U.S.NRC
United States Nuclear Regulatory Commission
Protecting People and the Environment

RIL-2021-XX

DRAFT - Integrated Human Event Analysis System for Human Reliability Data (IDHEAS-DATA)

Date Published:

Prepared by:
Jing Xing
Y. James Chang
Jonathan DeJesus Segarra

Research Information Letter
Office of Nuclear Regulatory Research

DISCLAIMER

ABSTRACT

The U.S. Nuclear Regulatory Commission (NRC) staff developed the Integrated Human Event Analysis System-General Methodology (IDHEAS-G) to address the staff requirements memorandum (SRM) M061020 on proposing human reliability analysis (HRA) models and guidance for the NRC to use. Human performance data is an essential element of IDHEAS-G. This report documents human performance and error data identified through an extensive literature review and the process used to generalize the data for IDHEAS-G. The report also supports SRM-090204B on the development and use of an HRA database.

The data documented in this report include operational and simulator data in the nuclear domain, operational data of human performance from non-nuclear domains, experimental data in the literature, expert judgment on human error probabilities (HEPs) in the nuclear domain, and others (e.g., statistical data, ranking, frequencies of errors or causal factors). The data are classified according to the NRC's IDHEAS-G methodology. Most of the data documented in this report have been generalized to develop the NRC's IDHEAS-ECA HRA method. IDHEAS-ECA, in conjunction with the data in IDHEAS-DATA, completes the scientific basis for the NRC's risk-informed decisionmaking processes when dealing with human reliability. The report will be expanded as new data become available. The data provide a basis for continuous improvements of HRA methods.

EXECUTIVE SUMMARY

The U.S. Nuclear Regulatory Commission (NRC) staff developed the Integrated Human Event Analysis System-General Methodology (IDHEAS-G) to address the staff requirements memorandum (SRM) M061020 on proposing human reliability analysis (HRA) models and guidance for the NRC to use. Human reliability data is an essential element of IDHEAS-G. This report documents human performance and error data identified through an extensive literature review and the process used to generalize the data for IDHEAS-G. The report also supports SRM-090204B on the development and use of an HRA database.

DHEAS-DATA is one of the three tasks the Human Factors and Reliability Branch, Division of Risk Analysis, Office of Nuclear Regulatory Research (RES/DRA/HFRB) performed to modernize the NRC's human reliability analysis (HRA) techniques with a solid scientific, technology inclusive foundation and a strong data basis. The three tasks included the development of IDHEAS-G (IDHEAS-general methodology) to provide the scientific foundation, IDHEAS-DATA for the data-basis, and IDHEAS-ECA (IDHEAS for Event and Condition Assessment) and IDHEAS-At Power (An Integrated Human Event Analysis System for Nuclear Power Plant Internal Events At-Power Application) HRA methods for applying HRA. The three tasks together support the reliability element of the NRC's Principles of Good Regulation and create a science-based framework for continuously improving human reliability analysis methods that support risk-informed decisionmaking at the NRC.

Human reliability is a significant contributor to overall plant risk, and HRA results directly affect the NRC's risk-informed decisions. Many conventional HRA methods were not developed with a strong data basis; therefore, their results can be associated with large uncertainties. From time to time, the uncertainties are large enough to affect the reliability of regulatory decisions. Further, many conventional HRA methods lack the data basis to support HRA applications for emerging technologies, such as for Diverse and Flexible Coping Strategies (FLEX) and digital instrumentation and control. The IDHEAS-series products address these issues by being human-centered (thus being expandable to novel situations) and data-based.

This report documents the human reliability and performance data collected through a large-scale literature review. The data were classified based on the scientific foundations described in IDHEAS-G and generalized to support the development of the IDHEAS-ECA method. The data were from various sources, including operational experience and studies of human reliability and performance in nuclear and non-nuclear domains. The large data diversity and quantity establish a strong data basis. The data generalization process and scientific foundation provide a sound process to include new HRA data. This report will be updated when more HRA data becomes available.

The data are generalized into 27 tables, referred to as IDHEAS-DATA TABLEs (IDTABLEs). IDTABLE-1 through IDTABLE-20 document the data related to the effects of the performance influencing factors (PIFs) documented in IDHEAS-G. IDTABLE-21 includes data associated with optimal human reliabilities. IDTABLE-22 concerns the combined effects of more than one PIF. IDTABLE-23 and IDTABLE-24 are data for assessing the uncertainty distribution of the time required to perform a task. The information documented in IDTABLE-23 and IDTABLE-24 are a small portion of the collected data. The NRC has begun work to analyze a much larger portion of the literature to support guidance development on specifying the uncertainty distributions of task completion times. IDTABLE-25 and IDTABLE-26 are information on task dependency and error recovery, respectively. Finally, IDTABLE-27 documents the situations where a high

percentage of human failures occurred. IDTABLE-27 helps HRA analysts understand the main drivers to human error to help them quickly perceive similar conditions in their analyses.

TABLE OF CONTENTS

Contents

ABSTRACT ... iii
EXECUTIVE SUMMARY ... v
TABLE OF CONTENTS ... vii
LIST OF FIGURES ... xi
LIST OF TABLES ... xi
ACRONYMS AND TERMS ... xiii
1 INTRODUCTION TO IDHEAS-DATA ... 1-1
1.1. Background ... 1-1
1.2. Purposes of this Report ... 1-2
1.3. Intended Use ... 1-2
1.4. Related NRC Documents ... 1-2
1.5. Organization of this Report ... 1-2
1.6. Status of the report ... 1-2
2 THE STRUCTURE AND DEVELOPMENT OF IDHEAS-DATA ... 2-1
2.1. IDHEAS-G Framework ... 2-1
2.2. Construct of IDHEAS-DATA ... 2-5
2.3. Identification and Review of Data Sources ... 2-7
2.4. Generalization of Human Error Data in IDHEAS-DATA ... 2-11
2.4.1. Generalizing Human Error Data to IDHEAS-DATA Base HEP Tables ... 2-12
2.4.2. IDHEAS-DATA PIF Weight IDTABLE-4 through IDTABLE-20 ... 2-14
2.4.3. IDTABLE-21 for the Lowest HEPs ... 2-15
2.4.4. IDTABLE-22 for PIF Interaction ... 2-18
2.4.5. IDTABLE-23 for Distribution of Time Needed in completing a human action ... 2-19
2.4.6. IDTABLE-24 for Modification to Time Needed ... 2-20
2.4.7. IDTABLE-25 for Dependency of Human Actions ... 2-21
2.4.8. IDTABLE-26 for Recovery of Human Actions ... 2-23
2.4.9. IDTABLE-27 for Main drivers to Human Failure Events ... 2-24
2.5. Integration of human error data to inform human error probabilities ... 2-26
2.5.1. Overview of an Application-specific IDHEAS method ... 2-26
2.5.2. The process of integrating human error data ... 2-28
2.5.3. Approaches of integrating human error data for IDHEAS-ECA ... 2-29
3 RESULTS ... 3-1

3.1. Overview of the Data Sources and Summary of the Generalized Data in IDHEAS-DATA 3-1
3.1.1. IDTABLE-1 for Scenario Familiarity 3-1
3.1.2. IDHEAS-DATA IDTABLE-2 for Information Completeness and Reliability...........3-4
3.1.3. IDHEAS-DATA IDTABLE-3 for Task Complexity 3-8
3.1.4. IDHEAS-DATA IDTABLE-4 for Workplace Accessibility and Habitability...........3-12
3.1.5. IDHEAS-DATA IDTABLE-5 for Workplace Visibility..3-13
3.1.6. IDHEAS-DATA IDTABLE-6 for Workplace Noise...3-16
3.1.7. IDHEAS-DATA IDTABLE-7 for Workplace Temperature 3-19
3.1.8. IDHEAS-DATA IDTABLE-8 for Resistance to Personnel Movement 3-22
3.1.9. IDHEAS-DATA IDTABLE-9 for System and Instrument & Control Transparency to Personnel..3-26
3.1.10. IDHEAS-DATA IDTABLE-10 for Human-System Interface.............................3-29
3.1.11. IDHEAS-DATA IDTABLE-11 for Portable Equipment, Tools, and Parts.........3-32
3.1.12. IDHEAS-DATA IDTABLE-12 for Staffing 3-34
3.1.13. IDHEAS-DATA IDTABLE-13 for Procedure, Guidance, and Instruction.........3-37
3.1.14. IDHEAS-DATA IDTABLE-14 for Training and Experience 3-39
3.1.15. IDHEAS-DATA IDTABLE-15 for Team and Organization Factors..................3-41
3.1.16. IDHEAS-DATA IDTABLE-16 for Work Process 3-43
3.1.17. IDHEAS-DATA IDTABLE-17 for Multitasking, Interruption, and Distraction ...3-44
3.1.18. IDHEAS-DATA IDTABLE-18 for Mental Fatigue...3-47
3.1.19. IDHEAS-DATA IDTABLE-19 for Time Pressure and Stress...........................3-48
3.1.20. IDHEAS-DATA IDTABLE-20 for Physical Demands......................................3-51
3.1.21. IDHEAS-DATA IDTABLE-21 for Lowest HEPs of the Cognitive Failure Modes .3-53
3.1.22. IDHEAS-DATA IDTABLE-22 for PIF Interaction 3-55
3.1.23. IDHEAS-DATA IDTABLE-23 for Distribution of Time Needed........................3-57
3.1.24. IDHEAS-DATA IDTABLE-24 for Modification of Time Needed.......................3-61
3.1.25. IDHEAS-DATA IDTABLE-25 for Dependency of Human Actions...................3-63
3.1.26. IDHEAS-DATA IDTABLE-26 for Recovery of Human Actions........................3-65
3.1.27. IDHEAS-DATA IDTABLE-27 for Main Drivers to Human Failure Events........3-66
3.2. Integration of the generalized data for IDHEAS-ECA 3-67
3.2.1. Assessing and organizing the datapoints...3-68
3.2.2. Detaching multi-component human error data 3-70
3.2.3. Estimating the lowest HEP 3-71
3.2.4. Reasonableness checking and Calibration of the estimated HEP 3-73
4 DISCUSSION AND CONCLUDING REMARKS 4-1

4.1. Generalization of human error data from various sources 4-1
4.2. Integration of the generalized data to inform IDHEAS-ECA 4-1
4.3. Limitations in the current status of IDHEAS-DATA...4-1
4.4. Perspectives of HRA data sources...4-1
4.5. Concluding Remarks 4-1
5 REFERENCES 5-1

LIST OF FIGURES

Figure 2-1 Illustration of IDHEAS-G Data Generalization and Integration 2-5
Figure 2-2 The Process of Generalize Human Error Data to IDEHAS-DATA.....................2-11
Figure 2-3 Illustration of HEP varying as a function of the measure of a PIF attribute 2-27
Figure 3-1 The human error rates for the lowest HEP of Failure of Detection.....................3-72

LIST OF TABLES

Table 2-1 Macrocognitive Functions and Their Basic Elements..2-2
Table 2-2 Performance-Influencing Factors in IDHEAS-G 2-3
Table 2-3 Sample of IDHEAS-DATA IDTABLE-3 – Base HEPs for Task Complexity........2-14
Table 2-4 Sample of IDHEAS-DATA IDTABLE-17 – Error Rates for PIF Multitasking, Interruptions, and Distractions 2-15
Table 2-5 Sample of IDHEAS-DATA IDTABLE-21 – Lowest HEP 2-17
Table 2-6 Sample of IDHEAS-DATA IDTABLE-22 – PIF Interaction.................................2-19
Table 2-7 Sample of IDHEAS-DATA IDTABLE-23 – Distribution of Time Needed 2-20
Table 2-8 Sample of IDHEAS-DATA IDTABLE-24 – Modification of Time Needed 2-21
Table 2-9 Sample of IDHEAS-DATA IDTABLE-25 – Dependency of Human Actions 2-23
Table 2-10 Sample of IDHEAS-DATA IDTABLE-26 – Recovery of Human Actions 2-24
Table 2-11 Sample of IDHEAS-DATA IDTABLE-27 – Main Drivers to Human Failure Events 2-25
Table 3-1 The operator response times in SGTR events [96]. 3-58
Table 3-2 Time needed analysis based on the example Table 3-1data 3-58
Table 3-3 The time to isolate the ruptured steam generator in actual events and simulated events. 3-59
Table 3-4 Comparing the response time of simple and complicated SGTR events 3-60
Table 3-5 The crew performance time in a basic SGTR event of a Westinghouse 3-loop PWR [98] 3-61
Table A1-1 Attribute Identifiers and Descriptions for PIF Scenario Familiarity 1-1
Table A1-2 IDHEAS-DATA IDTABLE-1 – Base HEPs for PIF Scenario Familiarity 1-1
Table A2-1 Attribute Identifiers and Descriptions for PIF Information Availability and Reliability 2-4
Table A2-2 IDHEAS-DATA IDTABLE-2 – Base HEPs for PIF Information Availability and Reliability 2-4
Table A3-1 Attribute Identifiers and Descriptions for PIF Task Complexity 3-1
Table A3-2 IDHEAS-DATA IDTABLE-3 – Base HEPs for PIF Task Complexity 3-2
Table A4-1 Attribute Identifiers and Descriptions for PIF Workplace Accessibility and Habitability 4-1
Table A4-2 IDHEAS-DATA IDTABLE-4 – PIF Weights for Workplace Accessibility and Habitability 4-1
Table A5-1 Attribute Identifiers and Descriptions for PIF Workplace Visibility........................5-1
Table A5-2 IDHEAS-DATA IDTABLE-5 – PIF Weights for Workplace Visibility 5-1
Table A6-1 Attribute Identifiers and Descriptions for PIF Workplace Noise............................6-1
Table A6-2 IDHEAS-DATA IDTABLE-6 – PIF Weights for Workplace Noise 6-1
Table A7-1 Attribute Identifiers and Descriptions for PIF Cold/Heat/Humidity........................7-1

Table A7-2 IDHEAS-DATA IDTABLE-7 – PIF Weights for Cold/Heat/Humidity 7-1
Table A8-1 Attribute Identifiers and Descriptions for PIF Resistance to Physical Movement .8-1
Table A8-2 IDHEAS-DATA IDTABLE-8 – PIF Weights for Resistance to Physical Movement 8-1
Table A9-1 Attribute Identifiers and Descriptions for PIF System and I&C Transparency to Personnel 9-1
Table A9-2 IDHEAS-DATA IDTABLE-9 – PIF Weights for System and I&C Transparency to Personnel 9-1
Table A10-1 Attribute Identifiers and Descriptions for PIF Human-System Interfaces........10-1
Table A10-2 IDHEAS-DATA IDTABLE-10 – PIF Weights for Human-System Interfaces ...10-1
Table A11-1 Attribute Identifiers and Descriptions for PIF Equipment and Tools 11-1
Table A11-2 IDHEAS-DATA IDTABLE-11 – PIF Weights for Equipment and Tools...........11-1
Table A12-1 Attribute Identifiers and Descriptions for PIF Staffing 12-1
Table A12-2 IDHEAS-DATA IDTABLE-12 – PIF Weights for Staffing................................12-1
Table A13-1 Attribute Identifiers and Descriptions for PIF Procedures, Guidelines, and Instructions 13-1
Table A13-2 IDHEAS-DATA IDTABLE-13 – PIF Weights for Procedures, Guidelines, and Instructions 13-1
Table A14-1 Attribute Identifiers and Descriptions for PIF Training 14-1
Table A14-2 IDHEAS-DATA IDTABLE-14 – PIF Weights for Training...............................14-1
Table A15-1 Attribute Identifiers and Descriptions for PIF Team and Organization Factors.15-1
Table A15-2 IDHEAS-DATA IDTABLE-15 – PIF Weights for Team and Organization Factors 15-1
Table A16-1 Attribute Identifiers and Descriptions for PIF Work Processes 16-1
Table A16-2 IDHEAS-DATA IDTABLE-16 – PIF Weights for Work Processes..................16-1
Table A17-1 Attribute Identifiers and Descriptions for PIF Multitasking, Interruptions, and Distractions 17-1
Table A17-2 IDHEAS-DATA IDTABLE-17 – PIF Weights for Multitasking, Interruptions, and Distractions 17-1
Table A18-1 Attribute Identifiers and Descriptions for PIF Mental Fatigue 18-1
Table A18-2 IDHEAS-DATA IDTABLE-18 – PIF Weights for Mental Fatigue 18-1
Table A19-1 Attribute Identifiers and Descriptions for PIF Time Pressure and Stress........19-1
Table A19-2 IDHEAS-DATA IDTABLE-19 – PIF Weights for Time Pressure and Stress ...19-1
Table A20-1 Attribute Identifiers and Descriptions for PIF Physical Demands 20-1
Table A20-2 IDHEAS-DATA IDTABLE-20 – PIF Weights for Physical Demands 20-1
Table A21-1 IDHEAS-DATA IDTABLE-21 – Lowest HEP..21-1
Table A22-1 IDHEAS-DATA IDTABLE-22 – PIF Interaction 22-1
Table A23-1 IDHEAS-DATA IDTABLE-23 – Distribution of Time Needed 23-1
Table A24-1 IDHEAS-DATA IDTABLE-24 – Modification to Time Needed to Complete a Human Action 24-1
Table A25-1 IDHEAS-DATA IDTABLE-25 – Instances and Data on Dependency of Human Actions 25-1
Table A26-1 IDHEAS-DATA IDTABLE-26 – Instances and Data on Recovery Actions......26-1
Table A27-1 IDHEAS-DATA IDTABLE-27 – Empirical Evidence on Main Drivers of Human Failure Events 27-1

