarized in the guidance for Task Context in Section 3.1.3?

61. Appendix C, Section C.1, Example 1, Worksheet A, Section A.4, Initial Assessment of PIFs

This section lists several performance-influencing factors (PIFs) that are judged to have no adverse effect ("no impact") on human failure event HFE1. Preceding comments on the main report and Appendix A address the need for analysts to document their rationale for these types of conclusions. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

JWS Note: My comments on worksheet Section A.1 and Section A.3 for this example analysis address my concern that analysts' premature judgments about the effects of specific PIFs do not provide an objective representation of the physical scenario progression or the context for HFE1. The assessments in Section A.4 are intended to systematically examine the objective information about the scenario and justify why only a subset of the PIFs are potentially important for the evaluation of HFE1.

JWS Note: This comment also applies to the lack of a documented rationale in Section A.4 for Example 2 and in Section A.4 for Example 3.

Why does this section not document the analysts' rationale for their conclusions that each PIF has no impact on HFE1 (i.e., based on assessments of the objective information about the scenario)?

62. Appendix C, Section C.1, Example 1, Worksheet B, Section B.1, HFE Definition; Worksheet C, Section C.2, Task Characterization; Worksheet E, Time Available

I am very confused about the relationship and consistency between the functional success criteria for HFE1 and the estimated time that is available to complete the actions to isolate the ruptured steam generator.

The event timeline in Worksheet A indicates that the operators typically isolate the ruptured steam generator about 20 to 30 minutes after the initiating event occurs. They then cool down the reactor coolant system (RCS) after the steam generator is isolated.

The Success Criterion for HFE1 in Section B.1 of Worksheet B is:

> "***Isolate the ruptured SG and control pressure below the SG PORV setpoint*** to prevent radiation leakage to the environment" [emphasis added]

These success criteria seem to imply that the actions to isolate the ruptured steam generator must be coordinated with the actions to initiate an RCS cooldown. They also imply that those actions must be completed before pressure in the ruptured steam generator reaches the lowest setpoint to open a power-operated relief valve (PORV) on that steam generator. Section B.1 indicates that this action "ends at completing the RCS cooldown and depressurization". Thus, it seems apparent that HFE1 involves actions to isolate the ruptured steam generator and start a cooldown before a PORV is challenged to open, and it includes actions to control the cooldown and depressurization. It is not apparent what specific RCS conditions determine successful completion of these actions (e.g., pressure maintained below the PORV opening setpoint, the temperature and pressure that permit startup of residual heat removal cooling, or some other end state that is defined in the PRA model).

The discussion of Time Available in Section B.1 indicates that HFE1 involves:

> "A time-critical action as the crew needs to ***isolate the ruptured SG before SG PORV opening***. The time available is ***2-3 hours***." [emphasis added]

The discussion of Special Requirements in Section C.2 of Worksheet C notes that:

> "The task needs to be performed ***before reaching the SG PROV*** [sic] ***setpoints***." [emphasis added]

Worksheet E indicates that the time available (T_{avail}) to complete HFE1 is:

> "2 - 3 hours in PRA models"

These descriptions and estimates do not seem to be internally consistent, or consistent with the thermal-hydraulic evolution of a steam generator tube rupture scenario. In particular, it is evident that the 2- to 3-hour time window that is listed in Worksheet B is used for the evaluation of P_t in Worksheet E. However, that time window does not seem to be consistent with the functional success criteria to isolate the steam generator before a PORV is challenged to open. Depending on the plant-specific design and thermal-hydraulic analyses for a 500-gpm tube rupture event, if the operators do not intervene, I would expect pressure in the ruptured steam generator to reach the PORV opening setpoint within about 20 to 60 minutes. Therefore, it seems that either the defined success criteria for HFE1 are not functionally accurate, or the 2- to 3-hour estimate for T_{avail} is excessively long.

Depending on the specific desired end state, a longer available time window may apply for the control actions that are needed to complete the cooldown and depressurization. However, if the success criteria for HFE1 require that the operators must isolate the steam generator and start the cooldown before a PORV is challenged to open, the time at which pressure reaches the PORV opening setpoint determines the most limiting value of T_{avail} for those actions, and it should be used for this analysis of HFE1. That time window is typically much shorter than 2 to 3 hours.

JWS Note: The discussion of HFE 3A in NUREG-2156 refers to the 2- to 3-hour time window. I do not know how, or whether, that estimate may have been used for the analyses in that report. (I did not study them for these comments.) However, the IDHEAS-ECA methodology explicitly accounts for that time in the evaluation of P_t. Explicit use of a time uncertainty analysis to quantify a contribution to the overall human error probability (HEP) is a rather new (and, for some analysts, a somewhat controversial) element of the IDHEAS-G methodology. Therefore, as an example of how the IDHEAS-ECA methodology is applied for an actual analysis, it is ***essential*** that the estimate for T_{avail} must be consistent with the functional success criteria for HFE1 and the actual tube rupture scenario thermal-hydraulic progression. Otherwise, readers may simply dismiss this example as a marginally-relevant academic exercise that does not fully demonstrate how the methodology is used for a realistic analysis.

How is the 2- to 3-hour estimate for T_{avail} in Worksheet B and Worksheet E consistent with the definition and functional success criteria for HFE1 in Worksheet B and Worksheet C?

In particular, do the success criteria for HFE1 require that the operators must isolate the ruptured steam generator and start the cooldown before a PORV is challenged to open?

If so, if the operators take no actions, what is the technical basis for the estimate that a steam generator PORV will not be challenged to open until 2 to 3 hours after the initiating event occurs?

In particular, if a 500-gpm tube rupture occurs at the nominal plant that is used for this example analysis, and the operators take no actions, when will the PORV with lowest pressure setpoint open on the ruptured steam generator?

63. Appendix C, Section C.1, Example 1, Worksheet B, Section B.2, Task Diagram and Identification of Critical Tasks

This section discusses operator actions to trip the reactor manually, which are specified in EOP-0. The analysts conclude that those actions are not critical to the success of HFE1, because "the system will automatically trip the reactor if the crew fails to manually trip it".

This seems to be an example of a "boundary condition" that applies for the scenario that involves HFE1. In particular, the Event Timeline in Section A.1 of Worksheet A indicates that an automatic reactor trip occurs at 4 minutes after the start of the scenario. Most PRA models explicitly evaluate the conditional probability that the reactor fails to trip automatically, resulting in a consequential ATWS scenario. The composite effects on human performance from coincident SGTR and ATWS conditions are certainly important for the evaluation of actions to isolate the ruptured steam generator and start a cooldown. Thus, it is evident that HFE1 applies for a scenario when the automatic reactor trip is successful. The discussion about operator actions to trip the reactor manually does not seem relevant to this scenario.

Of course, the evaluation of HFE1 must account for the time that is spent in EOP-0, but that is a different issue.

Why does this identification of critical tasks discuss operator actions to trip the reactor manually, including conclusions about their potential effects on HFE1?

64. Appendix C, Section C.1, Example 1, Worksheet C, Section C.2, Task Characterization

The information for Personnel in this section notes that:

> "***Adequate well-trained*** crew" [emphasis added]

The information for Procedure notes that:

> "EOP-0 and EOP-3 have been implemented in ***simulator training***. The procedures have been ***optimized*** based on training feedback. It is ***expected*** that the crew in the scenario ***will follow the procedures without any notable reason to deviate*** from the procedure instructions." [emphasis added]

The information for Multitasking notes that:

> "Parallel tasks are distributed by the two ROs handling the activities related to the primary and secondary systems, thus individual crew members are ***not impacted by multitasking***." [emphasis added]

These entries seem to involve assumptions or conclusions about the specific performance-influencing factors (PIFs) for "Staffing", "Training", "Procedures, Guidance, and Instructions", and "Multitasking, Interruption, and Distraction". In particular, they seem to imply that the analysts have already concluded that the "no impact" attribute applies for these PIFs for each of the cognitive failure modes (CFMs) in the critical task for HFE1.

The intent of Section C.2 in this worksheet is to describe factual information about the task that is then used in the PIF assessments that are summarized in Section C.3. In particular, the "Guidance for the Characterization of Critical Tasks" in Section 3.3.1 and Table 3-2 of the main report indicates that the information about Personnel should describe who performs the critical task and any special skills that are needed. The information about Procedures should list the specific procedures and other guidance that are intended to be used for the task. The information about Competing Goals and Multitasking should describe any other tasks that the personnel are required, or expected, to perform during the time that the critical task is needed, based on the scenario-specific context of everything that is occurring in the plant.

JWS Note: This comment also applies to similar assessments of specific PIFs in the information that is summarized in Section C.2 for HFE1-T1 and HFE1-T2 in Example 2.

Why does this entry include premature assumptions about the effects from specific PIFs for staffing, training, procedures, and multitasking?

Why does this entry not document the types of objective factual information that is summarized in the guidance for Characterization of a Critical Task in Section 3.3.1?

65. Appendix C, Section C.1, Example 1, Worksheet C, Section C.3, Assessment of PIFs, Editorial Comment

The assessment of "Task Complexity" for T1-CFM2 indicates that attribute C30 applies. Table B-3 in Appendix B indicates that attribute C30 applies to the cognitive failure mode (CFM) for Action Execution. Attribute C10 is the "no impact" condition for the Understanding CFM.

66. Appendix C, Section C.1, Example 1, Worksheet D, HEP Estimation; Worksheet E, Time Uncertainty Analysis of the HFEs

Worksheet D and Worksheet E do not show how the respective human error probabilities (HEPs) for P_c and P_t are calculated. Furthermore, the example does not show how P_c and P_t are combined to quantify the overall HEP for HFE1. Thus, this example does not fully demonstrate how the IDHEAS-ECA methodology and the guidance in this report are used to quantify an HEP.

JWS Note: The examples should not presume that analysts use the IDHEAS-ECA software and that the HEP calculations are "invisible" to the user. The examples should demonstrate how each HEP is calculated, so readers clearly understand the intent of the guidance and the quantification process.

Section C.3 of Worksheet C indicates that the "no impact" attributes apply for all performance-influencing factors (PIFs) that affect cognitive failure modes (CFMs) T1-CFM1 and T1-CFM2. It indicates that Task Complexity attribute C31 applies for T1-CFM4. Thus, it seems that P_c should be determined by the sum of the "no impact" HEP for T1-CFM1, plus the "no impact" HEP for T1-CFM2, plus the HEP for T1-CFM4 that applies for attribute C31.

The introduction to Appendix B indicates that the IDHEAS-ECA guidance recommends minimum HEPs of 1E-04 for Detection or Action Execution, and 1E-03 for Understanding, Decsionmaking, or Interteam Coordination. Table B-3 indicates that a "base" HEP of 1E-03 for Action Execution corresponds to Task Complexity attribute C31.

Based on these estimates, it seems that P_c should be:

$$P_c = HEP_{T1\text{-}CFM1} + HEP_{T1\text{-}CFM2} + HEP_{T1\text{-}CFM4} = 1E\text{-}04 + 1E\text{-}03 + 1E\text{-}03 = 2.1E\text{-}03$$

A very slightly lower value of approximately 2.099E-03 would be calculated if Boolean arithmetic is used to combine these three HEPs.

I think that it is very important for Worksheet D to describe how P_c is quantified so that analysts clearly understand how to treat the PIF attributes that have "no impact", how to use the tables in Appendix B, and how to combine the individual CFMs.

If T_{avail} is 2 hours, I confirmed that a normal distribution for T_{reqd} with a mean value of 18 minutes and a standard deviation of 5 minutes results in $P_t = 0$. However, Worksheet E should document that result (provided that the 2- to 3-hour estimate for T_{avail} actually applies for this scenario).

JWS Note: If T_{avail} is actually less than 38 minutes, P_t may be comparable to, or greater than

P_c. For example, if T_{avail} is 32 minutes, the value for P_t is slightly higher than my estimated value for P_c.

Is my estimate for P_c correct?

Why does Worksheet D not show how P_c is calculated?

Why does Worksheet E not show how P_t is calculated?

67. Appendix C, Section C.2, Example 2: Loss of Component Cooling Water and Reactor Coolant Pump Sealwater, General Comment

The description of this scenario and the functional success criteria for human failure event HFE1 are based on the information that is provided for Scenario 2 in NUREG-2156. I did not study the analyses of that scenario in NUREG-2156 as background for these comments. I occasionally referred to that report when I had questions about information that seems to be lacking from the scenario description, timeline, functional basis for the available time window, time required to perform the action, etc. I tried to avoid comments on issues that are related primarily to the lack of source information from NUREG-2156.

68. Appendix C, Section C.2, Example 2: Loss of Component Cooling Water and Reactor Coolant Pump Sealwater

The introduction to this section notes that:

> "The success criteria for crew responses are to trip the RCPs and to start the positive displacement pump (PDP) to provide RCP seal injection to prevent the seal water inlet temperature or the ***lower sealwater bearing temperature*** exceeds 230 degree F (to prevent an RCP seal failure)." [emphasis added]

Unless the reactor coolant pumps (RCPs) at the nominal plant for this example are very different from typical pumps, there is no physical parameter that corresponds to a "lower sealwater bearing temperature". In addition to the cited seal water inlet temperature, the RCP instrumentation typically monitors seal water return temperature. It also typically monitors temperatures of the motor bearings, the pump upper radial bearing, the pump lower radial bearing, and the thrust bearing. However, there is nothing that corresponds to a "lower sealwater bearing". That term is meaningless.

The text in Section 5.2 of NUREG-2156 uses the highlighted term. However, this example should not perpetuate that error. Of course, I do not know what specific indications are used to prompt these actions in the plant-specific procedures. The pump lower radial bearing is normally cooled by some of the seal injection water that flows down the pump shaft, past the bearing, and into the reactor coolant system. Thus, it is conceivable that the pump lower radial bearing temperature might be used as a criterion in the guidance to restore seal injection flow. However, since the analysis seems to focus on "seal temperature" as the relevant criterion, I suspect that the other temperature of interest is the seal water return temperature. It seems likely that the guidance instructs the operators to restore seal injection flow before the seal temperatures exceed 230 °F, due to concerns about thermal distortion of the seal assembly elastomer O-rings, rather than concerns about the lower radial bearing temperature. The introduction to this scenario and the narrative also indicate that the operators are apparently explicitly instructed to not restore seal injection flow if the seal temperatures are above 230 °F.

That guidance may be due to concerns about a possible thermal shock to the seal assemblies, or the lower radial bearing, when the relatively cold seal injection water is restored. In any case, the authors should confirm the actual parameter that is used in this guidance to improve plant operational credibility in the description of a rather extreme example scenario.

Of course, the temperatures of the pump bearings and motor bearings may be an important criterion in the guidance to trip the RCP when no component cooling water is available. However, that is a different issue than restoration of seal injection flow.

JWS Note: This comment also applies to the summary of the success criteria in the introduction to the Event Timeline in Section A.1 of Worksheet A, the Success Criteria in Section B.1 of Worksheet B for HFE1, and the Success Criteria in Section B.1 of Worksheet B for HFE2.

What is the actual second parameter (i.e., the physical component temperature) that is used for this criterion?

In particular, is the criterion based on seal water return temperature, pump lower radial bearing temperature, or some other plant-specific parameter?

69. Appendix C, Section C.2, Example 2: Loss of Component Cooling Water and Reactor Coolant Pump Sealwater

The introduction to this section notes that:

> "The criteria determine that operator actions are ***time critical***." [emphasis added]

A preceding comment addresses the use of the term "time critical" to characterize a human failure event (HFE). The analyses that are summarized in this example will determine whether the scenario-specific time constraints have a significant effect on personnel performance and the associated human error probabilities (HEPs). Therefore, it is premature to imply that these operator actions are "time critical", or that the evaluation of event timing is more important for this analysis than any other analysis.

Why does this summary contain this sentence?

Can it be simply removed from the report (i.e., rather than modified or elaborated further), without altering an analyst's basic understanding of the example?

70. Appendix C, Section C.2, Example 2, Worksheet A, Section A.1, Initiating Event

The summary appropriately identifies the loss of Distribution Panel 1201 as the relevant initiating event for this scenario. It also describes the personnel actions that are needed in response to that failure. Information about the expected crew response is certainly relevant to the scenario narrative and the integrated evaluation of personnel performance during the scenario. However, it is not pertinent to the description of the initiating event itself.

Why is this information about the expected crew response included in the identification of the initiating event for this scenario?

71. Appendix C, Section C.2, Example 2, Worksheet A, Section A.1, Event Timeline

The introductory paragraph to the timeline summarizes the success criteria for the two specific actions that are evaluated in this example. That information is not appropriate for this part of the scenario description.

The IDHEAS-ECA guidance indicates that the narrative should objectively describe the evolution of the scenario from an operational perspective. In particular, the scenario timeline should document all of the system responses and personnel actions that occur during the scenario, without focusing exclusively on the specific actions that are eventually analyzed. This comprehensive and objective summary of what is happening in the plant, and when it happens, is vital to the integrated context of the evolving scenario. A timeline that focuses on only information that is relevant to the specific actions of interest inappropriately biases analysts' understanding of the entire scenario context and their assessment of the modeled actions.

For example, the timeline indicates that Distribution Panel 1201 fails at time t = 2 minutes. It does not provide any information about the alarms and indications that result from that failure, or when they occur. This information is important, because the operators may initially focus on the unusual behavior of level in steam generator A, and their attention may remain focused on concerns about secondary heat removal for some time after the reactor trip occurs.