ACRONYMS AND TERMS

AC	alternating current
ADAMS	Agency wide Documents Access and Management System
ASP	accident sequence precursor (program)
CFM	cognitive failure mode
CT	critical task
D	*detection* (one of the five macrocognitive functions)
DM	*decisionmaking* (one of the five macrocognitive functions)
E	*action execution* (one of the five macrocognitive functions)
ECA	event and condition assessment
ELAP	extended loss of AC power
EOC	error of commission
EOL	end of life
EOO	error of omission
EOP	emergency operating procedure
FLEX	flexible and coping strategies
FSG	FLEX support guideline
HEP	human error probability
HFE	human failure event
HRA	human reliability analysis
HSI	human-system interface
IDHEAS	Integrated Human Event Analysis System
IDHEAS-DATA	Integrated Human Event Analysis System for Human Reliability Data
IDHEAS-ECA	Integrated Human Event Analysis System for Event and Condition Assessment
IDHEAS-G	General Methodology of an Integrated Human Event Analysis System
IDTABLE	IDHEAS-DATA TABLE
IHA	important human action
I&C	instrumentation and control
LOCA	loss-of-coolant accident
MCR	main control room
NPP	nuclear power plant

NRC	U.S. Nuclear Regulatory Commission
PDP	positive displacement pump
PIF	performance-influencing factor
PRA	probabilistic risk assessment
psig	pounds per square inch gauge
RCP	reactor coolant pump
RCS	reactor coolant system
RIL	Research Information Letter
RO	reactor operator
SACADA	Scenario Authoring, Characterization, and Debriefing Application
SDP	significance determination process
SSCs	structures, systems, and components
T	*interteam coordination* (one of the five macrocognitive functions)
TSC	technical support center
U	*understanding* (one of the five macrocognitive functions)
NRC	U.S. Nuclear Regulatory Commission
P_c	error probability due to CFMs
P_t	error probability due variability in T_{avail} and T_{reqd}
T_{avail}	time available
T_{reqd}	time needed
$\mu_{T_{avail}}$	mean of T_{avail}
$\sigma_{T_{avail}}$	standard deviation of T_{avail}
$\mu_{T_{reqd}}$	mean of T_{reqd}
$\sigma_{T_{reqd}}$	standard deviation of T_{reqd}

1 INTRODUCTION TO IDHEAS-DATA

1.1. Background

Probabilistic risk assessment (PRA) results and insights support risk-informed regulatory decision making. The U.S. Nuclear Regulatory Commission (NRC) continues to improve the robustness of PRA, including human reliability analysis (HRA) through many activities. To date, there have been about fifty HRA methods developed worldwide to estimate human error probabilities (HEPs) to support PRA. Yet, the use of empirical data for HEP estimation has been limited due to the lack of data and discrepancies in the formats of available data and the relevance to nuclear power plant operation. The lack of a strong data basis in HRA methods challenges the validity of HEP estimation.

The NRC staff developed the General Methodology of an Integrated Human Event Analysis System (IDHEAS-G)[1]. IDHEAS-G integrates the strengths in existing HRA methods, enhances the cognitive basis for HRA, and builds the capability of using human error data to improve HEP estimation. IDHEAS-G provides a hierarchical structure to analyze and assess the reliability of human actions. IDHEAS-G models human performance with five macrocognitive functions: *Detection, Understanding, Decisionmaking, Action execution, and Interteam coordination*. IDHEAS-G defines a set of cognitive failure modes (CFMs) for each macrocognitive function to describe the various ways of failing the macrocognitive function. IDHEAS-G also has a performance-influencing factor (PIF) structure that consists of a set of PIFs and their attributes to represent the context of a human event. IDHEAS-G analyzes an event in progressively more detailed levels: event scenario, human actions, critical tasks of the actions, macrocognitive functions and CFMs of the tasks, and PIFs and the associated attributes. This structure provides an intrinsic interface to generalize various sources of human error data for HEP estimation.

Along with the development of IDHEAS-G, the NRC staff developed IDHEAS-DATA, a data structure that generalizes and documents human error data from various sources into the IDHEAS-G CFMs and PIF attributes. The staff analyzed the source information of human error data reported in operational databases and literature, identified the CFMs and PIF attributes associated with the data, and documented the data according to the CFMs and PIF attributes. Developing IDHEAS-DATA has been a continuous effort as more human error data are identified from the literature and new data becomes available. The data, once sufficiently populated, can provide a basis for estimating HEPs.

In 2019, the NRC staff developed the IDHEAS for Event and Condition Assessment (IDHEAS-ECA) method based on IDHEAS-G. The first version of the IDHEAS-ECA method is documented in an NRC Research Information Letter (RIL), RIL-2020-02[2]. The method is to be used for HRA in the NRC's Events and Conditions Assessment (ECA) of nuclear power plants (NPPs). IDHEAS-ECA models human errors in a task with five CFMs, that is, the failure of the five macrocognitive functions in IDHEAS-G and has all the IDHEAS-G PIFs, but with fewer PIF attributes from IDHEAS-G for practical applications. IDHEAS-ECA uses a set of base HEPs and PIF weights to calculate HEPs of the CFMs of a human action for the given context. In developing IDHEAS-ECA, the NRC staff integrated the human error data populated in IDHEAS-DATA to estimate the base HEPs and PIF weights.

1.2. Purposes of this Report

This report describes the process used for generalizing human error data from various sources, summarizes the generalized data, and presents the generalized data in IDHEAS-DATA as of 2019. The purposes of this report are to:

(1) present the IDHEAS-DATA framework and the process of generalizing data into IDHEAS-DATA,

(2) share IDHEAS-DATA with the HRA community, and

(3) document the foundation of the base HEPs and PIF weights in IDHEAS-ECA.

1.3. Intended Use

The intended users of IDHEAS-DATA are NRC staff involved in PRA applications and researchers and HRA practitioners in the HRA community. This report provides the data foundation for IDHEAS-ECA for those who use IDHEAS-ECA and query the data basis. Also, IDHEAS-DATA can serve as the "hub" for HRA data exchanging and synthesis, which may be of interest to those who want to use human error data for HRA.

1.4. Related NRC Documents

Readers may acquire additional information in understanding IDHEAS-DATA and its use by obtaining and reading the following NRC documents:

- IDHEAS-G (NUREG-2198) [1]
- IDHEAS-ECA (RIL-2020-02) [2]
- Expert elicitation for FLEX HRA [3]

1.5. Organization of this Report

This report is organized as follows:

- Chapter 1 is a high-level introduction to IDHEAS-DATA.
- Chapter 2 describes the IDHEAS-DATA framework and the process of generalizing human error data to IDHEAS-DATA.
- Chapter 3 provides a summary of the data generalized in IDHEAS-DATA as of 2019.
- Chapter 4 discusses the limitations and uncertainties in IDHEAS-DATA as well as the pathways to improve data use in HRA.
- Chapter 5 has the references for the source articles of the data in IDHEAS-DATA.
- Appendix A1 through Appendix A20 present the generalized data in IDHEAS-DATA as of 2019, one for each PIF; Appendix A21 presents the human error data to inform the lowest HEPs; and Appendix A22 presents the data about the combined effects of multiple PIFs. Tables A23 and A24 concern assessing uncertainty of time needed. Tables A25, A26, and A27 cover dependency, recovery, and main drivers of performance.

1.6. Status of the report

The report is expected to be periodically updated as more human error data are generalized and new data become available. This DRAFT version of Appendix A presents human error data generalized in the 27 IDHEAS-DATA IDTABLEs. Note that the datapoints in the IDTABLEs

have not been independently verified for their accuracy and appropriateness. They are being made available to the public in this Research Information Letter only for the purpose of communicating information and demonstrating the data basis of IDHEAS-ECA. It is not recommended that these DRAFT IDTABLEs be used by HRA practitioners without first verifying the data validity.

2 THE STRUCTURE AND DEVELOPMENT OF IDHEAS-DATA

The lack of sufficient human reliability data has limited the empirical basis for HEP estimation in HRA methods. For a given context, the HEP of a human task can be calculated as the number of times the task fails divided by the total number of times the task is performed. Most HRA methods use a quantification model to estimate HEPs; the quantification models typically consist of base HEPs for a set of human failure modes or typical human tasks and PIF multipliers to adjust the base HEPs. In addition, many sources of human error data have not been used for HRA due to discrepancies in the formats of available data and relevance to the domain of the human performance that the HRA methods intended to model. Human error data are available from task performance in various domains, in different formats, and at a range of levels of details. Most of the human error data either cannot be directly used for HRA or they are formatted to support only one application-specific HRA method.

In the NRC's IDHEAS project, the NRC staff developed IDHEAS-G [1] as a general HRA methodology for developing application-specific HRA methods. The IDHEAS-G framework and its taxonomy of CFMs and PIFs are generic and flexible, so they were chosen to generalize human error data from various sources to IDHEAS-DATA. The NRC staff integrated the generalized data in IDHEAS-DATA to develop the IDHEAS-ECA HRA method [2] in 2019. This chapter will describe the process of generalizing human error data to IDHEAS-DATA.

IDHEAS-G incorporates advances made in cognitive and behavioral science in the past decades. IDHEAS-G has a macrocognition model with a basic set of CFMs, a PIF structure, and a quantification model to quantify the effect of PIFs on the HEP of a CFM. IDHEAS-G represents human failures with a basic set of CFMs and represents human event context with a set of PIFs. The IDHEAS-G quantification model calculates the HEP of a human action based on the CFMs and PIFs relevant to the action. The basic set of CFMs represents human failures at three levels of detail (i.e., failures of macrocognitive functions, failures of the processors in each macrocognitive function, and behaviorally-observable failure modes of the processors). The PIF structure represents the event context at two levels of detail: PIFs and their attributes. The underlying cognitive mechanisms can link CFMs and PIFs at any level of detail. Thus, IDHEAS-G is inherently capable of generalizing human error data of different task types and different levels of detail to inform HEP quantification. The CFMs and PIF structure together form a framework for generalizing human error data from various sources and integrating them to support the IDHEAS-G quantification model. The structured data can inform expert judgment, Bayesian estimates, or direct calculation of HEPs.

2.1. IDHEAS-G Framework

IDHEAS-G [1] implements its cognition model to the full span of the general HRA process. The HRA process of IDHEAS-G consists of four stages:

(1) Stage 1—Scenario analysis. The purpose of this stage is to understand the event and collect information about human actions from broad perspectives. This includes developing an operational narrative, analyzing the scenario context, and identifying important human actions (i.e., the ones considered in a PRA). IDHEAS-G provides a structured process to query and document the qualitative information used as the foundation of HEP quantification.

(2) Stage 2—Modeling of important human actions. The purpose of this stage is to model important human actions for structured analysis and HEP quantification. This includes identifying and characterizing critical tasks in an important human action, representing potential task failure with CFMs, and representing the context of the important human

action with PIFs. IDHEAS-G provides guidelines for task analysis, as well as a basic set of CFMs and a comprehensive taxonomy of PIFs from its cognition model.

(3) Stage 3—HEP quantification. The purpose of this stage is to estimate the HEP for important human actions. IDHEAS-G provides several approaches to HEP estimation, along with the human error data generalized in the IDHEAS-G framework.

(4) Stage 4—Integrative analysis. While Stages 2 and 3 analyze individual important human actions, Stage 4 analyzes all the important human actions as a whole. This includes addressing the dependencies between important human actions and documenting uncertainties in the event and its analysis. IDHEAS-G provides supplementary guidance for uncertainty analysis by consolidating existing guidelines.

The Macrocognition Model

The macrocognition model describes the cognitive and behavioral process of success or failure of a task. The model explains the cognitive process of human performance in applied work domains where human tasks are complex and often involve multiple individuals or teams. The model is described as follows:

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

Table 2-1 shows the basic cognitive elements (i.e., processors) for the macrocognitive functions. The detailed description of the elements can be found in Chapter 2 of the IDHEAS-G report [1].

Table 2-1 Macrocognitive Functions and Their Basic Elements

Detection	Understanding	Decisionmaking	Action Execution	Interteam Coordination
D1. Initiate detection – Establish the mental model for information to be detected **D2**. Select, identify, and attend to sources of information **D3**. Perceive, recognize and classify information	**U1**. Assess/select data **U2**. Select/adapt /develop the mental model **U3**. Integrate data with the mental model to generate the outcome of understanding (situational awareness, diagnosis, resolving conflicts)	**DM1**. Adapt the infrastructure of decisionmaking **DM2**. Manage the goals and decision criteria **DM3**. Acquire and select data for decisionmaking **DM4**. Make decision (judgment, strategies, plans)	**E1**. Assess action plan and criteria **E2**. Develop or modify action scripts **E3**. Prepare or adapt infrastructure for action implementation **E4**. Implement action scripts	**T1**. Establish or adapt interteam coordination infrastructure **T2**. Manage information **T3**. Maintain shared situational awareness **T4**. Manage resources **T5**. Plan interteam collaborative activities

D4. Verify and modify the outcomes of detection **D5**. Retain, document/record, or communicate the outcomes	**U4**. Verify and revise the outcome through iteration of U1, U2, and U3 **U5**. Export the outcome	**DM5**. Simulate or evaluate the decision or plan **DM6**. Communicate and authorize the decision	**E5**. Verify and adjust execution outcomes	**T6**. Implement decisions and commands **T7**. Verify, modify, and control the implementation

The Performance-Influencing Factor Structure

The PIF structure describes how various factors in the event context affect the success or failure of human tasks. PIFs affect cognitive mechanisms and increase the likelihood of macrocognitive function failure. The PIF structure is independent of HRA applications and systematically organizes PIFs to minimize inter-dependency or overlapping of the factors. The PIF structure is described as follows:

1. PIF category: PIFs are classified into four categories, corresponding to characteristics of environment and situation, systems, tasks, and personnel.
2. PIFs: Each category has high-level PIFs describing specific aspects of the environment and situation, systems, tasks, or personnel.
3. PIF attributes: These are the specific traits of a performance influencing factor. A PIF attribute represents a poor PIF state that challenges cognitive mechanisms and increases the likelihood of errors in cognitive processes.

Table 2-2 shows the PIFs within the four categories.

Table 2-2 Performance-Influencing Factors in IDHEAS-G

Environment and situation	System	Personnel	Task
• Work Location Accessibility and Habitability • Workplace Visibility • Noise in Workplace and Communication Pathways • Cold/Heat/Humidity • Resistance to Physical Movement	• System and Instrumentation and Control (I&C) Transparency to Personnel • Human-System Interface (HSI) • Equipment and Tools	• Staffing • Procedures, Guidelines, and Instructions • Training • Team and Organization Factors • Work Processes	• Information Availability and Reliability • Scenario Familiarity • Multi-Tasking, Interruptions and Distractions • Task Complexity • Mental Fatigue • Time Pressure and Stress • Physical Demands

The Human Error Probability Quantification Model

IDHEAS-G provides guidance on several ways to estimate HEPs, one of which is its HEP model to estimate the HEP of a human action. The estimation has two parts: estimating the error probabilities attributed to the CFMs (P_c) and estimating the error probability attributed to the uncertainties and variability in the time available and time needed to perform the HFE (P_t). The estimation of the HEP is the probabilistic sum of P_c and P_t:

$$P = 1 - (1 - P_c)(1 - P_t) \qquad (2.1)$$

In Equation (2.1), P is the probability of the HFE being analyzed (i.e., the HEP), and P_c and P_t have already been defined. Note the following:

- P_t can also be viewed as the probability that the time needed to perform an action exceeds the time available for that action, as determined by the success criteria. P_t assumes that actions are performed at a normal pace without complications and does not account for the increased likelihood of a human error due to time pressure. Time pressure is treated as a PIF and contributes to P_c.
- P_c assumes that the time to perform the HFE is sufficient. Sufficient time means that the HFE can be successfully performed within the time window that the system allows. If operators' responses are as trained, then the time available to complete the action is sufficient. P_c captures the probability that the human action does not meet the success criteria due to human errors made in the problem-solving process.

Estimation of P_c

P_c is the probabilistic sum of the HEPs of all the CFMs of the critical tasks in a human action. The probability of a CFM applicable to the critical task is a function of the PIF attributes associated with the critical task. The calculation of the probability of a CFM for any given set of PIF attributes, provided that all the PIF impact weights and base HEPs are obtained, is estimated as:

$$\begin{aligned} P_{CFM} &= P_{CFM_{Base}} \cdot \left(1 + \sum_{i=1}^{n} (w_i - 1)\right) \cdot C \cdot \frac{1}{Re} \\ &= \frac{P_{CFM_{Base}} \cdot \big(1 + (w_1 - 1) + (w_2 - 1) + \cdots + (w_n - 1)\big) \cdot C}{Re} \end{aligned} \tag{2.2}$$

The terms in Equation (2.2) are defined as follows:

- $P_{CFM_{Base}}$ is the base HEP of a CFM for the given attributes of the following three PIFs: *information availability and reliability*, *scenario familiarity*, and *task complexity*. $P_{CFM_{Base}}$ is also calculated as the probabilistic sum of the base HEPs for the three PIFs:

  $$P_{CFM_{Base}} = 1 - [(1 - P_{INF})(1 - P_{SF})(1 - P_{TC})] \tag{2.3}$$

  where P_{INF}, P_{SF}, and P_{TC} are the base HEPs for *information availability and reliability*, *scenario familiarity*, and *task complexity*, respectively.

- w_i is the PIF impact weight for the given attributes of the remaining 17 PIFs and is calculated as:

  $$w_i = \frac{ER_{PIF}}{ER_{PIF_{Base}}} \tag{2.4}$$

  where ER_{PIF} is the human error rate at the given PIF attribute and $ER_{PIF_{Base}}$ is the human error rate when the PIF attribute has no impact. The human error rates used in Equation (2.4) are obtained from empirical studies in the literature or operational databases that measured the human error rates while varying the PIF attributes of one or more PIFs. C is a factor that accounts for the interaction between PIFs, and it is set to 1 for the linear combination of PIFs impacts unless there are data suggesting otherwise.

- Re is a factor that accounts for the potential recovery from failure of a critical task, and it is set to 1 by default.

2.2. Construct of IDHEAS-DATA

Various sources of human error data provide instances of human errors, error rates (i.e., percent of errors), or task-related performance measures of human actions, tasks, or failure modes. The human error data are generally measured at a specific context. To use different sources of data together to inform HEPs, the NRC staff developed IDHEAS-DATA to generalize human error data into a common format. The construct of IDHEAS-DATA is based on IDHEAS-G.

IDHEAS-G is inherently capable of generalizing human error data of various sources because (1) IDHEAS-G can model any human task with its basic set of CFMs, (2) the CFMs are structured in different levels of details, and (3) the PIF structure models the context of a human action with high-level PIFs and detailed PIF attributes. Thus, the NRC staff used IDHEAS-G to develop the construct of IDHEAS-DATA to generalize various sources of human error data. For example, two data sources have human error data for different kinds of tasks and in different contexts, but the failure of the tasks can be represented with the applicable IDHEAS-G CFMs, and the context can be represented with the relevant PIF attributes. Thus, both data sources provide human error information with respect to the common sets of CFMs and PIF attributes. Generalization of human error data refers to the process of mapping the data source into the corresponding CFMs and PIFs. Figure 2-1 illustrates this approach.

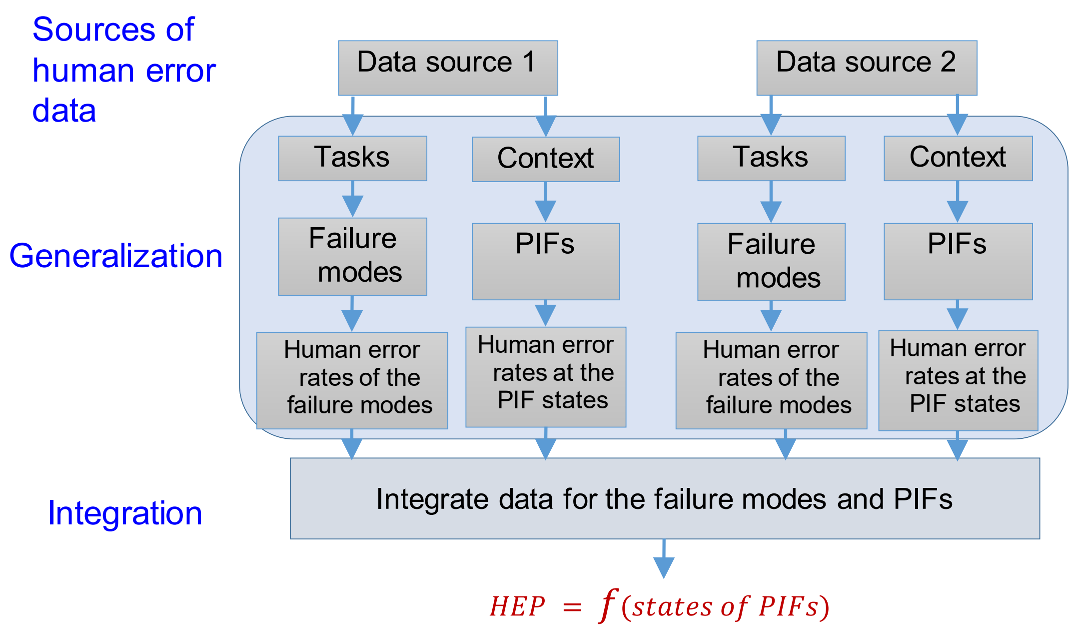

Figure 2-1 Illustration of IDHEAS-G Data Generalization and Integration

In addition to calculate P_c based on CFMs and PIFs, the IDHEAS-G HEP quantification model calculates P_t based on the time available and time needed for a human action. The HEP quantification model also addresses crediting recovery of human failures in an event. Moreover, IDHEAS-G has a dependency model to evaluate the effect of dependency between human actions on HEPs. IDHEAS-DATA is intended to document data sources in these areas as well.

Overall, IDHEAS-DATA includes 27 tables, referred to as IDHEAS-DATA TABLEs (IDTABLEs), each documenting the data in one element of IDHEAS-G. The IDTABLEs are listed as follows:

IDTABLE-1 – Base HEPs for Scenario Familiarity

IDTABLE-2 – Base HEPs for Information Availability and Reliability

IDTABLE-3 – Base HEPs for Task Complexity

IDTABLE-4 – PIF attribute weights for Workplace Accessibility and Habitability

IDTABLE-5 – PIF attribute weights for Workplace Visibility

IDTABLE-6 – PIF attribute weights for Noise in workplace and Communication Pathways

IDTABLE-7 – PIF attribute weights for Cold, Heat, and Humidity

IDTABLE-8 – PIF attribute weights for Resistance to Physical Movement

IDTABLE-9 – PIF attribute weights for System and I&C Transparency to Personnel

IDTABLE-10 – PIF attribute weights for Human-System Interfaces

IDTABLE-11 – PIF attribute weights for Equipment, Tools, and Parts

IDTABLE-12 – PIF attribute weights for Staffing

IDTABLE-13 – PIF attribute weights for Procedures, Guidelines, and Instructions

IDTABLE-14 – PIF attribute weights for Training

IDTABLE-15 – PIF attribute weights for Team and Organization Factors

IDTABLE-16 – PIF attribute weights for Work Processes

IDTABLE-17 – PIF attribute weights for Multi-tasking, Interruptions, and Distractions

IDTABLE-18 – PIF attribute weights for Mental Fatigue

IDTABLE-19 – PIF attribute weights for Time Pressure and Stress

IDTABLE-20 – PIF attribute weights for Physical Demands

IDTABLE-21 – Lowest HEPs of the CFMs

IDTABLE-22 – PIF Interaction

IDTABLE-23 – Distribution of Task Completion Time

IDTABLE-24 – Modification of Task Completion Time

IDTABLE-25 – Instances and Data on Dependency of Human Actions

IDTABLE-26 – Instances and Data on Recovery of Human Actions

IDTABLE-27 – Main Drivers to Human Failure Events

IDTABLE-1 to IDTABLE-IDTABLE-3 are Base HEP Tables. They document human error rates for base HEPs. The data of human error rates from various sources are analyzed for the applicable CFMs and relevant attributes of the three base PIFs.

IDTABLE-4 to IDTABLE-IDTABLE-20 are PIF Impact Tables. They document human error rates for the CFMs at different PIF attributes of the rest 17 PIFs. The data sources contain human error rates or task performance measures varying with specific PIF attributes. The attribute weight can be inferred from the data in which human error rates were measured as a PIF attribute was varied from a no or low impact status to a high impact status.

IDTABLE-21 is for Lowest HEPs of the CFMs. It documents human error rates when the tasks were performed under the condition that none of the known PIF attributes was present so that all the PIFs presumably had no impact on human errors. The data inform the lowest HEPs for the CFMs.

IDTABLE-22 is for PIF Interaction. It documents human error data on PIF interaction. The data are from the studies in which human error rates were measured as two or more PIF attributes varied independently as well as jointly. The data informs the PIF interaction factor C in the HEP quantification model (Equation (2.2)).

IDTABLE-23 is for Distribution of Task Completion Time, i.e., time needed to perform a human action. IDHEAS-G has a time uncertainty model that calculates P_t as the convolution of the distributions of time needed and time available. The data can be used to validate the IDHEAS-G time uncertainty model and inform the estimation of the time needed distribution.

IDTABLE-24 is for Modification to Task Completion Time. It documents empirical data on how various factors modify the time needed to complete a task. The IDHEAS-G time uncertainty model requires analysts to estimate the distribution of time needed for a human action. Many factors such as whether or environmental conditions can modify the center, range, and/or shape of the time distribution. IDTABLE-IDTABLE-24 provides the empirical basis for analysts to estimate the time needed distribution under different contexts.

IDTABLE-25 is for Dependency of Human Actions. It documents instances and empirical data on dependency between human actions. IDTABLE-IDTABLE-25 provides the technical basis and reference information for HRA analysts to evaluate dependency between human actions.

IDTABLE-26 is for Recovery of Human Actions. It documents instances of recovery actions. Currently, the IDHEAS-G HEP quantification model uses the factor Re to represent crediting recovery. The information can help HRA analysts to identify and assess and credit recovery actions.

IDTABLE-27 is for Main Drivers to Human Failure Events. It documents empirical evidence on main drivers to human failures in nuclear power plant events. The information should guide HRA analysts to capture the main drivers and to not overlook important drivers in human events.

The details of the IDTABLEs are described in later sections of this report.

2.3. Identification and Review of Data Sources

Since the 1950s, much human error data has been available in various work domains such as aerospace, aviation, manufacturing, and health care. Many cognitive behavioral studies produced human error data in controlled experimental contexts. Moreover, human performance data in nuclear power plant operations have become available in the last two decades. Several human performance databases have been developed to systematically collect operator performance data in NPPs for HRA. Such efforts include the SACADA database [3] developed by the NRC and the Human Reliability Data Extraction (HuREX) database [4] developed by the Korea Atomic Energy Research Institute. In addition, many HRA expert elicitation studies

produced expert judgment of HEPs for specific applications. While individual sources of human error data may not be enough to yield HEPs for all kinds of human tasks under a large breath of contexts, consolidating the available data and using the data together would yield more robust and valid HEPs.

Ideally, the data to inform HEPs would have the following features:

- The known numerator and denominator of human error rates are collected within the same context.
- Human error rates are measured repetitively to minimize uncertainties in the data.
- Human error rates are collected for a variety of personnel so that the data can represent average personnel or operators.
- Human error data are collected for a range of task types or failure modes and combinations of PIFs.

Such ideal data do not exist. However, these features can be used as criteria to evaluate real data for their applicability to HRA. Along with the development of IDHEAS-G, the NRC staff documented human error data in the literature and human performance databases. The data sources include the following categories:

A. Nuclear simulator data (e.g., SACADA) and operational data (e.g., German Maintenance human error data)

B. Operation performance data from other domains (e.g., air traffic control operational errors)

C. Experimental data reported in the literature

D. Expert judgment data

E. Inference data (statistical data, ranking, categorization, etc.)

The NRC staff examined the data for their ability to inform HEPs. The following are several types of human error data with examples to demonstrate if and how the data can be used to inform HEP estimation.

<u>Human error rates with known PIFs</u>

This type of data provides the numerator and denominator of human error rates for types of tasks performed in the same context or in a known range of contexts. Such data can inform the base HEPs for the CFMs (i.e., $P_{CFM_{Base}}$) relevant to the tasks. The following are two examples:

(1) Quantification of unsatisfactory task performance in NPP operator simulator training, as collected in the SACADA database by the NRC staff. The SACADA database was built with the same macrocognitive model as that in IDHEAS-G and collects operator task performance for different types of failures in various contexts. The different types of failures can be mapped to the detailed level CFMs in IDHEAS-G, and the various contexts can be mapped to the IDHEAS-G PIF attributes. Thus, the SACADA database can inform the base HEPs of IDHEAS-G CFMs and the quantitative effects of some PIF attributes.

(2) The analysis of human errors in maintenance operations of German NPPs. Preischl and Hellmich [4, 5] studied human error rates for various basic tasks in maintenance operations. The following are some example human error rates they reported:

- 1/490 for operating a circuit breaker in a switchgear cabinet under normal conditions
- 1/33 for connecting a cable between an external test facility and a control cabinet
- 1/36 for reassembly of component elements
- 1/7 for transporting fuel assemblies

This type of data from operational databases inherits uncertainties in the data collection process. For example, the definitions of human failure vary from one database to another, so caution is needed when aggregating human error rates from different sources.

Human error rates with unknown or mixed context

This type of data reports statistically calculated human error rates for specific tasks across a mixture of contexts. Such data cannot inform HEPs of the failure modes because neither the failure modes nor the context was specified. The data could represent the best or worst possible scenarios or the average scenario. This type of data can be used to validate the distribution of HEPs obtained by other means.

HEPs estimated through expert judgment

This type of data is not true human error data. They are generated through a formal expert elicitation process, representing the beliefs of the representative technical community on the likelihood of human failure for a given HRA application. Nevertheless, expert judgment has been widely used in risk-informed applications. The resulting estimates of HEPs bear validity and regulatory assurance if the judgment was obtained through a formal, scientifically founded expert elicitation process. This type of data can be used to inform the central tendency and range of HEPs for the context in which the expert judgment was made.