JWS Note: The discussion of "System Context" in Section A.3 of Worksheet A simply notes that "a lot of Train A indications were not available because of the DP-1201 failure".

An automatic reactor trip occurs at time t = 3 minutes. The timeline then focuses exclusively on system conditions and personnel responses that are directly associated with only the actions that are eventually evaluated in this analysis. In particular, the timeline does not identify any actions that are associated with personnel response to the initiating event (i.e., the manual actions that are summarized in the scenario introduction and perhaps others), or the times when those actions are performed. This information is potentially important to the example analysis. For example, if some of the initial responses continue after time t = 3 minutes, they may introduce distractions, conflicting priorities, or personnel constraints that affect performance of the specific modeled actions.

Why does the introduction to the timeline summarize the success criteria for the two specific actions that are evaluated in this example?

Why does the timeline focus exclusively on the system conditions and personnel responses that are directly related to only the modeled actions?

Why does the timeline not identify the alarms and indications that result from the initial power failure, and the times when they occur?

Why does the timeline not identify any actions that are associated with personnel response to the initiating event, or the times when those actions are performed?

Is it assumed that all of those actions are performed only during the 1-minute interval after Distribution Panel 1201 fails and before the reactor trip occurs?

If so, what is the basis for that assumption?

If some actions related to the initial power failure may persist after time t = 3 minutes, why do the timeline and the narrative not contain that information?

72. Appendix C, Section C.2, Example 2, Worksheet A, Section A.1, Event Timeline

JWS Note: This comment is related to the immediately preceding comment. However, it addresses a somewhat different concern.

The timeline lists the following actions between time t = 5 minutes and time t = 8 minutes.

"Start procedure ES-01"
"Detect no CCW or sealwater"
"Stop all RCPs"
"Start 0POP04-RC-0002 'Reactor Coolant Pump Off Normal' procedure"
"Start PDP"

These items are all related directly to the analysis of the specific actions that are evaluated in this example. The operators ***should*** perform these actions before reactor coolant pump (RCP) seal temperatures exceed 230 °F. However, they are not actually performed in the context of this scenario. They are the focus of this analysis, and they should not be listed explicitly in the scenario timeline.

The scenario timeline should identify the fact that RCP seal temperatures are expected to exceed 230 °F at some time between time t = 7 minutes and time t = 9 minutes. That is relevant objective technical information that supports the analysis.

Why does the timeline list these desired actions, which are the subject of this example analysis?

73. Appendix C, Section C.2, Example 2, Worksheet A, Section A.1, Event Timeline

The first cues to prompt the operator actions that are evaluated in this example occur at time t = 3 minutes, when the reactor trip occurs, component cooling water (CCW) is lost, and seal injection flow is lost. The timeline should also identify any other alarms and indications that may alert the operators to restore seal injection flow, and when they occur.

Are there any other alarms or indications that may alert the operators to restore seal injection flow?

If so, when do they occur?

74. Appendix C, Section C.2, Example 2, Worksheet A, Section A.2, Human Failure Event (HFE) Definition; Worksheet B for HFE1, Section B.1, HFE Definition; Worksheet B for HFE2, Section B.1, HFE Definition

Section A.2 of Worksheet A identifies the following human failure events (HFEs) for this analysis.

HFE1 Failure of the crew to trip the RCPs and to start the Positive Displacement Pump (PDP) to prevent RCP seal LOCA

HFE2 Crew starts the PDP after RCP temperature reaches 230 degrees F

Section B.1 of Worksheet B for HFE1 indicates that the success criteria for HFE1 are:

> "Trip the RCPs after the loss of CCW and start the PDP to provide seal injection before either sealwater inlet or lower sealwater bearing temperature exceeds 230 degrees F."

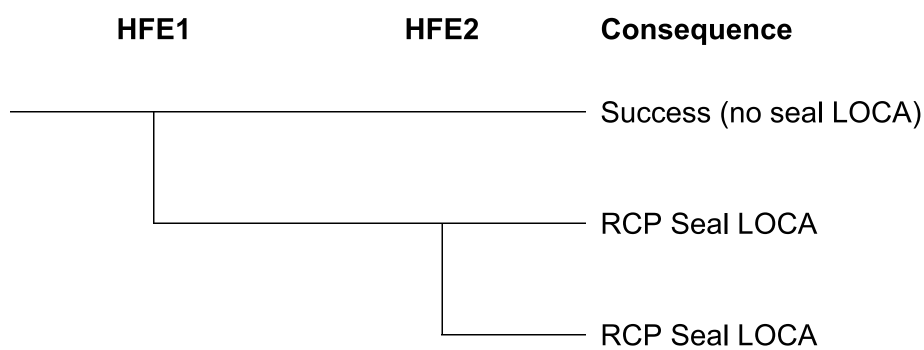
Section B.1 of Worksheet B for HFE2 indicates that the success criteria for HFE2 are:

> "NOT starting the PDP after either sealwater inlet or lower sealwater bearing temperatures exceeds 230 degrees F."

I understand the authors' desire to demonstrate the identification and evaluation of a possible "error of commission" (i.e., HFE2). However, I think that the examples in this report should demonstrate how a practical human reliability analysis is performed. In particular, I do not think that HFE2, as it is currently defined, is an appropriate example of how an "error of commission" should be defined, modeled, and evaluated in an actual PRA. In practice, an "error of commission" should be defined and evaluated only if performance of the identified actions changes the functional evolution of the scenario or affects possible subsequent human actions.

JWS Note: It is very difficult for many people to understand that the "***success***" state of an HFE applies when an action is ***not*** performed. It is much easier to understand HFE2 if the definition of that action answers the question "Do the operators start the PDP after seal temperatures exceed 230 °F?" If the answer is "Yes", the operators have made the "error of commission" and the "success" path from HFE2 applies. If the answer is "No", the operators do not start the PDP and the "failure" path from HFE2 applies. The following example and discussion use the current definition of HFE2. In other words, "***success***" of HFE2 means that the operators ***do not start*** the PDP after seal temperatures exceed 230 °F, and "***failure***" of HFE2 means that they ***do start*** the PDP.

Consider the following simple event tree for HFE1 and HFE2.



If the operators successfully trip the reactor coolant pumps (RCPs) and start the positive displacement charging pump (PDP) before seal temperatures exceed 230 °F (i.e., the success path from HFE1), no seal failures will occur. If the operators ***do not*** start the PDP ***before*** seal temperatures exceed 230 °F, and they ***do not*** start the PDP ***after*** seal temperatures exceed 230 °F (i.e., the success path from HFE2), no seal injection flow is available, and the seals will fail. If the operators ***do not*** start the PDP ***before*** seal temperatures exceed 230 °F, but they ***do start*** the PDP ***after*** seal temperatures exceed 230 °F (i.e., the failure path from HFE2), it seems that the seals will fail. In particular, if the actions that are modeled by HFE1 are functionally required to prevent failures of the seals, then it is irrelevant whether or not the operators start the PDP after seal temperatures exceed 230 °F. In practice, the only reason to identify, model, and evaluate HFE2 is if startup of the PDP after seal temperatures exceed 230 °F results in a functional change to the scenario or a change to the PRA models.

In principle, there may be many reasonable justifications for including HFE2 in the PRA models.

- It might be possible that seal failures can be prevented if seal injection flow is restored at some time after seal temperatures exceed 230 °F. In other words, the third sequence in the event tree might be Success. However, that situation would require changes to the functional definition of the "early" action to start the PDP in HFE1. It would also require more detailed thermal-hydraulic analyses of the RCP seals to determine the available time window for the "late" action in HFE2, and it would require a corresponding time uncertainty analysis to quantify P_t for that action.

- It might be possible that the seals will fail, regardless of whether or not the operators start the PDP at some time after seal temperatures exceed 230 °F. In other words, the event tree logic structure that is shown above applies. However, a thermal shock from the introduction of relatively cold seal injection water into the hot seals may accelerate their failure or change the conditional probability that a more severe LOCA occurs. For example, the seals may fail sooner or more severe seal damage may occur during the third sequence in the event tree, compared to the second sequence. (Contemporary PRA models often include a probabilistic evaluation for the timing and severity of RCP seal damage. More severe seal damage during the third sequence would affect the probabilities that are assigned for the consequential LOCA flow rates.) In that case, different success criteria for specific systems and personnel actions would apply during each sequence. The timing of some subsequent personnel actions may also be different. This situation would require more detailed thermal-hydraulic analyses of the RCP seals to determine how the introduction of PDP flow affects the timing or severity of seal damage.

- It might be possible that the seals will fail, regardless of whether or not the operators start the PDP at some time after seal temperatures exceed 230 °F. In other words, the event tree logic structure that is shown above applies. Introduction of PDP flow may not affect the timing or severity of the seal damage. However, subsequent functional progression of the scenario is affected by whether or not flow is available from the PDP. For example, the availability of makeup flow from the PDP might change the subsequent success criteria for some systems or personnel actions, or the scenario timing. In that case, different PRA event tree logic, fault tree logic, or basic event definitions would apply for the second sequence in the event tree when no PDP flow is available, compared to the third sequence when flow is available. That situation would also require supporting thermal-hydraulic analyses to determine when PDP flow must be started to affect the subsequent scenario progression, and it would require a corresponding time uncertainty analysis to quantify P_t for HFE2. (This consideration is conceptually similar to the preceding item, but it addresses potentially beneficial functional consequences from the availability of PDP flow after the "error of commission", rather than potentially adverse effects on the timing or severity of the seal LOCA.)

- It might be possible that the functional logic structure of the PRA event trees and fault trees is not affected by success or failure of HFE2. However, that "error of commission" introduces important effects on subsequent personnel cognitive performance. For example, it may be necessary to evaluate different performance-influencing factors (PIFs) or assign different PIF attributes for subsequent HFEX. The evaluation of HFEX1 would apply if the operators do not perform the "error of commission" (i.e., the second sequence in the event tree). The evaluation of HFEX2 would then account for specific PIFs that are affected adversely by the preceding "error of commission" (i.e., the third sequence in the event tree). Of course, this would require the definition of a new basic event for HFEX2, which is a change to the PRA logic models.

With this perspective, it might be reasonable to define and evaluate HFE2 in this example. However, the authors must provide a clear context and rationale for that decision, and they should explain how the PRA models should account for that action. Those considerations are very important elements of the integrated human reliability analysis process and the general IDHEAS methodology. Therefore, as a demonstration of the IDHEAS-ECA methods and guidance, the example should show how the entire analysis should be performed in practice (or at least describe how the PRA models should be changed to account for the effects from HFE2). The example should not simply focus on an exercise to quantify a nominal human error probability (HEP) for HFE2, without providing readers with a clear functional scenario context for that HFE and an understanding about how it will be used in the modified PRA model. The current information in Worksheet A and Worksheet B for HFE2 does not provide that context.

Considering these rather long comments, what is the functional reason for defining HFE2?

In particular, how does the action to start the PDP after seal temperatures exceed 230 °F change the scenario progression and the PRA models, compared to the scenario that applies if that action is not taken?

75. Appendix C, Section C.2, Example 2, Worksheet A, Section A.3, Scenario Context, System Context

The first bullet item in this section notes that:

> "The ***required*** equipment and instrumentation are ***available***" [emphasis added]

The third bullet item notes that:

> "The human-machine interface in the control room is well designed, but ***a lot of Train A indications were not available*** because of DP-1201 failure" [emphasis added]

JWS Note: A preceding comment addresses the premature conclusion that the human-machine interface is "well designed".

The "Guidance for Assessing the System Context" in Section 3.1.3 of the main report indicates that this section should objectively describe the physical status all equipment and indications that are relevant for analysts to understand the overall plant conditions when the modeled actions are needed. These bullet items do not seem to be consistent with that guidance. In particular, the first item contains an implied assumption or conclusion that the available equipment and instrumentation is fully adequate to support the desired actions. In other words, it seems to imply that the "no impact" attribute applies for the performance-influencing factors (PIFs) for "Information Availability and Reliability" and "System and I&C Transparency". It is not apparent why those conclusions are consistent with the observation that "a lot of Train A indications were not available".

Depending on the plant design (i.e., specific instrumentation AC and DC power supplies), it may also be possible that some additional relevant alarms and indications are disabled by the loss of power at Bus E1C, which coincides with the reactor trip at time t = 3 minutes.

What specific alarms and indications are disabled by the power failures at Distribution Panel 1201 and Bus E1C?

Do the de-energized failure modes for any of the affected instrumentation or indications introduce sources of confusion for understanding the plant status and diagnosing the need for the modeled actions (e.g., indications that fail high, fail low, fail as-is, etc.)?

Why does this section indicate that the required equipment and instrumentation are available, without a systematic evaluation of the PIF effects from the disabled alarms and indications (i.e., documented in Worksheet C)?

76. Appendix C, Section C.2, Example 2, Worksheet A, Section A.3, Scenario Context, System Context

This section notes that the operators must take manual control of several functions in response to the loss of power at Distribution Panel 1201 at time t = 2 minutes. That is certainly relevant information.

The introductory information for this analysis notes that the feedwater regulation valve for steam generator A remains fully open and cannot be operated manually. That failure results in high level in steam generator A, which then causes the automatic turbine trip and reactor trip at time t = 3 minutes. That valve failure is important for the scenario context. In particular, the operators may initially focus on the unusual behavior of level in steam generator A, and their attention may remain focused on concerns about secondary heat removal for some time after the reactor trip occurs.

The introductory information for this analysis also documents numerous failures that occur in coincidence with the reactor trip, or shortly thereafter. (The analysis apparently assumes that all failures occur at time t = 3 minutes.) All of those failures are relevant to the scenario context, and they should be summarized in this section of the worksheet. For example, it seems likely that the auxiliary feedwater (AFW) pump failure start will further focus the operators' attention on secondary heat removal and steam generator A in particular. The fault on Bus E1C and the mechanical failure of component cooling water (CCW) pump 1A, in combination with the unavailability of CCW pump 1B due to maintenance, disable all CCW flow. The fault on Bus E1C and the trip of charging pump 1B disable all reactor coolant pump (RCP) seal injection flow. These equipment failures are essential to an understanding of the reasons why the operators must trip the RCPs and restore seal injection flow from the positive displacement charging pump (PDP).

JWS Note: I do not know if the loss of power at Bus E1C affects another AFW pump. That is potentially important information for this scenario. However, it is unfortunately missing from the available plant technical information in NUREG-2156.

Why does this section not document the failure of the steam generator A feedwater regulation valve?

Why does this section not document the specific equipment failures that occur in coincidence with the reactor trip, or shortly thereafter?

77. Appendix C, Section C.2, Example 2, Worksheet A, Section A.3, Scenario Context, Personnel Context

This section provides only general information about the crew and their work process. Section A.1 indicates that the crew consists of a Shift Manager, Unit Supervisor, Shift Technical Advisor, and two Reactor Operators. The introductory information for this analysis and the summary of the System Context in this section note that the operators must take manual control of several functions in response to the loss of power at Distribution Panel 1201 at time t = 2 minutes.

This is a very complex scenario, and it evolves very quickly. It is very important for the analysts to understand who is performing what specific actions during this scenario. In particular, it is important to document what each Reactor Operator is controlling manually when the reactor trip occurs at time t = 3 minutes. Those activities do not necessarily cease completely when the trip occurs, and it may be necessary to continue some of them in parallel with the normal reactor trip responses. For example, the operator who normally has primary responsibility for monitoring the reactor coolant pumps (RCPs) may be engaged with other tasks that require attention for some time after t = 3 minutes. The other operator may be similarly occupied. Thus, it is important for analysts to understand how the available personnel are engaged as the scenario evolves, so the analysts can evaluate how the crew's preceding (and continuing) activities may affect their attention to the critical tasks for the modeled actions.

JWS Note: This information could be documented under the Task Context. However, since it pertains to the allocation of personnel resources, it is more relevant to the Personnel Context.

Why does this section not document who is performing what specific actions during this scenario (at least until time t = 3 minutes, and any actions which continue after that time)?

78. Appendix C, Section C.2, Example 2, Worksheet A, Section A.3, Scenario Context, Task Context; Worksheet B for HFE1, Section B.1, HFE Definition, Section B.2, Task Diagram and Identification of Critical Tasks, Critical Tasks; Worksheet D, HEP Estimation, HFE1; Worksheet E, Time Uncertainty Analysis of the HFEs

The summary of the Task Context in Section A.3 of Worksheet A notes that:

> "Time is critical for the required actions but the ***time available is only 7 to 9 minutes***, barely enough for the crews to perform all the needed actions." [emphasis added]

The summary of the Starting and Ending Point in Section B.1 of Worksheet B for HFE1 notes that:

> "***Starts at the failure of the DP-1201*** and ends at stopping the RCP and starting PDP." [emphasis added]

The summary of Time Available in Section B.1 of Worksheet B for HFE1 notes that:

> "***Time available for starting PDP is 7~9mins*** before the RCP temp exceeds 230 degrees F." [emphasis added]

The definition of Critical Task T2 in Section B.2 of Worksheet B for HFE1 notes that:

> "T2: Start the PDP before seal temp reaches 230 degrees F (in ***7~9 mins***)." [emphasis added]

The entry for Critical Task HFE1-T2 in Worksheet D notes that:

> "Start the PDP before seal temp reaches 230 degrees F (in ***7~9 mins***)." [emphasis added]

The entry for Time Available for HFE1-T2 in Worksheet E notes that:

> "The ***system time available is 7-9 minutes*** in the PRA model. However, operators need to perform HFE1-T1 first, so the time available for HFE1-T2 is the system time available subtracted by the time needed for HFE1-T1" [emphasis added]

JWS Note: Worksheet E does not document the quantification of P_t. That issue is addressed in a subsequent comment. However, it is evident that the intent is to use a total available time window of 7 to 9 minutes for the combined actions in HFE1-T1 and HFE1-T2. Worksheet E indicates that a nominal time of 1 minute is allocated for HFE1-T1. Therefore, it is evident that the analysts intend that an available time window of 6 to 8 minutes should be used to evaluate P_t for HFE1-T2.

I disagree with these estimates for the total amount of time that is available to perform the actions to trip the reactor coolant pumps (RCPs) and restore seal injection flow from the positive displacement charging pump (PDP).

A preceding comment addresses my concern that characterization of the available time window in this report does not seem to distinguish between the system time window (T_{SW}) and the time that is available to perform the desired action (T_{avail}). In particular, the IDHEAS-ECA guidance seems to assume that the delay time before the first relevant cue occurs (T_{delay}) is always negligible. That is not the case in this scenario. Furthermore, this example does not correctly characterize the nominal starting point (T_0) for the scenario timing analyses.

Section A.1 in Worksheet A documents the following timeline for this scenario.

t = 0 minutes	Nominal reference time
t = 2 minutes	Loss of power at Distribution Panel 1201
t = 3 minutes	Reactor trip, loss of component cooling water (CCW), loss of seal injection
t = 7 - 9 minutes	RCP seal temperatures reach 230 °F

According to the guidance in Section 5.3 of NUREG-2198 and Figure 5-2 of that report, the following times apply for the analysis of HFE1:

- Nothing happens prior to time t = 2 minutes. As noted correctly in Section A.1 of Worksheet A, the initiating event for this scenario is the loss of power at Distribution Panel 1201. Thus, the nominal T_0 starting point for T_{SW} is time t = 2 minutes.

- The loss of power at Distribution Panel 1201 does not affect CCW or RCP seal injection flow. Those functions are disabled by the equipment failures that coincide with the reactor trip at time t = 3 minutes. The loss of CCW and seal injection provides the first cue for the actions that are modeled by HFE1. Thus, T_{delay} for this scenario is 1 minute, and the time interval for T_{avail} begins at time t = 3 minutes.

- The scenario timeline indicates that RCP seal temperatures reach 230 °F at approximately time t = 7 to 9 minutes after the nominal reference time. That condition determines the end point for T_{SW}. Thus, the total available time window (T_{avail}) for the actions that are modeled by HFE1 is approximately 4 to 6 minutes (i.e., the interval from the first cue for those actions at t = 3 minutes until temperatures reach 230 °F at t = 7 - 9 minutes).

Explicit use of a time uncertainty analysis to quantify a contribution to the overall human error probability (HEP) is a rather new (and, for some analysts, a somewhat controversial) element of the IDHEAS-G methodology. Therefore, as an example of how the IDHEAS-ECA methodology is applied for an actual analysis, it is ***essential*** that the estimate for T_{avail} must be consistent with the functional success criteria for HFE1 and the IDHEAS-G guidance in NUREG-2198.

JWS Note: After I finished all of my comments on this example, I looked at Table 5-6 in Section 5.2.2 of NUREG-2156 to compare my estimated HEPs with the simulator performance of the four crews that were tested in that study. I then noticed that Table 5-6 indicates that RCP seal temperatures reached 230 °F between 7 minutes, 23 seconds, and 8 minutes, 30 seconds, after the loss of all CCW and seal injection. The timelines in NUREG-2156 Table 5-7 are consistent with those times. Therefore, it now seems evident that the Event Timeline that is shown in Section A.1 of Worksheet A is not consistent with the actual progression of this scenario, and that a total available time window (T_{avail}) of 7 to 9 minutes for the actions that are modeled by HFE1 may actually be correct (i.e., seal temperatures reach 230 °F at approximately time t = 10 to 12 minutes after time t = 0 in the Worksheet A timeline). However, my comments on this IDHEAS-ECA example are based on the supporting information for the analysis that is documented in this report, and they accurately account for that information. It is obviously essential that the scenario timeline that is used for any human reliability analysis must be consistent with the actual scenario progression.

Why does this analysis indicate that the total available time window (T_{avail}) for the actions that are modeled by HFE1 is 7 to 9 minutes, rather than 4 to 6 minutes?

Is the Event Timeline that is shown in Section A.1 of Worksheet A an accurate summary of the actual scenario progression?

79. Appendix C, Section C.2, Example 2, Worksheet A, Section A.3, Scenario Context

The following paragraph is provided at the end of this section.

> "Many things happening at the same time made it difficult to detect the priority items. Operators may experience multitasking, interruption, and distraction by the alarm cascade and required actions after the DP-1201 failure and the other concurrent failures especially related to SG water level control."

According to the guidance, Section A.3 should provide an objective factual summary of important information that determines the context of the scenario up until the time when the modeled actions should be performed. This paragraph contains analyst observations and judgments that are not appropriate for that objective summary. They are appropriate as supporting information to justify the performance-influencing factor (PIF) assessments that are documented in Worksheet C.

Why does Section A.3 of Worksheet A contain this paragraph?

80. Appendix C, Section C.2, Example 2, Worksheet B for HFE2, Section B.1, HFE Definition

The discussion of the Starting and Ending Point for this action notes that:

> "The action may begin as the crew enters ES-01 and recognizes that the seal flow is low; it ends as the crew starts the PDP after the seal temp exceeds 230 degrees F."

I disagree with the starting point that is cited in this summary. Human failure event HFE1 accounts for operator actions to start the positive displacement charging pump (PDP) before reactor coolant pump (RCP) seal temperatures reach 230 °F. If the operators do not start the PDP by that time, the analysis of HFE2 evaluates whether they will start the PDP after seal temperatures exceed 230 °F. Thus, the functional starting point for HFE2 is not when the crew enters ES-01. The time window for HFE2 begins when RCP seal temperatures reach 230 °F.

This summary indicates that the ending point for this "error of commission" occurs when the operators start the PDP. That is consistent with the way that the ending points are summarized in Section B.1 of Worksheet B for the other actions in these examples. However, I think that the functional constraint on the time window for this particular action is determined by the time when the RCP seals fail. In particular, the information for this analysis indicates that the operators are explicitly instructed to not start injection flow from the PDP after seal temperatures exceed 230 °F. That guidance is apparently due to concerns that a thermal shock from the introduction of relatively cold seal injection water into the hot seals may accelerate their failure or exacerbate the severity of seal damage. However, it seems that those concerns are not relevant after the seals fail.

JWS Note: In practice, the seals may fail progressively. Therefore, if injection flow is restored at some time after the onset of failure, but before the maximum damage is achieved, the thermal shock might make the damage worse. However, that fine distinction in the scenario timing is well beyond the fidelity of essentially all contemporary PRA models, which typically assign a nominal time at which seal damage occurs, with various probabilities for the extent of damage at that time.

Based on these considerations, I think that the functional system time window (T_{SW}) and the available time window (T_{avail}) for HFE2 begin when RCP seal temperatures reach 230 °F and end when the RCP seals fail.

Why does this summary indicate that the starting point for the action that is modeled by HFE2 begins when the crew enters ES-01, rather than when RCP seal temperatures reach 230 °F?

Why does this summary not document the fact that the "error of commission" that is modeled by HFE2 is relevant only if the operators start the PDP before the RCP seals fail?

81. Appendix C, Section C.2, Example 2, Worksheet B for HFE2, Section B.1, HFE Definition

JWS Note: This comment is related to the immediately preceding comment. However, it addresses a much more "philosophical" issue, so I kept it separate.

Section B.1 of this worksheet does not document a Time Available for HFE2. I think that it should.

The notion of an available time window (T_{avail}) is usually associated with the amount of time that is available for personnel to complete a desired action before that action is no longer functionally effective, or the scenario evolution in the PRA model otherwise alters the scope or the need for that action. However, the notion of an available time window also applies to the types of "errors of commission" that are evaluated by HFE2 in this example analysis. (The notion of a "window of opportunity", rather than an "available time window", might be easier to understand in this context.)

In particular, the analysis of HFE2 examines whether the operators will start injection flow from the positive displacement charging pump (PDP) after reactor coolant pump (RCP) seal temperatures exceed 230 °F. As discussed in the preceding comment, I think that the functional constraint for the relevant time window for this particular action is determined by the time when the RCP seals fail. To understand how the notion of T_{avail} applies for this analysis, consider the following possibilities.

- Suppose that the seals will fail within 30 seconds after temperatures reach 230 °F. The preceding failure of HFE1 in this scenario means that the operators did not start the PDP within approximately 4 to 6 minutes after the loss of all component cooling water (CCW) and seal injection flow. It seems that the conditional probability that they would start the PDP within the next 30 seconds may be rather small.

- Suppose that the seals will not fail until approximately 2 hours after temperatures reach 230 °F. In this case, it seems that there might be a much higher conditional probability that the operators would start the PDP at some time during this 2-hour interval. Of course, that probability also depends on how, and when, the crew might reach a decision to start the pump.

In this context, it seems relevant to quantify a value for T_{avail} for this example and to alert analysts to these types of considerations.

JWS Note: I have not thought carefully about how one might use the notion of T_{avail} to quantify a human error probability (HEP) for these types of actions. The calculation would certainly not be the same as that described for P_t in Section 5.1 of NUREG-2198 and Section 3.6 of this report. There are additional complicating considerations for that quantification, such as the possible "overlap" between T_{reqd} for HFE1, the end of T_{avail} for HFE1, and the start of the time window for HFE2. For example, if the operators intend to start the PDP, but their actions extend somewhat past the time when seal temperatures reach 230 °F, it may be very likely that they will continue those actions, despite the procedural guidance. In that case, the conditional probability that they start the PDP after temperatures exceed 230 °F might be rather high, even if the nominal value of T_{avail} is relatively short. Therefore, I think that the mathematical form of any time-related HEP calculation should be left to future research. However, I do think that it is worthwhile to document a value for T_{avail} in this example and to discuss these considerations as an input to a qualitative analysis of the time uncertainty effects on the HEP for HFE2.

Considering these comments, should Section B.1 of this worksheet provide an estimate for Time Available, with an appropriate discussion about what it means in the context of this analysis?

82. Appendix C, Section C.2, Example 2, Analysis of HFE2

Worksheet A and Worksheet B in this example define human failure event HFE2 to evaluate an "error of commission" for starting seal injection from the positive displacement charging pump (PDP) after reactor coolant pump (RCP) seal temperatures exceed 230 °F. Several preceding comments address the scenario context, definition, and characterization of HFE2.

The example does not contain a Worksheet C, Worksheet D, or Worksheet E for HFE2. Thus, the analysis of HFE2 is abruptly truncated, without demonstrating how the IDHEAS-ECA methodology and guidance can be used to evaluate the critical tasks, cognitive failure modes (CFMs), and performance-influencing factors (PIFs) for these types of actions, or quantify the likelihood that HFE2 occurs.

I think that the example should complete the analysis of HFE2. That analysis would provide important insights about the qualitative and quantitative evaluation of this particular action. It would also demonstrate that the IDHEAS framework and methodology can be used to evaluate "errors of commission", without the need for a separate methodology or guidance.

Why does this example not complete the analysis of HFE2?

83. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T1, Section C.1, Analysis of Cognitive Activities and Identification of Applicable CFMs

The summaries of the cognitive failure modes (CFMs) for Understanding and Decisionmaking note that:

> "*Understanding* - The task ***does not require Understanding*** activities because, ***from training, operators know*** to stop RCPs upon recognizing the simultaneous loss of CCW and RCP sealwater." [emphasis added]
>
> "*Decisionmaking* – The task ***does not require decisionmaking*** activities because, ***from training, operators know*** to stop RCPs upon recognizing the simultaneous loss of CCW and RCP sealwater." [emphasis added]

These entries seem to involve assumptions or conclusions about the attributes for the specific performance-influencing factor (PIF) for "Training" in the context of this scenario. In particular, they seem to imply that the analysts have already concluded that the "no impact" attribute applies for that PIF. Furthermore, they seem to imply that the analysts have concluded that adequate training, by itself, is sufficient to justify a conclusion that the macrocognitive functions for Understanding and Decision-Making are not needed for successful performance of the action to trip the reactor coolant pumps (RCPs). Preceding comments addresses my concerns about premature assumptions or conclusions about the relevance of specific PIFs or the status of specific PIF attributes, without an actual evaluation of those PIFs for each critical task, including justification for the analysts' conclusions.

The intent of Section C.1 in this worksheet is to provide an objective assessment of the CFMs that apply to each critical task from the perspective of the basic macrocognitive functions that are needed for successful human cognitive performance. The guidance in Section 3.3.2 and Table 3-3 of the main report emphasize that type of comprehensive, objective assessment. The relevance of each PIF and its scenario-specific attributes for each CFM are then evaluated in the analyses that are summarized in Section C.3 of the worksheet.

Of course, the analyses for HFE1-T1 may conclude that the "no impact" attribute for all relevant PIFs, including training, applies to the CFMs for Understanding and Decisionmaking during this scenario. However, that conclusion should be based on a comprehensive and systematic assessment, and it should not be applied prematurely to justify exclusion of these CFMs from further evaluation in Section C.3 of the worksheet.

JWS Note: The discussions of the Cognitive Activities in Section C.1 of Worksheet C in Example 1 provide good examples of the types of analyst considerations that should be documented in this section of the worksheet.

Why do the assessments of Understanding and Decisionmaking include premature assumptions or conclusions about the effects from the specific PIF for training to justify exclusion of these CFMs from further consideration?

Why do these entries not document the types of objective information that is summarized in the Guidance for Identifying the Applicable CFMs in Section 3.3.2?

84. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T1, Section C.2, Task Characterization

The discussion of Special Requirements in this section notes that:

> "The task needs to be performed ***within minutes*** after the loss of CCW and RCP sealwater." [emphasis added]

Section B.1 and Section B.2 in Worksheet B for HFE1 indicate that the Time Available for HFE-T1 is 1 minute.

Why does this summary of the Special Requirements for HFE-T1 not specify that the task must be performed within 1 minute after the loss of CCW and RCP seal water?

85. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T1, Section C.2, Task Characterization

The discussion of Cue in this section notes that:

> "The cues for starting the task include the alarms of CCW pump trip and ***PDP trip*** and the indications of no CCW flow and no sealwater. Alternatively, operators may also use the cue of ***loss of Bus C*** to recognize the loss of CCW and RCP sealwater ***based on their knowledge***." [emphasis added]

I do not think that this is an accurate characterization of the cues for the action to trip the reactor coolant pumps (RCPs). In particular:

- The introduction to this scenario indicates that centrifugal charging pump 1A is running when the initiating event occurs. The positive displacement charging pump (PDP) is not running. Therefore, it does not trip.

- The introduction to this scenario indicates that component cooling water (CCW) pump 1B is out of service when the initiating event occurs. The loss of power at Bus E1C disables CCW pump 1C and centrifugal charging pump 1A. Thus, the operators will not directly associate loss of power at Bus E1C with complete loss of CCW and RCP seal injection flow, because CCW pump 1A and centrifugal charging pump 1B are not disabled by that power failure. This is a very important element of this scenario.

- The introduction to this scenario indicates that all CCW flow and RCP seal injection flow are lost because CCW pump 1A experiences a mechanical failure and centrifugal charging pump 1B trips from an unspecified cause at approximately the same time that the reactor trip occurs. The operators must recognize that these additional equipment failures have occurred, or they must use other alarms and indications to prompt them to trip the RCPs and start seal injection flow from the PDP.

JWS Note: Charging pump 1B is cooled by CCW. Therefore, it may trip from loss of cooling. However, the introduction simply notes that it trips 2 minutes after it is started. In practice, that 2-minute delay may be important, and it may introduce further complications for the analysis of a scenario that is already very complex. The example analysis seems to assume that charging pump 1B fails at essentially the same time that the reactor trip occurs. That is a reasonable simplification.

JWS Note: This comment also applies to the discussion of Cue in Section C.2 of Worksheet C for Critical Task HFE1-T2. However, that discussion does not explicitly mention the PDP trip.

Why does this summary indicate that the salient cues for the operator action to trip the RCPs are a "PDP trip" and loss of power at Bus E1C?

What specific indications or distinct alarms alert the operators that all CCW and seal injection flow are lost at essentially the same time that the reactor trip occurs?

86. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T1, Section C.2, Task Characterization

The section indicates that there are no Competing Goals and Alternative Strategies for the action to trip the reactor coolant pumps (RCPs). The discussion of Multitasking acknowledges that the operators must "handle multiple, concurrent system malfunctions".

Preceding comments address considerations of the manual control actions that are needed in response to the loss of power at Distribution Panel 1201. In particular, the operators may initially focus on the unusual behavior of level in steam generator A, and their attention may remain focused on concerns about secondary heat removal for some time after the reactor trip occurs. That focus may be further enhanced by the fact that the loss of power at Distribution Panel 1201 also prevents one auxiliary feedwater (AFW) pump from starting when the reactor trip occurs.

JWS Note: I do not know if the loss of power at Bus E1C affects another AFW pump. That is potentially important information for this scenario. However, it is unfortunately missing from the available plant technical information in NUREG-2156.

JWS Note: This comment also applies to the discussion of these items in Section C.2 of Worksheet C for Critical Task HFE1-T2.