An example of an expert elicitation process used to estimate HEPs is the judgment of HEPs of the crew failure modes in the IDHEAS At-Power Application [6]. The method has 14 crew failure modes, which are a subset of IDHEAS-G behaviorally observable failure modes. A very limited set of PIF attributes is considered for each crew failure mode. An expert panel estimated the HEP distributions of the crew failure modes for the combinations of the PIF attributes.

This type of data has a limitation in that the full context in which the HEPs were estimated is often not well documented. Because expert judgment is typically elicited for a very specific domain of application and the expert panel consists of experienced domain experts, the expert panel makes its own assumptions about the context. For example, in the expert elicitation of HEPs for the IDHEAS At-Power Application [6], the expert panel assumed that NPP operators perform control room tasks by following procedures, and they would make a correct diagnosis with procedures as long as they have the right information. This assumption may not be true for tasks performed outside control rooms. Thus, caution is needed when generalizing expert judgment HEPs to other applications.

Quantification of PIF effects

Many sources present the changes in human error rates when varying the states of one or more PIFs. Such data can inform the quantification of PIF effects in the IDHEAS-G quantification model. The following are several examples:

- NUREG/CR-5572, “An Evaluation of the Effects of Local Control Station Design Configurations on Human Performance and Nuclear Power Plant Risk,” issued September 1990 [7], estimated the effects of local control station design configurations on human performance and NPPs. It estimated that HEP = 2 x 10^{-2} for ideal conditions and HEP = 0.57 for challenging conditions with poor HSIs and distributed work locations.
- Prinzo et al. [8, 9] analyzed aircraft pilot communication errors and found that the error rate increased nonlinearly with the complexity of the message communicated. The error rate was around 4 percent for an information complexity index of 4 (i.e., the number of messages transmitted per communication), 30 percent for an index of 12, and greater than 50 percent for indices greater than 20.
- Patten et al. [10] studied the effect of task complexity and experience on driver performance. The PIF states of the tasks manipulated in the experiment were low experience versus high experience, and low complexity versus high complexity. The mean error rates were 0.12, 0.21, 0.25, and 0.32 respectively for the four combinations of PIF states: low complexity and high experience, low complexity and low experience, high complexity and high experience, high complexity and low experience.

When documenting this type of data, the objective description of PIF states needs to be carefully considered. For example, the PIF state of “high complexity” in one data source can be referred to as “low complexity” in another data source. The NRC staff found that PIF attributes more accurately represent the actual context than the subjective assessment of “high” or “low” PIF states. In fact, using PIF attributes can make the definition for PIF states more objective.

PIF interaction

Most HRA methods treat the combined effects of PIFs on HEPs as the multiplication of the effects of the individual PIFs. Xing et al. [11] reviewed a limited set of cognitive literature in which human error rates were measured, as two or more PIFs varied independently and jointly. They observed that the combined effect of PIFs fits better to the addition than the multiplication of the individual PIF effects. In fact, the broad cognitive literature indicates that the combined effect is not simply the addition or multiplication of individual PIF effects. Instead, the interaction between PIFs may not fit to a single rule and can vary greatly for different combinations of PIFs. The interaction effect can be inferred from human error rates that are collected in a single study or database and with more than one PIF varying independently and jointly.

The significance or ranking of error types and causal factors

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

- Virovac et al. [12] analyzed human errors in airplane maintenance and found that the prevalent factors with frequent occurrence in human errors are communication (16 percent), equipment and tools (12 percent), work environment (12 percent), and complexity (6.5 percent).
- Kyriakidis et al. [13] analyzed U.K. railway accidents caused by human errors and calculated proportions of PIFs in the accidents. They reported that the most frequent

PIFs in the accidents were safety culture (19 percent), familiarity (15 percent), and distraction (13 percent).

The above examples are just a few of a large body of human error data we have documented so far. We also observed the consistency between the results obtained in controlled cognitive experiments and those from complex nuclear scenario simulation. Given the limited amount of nuclear operation data, the NRC staff generalized human error data in all the source categories and integrated them for estimating HEPs in complex nuclear scenarios.

2.4. Generalization of Human Error Data in IDHEAS-DATA

This section introduces the process of generalizing human error data. All the numeric values in this section are for demonstrating the process and their practical use in HRA applications is not recommended.

Human error data generalization is mapping the context and task from the data source onto the IDHEAS-G elements (e.g., CFMs and PIFs) and documenting them in the IDHEAS-DATA Tables. The process of data generalization is essentially the same as that of performing a qualitative HRA using IDHEAS-G. The following process, as illustrated in Figure 2-2, is adapted from IDHEAS-G for generalizing human error data:

- Analyzing the data source. This includes identifying the tasks of which human error information is reported, analyzing the context, characterizing the tasks and assessing the time uncertainties of the tasks.
- Mapping the data onto the IDHEAS-DATA structure. This includes representing the reported human errors of the tasks with applicable CFMs and representing the context of the tasks with PIF attributes.
- Analyzing recovery of human failures and dependency between human actions for events. Such information is often available in operational and simulation data.
- Documenting uncertainties in the data source and the mapping process.

The IDHEAS-G report (NUREG-2198) [1] has detailed guidance on the process above. Different elements of the process are tailored from IDHEAS-G for mapping human error data into different IDHEAS-DATA Tables.

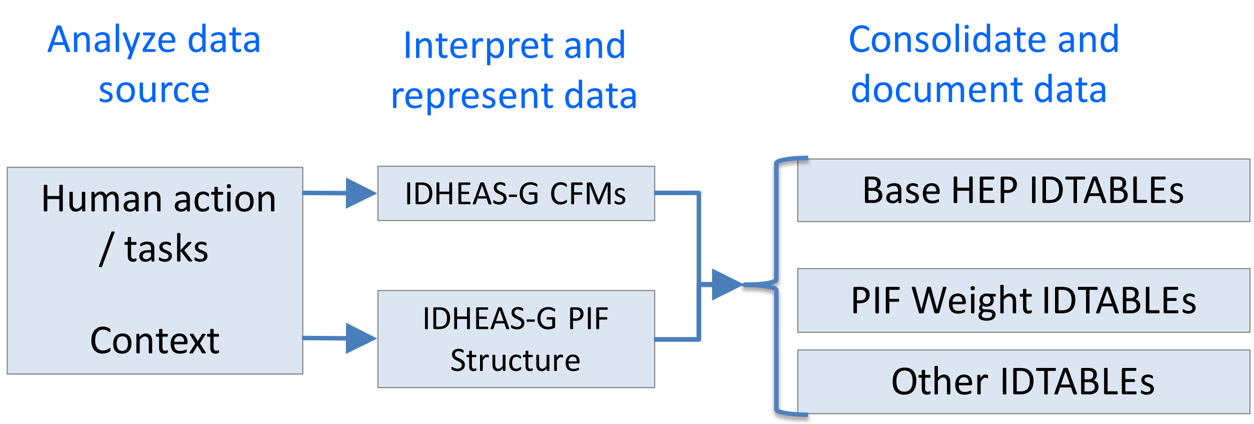

Figure 2-2 The Process of Generalize Human Error Data to IDEHAS-DATA

2.4.1. Generalizing Human Error Data to IDHEAS-DATA Base HEP Tables

IDTABLE-IDTABLE-1 to IDTABLE-IDTABLE-3 document human error rates for base HEPs. A base HEP is the error probability of a CFM under an attribute of the three base PIFs: *Scenario familiarity, Information availability and reliability,* and *Task complexity*. If one of these PIFs is present in the context of the tasks in a data source, the human error data reported in the data source are generalized and documented in the corresponding IDTABLE.

The following process is tailored to generalize human error data for the base HEPs:

(1) Analyze the data source. This includes identifying the tasks of which human error information is reported, analyzing the context, characterizing the tasks to identify cognitive activities involved in the tasks and time constraints when the tasks were performed.

(2) Map the human errors of the tasks to corresponding CFMs. The task characterization identifies cognitive activities involved in the tasks. The cognitive activities are then mapped to applicable IDHEAS-G CFMs. The mapping could be made to a single or multiple levels of CFMs: failure of macrocognitive functions, failure of processors, or detailed failure modes.

(3) Map the context to the relevant IDHEAS-G PIF attributes.

(4) Document the reported human error rates for the corresponding CFMs and PIF attributes in IDHEAS-DATA Base HEP Tables along with other items of context information.

(5) Evaluate and document uncertainties in the data source and mapping process.

<u>Structure of the Base HEP TABLEs</u>

A Base HEP IDTABLE documents human error data in the associated CFMs and PIF attributes. Each row of the TDTABLE is referenced as one datapoint, which may consist of one or several reported human error rates at different status of the PIF attribute. Each datapoint comes from one data source such as a technical report or a research paper, while one data source may contain multiple datapoints for the same or different IDTABLEs because the reported study may have examined human error rates for different tasks or different PIF attributes. The columns of the table document the following dimensions of information for every datapoint:

- Column 1: the base PIF attribute for the reported human error rates – The IDHEAS-DATA Tables use labels for PIF attributes. Appendix A1 provides the indices of the labels to the corresponding PIF attributes.[1]
- Column 2: the applicable CFMs of the reported human error data – The CFMs are labeled as D, U, DM, E, and T for *failure of Detection, Understanding, Decisionmaking, Action execution,* and *Interteam Coordination*. If the task for which the human error rates were reported contain more than one CFM, then the labels of all the applicable CFMs are presented in Column 2.

[1] Note that the labels are in two levels. The high-level labels are similar to those used in Appendix B of the IDHEAS-ECA report and in the IDHEAS-ECA Software. This is because the IDHEAS-G PIF attributes were consolidated into a concise set of the attributes in IDHEAS-ECA. The attributes in IDHEAS-DATA are essentially the same as those in IDHEAS-G.

- Column 3: human error rates – The human error rates reported in the data source. The error rates are percent of errors unless specified otherwise.
- Column 4: the tasks for which the human error rates were reported in the data source, along with the definition of the human errors measured for the tasks.
- Column 5: PIF attribute measure – The task-specific factor or variable used in the data source under which the tasks were performed and human error rates were measured.
- Column 6: Other PIFs that are also present in the tasks and uncertainties – In addition to the PIF attribute that were under the study, the context of the tasks in a data source may have other PIF attributes present during task performance; therefore, they would contribute to the reported error rates. Column 6 documents other PIF attributes that were present. In particular, Column 6 documents whether the tasks were performed under time constraints. Information about the time availability is important to infer the base HEPs from the reported human error data. If the time available is inadequate, then a reported human error rate corresponds the probabilistic sum of the base HEPs and the error probability due to inadequate time (P_t). Column 6 also documents the uncertainties in the data source and in the mapping to the CFMs and PIF attributes. The uncertainties would affect how the reported error rates are to be integrated to inform base HEPs.
- Column 7: The date source reference.

Next is an example to demonstrate the process of generalizing human error data to the Base HEP Tables. The data source is a report, "The Outcome of [Air Traffic Control] Message Complexity on Pilot Readback Performance," by Prinzo et al. [8, 9]. The study analyzed aircraft pilot communication errors and reported that the error rate increased nonlinearly with the complexity of the message communicated. The following is the process of generalizing the data to IDHEAS-DATA Base HEP IDTABLE-IDTABLE-3 for *Task complexity*.

<u>*Analyze the data source*</u>: Prinzo et al. [8, 9] — The task is that pilots listen to and read back messages from air traffic controllers. The pilots hold the information in their memory and read back at the end of the transmission. The cognitive activities involved are perceiving information and communicating it. The pilots perform the task individually without peer-checking, and the tasks are performed without time constraints. Readback errors are defined as misreading or missing key messages. Message complexity is defined as the number of key messages in one transmission. The study calculates percent of readback errors at different levels of message complexity from thousands of transmissions. Identified human error data for generalization: The readback error rates at different message complexity levels are identified as the data for this entry. <u>*Applicable CFMs*</u>: The CFM for readback errors is *failure of Understanding*. While the task is "listen to and readback messages," the cognitive activities required are identifying, comprehending, and relating all the key messages in one transmission. Those are the elements in the macrocognitive function *Understanding*. <u>*Relevant PIF attributes*</u>: The primary PIF is *Task complexity*. The attribute is C11, "the number of key messages to be kept." Another PIF present is the *Work Process* attribute, "Lack of verification or peer-checking." <u>*Other PIF attributes present*</u>: Some transmissions may be performed with the presence of other PIF attributes such as distraction, stress, or mental fatigue. Those PIFs were not

prevalent in the transmissions analyzed but could increase the overall error rates. Pilots' flying experience was not correlated with the error rates.

<u>*Uncertainties in the data and mapping*</u>: The source audio transmissions are mixture of normal and emergent operation.

The analysis results are documented in IDTABLE-3 as one datapoint. Table 2-3 shows the information documented for this datapoint. All the information items are in one row. The top two row has column numbers for referencing.

Table 2-3 Sample of IDTABLE-3 – Base HEPs for Task Complexity

1	2	3		4	5	6	7
PIF	**CFM**	**Error rates**		**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
C11	U	Number of messages	Error rate	Pilots listen to and read back key messages	Message complexity - # of key messages in one transmission	(Mixture of normal and emergent operation so other PIF attributes may exist)	[8, 9]
		5	0.036				
		8	0.05				
		11	0.11				
		15	0.23				
		17	0.32				
		>20	>0.5				

2.4.2. IDHEAS-DATA PIF Weight IDTABLE-4 through IDTABLE-20

IDTABLE-4 through IDTABLE-IDTABLE-20 document human error rates for the 17 PIFs other than the three base PIFs. A data source generalized to these IDTABLEs should have human error rates or task performance indicators measured at different status of one or more PIF attributes. The IDTABLEs contain datapoints at which the human error rates of a task were measured for one or more status of a PIF attribute (e.g., not-present vs. present, low vs. high). Such error rates can be used to infer the weight of the PIF attribute.

The process of generalizing human error data to a PIF Weight IDTABLE is the same as that for the Base HEP Tables. The structure of the PIF Weight IDTABLE is the same as that for the Base HEP Tables. A datapoint typically has more than one human error rate reported for different status of the PIF attribute, thus the third column "Human error data" for each row is typically split into multiple rows and columns for different PIF attribute status.

Each row of a PIF Weight IDTABLE documents one datapoint, containing the human error rate of a task for one PIF attribute and the related information in different columns. The column for error rates is typically split into several sub-rows and columns to record multiple error rates and the levels of the PIF attribute at which the errors were measured. If a data source has human error rates for more than one task, then the data for each task is documented either in a separate row or in different sub-columns of the error rate column. If a data source has error rates measured for more than one PIF attribute, then the data for every attribute is documented as a separate datapoint.

The next example demonstrates how to generalize human error data to a PIF Weight Table. The data source is the research paper, "Effects of Interruption Length on Procedural Errors," by Altmann et al. [14]. The study investigated effects of task interruption on procedural performance, focusing on the effect of interruption length on the rates of different categories of error at the point of task resumption. The following is the process of generalizing the data to IDTABLE-17 for *Multitasking, Interruption, and Distraction*.

Analyze the data source: The task [14] was that individual participants performed procedural sequences of computerized execution steps. The task required individuals memorizing the sequences. The study examined effects of interruption length on procedural performance parametrically across a range of practically relevant interruption durations—from about 3 seconds to about 30 seconds. The cognitive activities involved were executing sequential steps. The participants are well trained for the task. They performed the task individually without peer-checking and without time constraint. Performance errors are defined as loss of place in the procedure (*sequence errors*) and errors involving incorrect execution of a correct step after interruption (*non-sequence errors*)

Identify human error data for generalization: Both sequence and non-sequence error rates at different lengths of interruption are identified as the data for this entry.

Applicable CFMs: The CFM is *failure of action execution*.