Did the analysts for this example identify possible concerns about secondary heat removal as a potentially important competing priority or distraction for the crew's attention during the first few minutes of this scenario (i.e., from time t = 2 minutes until some time after the reactor trip occurs and steam generator levels are relatively stable)?

If not, why not?

If so, should the discussion of Competing Goals and Alternative Strategies, or the discussion of Multitasking, explicitly mention that possible competing priority or distraction?

87. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T1, Section C.3, Assessment of PIFs

As noted in the introduction to this section of the worksheet, the assessment that is documented in Section A.4 of Worksheet A concludes that the following seven performance-influencing factors (PIFs) are relevant during this scenario:

- Scenario Familiarity
- Task Complexity
- Human-System Interface
- Procedures, Guidance, and Instructions
- Training
- Multitasking, Interruption, and Distraction
- Mental Fatigue, and Time Pressure and Stress

The evaluation of cognitive failure mode CFM1 (Failure of Detection) for critical task HFE1-T1 addresses only the three PIFs for Scenario Familiarity; Task Complexity; and Multitasking, Interruption, and Distraction. It does not address the possible effects from the other four relevant PIFs. It is important that these examples should demonstrate systematic and comprehensive use of the IDHEAS-ECA methodology and guidance. Therefore, the evaluation of HFE1-T1-CFM1 should address all seven relevant PIFs. If the analysts conclude that the "no impact" attribute applies for each of the other four PIFs in the scenario-specific context of this particular cognitive failure mode (CFM), their rationale for each conclusion should be documented. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

JWS Note: This comment also applies to the incomplete evaluations of the PIFs for HFE1-T2-CFM1 and HFE1-T2-CFM2 in Section C.3 of Worksheet C for Critical Task HFE1-T2.

Why does this evaluation address the possible effects from only three PIFs?

In particular, why does this evaluation not address the possible effects from the other four relevant PIFs and document the analysts' rationale for the assigned attributes from each of those PIFs?

88. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T1, Section C.3, Assessment of PIFs

JWS Note: This comment is closely related to the immediately preceding comment. However, it addresses the simple assertion that the "no impact" attribute applies for all seven performance-influencing factors (PIFs). I decided to separate it, in order to further reinforce the notion that these types of simple "global" assertions are not consistent with the intent of a well-documented systematic analysis. It also identifies an editorial oversight.

Section C.3 of this worksheet identifies this cognitive failure mode (CFM) as HFE1-T1-***CFM2***. Section C.1 indicates that the relevant CFM applies to Action Execution. Therefore, this CFM should be identified as HFE1-T1-***CFM4***.

The evaluation of this CFM for critical task HFE1-T1 notes that:

> "All the PIFs are no impact. As long as the detection succeeds, stopping the RCP is a simple, one step activity and can be performed instantly."

As noted in preceding comments, it is important that these examples should demonstrate systematic and comprehensive use of the IDHEAS-ECA methodology and guidance. If the analysts conclude that the "no impact" attribute applies for each of the seven relevant PIFs in the scenario-specific context of this particular CFM, their rationale for each conclusion should be documented. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

JWS Note: This comment also applies to the "no impact" evaluation of all seven PIFs for HFE1-T2-CFM4 in Section C.3 of Worksheet C for Critical Task HFE1-T2.

Why does this evaluation not document the analysts' rationale for why the "no impact" attribute applies for each of the seven relevant PIFs in the scenario-specific context of this CFM?

89. Appendix C, Section C.2, Example 2, Worksheet C for Critical Task HFE1-T2, Section C.3, Assessment of PIFs

The evaluation of cognitive failure mode CFM1 (Failure of Detection) for critical task HFE1-T2 indicates that only attribute C1 applies for the performance-influencing factor (PIF) for Task Complexity.

Section C.3 of Worksheet C for Critical Task HFE1-T1 indicates that Task Complexity attributes C1 and C6 apply for the same cognitive failure mode for that action (i.e., HFE1-T1-CFM1).

Based on the available information, it seems that the crew uses the same cues to prompt their actions to trip the reactor coolant pumps (RCPs) and to start seal injection flow from the positive displacement charging pump (PDP). The operators must trip the RCPs within 1 minute after the loss of all component cooling water (CCW) and seal injection. They must start the PDP within approximately 4 to 6 minutes after the loss of all CCW and seal injection.

JWS Note: A preceding comment addresses the total available time window for these actions. I used my estimate in this comment. However, the intent of this comment does not depend on the specific time window. I mentioned it only to emphasize the fact that these actions are performed close in time, without any other evident intervening cues for HFE1-T2.

The scenario description and timeline do not identify any other compelling indications or alarms that occur after time t = 4 minutes (when the RCPs should be tripped) and before the RCP seal temperatures reach 230 °F, which may alert the operators to start the PDP. Therefore, I do not understand why attributes C1 and C6 apply for HFE1-T1-CFM1, while only attribute C1 applies for HFE1-T2-CFM1.

A preceding comment addresses my observation that the IDHEAS-ECA methodology does not describe how analysts should account for scenario-specific conditions that adversely affect multiple attributes for a particular PIF (i.e., like C1 and C6). Subsequent comments on Worksheet D for Critical Task HFE1-T1 and Worksheet D for Critical Task HFE1-T2 further address that issue in the context of quantifying the human error probabilities (HEPs) for HFE1-T1-CFM1 and HFE1-T2-CFM1.

JWS Note: In principle, I could speculate about possible reasons for this difference, but I will not. The point of this comment is that there is an apparent discrepancy in these assessments. If the analysts have a distinct rationale for why attribute C6 does not apply for HFE1-T2-CFM1, that rationale should be documented to avoid these types of observations and questions.

Considering the conclusion that Task Complexity attributes C1 and C6 apply for the evaluation of HFE1-T1-CFM1, why does only attribute C1 apply for the evaluation of HFE1-T2-CFM1?

90. Appendix C, Section C.2, Example 2, Worksheet D, HEP Estimation, Critical Task HFE1-T1

Worksheet D does not show how the human error probability (HEP) for P_c is calculated for critical task HFE1-T1. Thus, this example does not fully demonstrate how the IDHEAS-ECA methodology and the guidance in this report are used to quantify an HEP.

JWS Note: The examples should not presume that analysts use the IDHEAS-ECA software and that the HEP calculations are "invisible" to the user. The examples should demonstrate how each HEP is calculated, so readers clearly understand the intent of the guidance and the quantification process.

Critical task HFE1-T1 evaluates operator actions to trip the reactor coolant pumps (RCPs) within 1 minute after loss of all component cooling water (CCW) flow and all seal injection flow. Worksheet D indicates that the following performance-influencing factor (PIF) attributes apply for the two cognitive failure modes (CFMs) for this task.

T1-CFM1

- Scenario Familiarity SF3
- Task Complexity C1 and C6
- Multitasking, Interruption, and Distraction MT3

T1-CFM4

- All PIFs No Impact

Thus, it seems that the value of P_c for HFE1-T1 should be determined by the sum of the HEP for T1-CFM1 that applies for the combined effects from attributes SF3, C1, C6, and MT3, plus the "no impact" HEP for T1-CFM4.

A preceding comment addresses my observation that the IDHEAS-ECA methodology does not describe how analysts should account for scenario-specific conditions that adversely affect multiple attributes for a particular PIF (i.e., like C1 and C6 for T1-CFM1). Without further guidance, I have assumed that the effects from those attributes are additive (i.e., that the "base" HEPs from each attribute should be added to quantify the total HEP for T1-CFM1).

Calculation of P_c for T1-CFM1

Table B-1 in Appendix B indicates that the "base" HEP for Detection is 1.2E-02 when the "scenario is unfamiliar, rarely performed" condition applies for Scenario Familiarity attribute SF3.

Table B-3 in Appendix B indicates that the "base" HEP for Detection when Task Complexity attribute C1 applies depends on the number of competing signals. Section C.3 of Worksheet C for Critical Task HFE1-T1 does not document the analysts' assessment of this consideration for attribute C1. Therefore, the analysis documentation in Worksheet C is incomplete and deficient. For the purpose of this comment, based on the combined effects from the loss of power at Distribution Panel 1201, reactor trip, loss of power at Bus E1C, failure of the CCW pump, and failure of the centrifugal charging pump, I will assume that the scenario produces an "excessive amount (> 20)" of signals. The "base" HEP for Detection is 3E-01 when that condition applies for attribute C1.

Table B-3 in Appendix B indicates that the "base" HEP for Detection is 1E-01 when Task Complexity attribute C6 applies.

The PIF for Multitasking, Interruption, and Distraction determines a weighting factor that is applied to the "base" HEP. Table B-13 in Appendix B indicates that the PIF weight for Detection is 5 when the "moderate demanding" condition applies for attribute MT3.

According to Equation 3.4 and Equation 3.5 in the main report, it seems that P_c for T1-CFM1 should be:

$$P_{cT1\text{-}CFM1} = \{1 - [(1 - HEP_{T1\text{-}CFM1SF3}) * (1 - HEP_{T1\text{-}CFM1C1}) * (1 - HEP_{T1\text{-}CFM1C6})]\} * [1 + (w_{MT3} - 1)]$$

$$= \{1 - [(1 - 1.2E\text{-}02) * (1 - 3.0E\text{-}01) * (1 - 1.0E\text{-}01)]\} * [1 + (5 - 1)]$$

$$= \{3.78E\text{-}01\} * [5]$$

$$= 1.89$$

This is clearly wrong. Therefore, it seems that I have misinterpreted some fundamental element of the guidance for quantifying the HEP, the guidance for selection of the applicable "base" HEPs from Table B-1 and Table B-3, or the guidance for selection and use of the PIF weight from Table B-13.

I made an assumption about how the effects from Task Complexity attributes C1 and C6 should be combined. If I use only attribute C1 for Task Complexity, the result is:

$$P_{cT1\text{-}CFM1} = \{1 - [(1 - HEP_{T1\text{-}CFM1SF3}) * (1 - HEP_{T1\text{-}CFM1C1})]\} * [1 + (w_{MT3} - 1)]$$

$$= \{1 - [(1 - 1.2E\text{-}02) * (1 - 3.0E\text{-}01)]\} * [1 + (5 - 1)]$$

$$= \{3.08E\text{-}01\} * [5]$$

$$= 1.54$$

Thus, my assumption about how the effects from attributes C1 and C6 should be combined does not functionally affect the fact that my calculated HEP is unrealistic and not appropriate.

Calculation of P_c for T1-CFM4

The introduction to Appendix B indicates that the IDHEAS-ECA guidance recommends minimum HEPs of 1E-04 for Detection or Action Execution, and 1E-03 for Understanding, Decsionmaking, or Interteam Coordination.

Since CFM4 evaluates the Action Execution macrocognitive function,

$$P_{cT1\text{-}CFM4} = 1.0E\text{-}04$$

Based on my evident failure to correctly interpret and apply the guidance to quantify P_c for T1-CFM1, I think that it is very important for Worksheet D to describe how P_c is quantified, so that analysts clearly understand how to treat the combined effects from multiple attributes for a specific PIF, how to treat the PIF attributes that have "no impact", how to use the tables in Appendix B, how to apply the PIF weights, and how to combine the HEPs for each CFM to quantify the overall value of P_c for critical task HFE1-T1.

What specific condition from Table B-3 in Appendix B applies for Task Complexity attribute C1 during this scenario?

Why is that condition and its basis not documented in Section C.3 of Worksheet C for Critical Task HFE1-T1?

Where is the error in my calculation of P_c for T1-CFM1?

Why does Worksheet D not show how P_c is calculated for critical task HFE1-T1?

91. Appendix C, Section C.2, Example 2, Worksheet D, HEP Estimation, Critical Task HFE1-T2

Critical task HFE1-T2 evaluates operator actions to start reactor coolant pump (RCP) seal injection flow from the positive displacement charging pump (PDP) before RCP seal temperatures reach 230 °F.

Worksheet D indicates that the same performance-influencing factor (PIF) attributes apply for cognitive failure modes T1-CFM1 and T2-CFM1, except that only attribute C1 applies for Task

Complexity in the evaluation of T2-CFM1. The same "no impact" attributes apply for T1-CFM4 and T2-CFM4. The evaluation of P_c for critical task HFE1-T2 also includes T2-CFM2 for the Understanding macrocognitive function.

Please refer to the preceding comment and my estimate of the human error probability (HEP) for critical task HFE1-T1. For brevity in these comments, I did not try to quantify a value of P_c for T2-CFM2 or a total HEP for critical task HFE1-T2. As indicated in the preceding comment, if I account for only attribute C1, my estimated HEP for T2-CFM1 would be 1.54.

Why does Worksheet D not show how P_c is calculated for critical task HFE1-T2?

92. Appendix C, Section C.2, Example 2, Worksheet D, HEP Estimation

JWS Note: This comment is related to the immediately preceding comments. However, it addresses a broader concern about the IDHEAS-ECA guidance and, possibly, the guidance in NUREG-2198.

The preceding comments demonstrate my failure to use the IDHEAS-ECA guidance and the numerical values that are tabulated in Appendix B to quantify a reasonable human error probability (HEP) for critical task HFE1-T1. Use of the basic IDHEAS quantification model, the recommended "base" HEP values, and the performance-influencing factor (PIF) weights should never produce an estimated HEP that is greater than 1.0. Therefore, I hope that there is some fundamental flaw in my understanding or use of the guidance.

Of course, if the quantification process can actually produce HEP values that exceed 1.0, it is not appropriate to simply assume that the HEP is equal to 1.0 whenever that occurs. The same flaws in the basic quantification model, the guidance, or the numerical values would similarly produce inappropriate estimates for HEPs that are less than 1.0. Therefore, depending on the resolution of the preceding comments, this issue could represent a critical deficiency in either the basic IDHEAS quantification model (i.e., the equations in Section 4.4.1 and Section 4.4.3 of NUREG-2198, and in Section 3.5.1 of this report), or a lack of adequate testing, verification, and validation of the numerical values that are recommended in Appendix B of this report and how they are used in a variety of practical analyses.

Assuming that they correctly interpret the IDHEAS-ECA guidance, can analysts use the IDHEAS quantification model and the numerical values that are tabulated in Appendix B of this report to quantify HEPs that exceed 1.0?

If so, how will the authors of this report and NUREG-2198 resolve this issue?

93. Appendix C, Section C.2, Example 2, Worksheet E, Time Uncertainty Analysis of the HFEs, Critical Task HFE1-T1

Worksheet E does not show how the human error probability (HEP) for P_t is calculated for critical task HFE1-T1. Thus, this example does not fully demonstrate how the IDHEAS-ECA methodology and the guidance in this report are used to quantify an HEP.

Critical task HFE1-T1 evaluates operator actions to trip the reactor coolant pumps (RCPs) after loss of all component cooling water (CCW) flow and all seal injection flow. Worksheet E indicates that the time available for this action (T_{avail}) is 1 minute.

Worksheet E indicates that the time required to perform this action (T_{reqd}) is characterized by a normal probability distribution with a mean value of 6.5 minutes and a standard deviation of 1.5 minutes. It is noted that these times are derived from the simulator tests that are documented in NUREG-2156. Table 5-6 in NUREG-2156 lists the following crew times for the action to trip the RCPs (time after the loss of all CCW and seal injection):

- 4:49 (4.82 minutes), 6:45 (6.75 minutes), 7:29 (7.48 minutes), 8:39 (8.65 minutes)

The mean value of these times is 6.93 minutes. The 5th percentile of the assigned normal distribution is approximately 4 minutes, and the 95th percentile is approximately 9 minutes. Thus, the assigned distribution appropriately accounts for the observed range in T_{reqd}.

Based on the 1-minute value for T_{avail} and the uncertainty distribution for T_{reqd}, the value of P_t for critical task HFE1-T1 is essentially 1.0. (There is an extremely small probability that T_{reqd} is less than 1 minute.)

Is my calculation of P_t for critical task HFE1-T1 correct?

Why does Worksheet E not show how P_t is calculated for critical task HFE1-T1?

94. Appendix C, Section C.2, Example 2, Worksheet E, Time Uncertainty Analysis of the HFEs, Critical Task HFE1-T2

Critical task HFE1-T2 evaluates operator actions to start reactor coolant pump (RCP) seal injection flow from the positive displacement charging pump (PDP) before RCP seal temperatures reach 230 °F.

Worksheet E contains the following information about the time that is available to perform this action (T_{avail}).

> "The ***system time available is 7-9 minutes in the PRA model***. However, operators need to perform HFE1-T1 first, so the time available for HFE1-T2 is the system time available subtracted by the time needed for HFE1-T1." [emphasis added]

A preceding comment addresses the time when RCP seal temperatures reach 230 °F, compared to the nominal reference point for time t = 0 in the Event Timeline in Section A.1 of Worksheet A and the time when all component cooling water (CCW) flow and all seal injection flow are lost (i.e., at time t = 3 minutes in that timeline).

For the purpose of this comment, I will assume that RCP seal temperatures reach 230 °F approximately 7 to 9 minutes after loss of all CCW flow and all seal injection flow. That time is not consistent with the timeline in Worksheet A, but it is consistent with the event sequence timing that is documented in Table 5-7 of NUREG-2156.

The analysis in Worksheet E assumes that the actions to trip the RCPs (HFE1-T1) and to start the PDP (HFE1-T2) are performed sequentially. The available time window (T_{avail}) for critical task HFE1-T1 is 1 minute. Therefore, it seems that the analysis of P_t for critical action HFE1-T2 is based on a value for T_{avail} of approximately 6 to 8 minutes (i.e., the time from the end of the time window for HFE1-T1 until the seal temperatures reach 230 °F).