PIF attributes: The PIF being examined is *Multitasking, Interruption, and Distraction.* The attribute is "Interruption." The PIF *Work Process* attribute "Lack of verification or peer-checking" was present for all the human error data measured in the study.

Evaluate uncertainties in the data and mapping: This study is a well-controlled experimental study and there is no prevalent uncertainty involved.

The analysis results are documented in IDTABLE-17 as one datapoint. The sequence-error rates at different lengths of interruption are identified as the human error data for this datapoint. The post-interruption non-sequence errors, although not affected by interruption, is also documented for reference. The reported human error rates for the corresponding CFMs and PIF attributes are then documented along with other items of context information. Table 2-4 shows the information documented for this datapoint. All the information items are in one row. The top row has column numbers for referencing.

Table 2-4 Sample of IDTABLE-17 – Error Rates for PIF Multitasking, Interruptions, and Distractions

1	2	3			4	5	6	7
PIF	**CFM**	**Error rates (%)**			**Task (and error measure)**	**PIF measures**	**Other PIFs (and Uncertainty)**	**REF**
MT2	E	Interruption Length (s)	Sequence error	Non-sequence error	Individuals executed procedural steps of a computerized task. Performance errors are loss of place in the procedure (*sequence errors*) and errors involving incorrect execution of a correct step after interruption (*nonsequence errors*).	Interruption - Different interruption length (seconds). Baseline is no interruption.		[14]
		Baseline	2	2				
		3	4	2				
		13	10	2				
		22	14	2				

2.4.3. IDTABLE-21 for the Lowest HEPs

In the IDHEAS-G HEP quantification model, the lowest HEPs are used as the values for the base HEPs when none of the three base PIF attributes is present. The Lowest HEP IDTABLE documents datapoints of which the human error rate of a task is measured under the conditions that (1) none of the known PIF attributes are present or there is no prevalent known PIF

attribute present and (2) the number of times that the task was performed is substantially large so that the measured error rate is reliable. The human error rates measured under such conditions correspond to the lowest HEP that a CFM of tasks can achieve.

Ideally, data sources for lowest HEPs should meet the following conditions:

1) The error rates are measured from a sufficiently large number of times that the task is performed;
2) none of the attributes of the 20 PIFs is present or prevalent;
3) the task is performed without time constraint;
4) there is professional self-verification, peer-checking, and/or supervision for task performance;
5) the error rate is for a single CFM of a single task; and
6) the error rate is measured without recovery actions.

Hardly any data source can meet all the conditions above. When analyzing data sources for the lowest HEPs, it is important to annotate if any of these conditions is not met, such as whether there is lack of peer-checking or whether the task of which the error rates were measured had multiple applicable CFMs.

The structure of IDTABLE-21 is described as the follows:

- Column 1: The applicable CFMs of the reported human errors – The CFMs are labeled as D, U, DM, E, and T for *failure of Detection, Understanding, Decisionmaking, Action execution,* and *Interteam Coordination*. Note that the task may have multiple applicable CFMs.

- Column 2: Human error rates – The human error rates reported in the data source should meet most of the conditions for the lowest HEPs. If the range of an error rate was calculated or estimated in the data source, it should be documented as well to inform the integration of multiple data sources into the lowest HEPs.

- Column 3: Task and context - The task of which the human error rates are measured and the general context under which the task is performed.

- Column 4: Criteria assessment – Assessment of the human error data against the criteria of lowest HEPs. Five criteria are assessed: Adequate time available for performing the task, personnel's self-verification of task performance, Team verification (through peer-checking, independent checking / advising, and close supervision), recovery of human failure events, and presence of any PIF attribute. Each criterion assessed for "Yes," "No," Mixed Yes and No," and "Unknown."

- Column 5: Uncertainties – There are uncertainties in the data source and in the mapping to IDHEAS-G CFMs. In particular, if the number of the times the task was performed is not sufficiently large, the reported error rate may not represent the lowest HEP.

- Column 6: Source reference.

Next is an example to demonstrate how to generalize human error data to the Lowest HEP Table. The source of data is the research papers "Human error probabilities from operational experience of German nuclear power plants", Part I and Part II, by Preischl and Hellmich [4, 5]. The study collected human reliability data from the operational experience of German nuclear power plants to determine the number of times the task was performed in the past, as well as

the number of errors that occurred. The data source was the database of the German licensee event report system that collected the reportable events in German nuclear power plant installation work. The study reported error rates of many types of nuclear power plant maintenance tasks. This example only uses the datapoints for which the number of the task performed was greater than 1000 and no prevalent PIF attribute were reported. The following is the process of generalizing the data to inform lowest HEPs.

<u>Analyze the data source</u>: The tasks that maintenance personnel performed routine nuclear power plant maintenance. The cognitive activities involved were executing sequential steps. Participants are well trained for the task. They may perform the task with or without peer-checking. Most tasks should be performed without time constraints. Performance errors were defined as not performing steps of a task or incorrectly performing a task. The error rate is the number of times the error occurred divided by the number of times the same task type was performed. The data source provides both numbers for various task types.

<u>Identify human error data for generalization</u>: The human error data for this example are the rates extracted from the reported events in which no PIFs were identified.

<u>CFMs</u>: The CFM is f*ailure of Execution*.

<u>Evaluate uncertainties in the data and mapping</u>: It is unclear whether the tasks were performed with or without peer-checking. The reported events may or may not involve recovery actions. The definition of the errors was for task steps rather than a whole task; thus, the reported error rates could be higher than that for whole tasks if some tasks had errors in multiple steps.

The analysis results are documented in IDTABLE-21 as multiple datapoints. While the majority of datapoints have the CFM of *Failure of Execution*, two types of tasks were reading meters or reading instructions. The errors were incorrectly reading. This could be the CFM of *failure of Detection* or the CFM of *failure of Execution* because reading is a part of the execution. Table 2-5 shows the information documented from this data source for the lowest HEPs.

Table 2-5 Sample of IDTABLE-21 – Lowest HEP

1	2	3	4	5	6
CFM	**Error rate**	**Task and context**	**Criteria for lowest HEPs:** **TA - Time adequacy** **SelfV - Self verification** **TeamV – Team verification** **Rec - Recovery** **O - other factors** **(Y-Yes, N – No, M-Mixed Un-Unknown)**	**Uncertainty**	**REF**
E	8E-4 (1/1470)	Manually operating a local valve. Frequently performed task. Valve not operated, step in a sequence of different steps not remembered. - No known PIF exists	TA – Y, SelfV- Y, TeamV - Unknown Rec - Unknown	Error rates were for steps of a task. Most tasks performed may not have peer-checking. Some errors made may have been recovered so they did not get into the reporting system.	[4, 5]
E	8.9E-4 (7/8058)	Operating a control element on a panel, Wrong control element selected, - Similar controls within reach	TA – Y, SelfV- Y, TeamV - Unknown Rec - Unknown		
	8.78E-4 (1/1347)	59 Operation of a manual control at a Main Control Room (MCR) control (Task not remembered)	TA – Y, SelfV- Y, TeamV - Unknown		

		- Frequently performed task, part of professional knowledge, position of indicator lamps ergonomically unfavorably designed	Rec - Unknown		
E	1.04E-3 (2/2088)	Remembering professional knowledge, Remembered incorrectly. Part of frequently performed procedure	TA – Y, SelfV- Y, TeamV - Unknown Rec - Unknown		
E	1.03E-3 (3/3067)	Carrying out a sequence of tasks. Error were skipped steps. Frequently performed.	TA – Y, SelfV- Y, TeamV - Unknown Rec - Unknown		

2.4.4. IDTABLE-22 for PIF Interaction

The PIF Interaction TABLE documents datapoints of which the human error rates of a task were measured as two or more PIF attributes were varied independently and jointly. Each datapoint contains the human error rates under different status of the individual PIF attributes as well as the error rates under the combination of both PIF attributes. The weights of individual PIF attributes and the joint weight of the PIF attributes can thus be calculated from those error rates. The relationship between these weights would inform the quantitative aspect of PIF interaction. For example, if the two PIF attributes examined in a study have no interaction in their impacts on human error rates, then the combined weight is simply the sum of the individual weights. On the other hand, if there is interaction, the combined weight would not be the linear combination of the individual weights.

The structure of IDTABLE-22 is similar to that of PIF weight TABLEs but it has two PIF attributes in Column 2 “PIFs.” Each row is for one datapoint that represents human error rates of a task under individual and joint PIF attributes. The error rates of a datapoint are documented in sub-rows for the status of one PIF attribute and sub-columns for different status of another PIF attribute. A data source may contain multiple datapoints for different tasks or for different PIF combinations.

The following example demonstrates the process of generalizing human error data to the PIF Interaction TABLE. The source of data is the research paper about the effect of sustained acceleration (+Gz) and luminance on dial reading errors [15]. The following is the process of generalizing the data to IDHEAS-DATA PIF Interaction IDTABLE-22.

<u>Analyze the data source</u>: The task was that pilots with corrected normal vision and extensive centrifuge experience read aircraft instrument dials as the luminance (c/m^2) of dials and degree of acceleration varied. The macrocognitive function required for the task was *Detection*. Participants performed the task individually without peer-checking. Performance errors were measured as the percent of misreading dials.

<u>CFMs</u>: The CFM is *failure of Detection*.

<u>PIF attributes</u>: The two PIF attributes were VIS1 “Target or object luminance” of PIF *Workplace Visibility* and PR1 “Resistance to personnel movement” of PIF *Physical Resistance*.

<u>Evaluate uncertainties in the data and mapping</u>: It is unclear whether the task was performed under time constraint and what HSI attributes might have been present. The PIF *Work Process* attribute “Lack of verification or peer-checking” was present in all the error data measured.

The analysis results are documented in IDTABLE-22 as one datapoint, as shown in Table 2-6.

Table 2-6 Sample of IDTABLE-22 – PIF Interaction

1	2	3			4	5	6	7
CFM	**PIFs**	**Error rates**			**Task**	**PIF2 measure**	**Other Factors And Uncertainty**	**REF**
D	PIF1-VIS, PIF2 – PR	PIF1 \ PIF2	2G	4G	Pilots read aircraft instrument dials as the luminance (c/m^2) of dials and degree of acceleration (+Gx) vary. Errors are percent of misreading dials.	VIS-Luminance PR – Acceleration	Maybe time constraint	[15]
		150	7	7				
		15	7	15				
		1.5	10	20				
		0.15	20	45				
		0.015	50	63				

2.4.5. IDTABLE-23 for Distribution of Time Needed in completing a human action

The IDHEAS-G HEP model considers that the HEP of an important human action consist of P_t, the error probability attributing to time availability and P_c the error probability attributing to cognitive failure modes. P_t is calculated as the convolution of the distributions of time available for the action and time needed to complete the action. HRA analysts use available operational data and their engineering judgment to estimate the distribution of time needed. IDTABLE-23 documents time distributions of professional personnel performing important human actions. The information is used to develop guidance and inform HRA analysts about the estimation of the distribution of time needed.

The time distribution reported in data sources can come with various formats, e.g., mean and standard deviation, low and upper bounds of the time variation, the actual time spent for completing a human action, or histograms of the time spent. IDTABLE-23 documents time distribution in data sources. A datapoint should capture the information about the distribution in a data source, such as mean, standard deviation, range, sample size, etc. The structure of IDTABLE-23 is the following:

- Column 1: Scenario, human actions or tasks, and prevalent cognitive activities involved – This column documents the human action or task and the scenario under which the action was performed. It should be noted if the action is procedure based. Personnel performing the actions should also be noted unless by default they are nuclear power plant operators or well trained, experienced professionals. This column also documents the prevalent cognitive activities contributing to the time needed.

- Column 2: Distribution of time needed to perform the action – This column documents the actual time information as it is reported in the data source. It should be annotated if the time spent for the action was inadequate for personnel to complete the action.

- Column 3: Uncertainties in the data source – This column documents the time uncertainties that may cause variation and affect the distribution of time needed.

- Column 4: Reference to the data source.

The following example demonstrates the documentation of time distribution in IDTABLE-23. The data source is the U.S. HRA Empirical Study [16]. Four crews from a U.S. plant performed three scenarios on simulators. This example only documents the time data for Scenario 2, Component Cooling Water (CCW) and Reactor Coolant Pump (RCP) sealwater. The data of the four crews' task performance time are documented in IDTABLE-23 as one datapoint, as shown in Table 2-7. The time variation like shown in this datapoint can be used to develop the guidance on estimating time uncertainty distribution.

Table 2-7 Sample of IDTABLE-23 – Distribution of Time Needed

<table>
<tr><th>1</th><th>2</th><th>3</th><th>4</th></tr>
<tr><th>Scenario and Action a</th><th>Time needed
M – Mean, SD – Standard deviation, Range – [Min, Max], N – sample size</th><th>Uncertainty</th><th>REF</th></tr>
<tr><td>Scenario – Internal at-power event
Actions - Stop RCPs and Start PDP in loss of CCW and RCP sealwater
Personnel - 4 crews of NPP operators.
Cognitive activities –
D – Detect loss of CCW and RCP sealwater
U – Diagnose the need of starting PDP
E – Execute procedures</td><td><table>
<tr><th>Tasks</th><th colspan="4">Time (min) for each crew</th></tr>
<tr><td>Start of scenario</td><td>0</td><td>0</td><td>0</td><td>0</td></tr>
<tr><td>Reactor trip</td><td>3</td><td>3</td><td>3</td><td>3</td></tr>
<tr><td>Loss of CCW</td><td>3</td><td>3</td><td>3</td><td>3</td></tr>
<tr><td>Start procedure E-0</td><td>3</td><td>3</td><td>3</td><td>3</td></tr>
<tr><td>Start procedure</td><td>8</td><td>8.5</td><td>9.6</td><td>7</td></tr>
<tr><td>Detect no CCW</td><td>9</td><td>9</td><td>7</td><td>9</td></tr>
<tr><td>Trip all RCPs</td><td>11.5</td><td>9.5</td><td>7.6</td><td>10.5</td></tr>
<tr><td>Start "RCP-</td><td>10</td><td>13</td><td>13</td><td>-</td></tr>
<tr><td>Start PDP</td><td>-</td><td>-</td><td>-</td><td>-</td></tr>
</table>
The distribution for the time needed from "Reactor trip" to "Trip all RCPs" is: M=9.6, SD=1.5, Range=[7.6, 11.5], N=4</td><td>Unfamiliar scenario – simultaneous loss of CCW and RCP sealwater is rare and was not in training.</td><td>[16]</td></tr>
</table>

2.4.6. IDTABLE-24 for Modification to Time Needed

TABLE-24 documents the effects of time uncertainty factors on time needed for completing human actions. Many factors can affect task completion time. These factors contribute to the uncertainty in time distribution. IDHEAS-G provides a list of prevalent time uncertainty factors, as shown in its Table 5-2 [1]. Note that there could be additional factors affecting time needed. In fact, most PIF attributes modify task completion time. IDTABLE-24 is open to any factor that can influence time distribution.

The most useful data for IDTABLE-24 would be operational data from tasks performed by licensed professional personnel. However, while with high fidelity, operational data typically do not systematically record action performance time under different factors. On the other hand, extensive experimental literature reports task completion times with varying time uncertainty factors or PIF attributes. A data source for IDTABLE-24 should have task completion times under at least two different states of a time uncertainty factors or PIF attributes to inform the effect of the factor on task completion time.

The structure of IDTABLE-24 is as follows. Each row of the IDTABLE is referenced as one datapoint, which may consist of one or several reported human error rates at different states of the PIF attribute. Each datapoint comes from one data source such as a technical report or a research paper, while one data source may contain multiple datapoints for the same or different IDHEAS-DATA Tables because the reported study may have examined human error rates for different tasks or different PIF attributes. The columns of the IDTABLE document the following dimensions of information for every datapoint:

- Column 1: the applicable CFMs of the task or human action being studied. If the task completion time is reported for an event in which applicable CFMs cannot be distinguished, then this column is filled with “Unsp” for unspecified CFMs.
- Column 2: The PIF or other time uncertainty factor that modifies the task completion time.
- Column 3: This column documents the task completion time information under the variation of the PIF or time uncertainty factor.
- Column 4: the tasks of which the completion time was reported in the data source.
- Column 5: The factor or variable used in the data source under which the tasks were performed, and task completion time was measured.
- Column 6: Note – this column annotates comments on the data source
- Column 7: The date source reference.