Worksheet E does not indicate whether a probability distribution is used to characterize the uncertainty in T_{avail}, or if a single point-estimate value is used. For the purpose of this comment, I will perform simple calculations that apply for 6-minute, 7-minute, and 8-minute values for T_{avail}.

Worksheet E contains the following information about the time that is required to perform this action (T_{reqd}):

> "After operators performed HFE1-T1, they already detected the loss of CCW and sealwater. Then, the time needed for the task is the time to enter the RCP procedure and start the PDP. The mean time for detecting the loss of CCW and sealwater is ***5.5min*** with a deviation of 1.5min. The mean time taken from detecting the loss of CCW and sealwater to entering the RCP procedure is approximately ***3min with a standard deviation 2min***, based on operator simulator performance data in NUREG-2156." [emphasis added]

Based on this discussion, it is somewhat difficult to understand the intended uncertainty distribution for T_{reqd}. I think that the highlighted value of 5.5 minutes is a typographical error, and it should be 6.5 minutes (i.e., the mean T_{reqd} for critical task HFE1-T1). The analysis seems to indicate that the operators spend the time that is required for critical task HFE1-T1 to identify the loss of CCW and seal injection. After those initial identification activities are performed, the analysis seems to indicate that the additional time that is required to consult the relevant procedure and start the PDP is characterized by a normal distribution with a mean value of 3 minutes and a standard deviation of 2 minutes. Thus, this seems to be the intended estimate of the additional T_{reqd} for critical task HFE1-T2.

Worksheet E indicates that these times are derived from the simulator tests that are documented in NUREG-2156. However, Table 5-6 and Table 5-7 in NUREG-2156 do not list the crew times for the action to start injection flow from the PDP. They simply indicate that none of the four crews started the PDP before RCP seal temperatures reached 230 °F. Therefore, I do not understand the basis for the assigned probability distribution for T_{reqd} for critical task HFE1-T2.

For the purpose of this comment, I will assume that T_{reqd} for critical task HFE1-T2 is characterized by a normal distribution with a mean value of 3 minutes and a standard deviation of 2 minutes.

Based on these estimates, I calculated the following values for P_t for critical task HFE1-T2.

- If T_{avail} = 6 minutes, P_t = 6.68E-02
- If T_{avail} = 7 minutes, P_t = 2.28E-02
- If T_{avail} = 8 minutes, P_t = 6.20E-03

As noted in NUREG-2156, none of the four crews started the PDP before RCP seal temperatures reached 230 °F.

Is the value of T_{avail} for critical task HFE1-T2 approximately 6 to 8 minutes after the end of the time window for critical task HFE1-T1?

Is T_{avail} for critical task HFE1-T2 characterized by a probability distribution or a point-estimate value?

Is T_{reqd} for critical task HFE1-T2 characterized by a normal probability distribution with a mean value of 3 minutes and a standard deviation of 2 minutes, after the operators complete critical task HFE1-T1?

If so, what is the basis for that estimate?

In particular, how is it derived from the crew response times that are listed in Table 5-6 and Table 5-7 of NUREG-2156?

Are my understanding of T_{avail} and T_{reqd} consistent with the intent of the information in Worksheet E?

Based on my understanding of T_{avail} and T_{reqd}, are my calculations of P_t for critical task HFE1-T2 correct?

Is the total HEP for critical task HFE1-T2 (i.e., the sum of P_c and P_t) consistent with the simulator experience that is summarized in NUREG-2156?

Why does Worksheet E not show how P_t is calculated for critical task HFE1-T2?

95. Appendix C, Section C.3, Example 3: Human Actions of Implementing FLEX Strategies in a Beyond-Design-Basis Event, General Comment

The introduction to this example indicates that the scenario and supporting information for this analysis are from "NRC's 2018 FLEX-HRA Expert Elicitation". I obtained a copy of the following report that was released publicly to support a September 23, 2020 meeting of the ACRS Subcommittee on Reliability and Probabilistic Risk Assessment.

- "Utilization of Expert Judgment to Support Human Reliability Analysis of Flexible Coping Strategies (FLEX)", Volume 1 (draft)

I did not study that report before I prepared these comments. Therefore, my comments are limited to the information that is documented in this example. I will comment on the referenced report separately.

96. Appendix C, Section C.3, Example 3, Worksheet A, Section A.1, Event Description, Initiating Event, Initial Conditions

The summary of the Initiating Event notes that:

> "An external hazard caused ***flooding*** in a single-unit nuclear power plant and led to damage of plant systems (see initial conditions below)." [emphasis added]

The summary of the Initial Conditions notes that:

- "The external hazard caused an ***ELAP event*** immediately after the plant is impacted." [emphasis added]

- "***Some plant systems***, equipment, and structures that ***do not have direct impact on plant safety*** were damaged." [emphasis added]

- "The ***indications*** of the plant parameters key to responding to the event to protect plant safety ***were available***." [emphasis added]

- "***Debris on the FLEX generator transportation route*** needs to be removed to bring the FLEX equipment to its designated setup location." [emphasis added]

- "Some of the work areas were ***flooded*** but accessible for work." [emphasis added]

I am very confused about this characterization of the scenario.

The introductory summary of the event scenario in this worksheet indicates that the scenario is initiated by a complete "blockage of intake". The intake blockage disables all external cooling water (e.g., main condenser circulating water, non-safety service water systems, and safety-related service water systems). It does not cause flooding of the plant, with consequential flooding damage to any equipment in the plant. Operating equipment overheats and either trips or is damaged by the loss of cooling water. Standby systems that rely on service water or component cooling water are functionally disabled. The intake blockage does not affect offsite power supplies, and it does not disable in-plant electrical equipment (except for the effects on switchgear room ventilation and diesel generator cooling). All onsite power is available, and it will remain available if personnel open the switchgear room doors.

Based on the summary of the Initiating Event and the Initial Conditions, it is obvious that the scenario that is evaluated in this example analysis has ***absolutely no relevance*** to the scenario that is summarized in the introduction to this worksheet.

Why does Worksheet A contain the extensive introductory summary of a scenario that has absolutely no relevance to the scenario that is evaluated in the example analysis?

Is the introductory scenario summary derived from "NRC's 2018 FLEX-HRA Expert Elicitation"?

How does "NRC's 2018 FLEX-HRA Expert Elicitation" pertain to the scenario that is evaluated in this example?

97. Appendix C, Section C.3, Example 3: Human Actions of Implementing FLEX Strategies in a Beyond-Design-Basis Event, General Comment

My subsequent comments on this example completely disregard the summary of the scenario that is presented in the introduction to Worksheet A. In particular, I will assume that the initiating event is a severe external flood that causes a loss of offsite power. The flooding event also apparently causes physical damage to some plant equipment (i.e., due to inundation) or functional damage to some plant systems (e.g., due to loss of external cooling water). My comments will address any lack of information that is needed for analysts to understand the causes and extent of the consequential damage from the flood, and to understand the scenario progression to the time when the modeled actions are needed.

98. Appendix C, Section C.3, Example 3, Worksheet A, Section A.1, Event Description, Initial Conditions

This section notes that:

- "The ***reactor trips automatically*** immediately after the external hazard impacts the plant." [emphasis added]

- "The external hazard caused an ***ELAP event*** immediately after the plant is impacted." [emphasis added]

- "***Some plant systems***, equipment, and structures that ***do not have direct impact on plant safety*** were damaged." [emphasis added]

- "The ***indications*** of the plant parameters key to responding to the event to protect plant safety ***were available***." [emphasis added]

- "***Debris on the FLEX generator transportation route*** needs to be removed to bring the FLEX equipment to its designated setup location." [emphasis added]

- "Some of the work areas were ***flooded*** but accessible for work." [emphasis added]

According to the "Guidance for Developing the Scenario Narrative" in Section 3.1.1 of the main report, this section of the worksheet should describe the plant status when the initiating event occurs, including any relevant information about staffing, personnel allocations, etc.

The items that are listed above do not describe the initial conditions for this scenario. They summarize some information about the scenario progression. That information should be described in other sections of this worksheet.

JWS Note: Depending on the specific information that is provided in subsequent sections of this worksheet, I will comment separately on any lack of information that is needed for analysts to understand the specific effects that are listed above.

Why does this section not document the initial plant conditions when the scenario occurs, according to the guidance in Section 3.1.1?

99. Appendix C, Section C.3, Example 3, Worksheet A, Section A.1, Event Description, Boundary Conditions (i.e., Assumptions) for HEP Estimation

Section A.1 of this worksheet is intended to provide the information that is needed for analysts to understand the event scenario progression until the time when the modeled personnel actions are needed. The IDHEAS-ECA guidance indicates that the scenario narrative should objectively describe the evolution of the scenario from an operational perspective.

This section lists several items that are more appropriate for the Scenario Context that is summarized in Section A.3 of this worksheet, the human failure event (HFE) definitions in Worksheet B, and the evaluations that are documented in Worksheet C. For example, several items prematurely address considerations and assumptions that apply to the specific HFEs that are eventually analyzed in this example. However, those HFEs are not identified until Section A.2 of the worksheet is completed, and they are not fully defined until Worksheet B is completed. Furthermore, some of the "boundary conditions" listed in this section contain analyst assessments of specific performance-influencing factors (PIFs). Those assessments should be documented and justified in Worksheet C for each relevant critical task. Therefore, it is premature and inappropriate to discuss assumptions that pertain to the analyses of specific

HFEs as part of the "boundary conditions" for the operational scenario narrative.

According to the "Guidance for Developing the Scenario Narrative" in Section 3.1.1 of the main report, this section of the worksheet should describe only information that clarifies the scope and constraints of the scenario in the context of the PRA models, assumptions about the status of specific plant equipment (e.g., unavailable due to maintenance), etc.

JWS Note: To keep this comment relatively brief, I did not list or discuss the specific items that pertain to assumptions about the HFE evaluations, rather than the boundary conditions for the scenario narrative. I think that most of the items listed under "System and Environment", the last two items under "Personnel", and both items under "Human Actions" are not appropriate for these "boundary conditions".

Why does this section prematurely describe several assumptions that pertain to the analyses of specific HFEs, before those HFEs are identified and defined in the context of the evolving scenario?

100. Appendix C, Section C.3, Example 3, Worksheet A, Section A.1, Event Description, Scenario Timeline

The extremely abbreviated timeline that is shown in this section does not contain sufficient information about the actual scenario progression for analysts to understand the scenario-specific context for the personnel actions that are evaluated in this example. In particular, the scenario timeline should objectively describe what happens in the plant, and when it happens, from an operational perspective. The "Guidance for Developing the Scenario Timeline" in Section 3.1.1 of the main report summarizes the types of information that should be included in the timeline. The timeline should document events that span the time from the initiating event until at least the time when the desired personnel actions should be performed, and perhaps beyond that time as needed to document the occurrence of plant conditions which determine the available time windows for those actions.

For example, the timeline simply notes that an "ELAP" (extended loss of AC power) occurs at time t = 0. In practice, supervisory personnel are supposed to formally declare that an ELAP condition applies only after they survey the plant status, estimate the likelihood that AC power can be restored within a specified time interval (which may vary from plant-to-plant), and conclude that power will not be restored within that interval. The ELAP declaration then triggers several specific actions, including the deployment and connection of FLEX equipment. Thus, in practice, the ELAP declaration will never coincide with the scenario initiating event (i.e., the arrival and immediate effects from the external flood).

The scenario timeline should objectively and comprehensively describe the specific damage throughout the plant that is caused directly by the external flood. It should also document all subsequent system responses, equipment failures, and when they occur. For example, the timeline does not document why no power is available from the installed emergency diesel generators (EDGs) or the station blackout (SBO) diesel generator, and it does not indicate when those power supplies are disabled. Without that information, analysts cannot understand why the scenario involves loss of all onsite AC power, or why power cannot be restored within the plant-specific time window for formal declaration of the ELAP condition. Furthermore, they cannot understand other important features of the scenario context that may affect personnel performance during the time interval before the desired actions are needed, or during the interval when they should be performed (e.g., disabled instrumentation and indications, when the indications are lost, progressive loss of specific DC power supplies, conflicting priorities or distractions from the need to cope with plant-wide damage, etc.).

After the "ELAP" is listed at time t = 0, the timeline notes that the operators subsequently declare that an ELAP condition exists at time t = 1 hour. It then simply indicates that personnel are expected to perform several actions that are related to the deployment and connection of FLEX equipment during the interval from time t = 1 hour until time t = 6 hours. According to Section A.2 of this worksheet, the example analysis evaluates whether the ELAP condition is declared within the relevant time window, and the analysis evaluates some of the specific actions that are listed during the 5-hour interval after time t = 1 hour.

The timeline also contains inappropriate information about specific performance-influencing factors (PIFs) that affect personnel response (e.g., the analysts' observations that the FLEX actions are based on procedural instructions, and those instructions are available). Section A.4 of this worksheet should provide the analysts' initial assessment of the relevant scenario-specific PIFs. Worksheet C should document the evaluation of those PIFs for each relevant critical task.

JWS Note: Section A.3 of this worksheet indicates that the scenario context for the actions that are evaluated in this example apparently involves successful completion of some FLEX-related activities. For example, Action 1 to deploy, connect, and control the portable pumps and Action 3 to refill water tanks may be performed before, during, or after modeled Action 4 to declare an ELAP condition and Action 2 to deploy, connect, and control the portable generators. The timeline should document when Action 1 begins and when it is completed, so that analysts can consider possible competing demands for the available personnel when they evaluate Action 2. It is also apparent that the initial DC load shed that is mentioned in Action 5 is performed before the ELAP is declared. Therefore, the timeline should document when that action begins and when it is completed. The deep DC load shed that in mentioned in Action 5 is apparently performed after the ELAP is declared. Therefore, the timeline should also identify that action, so that analysts know that it is being performed in parallel with Action 2.

Why does the scenario timeline not provide an objective, operational summary of what happens during the scenario and when it happens, according to the guidance in Section 3.1.1?

101. Appendix C, Section C.3, Example 3, Worksheet A, Section A.2, Identification of Important Human Actions

This section indicates that the example analysis evaluates only Action 2 to deploy, connect, and control the portable generators, and Action 4 to declare an ELAP (extended loss of AC power).

It is important for analysts to document how the scenario-specific context for the two modeled actions accounts for the other three listed actions. In particular, at this stage of the analysis process, it is appropriate, and necessary, to document the analysts' assumptions about those actions. For example, the timing of those actions and the specific personnel needed to perform them may affect the analysts' evaluations of the modeled actions.

JWS Note: The immediately preceding comment addresses how the scenario timeline in Section A.1 should document Action 1 and the two activities that are noted in Action 5.

Why does this section not explicitly document the analysts' assumptions about performance and completion of the three actions that are not evaluated in this example?

102. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Environment Context, Action 2; Worksheet B for HFE2 (Action 2), Section B.1, HFE Definition

The Environment Context for Action 2 in Section A.3 of Worksheet A notes that:

- "Worksite accessibility and habitability are not affected ***after debris removal***." [emphasis added]

Action 2 involves personnel actions to deploy, connect, and control the portable generators. The summary of the environmental context for those actions seems to imply that any debris in the portable generator transportation path has already been cleared when the time window for Action 2 begins. If that is true, this is a very important assumption and functional constraint for the example analysis. It should be documented explicitly in Section A.2. The timeline in Section A.1 should also show when the actions to clear the debris begin and when they are completed.

If the scope of Action 2 includes clearing the debris, then the Environment Context should briefly describe the types and amounts of debris that must be removed from the transportation pathway.

Section B.1 of Worksheet B for HFE2 (Action 2) notes that:

> "The debris in the transportation route ***needs to be removed*** before the action can be performed. Removing debris is considered as a ***separate important human action*** because it is performed by a different group of people and affects the deployment of all FLEX equipment." [emphasis added]

The need to clear debris is not mentioned in any of the analyses in Worksheet C, Worksheet D, or Worksheet E for this example. In particular, in Worksheet E, the Time Available and the Time Needed for HFE2 do not mention whether, or how, those estimates account for the time that is needed to clear debris, or how the debris removal time affects the quantification of P_t for HFE2.

Based on this evidence, it seems apparent that the scope of Action 2, as evaluated in this example, does not include the activity to clear the debris before the portable generator can be deployed. That activity is obviously an "important human action" which affects the timing of Action 2 and the analysts' evaluation of the integrated scenario. Therefore, it seems evident that the action to clear the debris should be identified in Section A.2 of this worksheet, the analysts' assumptions about that action should be documented, and it should be shown explicitly in the scenario timeline in Section A.1.

JWS Note: The discussion of Situation Context notes that "personnel need to assess site damage". It also notes that an "error or delay in one FLEX action could have ripple effects on the other FLEX actions".

Does the scope of Action 2 include personnel actions to clear debris from the transportation path for the portable generator?

If Action 2 includes the functional requirements and the time that is needed to clear the debris, why does the Environment Context not summarize the types and amounts of debris that must be removed?

If the scope of Action 2 for this example analysis does not include actions to clear the debris, why are those actions not identified in Section A.2 of this worksheet and shown in the timeline in Section A.1, including the analysts' assumptions about when those actions begin and when they are completed?

103. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Environment Context, Action 4

The information for Action 4 notes that Main Control Room (MCR) lighting is provided by emergency lights. It does not mention the status of MCR ventilation and cooling. Depending on the plant-specific MCR heatup analyses, the room temperature or humidity may reach levels that adversely affect human cognitive performance, especially for actions that are needed after extended times.