The following example demonstrates the documentation of time needed in IDTABLE-24. The source of data is the research paper by Berg et. al. [17] that examined the effects of Visual Distractions on Completion of Security Tasks. The following is the process of generalizing the data to IDTABLE-24.

<u>Analyze the data source</u>: 169 subjects (mostly technical, navy, male college students) performed a security-critical task (Bluetooth Pairing) while static or flicking colored visual distractors were present versus absent. The task required the subjects to read, compare, and confirm Bluetooth numbers. The subjects practiced the task then performed the task in an unattended environment mimicking the real job context. Participants perform the task individually without time constraint.

<u>Time Uncertainty factors</u>: The factor varied in the study is the presence vs. absence of visual distraction as well as different types of visual distraction.

The results of the analysis are documented in IDTABLE-24 as one datapoint, as showing in Table 2-8.

Table 2-8 Sample of IDTABLE-24 – Modification of Time Needed

CFM	PIF or Time-Factor	Task completion time (mean and SD)		Task	PIF or Time Factor measure	Note	REF
		Factor-Lo	Factor-Hi				
D	MT1	35(12)s	88(25)s	Security-critical detection task requiring reading, comparing, and confirming Bluetooth numbers.	Lo – No distraction Hi – static red visual stimuli for distraction	169 college students	[17]
D	MT1	35(12)s	90(16)s	Security-critical detection task	Lo – No distraction Hi - flicking red visual stimuli for distraction	169 college students	[17]

2.4.7. IDTABLE-25 for Dependency of Human Actions

IDHEAS-G proposes a dependency model to evaluate the dependency between two important human actions. The dependency model identifies the types of dependency, evaluates how the dependency changes the context of the subsequent action, and re-estimates the HEP of the

action based on the changes of the context. This model is different from the traditional HRA methods that evaluate dependency based on context similarity of the two actions. IDTABLE-25 documents empirical evidence of dependency in operational or simulated NPP events to establish the technical basis for dependency evaluation. The structure of IDTABLE-25 is as follows:

- Column 1: Dependency type – IDHEAS-G dependency model defines three types of dependency: consequential dependency (SD), resource-sharing dependency (RSD), and cognitive dependency (CD).
- Column 2: Brief narrative of the scenario, human actions, and consequence of the dependency. Also, documented in this column is brief explanation on why the narrative is categorized as the dependency type in column 1.
- Column 3: Reference of the information source. Note that the primary sources of information are the event reports, accident sequence precursor (ASP) and significance determination process (SDP) analysis reports, operational experience review, and reports on operator performance simulation.

The IDHEAS-G dependency model calculates the effect of dependency on HEPs based on the changes in the context of the action due to dependency. IDHEAS-G models the changes of the context in terms of human action feasibility, time availability (time needed and time available), new or different critical tasks, new or different CFMs, and changes in applicable PIF attributes. IDTABLE-25 should document the changes in the context of the subsequent human action due to its dependency on the failure of the previous action. However, making proper context judgment requires event details. Analyzing the changes of context may not be viable due to the lack of context information details in data sources. At present, IDTABLE-25 provides empirical information to verify IDHEAS-G dependency model and to inform HRA in identifying types of dependency.

Presented next is an example demonstrating the generalization of empirical information of dependency in NPP events to IDTABLE-25. The example is from the report "Review of Human Error Contribution to Operational Events — Summary Report" [18]. In this study, precursor data from the Accident Sequence Precursor (ASP) Program during the Fiscal Year 2000–2004 period was reviewed to identify the kinds of human errors that are associated with precursor events. The report analyzed many risk precursor events and identified the types of human errors. This example used one of the events documented in the report. The following is the process generalizing the source information to IDTABLE-25.

<u>Analyze the data source/Narrative of the scenario and event</u>: (NRC Integrated Inspection Report 05000528/2004003, 05000529/2004003 [19]) Simultaneous testing of the atmospheric dump valve and boron injection systems resulted in a loss of letdown event on high regenerative heat exchanger temperature. The letdown event occurred because operations personnel were using a single charging pump for the boron injection test and using excess letdown to accommodate a plant heat-up following atmospheric dump valve testing. The combination of activities resulted in pressurizer level exceeding the TS limit of 56 percent. This issue involves human performance crosscutting aspects associated with poor decision making, questioning attitude, awareness of plant conditions, and communications between personnel performing concurrent evolutions. <u>Dependency analysis</u>: operators elected to perform a combination of surveillance tests that caused a loss of letdown and pressurizer level transient. This is the resource-sharing dependency (RSD). Simultaneous tests of the atmospheric dump valve and the boron injection system demanded the charging flow exceeded the charging pump capacity.

The results of the analysis are documented in IDTABLE-25 as one datapoint, as shown in Table 2-9.

Table 2-9 Sample of IDTABLE-25 – Dependency of Human Actions

1	2	3
Type	**Scenario and event narrative**	**Ref**
Resource-sharing dependency	**A pressurizer level transient above Technical Specification limits** Simultaneous testing of the atmospheric dump valve and boron injection systems resulted in a loss of letdown event on high regenerative heat exchanger temperature. The letdown event occurred because operations personnel were using a single charging pump for the boron injection test and using excess letdown to accommodate a plant heat-up following atmospheric dump valve testing. The combination of activities resulted in pressurizer level exceeding the Technical Specification limit of 56 percent. **Explanation.** Simultaneous tests of the atmospheric dump valve and the boron injection system demanded the charging flow to exceed the charging pump capacity.	[19]

2.4.8. IDTABLE-26 for Recovery of Human Actions

IDTABLE-26 collects empirical information from NPP human events on recovery actions. The documented events can establish a technical basis for modeling and crediting recovery actions. IDTABLE-26 has the following structure:

- Column 1: Narrative of the recovery action – This column documents a brief narrative of the scenario, the action to be recovered, the recovery action, whether the recovery action was a success or failure, and prevalent context of the recovery action, such as if the recovery action is skill-of-the-craft.
- Column 2: Notes – This column documents the information about the factors that make the recovery action feasible, factors affecting the success of the recovery, and dependency with the human action to be recovered. Any reported likelihood or chances of the recovery action should also be documented in this column.
- Column 3: Reference of the information source – Note that the primary sources of information are the event reports, ASP/SDP analysis reports, operational experience reviews, and reports on operator performance in simulators.

In principle, the reliability of a recovery action is determined by its CFMs and associated PIF attributes, and it is subject to the dependency with the human action to be recovered. IDTABLE-26, as it is now, does not explicitly collect information on reliability of recovery actions. Making proper contextual judgment of applicable CFMs, PIF attributes, and dependency requires event details that may not be available in the data sources. Also, because IDHEAS-G, as of now, does not have a matured model to assess and credit recovery actions, IDTABLE-26 only documents narrative information without characterizing recovery actions.

The following example, from Reference [20], demonstrates the generalization of empirical information of a recovery action in an NPP human event to IDTABLE-26.

Narrative of the scenario and event: In the course of the startup of the plant, it was discovered that the isolation valves in each of the three high pressure safety injection lines to the cold legs of the primary circuit were in the closed position. Their power supplies were disconnected. One day before startup, a leak-tight test of the check (isolation) valves in the high-pressure injection system was performed. The test requires that the isolation valves should be closed but not disconnected from the electrical power supply. The test procedure did not provide specific instructions to restore or to verify the proper line-up of the safety after

the test. The day following the completion of the test, the operators verified the line-up of the safety injection system as instructed in operating procedures. Failure of the important human action to be recovered and recovery actions: The failed important human actions are the omission to re-establish the required line up of the system after the leak-tightness test and the disconnection of the valves' electric power supply without instruction. The recovery action is the operator's verification of the safety injection system line-up in accordance with operating procedures when changing technical specification modes during startup. Feasibility of the recovery action: The recovery action is feasible because the system line-up verification was directed by procedures. Potential dependency between the action being analyzed and its recovery action: In this case, there is no dependency between the action being analyzed and its recovery action because the recovery action was performed a day later and it is likely that the safety system line-up verification was performed by different operators than the one that performed the test using different procedures.

The results of the analysis are documented in IDTABLE-26 as one datapoint, as showing in Table 2-10.

Table 2-10 Sample of IDTABLE-26 – Recovery of Human Actions

1	2	3
Narrative of the recovery action	**Notes**	**Ref**
In the course of the startup of the plant, it was discovered that the isolation valves in each of the three high pressure safety injection lines to the cold legs of the primary circuit were in the closed position. Their power supplies were disconnected. One day before startup, a leak-tight test of the check (isolation) valves in the high-pressure injection system was performed. The test requires that the isolation valves should be closed but not disconnected from the electrical power supply. The test procedure did not provide specific instructions to restore or to verify the proper line-up of the safety after the test. The day following the completion of the test, the operators verified the line-up of the safety injection system as instructed in operating procedures.	The recovery action of the operator's verification of the safety injection system line-up is feasible because it was directed by procedures. No dependency between the failed action and its recovery action because the recovery action was performed a day later, and it is likely that the safety system line-up verification was performed by different operators than the one that performed the test using different procedures. Also, Reference [20] analyzed 17 human failure events. Eleven events occurred in the outage phase, and 5 of these during start up. Another might be during power operation. Scheduled periodical tests detected most (9) of the events. In 5 events, the deficiencies occurred on demand and 3 deficiencies were detected by chance. This reference provides a data point of error recovery in maintenance surveillance tests as 0.7 (= 12/17).	[20]

2.4.9. IDTABLE-27 for Main drivers to Human Failure Events

IDHEAS-G models context of a human action with a comprehensive set of PIF attributes. The main drivers to human failure events are the specifics of situations or context that more likely leads to failure or leads to high HEPs. In the IDHEAS-G framework, the main drivers are the contexts that results in the PIF attributes of high base HEPs or large PIF weights. IDTABLE-27 shows empirical evidence on specifics of situations or context that are the main drivers to human failure in operational or simulated events. It also represents the main drivers in PIF attributes. The information in IDTABLE-27 can assist HRA analysts to capture main drivers in human events and represent them with proper PIF attributes.

The data sources in IDTABLE-27 are primarily from the nuclear domain. The main data sources for IDTABLE-27 can be from analysis of LERs, human event analysis reports, human

performance data of real operation, simulator training, simulation studies, as well as literature on human error analysis and root cause analysis. A datapoint in IDTABLE-27 documents operator performance of a human event or a certain type of human actions. The datapoints in IDTABLE-27 will demonstrate the operational expression of the PIF attributes with high base HEPs or PIF weights. The datapoints serve as the linkage between context and PIF attributes. Such linkage can assist HRA analysts to avoid overlooking main drivers and to support the evaluation of PIF attributes. The following is the structure of IDTABLE-27.

- Column 1: The CFMs that occurred in the event. If the event involved a complex scenario, it might have multiple CFMs or the CFMs could not be specified from the information available.
- Column 2: PIFs or PIF attributes representing the main drivers of the human failure in the event. Sometimes detailed information for analyzing specific attributes may not be available in the data source, thus that the main drivers can only be represented at the PIF level.
- Column 3: Human error rates: This column documents the human error rate of the event or the type of the events if the error data is available. Many data sources such as case studies or analysis of individual events do not have any numeric data on human error rates relevant to the main drivers.
- Column 4: Narrative of the human event and main drivers: This includes a brief description of the human event and main drivers of the human failure as well as the event context and considerations of representing the main drivers in CFMs and PIFs.
- Column 5: The data source reference.

The below example demonstrates the generalization of empirical information in IDTABLE-27. The example is from the International HRA Benchmarking Study.

In the study, 7 out of 10 crews failed HFE1B, i.e., initiate bleed and feed cooling before steam generator (SG) dry-out in the complex Loss of Feed Water (LOFW) scenario. One of the main drivers to the HFE was that the SG water level indicators had misleading information, caused by the fact that the scenario had a steam generator tube rupture and a water leak. The information about water leading was masked by the indications of the tube rupture. In the study, 14 HRA analyst teams were given the material package including the scenario description and procedures. They identified the main drivers to the human failure events in the scenario and performed HRA using various HRA methods. Most HRA analyst teams did not identify information masking as a main driver to the human failure events and subsequently they predicted much lower HEPs of the HFE compared to the 7 out 10 crews failing the event. The result of the analysis was documented in IDTABLE-27, as showing in Table 2-11.

Table 2-11 Sample of IDTABLE-27 – Main Drivers to Human Failure Events

1	2	3	4	5
CFM	**PIFs**	**Error rates**	**Narrative of the event and main drivers to human failures**	**Ref**
U	SF3, INF6	0.7 (7/10)	<u>Main Drivers:</u> Inadequate knowledge, <u>key information was cognitively masked.</u> This is HFE1B, initiate bleed and feed before steam generator (SG) dryout in the complex Loss of Feed Water (LOFW) scenario, in the International HRA Benchmarking Study. The following are from section 2.3.2 of volume 3 of The International Benchmark Study report series: • The complex scenario contained multiple issues, including degraded condensate pump and failures of two SGs' wide range (WR) level indications). The first issue was that one condensate pump was successfully running at the beginning, leading the crew to depressurize	[16, 21-23]

			the SGs to establish condensate flow. However, the running condensate pump was degraded and gave a pressure so low that the SGs became empty before the pressure could be reduced enough to successfully inject water. • The procedure step to depressurize is complicated, and this action both kept the crew busy and gave them a concrete chance to re-establish feedwater to the SGs. The crews were directed by procedure FR-H.1 to depressurize the SGs to inject condensate flow. • Two of the three SGs had WR level indicators malfunction that would incorrectly show a steady (flat) value somewhat above 12% when the actual level would be 0% due to the water leaking. The two failing SG levels both indicated a level above the 12% criterion to start Bleed & Feed. To follow the criterion, the crews had to identify and diagnose the indicator failures, since the criterion, interpreted literally, would never be met.	

2.5. Integration of human error data to inform human error probabilities

Integration of the generalized data in IDHEAS-DATA to inform HEP estimation depends on the specific HRA method or application. IDHEAS-G describes several ways of using human error data for HEP quantification: Using the data as the basis for expert judgment of HEPs, using the data to derive basic parameters needed for calculating HEPs in a HEP model, or calculating HEPs from the data using statistic regression. This section describes the process of integrating the data in IDHEAS-DATA to provide the base HEPs and PIF weights needed for calculating HEPs in IDHEAS-ECA method. Chapter 3 of this report presents several examples of integrating the data for HEP quantification in IDEHAS-ECA.

2.5.1. Overview of an Application-specific IDHEAS method

IDHEAS-G provides the basic framework for qualitative analysis and HEP quantification. It has a basic set of CFMs, a comprehensive set of PIF attributes, several ways of estimating HEPs including a HEP quantification model, but it does not offer HEP calculation. An application-specific IDHEAS method is derived from IDHEAS-G. It uses a limited subset of IDHEAS-G CFMs and PIF attributes specific for the given HRA application and it can generate HEP estimates. Two application-specific IDHEAS-methods have been developed: IDHEAS-AtPower Application for internal at-power NPP events and IDHEAS-ECA for Event and Conditions Assessment.

An Application-specific IDHEAS method should have the following three elements derived from IDHEAS-G:

1) A set of application-specific CFMs and PIF attributes

IDHEAS-G offers three levels of CFMs, 20 PIFs, and the attributes of every PIF. An application-specific method may choose to use a subset of the CFMs and PIFs. IDHEAS-ECA uses the 5 high-level CFMs, i.e., failure of the five macrocognitive functions, and uses all 20 PIFs but condenses the attributes to a smaller set.