If the scenario-specific context for this example analysis includes assumptions about personnel actions to open MCR doors or provide some other form of alternative ventilation and cooling, those actions and assumptions should be documented in Section A.2 of this worksheet. The actions should also be shown in the scenario timeline, so that analysts can appropriately account for how their requirements and timing are integrated into the overall scenario evolution.

Why does the Environment Context for Action 4 not document the status of MCR ventilation and cooling?

104. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, System Context

This summary simply notes that the human-system interface is "well-designed and not damaged", and all needed parts and tools are "available or accessible".

This summary is not sufficient for analysts to understand the actual status of plant systems and the status of displays and indications in the Main Control Room (MCR) when the modeled actions are needed. It is not consistent with the "Guidance for Assessing the System Context" in Section 3.1.3 of the main report.

In particular, this summary does not document the specific damage throughout the plant that is caused directly by the external flood. It should also document all subsequent system responses and equipment failures. For example, the summary does not document the fact that no AC power is available. It also does not document whether any non-safety DC power supplies are lost due to battery depletion before the modeled actions are needed. The summary of Action 5 in Section A.2 of this worksheet indicates that operators initially de-energize some DC loads before Action 4 is needed. The System Context should document the effects from the AC power failures, DC battery depletion (if any), and DC load shedding on displays and indications in the MCR, regardless of whether the affected indications are directly needed to support the modeled actions. It is essential for analysts to have a clear understanding of the entire plant status, including the actual situation in the MCR, when the modeled actions are needed.

JWS Note: The summary of the Task Context in this section of the worksheet indicates that "information needed for FLEX actions may not be readily available and may be presented with a large uncertainty because of unclear status of equipment damage".

Why does this entry not fully document the actual status of plant systems and MCR indications when the modeled actions are needed?

105. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Personnel Context

JWS Note: Preceding comments address premature assessments of specific performance-influencing factors (PIFs) in the Scenario Context summaries in Section A.3 of Worksheet A. Those comments also apply to this summary of Personnel Context. My comments on this summary pertain to the scope and content of objective information that is needed for analysts to understand this element of the scenario context.

This section notes that:

- "Staffing – minimum required staffing on site"

The simple observation that the "minimum required staffing" is available is not sufficient for analysts to understand the actual personnel context for this scenario. According to the "Guidance for Assessing the Personnel Context" in Section 3.1.3 of the main report, this summary should identify the specific personnel who are available to perform all of the actions that are needed to cope with the plant-wide damage, their skill sets, and scenario-specific demands that affect their availability to perform the modeled actions when they are needed.

For example, considering only the five actions that are summarized in Section A.2 of this worksheet, Action 4 involves the operating crew in the Main Control Room. Action 5 involves operators who locally de-energize DC loads. Action 1 and Action 3 involve personnel who deploy, connect, and control the portable pumps. Action 2 involves personnel who deploy, connect, and control the portable generators. As discussed in a preceding comment, personnel must also clear debris from the transportation pathways to facilitate Action 1 and Action 2. It is very likely that personnel are also involved with other local activities throughout the plant that are needed to cope with the site-wide damage. Furthermore, if the analysis in this example accounts for personnel other than the plant operators to perform some of the needed actions, the Personnel Context should explicitly identify those people and summarize what they are doing before the modeled actions are needed (e.g., maintenance personnel, radiation protection personnel, security personnel, or others).

Why does this entry not fully document the actual number of people who are onsite when the flood occurs, the specific personnel who may participate in the modeled actions, their skill sets, and how they are engaged when the modeled actions are needed?

106. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Personnel Context

The summary of "Interteam Coordination" does not document the specific coordination that is needed for Action 2 to deploy, connect, and control the portable generators, and for the actions to clear debris before Action 2 can begin. In particular, this summary should provide an objective description of who performs each activity, how those activities are performed (e.g., sequentially, in parallel, or a mixture), and how the activities are coordinated by supervisors in the Main Control Room (MCR) and among the local personnel. For example, different personnel may clear the debris, transport the generators, connect the generators to the electrical buses, and operate the generators after they are installed. This summary should describe how those various activities and personnel are coordinated, and it should describe who leads that coordination.

The summary should also describe how the interteam coordination needs for Action 1 may affect the performance or timing of Action 2. This consideration is important if the actions are performed in parallel. It is especially important if some of the same personnel are involved in both actions.

Why does this summary not document how the specific activities for Action 2 and the activities for debris removal are coordinated by supervisors in the MCR and among the local personnel who perform those actions?

Do the interteam coordination needs for Action 1 affect the performance or timing of Action 2?

107. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Personnel Context

The first sub-bullet item under "Coordination" notes that one consideration for "Interteam Coordination" is:

- "interdependence with other stakeholders (e.g., ***evacuation of the surrounding populations and working with firefighters***)" [emphasis added]

It is not apparent why these particular considerations are relevant to the scenario context for the actions that are evaluated in this example, at the time when those actions are needed. It is very likely that a shift supervisor will be engaged in discussions with the corporate emergency response organization, the NRC, local and regional emergency responders, and municipal or state officials who are responsible for implementing emergency planning guidelines. Those activities are certainly relevant to the scenario context (e.g., as part of the roles and responsibilities for specific personnel).

In particular, it is not apparent why this summary specifically highlights "evacuation of the surrounding populations". I thought that the primary functional intent of the FLEX equipment and actions is to prevent core damage and to ensure that containment functions are maintained. Of course, the FLEX equipment may also be useful for mitigating the event progression after core damage occurs. However, in the context of the scenario that is evaluated in this example, it seems that deployment and connection of the portable pumps and generators is intended to prevent core damage. Thus, depending on the plant-specific emergency plan, it seems that considerations of the need to evacuate members of the local populace may be premature at the time when the modeled actions are needed (i.e., starting at about 1 hour after the flood occurs).

Furthermore, it is not apparent why coordination with outside firefighters is needed for this flooding scenario, unless the analysis accounts for those external resources to help mitigate the effects from the site-wide flooding damage. The scenario summary does not identify any consequential fires that must be extinguished. Thus, it seems that offsite firefighters might only be needed to assist with pumping out flooded plant locations, or perhaps to help with the deployment and connection of the FLEX portable pumps. Of course, if that is the case, the Personnel Context should identify those firefighters and what particular actions they perform. The scenario timeline should also document when they arrive onsite.

Why does this item explicitly emphasize "evacuation of the surrounding populations" and coordination of "firefighters"?

How are these specific requirements related to the scenario context for the modeled actions, at the time when those actions are needed?

Should this summary only document the expected needs for shift supervisor discussions with offsite personnel (e.g., corporate emergency response organization, the NRC, local and regional emergency responders, municipal and state officials, etc.), without citing the specific highlighted coordination requirements?

108. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Personnel Context

The second sub-bullet item under "Coordination" notes that one consideration for "Interteam Coordination" is:

- "difficulty in anticipating events for unexpected action sequences (e.g., no one had anticipated that an air compressor would be needed to open the venting valve remotely, and sourcing one significantly delayed the venting)"

This item seems to be more relevant to the summary of the Task Context, rather than an example of a special consideration of Interteam Coordination for the modeled actions.

Furthermore, it is not apparent why the procurement and installation of a portable air compressor is relevant to the actions to deploy, connect, and control the portable generators that are evaluated by Action 2 in this example. It seems that the air compressor is needed to open a vent valve that is used to depressurize a system, so that low pressure makeup flow can be established in Action 1. (I do not know if that system is the reactor coolant system, a steam generator, or some other system.) However, if that is the case, then those activities to procure and install the air compressor (regardless of whether or not they were anticipated) should be documented as part of the Task Context for Action 1, and they should be evaluated as one of the critical tasks for that action. Of course, the analysts' assessment of the scenario-specific performance-influencing factors (PIFs) for that task should account for possible unexpected conditions that require enhanced coordination. The quantification of P_t for Action 1 should also account for uncertainties about the additional time delay. However, the scope of this example analysis does not include Action 1. Therefore, unless this particular observation about the air compressor affects the scenario context for the modeled actions, it is not apparent why it is mentioned in this summary.

If this requirement to procure and install the air compressor is relevant to the scenario context for the modeled actions, the discussion in this section of the worksheet and in the Task Context should clarify why it is relevant.

Why is the unexpected need to procure and install an air compressor relevant to the scenario context for the actions that are evaluated in this example?

109. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Personnel Context

The second major bullet item under "Interteam Coordination" notes that:

"the MCR crew has the initial ultimate decision-making authority which is transferred to the technical support center ***after the center is in operation***." [emphasis added]

The third major bullet item under "Interteam Coordination" notes that:

"The MCR crew coordinates the event mitigation activities before the TSC is in operation. ***After the TSC is in operation***, the TSC, MCR, and Offsite Support Center coordinate the event mitigation activities and manpower assignments." [emphasis added]

The Scenario Timeline in Section A.1 of this worksheet indicates that "offsite emergency response personnel" arrive onsite at approximately 6 hours after the flood occurs, which is after the actions that are evaluated in this example should be completed. The Situation Context in this section of the worksheet also indicates that only those personnel who are onsite when the initiating event occurs are available to perform the needed actions.

The Technical Support Center (TSC) is typically staffed by offsite personnel. Therefore, based on the information in the Scenario Timeline and the Situation Context for this scenario, it does not seem that the TSC is activated or staffed at the time when the modeled actions are needed (i.e., beginning at about 1 hour after the initiating event and continuing for about 5 more hours). Thus, it is not apparent why this summary discusses the transfer of decision-making authority to personnel in the TSC, or why it is necessary (or possible) to coordinate the modeled actions through the TSC.

If the analysis accounts for activation and staffing of the TSC at some time before the modeled actions are needed, or while they are being performed, the Personnel Context should describe who is available in the TSC and what roles they perform to support the modeled actions. The scenario timeline should also document when those personnel arrive onsite and when the TSC is operational.

JWS Note: The discussion of "Interteam Coordination Consideration" in Section C.2 of Worksheet C for Critical Task HFE2-T1 indicates that the action to transport the portable generator is coordinated with personnel in the TSC. The discussion of "Teamwork and Organizational Factors" in Section C.3 of Worksheet C for Critical Task HFE2-T2 indicates that the action to connect the portable generator is coordinated with personnel in the TSC.

Is the TSC staffed and operational at the time when the modeled actions are needed?

If not, why does this summary discuss coordination with the TSC and the Offsite Support Center?

If the TSC is staffed and operational before the modeled actions are needed, or while they are being performed, why do the scenario timeline and the Personnel Context not document when the TSC is activated and who is available to assist with the modeled actions?

110. Appendix C, Section C.3, Example 3, Worksheet A, Section A.3, Scenario Context, Task Context

This summary does not document objective factual information about the cognitive and physical tasks that are involved in Action 4 or Action 2. It does not provide sufficient information for analysts to identify the critical tasks for each action or to evaluate how the scenario-specific performance-influencing factors (PIFs) affect the relevant cognitive failure modes (CFMs) for each task. This summary should describe what must to be done to successfully perform each modeled action, according to the "Guidance for Assessing the Task Context" in Section 3.1.3 of the main report.

Why does this summary not provide objective information about what must to be done to successfully perform each modeled action, including the indications, procedures, other guidance, special tools and equipment, etc. that are used for those actions?

111. Appendix C, Section C.3, Example 3, Worksheet B for HFE1 (Action 4), Section B.2, Task Diagram and Identification of Critical Tasks

The summary of the Task Diagram includes the following tasks that are part of the declaration of the ELAP condition:

- "Plan the various ELAP-related activities."
- "Prioritize resources for ***recovering the EDG*** or performing load shed." [emphasis added]

It is not apparent why these tasks are part of the cognitive activities that are needed to determine that AC power will not be restored and to declare that an ELAP condition applies. It seems that they may be associated with implementation of the actions that are prompted by the ELAP declaration (e.g., shedding selected DC loads, deploying the FLEX portable pumps and generators, etc.). Of course, uncertainty about when, or if, AC power might be restored from the offsite sources, the installed emergency diesel generators (EDGs), or the station blackout (SBO) diesel generator will affect the timing of the decision to declare the ELAP. That decision might also be affected by the supervisors' knowledge of how many people are available to perform the needed actions and how they are engaged during the first hour after all AC power fails. However, it seems that the specific planning and prioritization activities that are noted in these bullet items are associated primarily with post-ELAP cognitive performance, rather than activities which affect the ELAP decision or its timing.

Are these cognitive activities associated with determination that an ELAP condition applies, or are they associated with implementation of the actions that are prompted by the ELAP declaration?

If the second bullet item is relevant to action HFE1, should it refer to deployment of the FLEX portable generators, rather than recovery of power from the EDGs?

112. Appendix C, Section C.3, Example 3, Worksheet B for HFE2 (Action 2), Section B.1, HFE Definition; Worksheet E, Time Uncertainty Analysis of the HFEs

Section B.1 of Worksheet B for HFE1 (Action 4) indicates that the operators must declare that an ELAP condition applies within 1 hour after the loss of all AC power. It notes that:

> "In the event, the MCR crew needs to timely declare an ELAP event because the essential dc power will only last for ***4 hours***." [emphasis added]

Section B.1 of Worksheet B for HFE2 confirms that the success criteria for action HFE2 are:

> "The success criterion of the action is correctly operating FLEX generator to power the 480 VAC emergency buses ***before the depletion of dc power***." [emphasis added]

Section B.1 of Worksheet B for HFE2 also notes that:

> "The action (deploying FLEX generator to charge essential batteries) is one of the few FLEX actions to be ***initiated right after ELAP is declared*** based on the ELAP procedure." [emphasis added]

The example analysis seems to be based on the assumption that all AC power fails immediately after the flood occurs (i.e., at essentially time t = 0). Thus, the functional time window (or the "system time window") for action HFE2 ends when the batteries are depleted at time t = 4 hours. The earliest start of the time window that is available to perform action HFE2 is determined by the time when the ELAP condition is declared.

If the ELAP condition is declared at time t = 1 hour, it seems that the amount of time that is available to perform action HFE2 is 3 hours. Of course, if the ELAP condition is declared sooner, more time is available to perform action HFE2. Thus, the amount of time that is available to perform action HFE2 depends on the uncertainty in the amount of time that is needed to perform action HFE1.

JWS Note: The definition of action HFE1 in Section B.1 of Worksheet B for HFE1 notes that declaration of the ELAP condition also initiates actions to shed DC loads. Section A.2 of Worksheet A indicates that those actions are included in Action 5, which is not evaluated in this example analysis. Success of the deep load shed will extend the battery life beyond 4 hours, and thus also extend the amount of time that is available to perform action HFE2. Because this example analysis does not evaluate Action 5, I used the 4-hour battery depletion time to define the end of the functional time window for action HFE2. Section M.2.6 of Appendix M in NUREG-2198 includes a short discussion of how to account for different battery depletion times in that example quantification of P_t. Please also refer to my September 24, 2020 comments on that example.

JWS Note: Worksheet E does not document the uncertainty distribution for the amount of time that is needed to declare the ELAP condition. In practice, premature declaration that an ELAP condition applies will start a sequence of actions that may limit the operators' options to recover from the initial power failure. Therefore, while it might seem advantageous to declare the ELAP condition as soon as possible in the context of this particular PRA model, in practice, there may be a rather strong incentive to wait as long as possible.

Section B.1 of Worksheet B for HFE2 notes further that:

> "The debris in the transportation route needs to be removed ***before the action can be performed***. Removing debris is considered as a separate important human action because it is performed by a different group of people and affects the deployment of all FLEX equipment." [emphasis added]

Thus, the available time window for the tasks that are modeled by action HFE2 does not actually begin until the ELAP condition is declared ***and*** the debris has been cleared from the portable generator transportation pathway.

In other words, the available time window for action HFE2 must also account for the amount of time that is needed to clear the debris, before personnel can begin to move the generators. In practice, it seems that there may be uncertainty about when the actions to clear the debris will begin and substantial uncertainty about how much time is needed to clear the debris. Thus, the amount of time that is available to perform action HFE2 also depends on the uncertainty in the total amount of time that is needed to clear the debris.

Let

$T_{avaiHFE2}$ = Amount of time that is available to perform action HFE2

$T_{reqdHFE1}$ = Amount of time that is needed to perform action HFE1

T_{reqdDC} = Amount of time that is needed to clear the debris

If the actions to clear the debris begin when the ELAP condition is declared, it seems that the ***minimum*** value of $T_{availHFE2}$ is:

$T_{availHFE2(min)} = 4 - (T_{reqdHFE1} + T_{reqdDC})$

For example, if $T_{reqdHFE1}$ is 1 hour and T_{reqdDC} is 2 hours, $T_{availHFE2(min)}$ is 1 hour.

If the actions to clear the debris begin immediately at time t = 0, it seems that the ***maximum*** value of $T_{availHFE2}$ is:

$T_{availHFE2(max)} = 4 - (T_{reqdHFE1}$ or $T_{reqdDC})$, depending on whichever time is longer

For example, if $T_{reqdHFE1}$ is 1 hour and T_{reqdDC} is 2 hours, $T_{availHFE2(max)}$ is 2 hours.

Of course, in each calculation, the values for $T_{reqdHFE1}$ and T_{reqdDC} must account for the uncertainties in the amount of time that is needed to perform each action.

Section B.1 of Worksheet B for HFE2 does not document the amount of time that is available for action HFE2. Worksheet E indicates that the Time Available for action HFE2 is 4 hours. Based on the discussion in this comment, that 4-hour estimate is clearly wrong. (It would apply only if the operators instantaneously declare an ELAP condition, and all debris is removed instantaneously. The scenario summary indicates that debris must be cleared. Therefore, the possibility that there is no need to remove any debris does not apply during this particular scenario.) Use of the 4-hour estimate for $T_{availHFE2}$ will result in an inappropriately optimistic evaluation of P_t for action HFE2.

JWS Note: Section M.2.6 of Appendix M in NUREG-2198 includes a short discussion of how to account for the time that is needed to remove debris in that example quantification of P_t. Please also refer to my September 24, 2020 comments on that example.

What is the uncertainty in the amount of time that is needed to perform action HFE1?

When do the actions to clear the debris begin?

What is the basis for that starting time?

How much time is needed to clear the debris?

What is the basis for that time, and what is the uncertainty in that estimate?

What is the available time window (T_{avail}) to perform action HFE2?

113. Appendix C, Section C.3, Example 3, Worksheet B for HFE2 (Action 2), Section B.2, Task Diagram and Identification of Critical Tasks

This section notes that:

> "Deploying the FLEX generator to power the 480 VAC emergency buses starts with the MCR giving the order to the ***OSC manager*** to deploy a team to implement the order." [emphasis added]

I could not find a prior reference to the "OSC" or the "OSC manager" in the summary of this scenario or in the Personnel Context that is provided in Section A.3 of Worksheet A.

The Scenario Timeline in Section A.1 of Worksheet A indicates that "offsite emergency response personnel" arrive onsite at approximately 6 hours after the flood occurs, which is after the actions that are evaluated in this example should be completed. The Situation Context in Section A.3 of Worksheet A also indicates that only those personnel who are onsite when the initiating event occurs are available to perform the needed actions.

JWS Note: The discussion of "Cue and Supporting Information" in Section C.2 of Worksheet C for Critical Task HFE2-T2 also notes that "the OSC specifies the individuals to perform the task".

What is the "OSC", and who is the "OSC manager"?

In particular, are the "OSC" and the "OSC manager" part of the normal complement of personnel who are onsite when the flood occurs?

114. Appendix C, Section C.3, Example 3, Worksheet C for Critical Task HFE1-T1, Section C.1, Analysis of Cognitive Activities and Identification of Applicable CFMs

The summary of Cognitive Activities discusses only the cognitive failure mode (CFM) for Decisionmaking. It does not document the analysts' evaluations of the CFMs for Detection, Understanding, Action Execution, and Interteam Coordination.

It is very important that the examples in this report demonstrate systematic and comprehensive application of the IDHEAS-ECA methods and guidance. The intent of Section C.1 in this worksheet is to provide an objective assessment of the CFMs that apply to each critical task from the perspective of the basic macrocognitive functions that are needed for successful human cognitive performance. The "Guidance for Identifying the Applicable CFMs" in Section 3.3.2 and Table 3-3 of the main report emphasize that type of comprehensive, objective assessment. This documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

JWS Note: The discussions of the Cognitive Activities in Section C.1 of Worksheet C in Example 1 provide good examples of the types of analyst considerations that should be documented in this section of the worksheet.

JWS Note: This comment also applies to the incomplete assessments in Section C.1 of Worksheet C for Critical Task HFE2-T1, Section C.1 of Worksheet C for Critical Task HFE2-T2, and Section C.1 of Worksheet C for Critical Task HFE2-T3.

Why does this section not document the analysts' assessments of the CFMs for Detection, Understanding, Action Execution, and Interteam Coordination, according to the guidance in Section 3.3.2?

115. Appendix C, Section C.3, Example 3, Worksheet C for Critical Task HFE1-T1, Section C.2, Task Characterization

The discussion of Cue in this section notes that:

> "The cue to start the task is explicitly stated in the ***SBO procedure***" [emphasis added]

This is not an appropriate characterization of the relevant cues to begin the operators' process to determine whether an ELAP condition should be declared, and to decide when that declaration should be made. In particular, this summary inappropriately implies that the operators are successfully using the station blackout (SBO) procedure immediately after the flood occurs at time t = 0. It also inappropriately implies that the procedure itself provides all of the needed cues to prompt the decision to declare an ELAP (i.e., that no other information about the plant is needed).

To make the decision to declare an ELAP, the operators must first recognize that no AC power is available. They must work their way through the available Emergency Operating Procedures to enter the SBO procedure. They must then use the guidance in that procedure to help them to determine whether the ELAP declaration is warranted. In practice, it is likely that their final decision will also depend on information that is received from personnel who survey the extent of the plant damage and provide feedback about the likelihood that power may be restored from the offsite sources, the emergency diesel generators (EDGs), or the SBO diesel generator. All of those activities are relevant to successful completion of critical task HFE1-T1 and the time that is needed to perform that task.

According to the "Guidance for the Characterization of Critical Tasks" in Section 3.3.1 and Table 3-2 of the main report, this section of the worksheet should identify the actual cues that are available to prompt the desired action. In addition to specific indications and alarms in the Main Control Room, those cues may also involve oral information about the plant status from local operators or other personnel. The summary should document estimates for the times when the oral feedback may occur and the basis for those estimates, since that information will affect the timing of the decision to declare the ELAP.

JWS Note: This comment also applies to the related discussion of Information in this section of the worksheet. It also applies to the Procedure summary, which identifies only the SBO procedure.

Why does this assessment inappropriately presume that the operators successfully enter the SBO procedure immediately after the flood occurs?

Why does this entry not document the actual cues that the operators use to determine that no

AC power is available, and that it is unlikely that power will be restored within one hour (including oral feedback from local operators or other personnel)?

Why does the Procedure summary not list all of the procedures that the operators use for this decision, including the procedures which should lead them to enter the SBO procedure?

116. Appendix C, Section C.3, Example 3, Worksheet C for Critical Task HFE1-T1, Section C.3, Assessment of PIFs

Section A.4 of Worksheet A concludes that the following seven performance-influencing factors (PIFs) are relevant during this scenario:

- Environmental Factors
- Scenario Familiarity
- Information Availability and Reliability
- Task Complexity
- Procedures, Guidance, and Instructions
- Multitasking, Interruption, and Distraction
- Mental Fatigue, and Time Pressure and Stress

The evaluation of cognitive failure mode CFM3 (Failure of Decisionmaking) for critical task HFE1-T1 addresses only the PIFs for Information Availability and Reliability, and Task Complexity. It does not address the possible effects from the other five relevant PIFs. It is important that these examples should demonstrate systematic and comprehensive use of the IDHEAS-ECA methodology and guidance. Therefore, the evaluation of HFE1-T1-CFM3 should address all seven relevant PIFs. If the analysts conclude that the "no impact" attribute applies for each of the other five PIFs in the scenario-specific context of this particular cognitive failure mode (CFM), their rationale for each conclusion should be documented. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis.

JWS Note: This comment also applies to the incomplete evaluations of the PIFs for HFE2-T1-CFM4 in Section C.3 of Worksheet C for Critical Task HFE2-T1, HFE2-T2-CFM4 in Section C.3 of Worksheet C for Critical Task HFE2-T2, and HFE2-T3-CFM4 in Section C.3 of Worksheet C for Critical Task HFE2-T3.

Why does this evaluation address the possible effects from only two PIFs?

In particular, why does this evaluation not address the possible effects from the other five relevant PIFs and document the analysts' rationale for the assigned attributes from each of those PIFs?

117. Appendix C, Section C.3, Example 3, Worksheet C for Critical Task HFE2-T1, Section C.3, Assessment of PIFs; Worksheet C for Critical Task HFE2-T2, Section C.3, Assessment of PIFs; Worksheet C for Critical Task HFE2-T3, Section C.3, Assessment of PIFs; Worksheet A, Section A.4, Initial Assessment of PIFs

The evaluation of cognitive failure mode CFM4 (Failure of Action Execution) for critical task HFE2-T1 documents the analysts' conclusion that attribute TF1 applies for the Teamwork and

Organizational Factors performance-influencing factor (PIF).

The evaluation of cognitive failure mode CFM4 for critical task HFE2-T2 documents the analysts' conclusion that attribute TF2 applies for that cognitive failure mode (CFM). The evaluation of cognitive failure mode CFM4 for critical task HFE2-T3 does not document an assessment of the Teamwork and Organizational Factors PIF.

Section A.4 of Worksheet A concludes that seven PIFs are relevant to the actions that are evaluated during this scenario. Those PIFs are listed in the immediately preceding comment. That list does not include the PIF for Teamwork and Organizational Factors. Since that PIF is excluded from the list in Worksheet A, the analysts apparently did not evaluate its relevance to the other critical tasks and CFMs that are analyzed for each action in this example. Those incomplete assessments do not appropriately demonstrate how the IDHEAS-ECA methods and guidance should be used to perform a systematic, comprehensive, and well-documented evaluation of the scenario-specific factors that may affect human performance.

JWS Note: A preceding comment addresses the lack of documentation in Section A.4 of Worksheet A to justify why each excluded PIF is not relevant to any of the actions that are evaluated during this scenario.

JWS Note: This comment also applies to the evaluation of CFM4 for critical task HFE2-T2 and the evaluation of CFM4 for critical task HFE2-T3. Those evaluations conclude that Training attribute TE4 applies for HFE2-T2-CFM4, and attribute TE2 applies for HFE2-T3-CFM4. Section A.4 of Worksheet A does not list Training as a relevant PIF for this scenario.

For example, it is not apparent why the Teamwork and Organizational Factors PIF is not relevant (or why the "no impact" attribute applies) for HFE2-T3-CFM4 and for HFE1-T1-CFM3 (e.g., for needed local feedback about the plant status before the ELAP declaration is made). It is also not apparent why Training is not relevant (or why the "no impact" attribute applies) for HFE1-T1-CFM3 and HFE2-T1-CFM4.

Why does the list of relevant PIFs in Section A.4 of Worksheet A not include Training, and Teamwork and Organizational Factors?

Why do the evaluations in Section C.3 of Worksheet C for each critical task and CFM in this example analysis not document the analysts' assessments of the PIF attributes for Training, and Teamwork and Organizational Factors?

118. Appendix C, Section C.3, Example 3, Worksheet C for Critical Task HFE2-T2, Section C.3, Assessment of PIFs

The evaluation of Procedures, Guidance, and Instructions notes that:

> "The procedure for aligning buses and connecting the generator ***may not*** have adequate detail. The specifications on some steps ***may not*** match the situation." [emphasis added]

The IDHEAS methodology and guidance emphasize the importance of a realistic evaluation of the factors that affect human performance in the context of the specific event scenario. Therefore, the example analysis should evaluate the applicability and clarity of the actual procedures, guidance, and instructions that would be used during this scenario. The highlighted phrases inappropriately imply that this element of the analysis is speculative, and that it does not involve a thorough evaluation of the procedures and guidance that are actually used.

JWS Note: This comment also applies to the evaluation of Multitasking, Interruption, and Distraction in Section C.3 of Worksheet C for Critical Task HFE2-T3, where it is noted that "this assessment is hypothetical and HRA analysts should identify the potential other tasks that may cause MT2".

Why does this summary imply that this assessment does not account for the actual procedures, guidance, and instructions that would be used during this scenario?

In particular, what specific elements of the actual procedures, guidance, and instructions justify the analysts' conclusion that the PG3 attribute applies during this scenario?

119. Appendix C, Section C.3, Example 3, Worksheet C for Critical Task HFE2-T3, Section C.3, Assessment of PIFs

The evaluation of Training notes that:

> "(Note: This is based on the assumption in 2018 that FLEX equipment is not included in the plant's implementation of the Maintenance Rule. Had the equipment been included in the Maintenance Rule implementation, the generator would have been periodically tested for maintenance purposes and the equipment operators would start and run the generator in testing. In that case the Training PIF would be considered as ***No Impact***.)" [emphasis added]

This note contains an inappropriate speculation about the possible effects from Training during this specific event scenario. It also implies that analysts should assign the "no impact" attribute for Training, simply because a particular piece of equipment is included in a plant's Maintenance Rule program. I disagree very strongly with that assertion.

For example, compliance with the plant's Maintenance Rule program may simply require that personnel test the portable generator periodically to confirm that it starts and runs. The tests may not require that the generator must be connected to the actual plant buses and used to supply their respective electrical loads. Furthermore, it is very unlikely that the tests would require that personnel must control the generator to prevent possible overloading or other electrical transients that may occur as various loads are connected and disconnected during the evolution of this specific scenario. Thus, without knowledge of the actual testing program, it is not appropriate to speculate about the effectiveness of Training for the specific tasks that are needed to successfully perform action HFE2-T3. Furthermore, it is certainly not appropriate for "NRC-approved" guidance to imply that inclusion of equipment in a plant's Maintenance Rule program ensures that Training is always perfect for every conceivable scenario that may be evaluated in a PRA, without a critical scenario-specific assessment of that training.

Why does this assessment imply that the "no impact" attribute for Training applies universally, simply because a particular piece of equipment is included in a plant's Maintenance Rule program?

Why does this assessment contain this speculative note, without a thorough evaluation of the actual plant-specific training for the personnel actions to operate the portable generator and to control its operation during the scenario-specific variations of applied loads?

120. Appendix C, Section C.3, Example 3, Worksheet D, HEP Estimation, Critical Task HFE1-T1

Worksheet D does not show how the human error probability (HEP) for P_c is calculated for critical task HFE1-T1. Thus, this example does not fully demonstrate how the IDHEAS-ECA methodology and the guidance in this report are used to quantify an HEP.

JWS Note: The examples should not presume that analysts use the IDHEAS-ECA software and that the HEP calculations are "invisible" to the user. The examples should demonstrate how each HEP is calculated, so readers clearly understand the intent of the guidance and the quantification process.

Critical task HFE1-T1 evaluates the shift supervisor's actions to declare that an ELAP condition applies within 1 hour after the loss of all AC power.

Worksheet D indicates that the following performance-influencing factor (PIF) attributes apply for cognitive failure mode HFE1-T1-CFM3.

- Information Availability and Reliability INF1
- Task Complexity C25

Thus, it seems that the value of P_c for critical task HFE1-T1 should be determined by the sum of the "base" HEPs for cognitive failure mode CFM3 that apply for the combined effects from attributes INF1 and C25.

Table B-2 in Appendix B indicates that the "base" HEP for Decisionmaking is 5E-02 when the "information is moderately incomplete" condition applies for Information Availability and Reliability attribute INF1.

Table B-3 in Appendix B indicates that the "base" HEP for Decisionmaking is 1.4E-01 when Task Complexity attribute C25 applies.

According to Equation 3.5 in the main report, it seems that P_c for critical task HFE1-T1 should be:

$$P_{cHFE1\text{-}T1} = 1 - [(1 - HEP_{T1\text{-}CFM3INF1}) * (1 - HEP_{T1\text{-}CFM3C25})]$$

$$= 1 - [(1 - 5.0E\text{-}02) * (1 - 1.4E\text{-}01)]$$

$$= 1.83E\text{-}01$$

Is my estimate for P_c correct?

Why does Worksheet D not show how P_c is calculated for critical task HFE1-T1?

121. Appendix C, Section C.3, Example 3, Worksheet D, HEP Estimation, Critical Task HFE2-T1

Critical task HFE2-T1 evaluates personnel actions to transport the portable generator from its storage building to the location where it is connected to the in-plant electrical buses.

Section C.3 of Worksheet C for HFE2-T1 indicates that the following performance-influencing factor (PIF) attributes apply for cognitive failure mode HFE2-T1-CFM4: SF3, C38, ENV11, and TF1. Worksheet D lists only SF3, ENV11, and TF1.

My estimate for P_c accounts for the following PIF attributes for cognitive failure mode HFE2-T1-CFM4.

- Scenario Familiarity SF3
- Task Complexity C38
- Environmental Factors ENV11
- Teamwork and Organizational Factors TF1

Thus, it seems that the value of P_c for critical task HFE2-T1 should be determined by the sum of the "base" HEPs for cognitive failure mode CFM4 that apply for the combined effects from attributes SF3 and C38, modified by the combined weights for ENV11 and TF1.

Table B-1 in Appendix B indicates that the "base" HEP for Action Execution is 3.3E-02 when the "scenario is unfamiliar, rarely performed" condition applies for Scenario Familiarity attribute SF3.

Table B-3 in Appendix B indicates that the "base" HEP for Action Execution is 5E-02 when Task Complexity attribute C38 applies.

The PIF for Environmental Factors determines a weighting factor that is applied to the "base" HEP. Table B-4 in Appendix B indicates that the PIF weight for Action Execution is 2 when Environmental Factors attribute ENV11 applies.

The PIF for Teamwork and Organizational Factors determines a weighting factor that is applied to the "base" HEP. Table B-11 in Appendix B indicates that the PIF weight for Action Execution is 2 when Teamwork and Organizational Factors attribute TF1 applies.

According to Equation 3.4 and Equation 3.5 in the main report, it seems that P_c for critical task HFE2-T1 should be:

$$P_{cHFE2\text{-}T1} = \{1 - [(1 - HEP_{T1\text{-}CFM4SF3}) * (1 - HEP_{T1\text{-}CFM4C38})]\} * [1 + (w_{ENV11} - 1) + (w_{TF1} - 1)]$$

$$= \{1 - [(1 - 3.3E\text{-}02) * (1 - 5.0E\text{-}02)]\} * [1 + (2 - 1) + (2 - 1)]$$

$$= \{8.14E\text{-}02\} * [3]$$

$$= 2.44E\text{-}01$$

Why does Worksheet D not include Task Complexity attribute C38 for cognitive failure mode HFE2-T1-CFM4?