2) Quantitative measures of PIF attributes

Most PIF attributes are continuous variables. Figure 2-3 illustrates that the HEP of a task varies as a continuously nonlinear function of the measure of a PIF attribute.

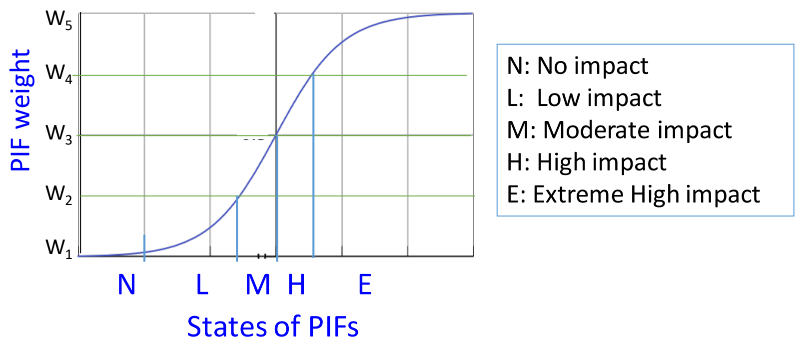

Figure 2-3 Illustration of HEP varying as a function of the measure of a PIF attribute

However, many PIF attributes may not be modeled as continuous variables because there may not be enough data to support a continuous relationship between a PIF attribute and its impact on HEPs. In addition, it can be challenging for HRA analysts to quantify a PIF attribute on a continuous scale. For example, workplace luminance varies continuously so does its impact on HEPs, and the luminance does not affect HEPs within a certain range. However, HRA analysts may only have the information of “good visibility” or “poor visibility” regarding workplace luminance. Thus, an application-specific IDHEAS-method would need to specify how to quantitatively represent PIF attributes. IDHEAS-ECA uses the following combination of ways:

- Multiple discrete scales from 1-10 with anchoring for the scales of 1, 5, and 10.
- Several subjective levels such as low, medium, high, extremely high with explanation for each level.
- Binary states, the presence versus absence of the attribute.

The selection of a quantification format for a PIF attribute is informed by the data available and the extent that the PIF attribute changes HEPs. The datapoints in IDTABLE-1 through IDTABLE-20 were used to define PIF attribute measures and relate these measures to base HEPs or PIF weights. For example, the datapoints in IDTABLE-6 on PIF Cold, Heat, and Humidity in Workplace show that the effect of cold on HEPs continuously vary with work environment temperature. However, coldness within habitable temperature ranges increases the HEP 1.1-2 times, while the effect can be up to 3-5 times in the extreme cold environment. Therefore, the attribute can be represented with two states: cold and extremely cold.

3) HEP quantification

An application-specific IDHEAS method may choose to quantify HEPs through expert judgment, modeling, or statistic regression of available human error data. IDHEAS-ECA uses an HEP quantification model as follows (details described in Section 2.1)

- The HEP of an important human action is the probabilistic sum of Pt and Pc.
- P_{CFM} is the probability of a CFM. The calculation of P_{CFM} for any given set of PIF attributes is estimated as:

$$P_{CFM} = P_{CFM\,Base} \cdot \left(1 + \sum_{i=1}^{n}(w_i - 1)\right) \cdot \frac{1}{Re}$$

$P_{CFM\,Base}$ is the base HEP of a CFM, w_i is the PIF impact weight for a PIF attribute, Re is a factor that accounts for the potential recovery from failure of a critical task, and it is set to 1 by default.

The following are needed to use this model for calculation of HEPs for a given set of PIF attributes:

- Five lowest HEPs for the five CFMs;
- The base HEPs of the five CFMs at the various states of every attribute of the base PIFs;
- The PIF weights for every CFM and every attribute of the remaining 17 PIFs as the PIF attributes vary from the no-impact state to a poor state.

The above parameters needed for IDHEAS-ECA were derived from IDHEAS-DATA. The next section describes the process of integrating the generalized data in IDHEAS-DATA to infer those parameters.

2.5.2. The process of integrating human error data

The generalized human error datapoints in IDTABLE-1 thorough IDTABLE-21 can be referred to as the following:

- Single-component datapoints – A datapoint has the error rate for a single CFM with the presence of a single PIF attribute;
- Multi-component datapoints – A datapoint has the error rate for more than one CFMs, or with the presence of more than one PIF attribute.
- Bounding datapoints – A datapoint has the error rates calculated or estimated from whole events or scenarios. Such error rates are for the combination of multiple CFMs and PIF attributes. Thus, the effect of a PIF attribute on individual CFMs is inseparable in the human error data. Such datapoints cannot be directly used for calculating the base HEPs and PIF weights, but they can be used to anchor or bound the estimated HEPs or PIF weights.

The process of integrating human error data is described as follows:

1) Use single-component data to make initial estimation of the base HEPs and PIF weights;
2) Use the initial estimation to detach multi-component data into single-component ones. For example:
    - A datapoint has the error rate of a task that requires *Understanding* and *Decisionmaking.* The reported error rate is thus divided by two for each CFM unless the data source has information suggesting otherwise.
    - A datapoint has the error rate for the presence and absence of a base PIF attribute while the task was performed with time constraints. Therefore, the error rate is the probabilistic sum of ***Pt*** and ***Pc***. ***Pt*** can be estimated as the error rate for the absence of the PIF attribute subtracted by the lowest HEP for the CFM, then ***Pc*** for the presence of the PIF attribute is the reported error rate subtracted by the estimated***.*** Otherwise, if the data source suggests that the time availability is different for the presence vs. absence of the PIF attribute, then ***Pt*** needs to be adjusted accordingly.
    - If the multi-component error rates cannot be detached, they can be used for the range of the base HEPs or PIF weights. For example, if a datapoint has an error rate measured at the presence of two PIF attributes and the data source does not have information about the contribution of each individual attribute, then the PIF weight calculated from the error rate corresponds to the combined weight of the two attributes, thus the weights of the two attributes should be less or at most equal to the calculated weight.
3) Integrate all the data available from the single-component and detached multi-component datapoints to estimate the range and mean of a base HEP or PIF weight.

4) Use the unspecific datapoints to calibrate the estimated HEPs and PIF weights.
5) Use the mean values as the new initial estimation to iterate the process 2), 3), and 4) until the obtained mean values represent the breath of the available data.

Theoretically, the above process could be done with multi-variable fitting or statistic regression methods. However, given the limited sample size of the available data and the large number of variables (base HEPs, different scales or states of PIF attributes), the parameters obtained through multi-variable fitting would be unstable and highly dependent of the choice of the initial estimation. The NRC staff manually performed the data integration for IDHEAS-ECA in 2019.

The critical step in the process is detaching multi-component datapoints. It requires a thorough understanding of the data sources. Often, it requires reading additional research papers by the same authors or the papers on similar topics by other authors to fully understand the task performed and the variables involved in the study.

To detach multi-component data, the lowest HEPs of the CFMs were first estimated from IDTABLE-21. Using the lowest HEPs, the multi-component datapoints in IDTABLE-1, IDTABLE-2, IDTABLE-3 for the three base PIFs were detached and the base HEPs were then estimated. With the estimated base HEPs, the multi-component datapoints in IDTABLE-4 through IDTABLE-20 were detached using the iterative process described above.

2.5.3. Approaches of integrating human error data for IDHEAS-ECA

There are mathematical or statistical approaches for dealing with uncertain, aggregated, and/or truncated/censored data. Those approaches can be as simple as calculating the mean of the numeric values of a data set or the weighted average by some weighting rules, or as sophisticated as multi-variable fitting. However, the confidentiality in integrating a set of data to generate a single representative value or probabilistic distribution depends on the sample size and quality of the data set. For example, if the numeric values of the data are not continuously distributed, the mean of the numeric values does not represent the center of distribution of the data set.

As of 2019, the data generalized in IDHEAS-DATA were limited. Even when there were multiple datapoints for one HEP or PIF weight, they did not constitute a continuous distribution. Moreover, some PIF attributes had no datapoint generalized. Therefore, when the NRC staff integrated the data for the IDHEAS-ECA method in 2019, they applied several approaches depending on the availability of the generalized data. The NRC staff used aggregation, interpolation, reasoning, and engineering judgment on a case by case basis to generate the lowest HEPs of the CFMs, base HEPs, and the PIF weights in IDHEAS-ECA. The following are the descriptions of the approaches used in the integration:

1) Aggregation of multiple datapoints for a base HEP or PIF weight

The human error data were first evaluated for practicality and uncertainties in the source documents. NPP operational data that were systematically collected for HRA had the highest practicality. The following categories of data sources have the practicality from high to low:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

The single-component high-practicality data were first used to anchor a base HEP or PIF weight and other datapoints were used to adjust the uncertainties in the high-practicality datapoints. If there was no high-practicality NPP operational data, the mean of the datapoints were used as the initial estimation.

2) No single-component data exclusive for a base HEP or PIF weight, but there were multi-component datapoints on the combined effects of several CFMs and/or PIF attributes

When there were multiple datapoints with combined effects of two or more CFMs, PIF attributes, and/or time constraints, detaching was performed using the initial estimations of a base HEP or PIF attribute weight. When there were only a few datapoints or a variety of CFMs and PIFs involved in the datapoints, the range of the combined base HEPs or PIF attribute weights was calculated and the middle of the range was assigned the base HEP or PIF weight.

3) No datapoint for a PIF weight

The available data in the IDHEAS-DATA do not have numeric human error information for many attributes in the PIFs such as Work Process or Teamwork and Organizational Factors. Yet, there have been studies demonstrating that those attributes impact human performance in measures other than human error rates, such as increasing personnel' workload or reducing situational awareness. We assigned the PIF weight as 1.1 or 1.2 for those attributes, pending for future updates as relevant human error data become available. The rules used are the following:

i) There are data sources showing detrimental effects of a PIF attribute on some task performance measures but the relation between the task performance measure and human error rates could not be determined, the PIF attribute weight was assigned as 1.2.
ii) There are data sources showing quantitative detrimental effect of a PIF attribute on task performance (e.g., through subjective rating, observations, or root causal analysis) but there was not task performance data available, the PIF attribute weight was assigned as 1.1.

4) Consistency checking and adjustment with benchmark values

After the initial base HEPs and PIF weights are developed, they are checked for internal consistency against the literature that ranks the likelihood of certain types of human errors and the contribution of various PIFs. We also used reported rates of human events and estimated HEPs from the NRC 2018 FLEX HRA expert elicitation as benchmarks to check and adjust some base HEPs and PIF weights within their uncertainty ranges.

Chapter 3 the RESULTS section will present several examples to demonstrate how these approaches were used for obtaining the base HEPs or PIF weights in IDHEAS-ECA.

As more sources of data are generalized to IDHEAS-DATA, there will be multiple datapoints of various sources for a PIF attribute. Before using the data to inform HEP estimation, the context and uncertainties of the data should be evaluated for their reliability and relevance to the HRA application of interest. For example, if the HRA application is for a well-trained crew implementing EOPs in an NPP control room, the analyst may choose to use only the data collected from NPP operator training simulation and not use the data from cognitive experiments in which tasks were performed by college students. However, if there is no NPP operation data

available, then using data from other domains is better than not using any data to inform the HEPs of NPP operation.

In summary, Chapter 2 describes how IDHEAS-G is used as a framework for generalizing human error data of various sources. Human error data and empirical information are generalized into 27 IDTABLEs. The generalized data can inform base HEPs, PIF weights, and other elements in any IDHEAS applications that use the IDHEAS-G quantification model. For every human error data source, the task performance errors are mapped to IDHEAS-G CFMs, and the context of task performance is mapped to IDHEAS-G PIF attributes. Specifically, IDHEAS-G are in the same framework as the SACADA database; thus, it is relatively straightforward to use SACADA data for the HEP estimation in IDHEAS-G. Engineering judgment is still needed to map the data sources to IDTABLEs. Thus, every TABLE specifically documents uncertainties in data sources as well as in the generalization process. Growing experience and lessons learned in generalizing human error data should be captured to improve process.

3 RESULTS

This chapter describes the data generalized in IDHEAS-DATA as of June 2020 and demonstrates the integration of the data into the base HEPs and PIF weights for IDHEAS-ECA. The 27 IDHEAS-DATA Tables contain the generalized data and are presented in the appendices. Section 3.1 presents an overview of the data sources and the summary and observations of the generalized data. Section 3.2 demonstrates the integration of the data in IDHEAS-DATA to inform HEPs in HRA and shows the step-by-step process (described in Chapter 2) to estimate the base HEPs and PIF weights.

3.1. Overview of the Data Sources and Summary of the Generalized Data in IDHEAS-DATA

Section 3.1 has 27 subsections, one for each IDHEAS-DATA IDTABLE. Each subsection introduces the IDTABLE, presents an overview of the data sources, and summarizes the data generalized. Most subsections also discuss the gaps in the data generalized and perspectives of expanding the data sources.

3.1.1. IDTABLE-1 for Scenario Familiarity

Introduction to PIF Scenario Familiarity

When a scenario is familiar to personnel, it has predictable event progression and system dynamics, and it does not bias personnel's understanding of what is happening. Unfamiliar scenarios can pose challenges to personnel in understanding the situation and making decisions. In addition, compared to familiar scenarios, responses to unfamiliar scenarios could entail greater uncertainties in detecting information, executing actions, and coordinating interteam activities. In unfamiliar scenarios, personnel are more likely to perform situation-specific actions not specified in the procedures.

The following are the identifiers and short descriptions of the attributes for Scenario Familiarity. The details of the attributes can be found in Table A1-1 of Appendix A1.

- SF0 - No-impact, frequently performed tasks in well-trained scenarios, routine tasks
- SF1 - Unpredictable dynamics in known scenarios
- SF2 - Unfamiliar elements in the scenario
- SF3 - Scenario is unfamiliar
- SF4 - Bias, preference for wrong strategies, or mismatched mental models

Summary of the Data Sources

The data generalized for this PIF are presented in Table A1-2 of Appendix A1, IDTABLE-1. The data sources for Scenario Familiarity are organized into the following categories:

A. Operational data and simulator data in the nuclear domain

B. Operational data of human performance from non-nuclear domains

C. Experimental data in the literature

D. Expert judgment of HEPs in the nuclear domain

E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – Preischl and Hellmich [4, 5] analyzed German NPP maintenance event database and identified human errors in the events by types of tasks and PIFs. Several factors correspond to Scenario Familiarity attributes, such as tasks being "frequently performed," "rarely performed," and "extremely rarely performed." The study presented 67 human error rates for different types of tasks under different combinations of PIFs. The error rates were calculated as the number of times the errors were made divided by the number of times the tasks were performed. Another data source in Category A is the SACADA database [24-26], which collects NPP operators' task performance data in simulator training for requalification examination. Using the SACADA data available until April 2019, Chang calculated the rates of unsatisfactory performance ("UNSAT") for training objective tasks when a situational factor is checked versus not checked. The UNSAT rates are generalized in IDTABLE-1 for the applicable CFMs of the tasks and PIF attributes representing the situation factors. For example, the UNSAT rate for diagnosis tasks is 1.2E-1 and the UNSAT rate for decisionmaking is 1.1E-2 where the familiarity factor in SACADA was characterized as "Anomaly" among the three available options (Standard, Novel, and Anomaly) The generalized data points are shown in the following:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
SF3.1	U	1.2E-1 (8/69)	NPP operators diagnose in simulator training	Anomaly scenario	(Other PIFs may exist)	[26]
SF3.1	DM	1.1E-2 (1/92)	NPP operators decisionmaking in simulator training	Anomaly scenario	(Other PIFs may exist)	[26]