Is my estimate for P_c correct?

Why does Worksheet D not show how P_c is calculated for critical task HFE2-T1?

122. Appendix C, Section C.3, Example 3, Worksheet D, HEP Estimation, Critical Task HFE2-T2

Critical task HFE2-T2 evaluates personnel actions to connect the portable generator to the in-plant electrical buses.

Worksheet D indicates that the following performance-influencing factor (PIF) attributes apply for cognitive failure mode HFE2-T2-CFM4.

- Scenario Familiarity SF3
- Task Complexity C32 and C37
- Procedures, Guidance, and Instructions PG3
- Training TE4
- Teamwork and Organizational Factors TF2

Thus, it seems that the value of P_c for critical task HFE2-T2 should be determined by the sum of the "base" HEPs for cognitive failure mode CFM4 that apply for the combined effects from attributes SF3, C32, and C37, modified by the combined weights for PG3, TE4, and TF2.

A preceding comment addresses my observation that the IDHEAS-ECA methodology does not describe how analysts should account for scenario-specific conditions that adversely affect multiple attributes for a particular PIF (i.e., like C32 and C37 for T2-CFM4). Without further guidance, I have assumed that the effects from those attributes are additive (i.e., that the "base" HEPs from each attribute should be added to quantify the total HEP for T2-CFM4).

Table B-1 in Appendix B indicates that the "base" HEP for Action Execution is 3.3E-02 when the "scenario is unfamiliar, rarely performed" condition applies for Scenario Familiarity attribute SF3.

Table B-3 in Appendix B indicates that the "base" HEP for Action Execution is 5E-03 when Task Complexity attribute C32 applies.

Table B-3 in Appendix B indicates that the "base" HEP for Action Execution is 1E-02 when Task Complexity attribute C37 applies.

The PIF for Procedures, Guidance, and Instructions determines a weighting factor that is applied to the "base" HEP. Table B-9 in Appendix B indicates that the PIF weight for Action Execution is 2.2 when Procedures, Guidance, and Instructions attribute PG3 applies.

The PIF for Training determines a weighting factor that is applied to the "base" HEP. Table B-10 in Appendix B indicates that the PIF weight for Action Execution is 6.1 when Training attribute TE4 applies.

The PIF for Teamwork and Organizational Factors determines a weighting factor that is applied to the "base" HEP. Table B-11 in Appendix B indicates that the PIF weight for Action Execution is 1.5 when Teamwork and Organizational Factors attribute TF2 applies.

According to Equation 3.4 and Equation 3.5 in the main report, it seems that P_c for critical task HFE2-T2 should be:

$$P_{cHFE2\text{-}T2} = \{1 - [(1 - HEP_{T2\text{-}CFM4SF3}) * (1 - HEP_{T2\text{-}CFM4C32}) * (1 - HEP_{T2\text{-}CFM4C37})]\} * [1 + (w_{PG3} - 1) + (w_{TE4} - 1) + (w_{TF2} - 1)]$$

$$= \{1 - [(1 - 3.3E\text{-}02) * (1 - 5.0E\text{-}03) * (1 - 1.0E\text{-}02)]\} * [1 + (2.2 - 1) + (6.1 - 1) + (1.5 - 1)]$$

$$= \{4.75E\text{-}02\} * [7.8]$$

$$= 3.70E\text{-}01$$

Is my estimate for P_c correct?

Why does Worksheet D not show how P_c is calculated for critical task HFE2-T2?

123. Appendix C, Section C.3, Example 3, Worksheet D, HEP Estimation, Critical Task HFE2-T3

Critical task HFE2-T3 evaluates personnel actions to operate the portable generator and control its loading.

Worksheet D indicates that the following performance-influencing factor (PIF) attributes apply for cognitive failure mode HFE2-T3-CFM4.

- Training TE2
- Multitasking, Interruption, and Distraction MT2

The assessments in Section C.3 of Worksheet C for Critical Task HFE2-T3 do not identify any adverse effects on the Action Execution cognitive failure mode CFM4 from the three "base" PIFs (i.e., Scenario Familiarity, Information Availability and Reliability, and Task Complexity). Thus, it seems that the value of P_c for critical task HFE2-T3 should be determined by "base" HEP for cognitive failure mode CFM4 when the "base" PIFs have "no impact", modified by the combined weights for TE2 and MT2.

The introduction to Appendix B indicates that the IDHEAS-ECA guidance recommends minimum HEPs of 1E-04 for Detection or Action Execution, and 1E-03 for Understanding, Decsionmaking, or Interteam Coordination.

The PIF for Training determines a weighting factor that is applied to the "base" HEP. Table B-10 in Appendix B indicates that the PIF weight for Action Execution is 1.5 when Training attribute TE2 applies.

The PIF for Multitasking, Interruption, and Distraction determines a weighting factor that is applied to the "base" HEP. Table B-13 in Appendix B indicates that the PIF weight for Action Execution is 2.8 when the "moderate interruptions" condition applies for attribute MT2.

According to Equation 3.4 and Equation 3.5 in the main report, it seems that P_c for critical task HFE2-T3 should be:

$$P_{cHFE2\text{-}T3} = HEP_{T3\text{-}CFM4\text{-}NI} * [1 + (w_{TE2} - 1) + (w_{MT2} - 1)]$$

$$= (1.0E\text{-}04) * [1 + (1.5 - 1) + (2.8 - 1)]$$

$$= (1.0E\text{-}04) * [3.3]$$

$$= 3.3E\text{-}04$$

Is my estimate for P_c correct?

Why does Worksheet D not show how P_c is calculated for critical task HFE2-T3?

124. Appendix C, Section C.3, Example 3, Worksheet D, HEP Estimation, HFE2: Fail to Use FLEX Generator

Action HFE2 will fail if any one of the three critical tasks HFE2-T1, HFE2-T2, or HFE2-T3 is not performed correctly. Thus, it seems that P_c for action HFE2 should be:

$$P_{cHFE2} = 1 - (1 - P_{cHFE2\text{-}T1}) * (1 - P_{cHFE2\text{-}T2}) * (1 - P_{cHFE2\text{-}T3})$$

$$= 1 - (1 - 2.44E\text{-}01) * (1 - 3.70E\text{-}01) * (1 - 3.3E\text{-}04)$$

$$= 5.24E\text{-}01$$

Is my estimate for P_c correct?

Why does Worksheet D not show how the total P_c is calculated for action HFE2?

125. Appendix C, Section C.3, Example 3, Worksheet D, HEP Estimation, HFE2: Fail to Use FLEX Generator, General Comment

If any of my estimates of P_c for critical tasks HFE2-T1, HFE2-T2, and HFE2-T3 are wrong, it is apparent that I did not correctly interpret the basic guidance for how to quantify P_c or how to use the numerical values from the tables in Appendix B. That situation would further reinforce why I think that it is very important for Worksheet D to show how P_c is quantified for each example, so that analysts clearly understand how to quantify the "base" human error probability (HEP) from the three "base" performance-influencing factors (PIFs), how to treat the combined effects from multiple attributes for a specific PIF, how to treat the PIF attributes that have "no impact", how to use the tables in Appendix B, how to apply the PIF weights, how to combine the HEPs for all cognitive failure modes (CFMs) for each critical task, and how to quantify the overall value of P_c.

126. Appendix C, Section C.3, Example 3, Worksheet E, Time Uncertainty Analysis of the HFEs, Action HFE1

Action HFE1 evaluates the shift supervisor's actions to declare that an ELAP condition applies.

Worksheet E indicates that the Time Available (T_{avail}) for this action is 1 hour.

Worksheet E indicates that the amount of time that is needed to perform the action (T_{reqd}) is represented by a five-point discrete probability distribution. However, the actual distribution is not documented in the worksheet. Thus, analysts cannot understand how P_t is quantified for action HFE1.

In practice, premature declaration that an ELAP condition applies will start a sequence of actions that may limit the operators' options to recover from the initial power failure. Therefore, while it might seem advantageous to declare the ELAP condition as soon as possible in the context of this particular PRA model, in practice, there may be a rather strong incentive to wait as long as possible. Thus, it is very important that Worksheet E should clearly document the basis for the assigned times and probabilities for T_{reqd}.

Action HFE2 evaluates personnel actions to transport, connect, and operate the FLEX portable generator after the ELAP condition is declared. A preceding comment on Section B.1 of Worksheet B for HFE2 (Action 2) addresses estimation of the amount of time that is available (T_{avail}) to perform action HFE2. The evaluation of P_t for action HFE1 is determined by the uncertainty in T_{reqd} for that action. However, the time that remains available to perform action HFE2 also depends on when the ELAP condition is declared. Thus, the uncertainty distribution for T_{reqd} for action HFE1 also affects the uncertainty distribution for T_{avail} for action HFE2.

Why does Worksheet E not document the uncertainty distribution for T_{reqd} for action HFE1, including the basis for the assigned times and probabilities?

Why does Worksheet E not show how P_t is calculated for action HFE1?

127. Appendix C, Section C.3, Example 3, Worksheet E, Time Uncertainty Analysis of the HFEs, Action HFE2

Action HFE2 evaluates personnel actions to transport, connect, and operate the FLEX portable generator. The functional time window for action HFE2 begins when the ELAP condition is declared, and it ends when the station batteries are depleted at time t = 4 hours after the loss of all AC power. The example analysis seems to be based on the assumption that all AC power fails immediately after the flood occurs (i.e., at essentially time t = 0).

Worksheet E indicates that the Time Available (T_{avail}) for this action is 4 hours.

A preceding comment on Section B.1 of Worksheet B for HFE2 (Action 2) addresses estimation of T_{avail} for action HFE2. The 4-hour estimate that is shown in this worksheet is clearly wrong. It would apply only if the operators instantaneously declare an ELAP condition, and all debris is instantaneously removed from the generator transportation pathway. As noted in that comment, the actual uncertainty distribution for T_{avail} depends on the amount of time that is needed to perform action HFE1, the time when the actions to clear the debris begin, and the amount of time that is needed to clear the debris.

For the purpose of this comment, I will use three point-estimate values for T_{avail} for action HFE2.

- Case 1: Debris removal begins at time t = 0, the ELAP condition is declared at time t = 1 hour, and all debris is removed before that time. In this case, T_{avail} = 3 hours.

- Case 2: Debris removal begins at time t = 0, the ELAP condition is declared at time t = 1 hour, and 2 hours are needed to clear the debris. In this case, T_{avail} = 2 hours.

- Case 3: The ELAP condition is declared at time t = 1 hour, debris removal begins when the ELAP condition is declared, and 2 hours are needed to clear the debris. In this case, T_{avail} = 1 hour.

Worksheet E indicates that the time that is needed (T_{reqd}) to transport and stage the generator is "45 +/- 15 mins". It also indicates that T_{reqd} to connect the generator and start its operation is "30 +/- 15 mins". The worksheet does not document the bases for these estimates. It also does not document the form of the uncertainty distribution for each estimate, or how the estimates are combined to derive the total uncertainty distribution for T_{reqd} for action HFE2.

For the purpose of this comment, I will assume that each of the times is characterized by a normal probability distribution and that the "+/-" values represent the standard deviation in each estimate. The sum of those distributions is a normal probability distribution for T_{reqd} that has a mean value of 75 minutes (1.25 hours) and a standard deviation of approximately 21 minutes (0.35 hour).

Based on these estimates, I calculated the following values for P_t for action HFE2.

- If T_{avail} = 3 hours, P_t = 0
- If T_{avail} = 2 hours, P_t = 1.69E-02
- If T_{avail} = 1 hour, P_t = 7.60E-01

These estimates show why it is very important that Worksheet E should document the uncertainty distributions for T_{avail} and T_{reqd}, including the technical bases for each distribution, and it should show how P_t is quantified for this action.

What is the available time window (T_{avail}) to perform action HFE2?

What is the basis for the uncertainty distribution for T_{avail}?

What is the total amount of time that is needed (T_{reqd}) to perform action HFE2?

What is the uncertainty distribution for T_{reqd}, and what is the basis for that distribution?

Why does Worksheet E not show how P_t is calculated for action HFE2?

128. Appendix D, Introduction to the IDHEAS-ECA Software, Calculate P_c

Item 1.1 in the Display column in Table D-1 notes that:

> "***Three critical tasks*** and checkboxes to include and exclude the critical tasks" [emphasis added]

This seems to indicate that the software allows an analyst to specify a maximum of three critical tasks for each modeled human failure event (HFE).

Is that correct?

If so, why is the software limited to only three critical tasks per HFE?

129. Appendix D, Introduction to the IDHEAS-ECA Software, Calculate P_t

This section indicates that one million Monte Carlo samples are used for the quantification of P_t. It is also noted that:

> "For certain distribution combinations, one million samples may produce slightly different P_t results each time. The differences are ***considered to have negligible effects on the HEP***." [emphasis added]

In practice, one million samples should be adequate to achieve reasonable convergence in the results for most analyses. However, it is certainly true that the results from some combinations of very broad distributions may exhibit variability from run-to-run. I do not understand why the software authors assert that these differences will always have "negligible effects on the HEP". I also do not understand why this statement is needed in this rather high-level summary of the software and its basic functions. For example, the detailed software users' guide should alert analysts to the types of situations that may produce significant run-to-run variability, so they can test for the numerical effects on the estimate for P_t (e.g., by performing several runs). The users' guide should also provide recommended techniques to achieve better convergence in the results if the run-to-run variability has a potentially significant effect on P_t.

What is the technical basis for the software authors' assertion that run-to-run variability in the Monte Carlo results will always have "negligible effects on the HEP"?

Why is this assertion included in this high-level summary of the software and its basic functions?

What guidance does the detailed software users' guide contain to alert analysts to situations when the run-to-run variability may have a potentially significant effect on P_t?

Does the users' guide recommend techniques to achieve better convergence in those situations?

130. Appendix D, Introduction to the IDHEAS-ECA Software, Documentation

The introduction to this appendix notes that:

> "The software is used to calculate the HEP of an HFE and to ***document the calculation***." [emphasis added]

The introductory summary of the Documentation function notes that:

> "All parameters that the analyst entered to calculate P_c and P_t and the ***other relevant information*** are documented in a rich text file to be integrated in the overall analysis document." [emphasis added]

The Documentation section notes that:

> "***After the analyst completes the HEP calculation*** for an HFE, the software provides two options for documentation. The first option is to generate a document in rich text format that

> has all ***parameters*** specified to calculate the HFE's HEP, the HEPs (i.e., HEPs of CFM, critical tasks, Pc, Pt, and HFE), and the ***other relevant information (e.g., HEP impact of each PF attribute)***. This option provides a convenient way for the analyst to integrate to the information into the final analysis report." [emphasis added]

This summary seems to address only an archival catalog of the parameter values that are used for each human error probability (HEP) calculation.

Several preceding comments address the fact that it is essential for analysts to clearly document the rationale for their decisions throughout the analysis process. That documentation is very important for understanding the reasons for analyst-to-analyst variability when the methodology is used, and it is important for reviewers' understanding of the technical bases for each analysis. For example, analysts should document their rationale for selecting the specific cognitive failure modes (CFMs) that apply for each critical task, including their rationale for why any excluded CFMs are not relevant. Analysts should also document their rationale for selecting each applicable attribute for every performance-influencing factor (PIF) that is relevant to the quantification of the "base" HEP, and their rationale for selecting the attributes for the applied PIF weights (including the reasons why a "no impact" attribute applies). Finally, analysts should document the technical bases for their estimates of the amount of time that is available for an action (T_{avail}) and the amount of time that is needed to perform the action (T_{reqd}), including the bases for the range and shape of the uncertainty distributions for those estimates.

In other words, the analysis documentation should not only contain an archive of ***what*** specific parameter values were used for each HEP calculation, it should more importantly also contain the analysts' rationale for ***why*** those particular values were selected.

How does the software facilitate this necessary documentation?

For example, does the software require that an analyst must complete a text-based documentation summary for every selected value before that value can be used in the quantification process?

131. Appendix D, Introduction to the IDHEAS-ECA Software, Quantification of Uncertainty

The material in this appendix does not mention how, or whether, the software facilitates the quantification of uncertainty.

The summary for "Calculate P_t" indicates that the software accepts a variety of input uncertainty distributions for the amount of time that is available for an action (T_{avail}) and the amount of time that is needed to perform the action (T_{reqd}). A Monte Carlo sampling routine is used for the convolution of those distributions to quantify P_t.

For many practical IDHEAS-ECA applications, the value of P_t may be quite small (or even zero). The uncertainty in the overall human error probability (HEP) is then determined by the uncertainty in P_c. However, it is also important to characterize and quantify the uncertainty in P_c, even if the contribution from P_t is rather large. The summary for "Calculate P_c" does not indicate whether the software accepts input uncertainty distributions for the "base" HEPs from Table B-1 through Table B-3 of Appendix B or input uncertainty distributions for the performance-influencing factor (PIF) weights from Table B-4 through Table B-15. It is also not apparent whether the software contains the necessary Monte Carlo sampling algorithms to combine those distributions and to quantify the uncertainty in P_c, according to Equation 3.4 and Equation 3.5 in the main report.

Does the IDHEAS-ECA software accept input probability distributions for quantification of the uncertainty in P_c, and does it contain the necessary algorithms to combine those distributions according to the quantification formulas?

If not, why does the software not facilitate the quantification of uncertainty at each level of the calculation for P_c (e.g., for each cognitive failure mode, each critical task, and the overall HEP)?