Category B – Human error data from operational performance relevant to Scenario Familiarity are available in medicine dispensing, aviation, railroad maintenance, and oil ship control industries. Those are operational data measured from professional personnel. For example, the study of target monitoring for collision avoidance in simulated oil ship control [27] reported that the error rates for detecting collisions is 1.4E-2 for alerting targets in normal responses and 1.1E-1 for alerting targets in emergency responses. The error rates were measured while the ship operators performed dual tasks. The generalized data points are shown as follows:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
SF1.1	D	1.4E-2	Collision avoidance and target monitoring in simulated ship control	Alerting target, normal response	Dual task	[27]
SF1.1	D	1.3E-2	Collision avoidance and target monitoring in simulated ship control	Alerting target, routine response	Dual task	[27]
SF1.1 & SF2.1	D	1.06E-1	Collision avoidance and target monitoring in simulated ship control	Alerting target, emergency response	Dual task, (Time urgent)	[27]

Category C – There are limited experimental studies measuring human error rates under Scenario Familiarity because it needs professional personnel to be familiar with scenarios. One experimental study [28] examined the predictability of scenarios on diagnosing patterns and personnel using structured information to guide diagnosis; the reported error rate was the inaccuracy of diagnosing patterns. The CFM applicable to diagnosis errors is Failure of Understanding. The applicable PIF attribute is SF1.2 "Unpredictable dynamics." In addition, PIF

Task Complexity was also involved in the tasks. The generalized data points are shown in the following:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
SF0	U	4E-2	diagnosing a pattern; personnel uses structured information to guide diagnosis	Predictive situation	Task complexity	[28]
SF1.2	U	1.2E-1	diagnosing a pattern; personnel uses structured information to guide diagnosis	Unpredictive situation	Task complexity	[28]

Category D – The expert judgment for IDHEAS At-Power Application [6] estimated the HEPs of 14 crew failure modes under different combinations of relevant situational factors. Six domain experts from US nuclear regulatory and industry followed a formal expert elicitation procedure specified in the NRC's expert elicitation guidance [29] to estimate the HEPs of crew failure modes for licensed crew performing EOPs in MCRs. The expert panel estimated the HEPs of a crew failure mode given the situation factors affecting the failure mode. For example, the HEP estimated for the failure mode "Failure of attending to the source of information" is 4E-3 under the condition "Poor familiarity with the Source of information," and the HEP for "dismiss/discount critical data" in situation assessment is 2.5E-1 under the condition that personnel formed biases on the situation. These are generalized as two datapoints in IDTABLE-1 as follows:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
SF2	D	4E-3	Attend to source of information (HEP)	Poor familiarity with the Source	Crew with peer-checking	[6]
SF4	U	2.5E-1	Situation assessment in EOP (HEP of Critical Data Dismissed / Discounted)	Inappropriate Bias formed	Crew with peer-checking	[6]

Category E – No datapoint from this category was generalized.

Summary of Human Error Data for Scenario Familiarity

The generalized human error data are summarized according to the CFMs. The ranges of the generalized error rates for the CFMs were examined. The numbers are directly from IDTABLE-1 without detaching the effects of other CFMs and PIFs. The ranges show the general trends of the HEPs.

- Failure of Detection (D) – The error rates for Failure of Detection vary in the range of 5E-4 to 0.2 as the PIF attributes vary from SF1 "Unpredictable dynamics in known scenarios" to SF4 "Bias, preference for wrong strategies, or mismatched mental models." An exception is that the error rate under biases and inadequate time is 0.5, based on medicine dispensing data.
- Failure of Understanding (U) - The error rates vary in the range of 1E-3 to 0.25.

- Failure of Decisionmaking (DM) - The error rates vary in the range of 1E-3 to 0.1. An exception is that the error rate under biases and inadequate time is 0.5, based on medicine dispensing data.
- Failure of Action Execution (E) – There are many datapoints for this CFM from the analysis of German NPP maintenance event report database. The error rates range from 1E-4 for frequently performed simple tasks to 0.33 for “extremely rarely performed” tasks.
- Failure of Interteam Coordination (T) – There are no generalized data points for this CFM. Most studies on the interteam coordination only have qualitative results.

Observations from the Generalized Data

The following are some observations from the data in IDHEAS-DATA IDTABLE-1.

- Several sets of datapoints that have SF3 “Scenario is unfamiliar” varied from “frequently, routinely performed” to “rarely or extremely rarely performed.” The error rates in these sets of datapoints can vary from 1E-4 to 0.33. That is up to three orders of magnitude. It is an evidence that the PIF Scenario Familiarity is a base PIF such that it alone can drive the HEPs from the lowest to a very high value.
- Several datapoints are for SF0. These datapoints provide the basis for the lowest HEP when none of the PIF attributes have an impact on the HEP, i.e., none of the PIF attributes are present. These datapoints belong to IDHEAS-DATA IDTABLE-21 for the lowest HEPs. Yet, having them here serves as the references for the HEP values of the attributes.
- There are eight datapoints from the expert judgment in IDHEAS-AtPower Application (NUREG-2199, Vol. 1) [6]. The expert judgment was made for actions performed by well-trained, licensed crews using emergency operating procedures (EOPs) in NPP control rooms. Thus, the estimated HEPs are implicitly assumed for the tasks performed with good peer-checking and supervision. This assumption does not apply to many other datapoints in IDTABLE-1.

The data sources identified from Category A or B are limited for the CFM Failure of Decisionmaking. On the other hand, there is a large volume of task performance data involving decisionmaking in Category C data sources. Extensive research has shown that Scenario Familiarity is essential for correct decisionmaking. Yet, most experimental studies use task performance measures instead of error rates. Such data sources were not selected for generalization to IDTABLE-1 because it is difficult to derive the quantitative relationship between the reported task performance measures and human error rates.

3.1.2. IDHEAS-DATA IDTABLE-2 for Information Completeness and Reliability

Introduction to the PIF Information Completeness and Reliability

Personnel need information to perform tasks. Information is expected to be complete, reliable, and presented to personnel in a timely and easy-to-use manner. Large amounts of information in operation are expected to be preprocessed and organized for personnel. Yet, information in event scenarios could be incomplete, unreliable, untimely, or incorrect. Personnel receive information via sensors, instrumentation, alarms, oral communication, local observation, or other means. Information that is obtained from sensors and instrumentation are usually presented to personnel with the human-system interface (HSI) such as indicators and displays. There are

situations that local observations and oral transmittal of information are the only available options to obtain information.

Personnel rely on key information to understand the situation and make decisions. In event scenarios, key information may be unavailable, unreliable, or even misleading. For example, sensors or indicators may be unreliable or display incorrect values (e.g., damaged or degraded while appearing to be working, false alarms in design, out-of-range, or inherently unreliable sources). There could also be flaws in system state indications, e.g., an indicator shows the demanded position of a component or control function rather than the actual equipment status. (An example was the pressurizer pressure operated relief valve indications at Three-Mile Island, which showed that the valves were closed, while one of those was not closed.)

This PIF is defined as the availability and reliability of key information in personnel's performing the macrocognitive functions of Understanding and Decisionmaking, thus the PIF affects the CFMs Failure of Understanding and Failure of Decisionmaking. The effect of information quality on other CFMs, Failure of Detection, Action Execution, and Interteam Coordination are modeled by other PIFs such as Task Complexity or HSI.

The following are the classes of attributes for Information Completeness and Reliability. The full set of attributes can be found in Appendix A.

INF0 No impact – Key information is reliable and complete
INF1 Key information is incomplete
INF2 Information is unreliable

Summary of the Data Sources

The data generalized for this PIF are presented in IDHEAS-DATA IDTABLE-2. The following categories are used to overview data sources for Information Completeness and Reliability:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – SACADA database collected operators' performance on diagnosis and decisionmaking in simulator training for requalification examinations. Based on the SACADA data available by April 2019, the UNSAT rates of Diagnosis and Decisionmaking are generalized as IDHEAS-DATA IDTABLE-2 datapoints for the corresponding PIF attributes. Because other PIF attributes (having negative effects on performance) may also exist in the datapoints, the calculated UNSAT rates from SACADA data could be higher than the case of no presence of other PIF attributes. Other NPP simulation data sources are the International HRA Empirical Study and the US HRA Empirical Study.

The International Empirical Study [23] had scenarios where the cues of the problem were difficult to detect. That can be represented by the PIF attribute INF1.5 "Information is largely incomplete - Key information is masked or key indication is missing." The complex LOFW scenario started with a loss of feedwater. The condensate pump used for feedwater injection was successfully running, leading the crew to depressurize the SGs to establish condensate flow. However, the running condensate pump was degraded and gave a pressure so low that the SGs became empty before the pressure could be reduced enough to successfully inject

water. Therefore, the SGs' water levels started to decrease. Another scenario complication was two of the three SGs' wide range (WR) level indications failed. One failed at 16% and another 14%. After failing, their indications remained constant (at 16% and 14%). The operator needed to initiate feed-and-bleed cooling when two SGs' WR indications fell below 12%. In this study, 7 out of 10 crews failed to initiate feed-and-bleed when the criteria was reached.

PIF	CFM	Error rates	Task (and error measure)	PIF Measure	Other PIFs (and Uncertainty)	REF
Inf1.5	U	0.7 (7/10)	Initiate feed-and-bleed	Two SG WR levels were indicated incorrectly	Operators were busy with trying to depressurize SGs	[23]

Category B – The data sources identified in non-nuclear domains are from operational reports and studies of high-fidelity simulations of human performance, such as pilots flying simulators, air traffic controllers controlling traffic, licensed drivers avoiding collision, physicians' diagnosing, and pharmacists dispersing medicines. For example, Sarter et. al. [30] studied pilots' decisionmaking of deicing with different information displays: untimely information with a baseline display, timely information with an additional status display, and 30% unreliable information on the status display. The failure to prevent stall was 7.9%, 20.6%, and 73.6% for each situation. Pilots performed the simulated tasks under time pressure. Note that the aircraft stall has two applicable CFMs: Failure of Understanding and Failure of Decisionmaking. The data are generalized in DIHEAS-DATA IDTABLE-2 as following:

PIF	CFM	Error rates	Task (and error measure)	PIF Measure	Other PIFs (and Uncertainty)	REF
Inf0	U & DM	7.9E-2	Pilots in flight deicing (Percentage of early buffet)	Accurate information timely with status displays	Inadequate time	[30]
Inf1.1	U & DM	2.06E-1	Pilots in flight deicing (Percentage of early buffet)	Accurate information not timely without status displays	Inadequate time	[30]
Inf2.6	U & DM	7.36E-1	Pilots in flight deicing (Percentage of early buffet)	(30%) inaccurate information on status displays	Inadequate time	[30]

Category C – There are many experimental studies about information availability or reliability on human performance of tasks requiring understanding the situation or making decisions. The controlled experimental studies measured human error rates while systematically varying the level of information availability or reliability. Only a few data sources from this category have been generalized to IDHEAS-DATA IDTABLE-2 so far. One example is the study by Albantakis and Deco [31] that measured college students 2- and 4-alternative- choice decisionmaking errors while systematically varying the percent of information coherence or consistency. The result showed that the error rates varied with the percent of information coherence in a logistic function. The result is generalized in IDHEAS-DATA IDTABLE-2 as one datapoint, but it consists of continuously varying error rates.

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF

Inf2.4	DM	Sigmoid function 0-0.4	Students make 2-alternative choices	100% to 10% of information coherence		[31]
Inf2.4	DM	Sigmoid function 0-0.6	Students make 4-alternative choices	100% to 10% of information coherence		[31]

Category D – In the expert judgment for IDHEAS At-Power Application (NUREG-2199, Vol. 1)[6], six domain experts from the NRC and nuclear industry followed a formal expert elicitation procedure specified in the NRC's expert elicitation guidance [29] to estimate the HEPs of crew failure modes for licensed crews performing EOPs in MCRs. For example, the expert panel estimated the HEPs for the failure mode "failed to use alternative source of information" in situation assessment under the conditions "Primary source of information NOT obviously Incorrect" and "Primary source of information obviously Incorrect." The estimated HEPs are 0.012 and 0.32, respectively. These are generalized as two datapoints in IDHEAS-DATA IDTABLE-2 as following:

PIF	CFM	Error rates	Task (and error measure)	PIF Measure	Other PIFs (and Uncertainty)	REF
Inf2.3	U	1.2E-2	MCR critical tasks with EOPs (failed to use alternative source of information)	Primary source of information obviously Incorrect	Licensed crew with peer-checking	[6]
Inf2.6	U	3.2E-1	MCR critical tasks with EOPs (failed to use alternative source of information)	Primary source of information NOT obviously Incorrect	Licensed crew with peer-checking	[6]

Category E – The source data in this category are not generalized.

Summary of Human Error Data for Information completeness and reliability

The generalized human error data are summarized according to the cognitive failure modes (CFMs). The range and trends of the generalized error rates for the CFMs are roughly examined. The numbers are directly from IDHEAS-DATA IDTABLE-2 without detaching the effects of other CFMs and PIFs, thus they cannot be used for inferring the HEPs. Nevertheless, the ranges show the general trends of the HEPs.

- Failure of Understanding (U) - The error rates vary in the range of 3.3E-3 to 0.9. The lowest error rate was from the expert judgment of HEP for the situation "indications not reliable" (NUREG-2199), and the highest error rate was NPP crews failing to diagnose the ISLOCA due to information being masked (International Study).

- Failure of Decisionmaking (DM) - The error rates vary in the range of 4.5E-2 to 0.89, except for an experimental study in which the error rate continuously varied from 0 to 0.6. The lowest error rate of 4.5E-2 in operational data was for making incorrect task plans in the maintenance of a cable production process due to information not being organized or missing; the highest error rate of 0.89 is from pilots deicing decisions while 30% of the key information was unreliable [30].

Observations from the Generalized Data
Several observations were made from the generalized data in IDHEAS-DATA IDTABLE-2.

- The human error rates in the datapoints from the same data source varied from the range of E-3 to close to 1. This variation is evidence that the PIF, Information

Completeness and Reliability, is a base PIF. That PIF alone can drive the HEPs from very low to very high values.

- The experimental study of varying information coherence from 100% to 10% resulted in error rates from nearly zero to 0.6. The resulted error rates varied as a logistic function of the percent of information coherence. This kind of logistic function between human error rates and the measure of a PIF attribute has been reported in many experimental studies, as shown in IDTABLE-2.
- The experimental study of varying information coherence shows that the error rates of decisionmaking is nearly zero when the information is reliable (100% coherence) regardless of if the decision is a 2-alternative or 4-alternative choice. When the information is not reliable, the error rates are higher for the 4-alternative choices than the 2-alternative choices because more choices add uncertainty to the decisionmaking process.

Extensive data sources, as shown in IDTABLE-2, are available for the PIF Information Availability and Reliability. The generalized data shows that the PIF is essential for situation assessment and decisionmaking. There are a large volumes of human error data on this PIF in controlled experimental studies. Only a few examples from that category of data sources were generalized given that there are already many data points in Category A and B. Nevertheless, more studies on this PIF with NPP operators are desirable to calibrate the effects of individual PIF attributes on the HEPs of Understanding and Decisionmaking.

3.1.3. IDHEAS-DATA IDTABLE-3 for Task Complexity

Introduction to the PIF Task Complexity

Task Complexity, also referred as cognitive complexity, measures task demand for cognitive resources (e.g., working memory, attention, executive control). Nominal complexity refers to the level of complexity that is within the capability limits of cognitive resources thus does not overwhelm personnel. The cognitive complexity of a task has two parts: the complexity in processing the information to achieve the macrocognitive functions of the task, and the complexity in developing and representing the outcomes to meet the task criteria. For example, a task is to monitor a set of parameters, and the outcome is to identify the parameters outside a certain range or determine the trends of the parameters. The latter imposes higher cognitive demands on personnel's working memory; thus, it is more complex. Complexity is characterized by the quantity, variety, and relation of the items to be processed or represented in a task [32, 33].

There are over 30 attributes for Task Complexity. They are grouped by the macrocognitive function they impact. The following are the identifiers and short descriptions of the attribute groups. The full set of attributes can be found in Appendix A.

- C1 - C7 Detection complexity
- C10 - C16 Understanding complexity
- C20 - C28 Decisionmaking complexity
- C30 - C39 Execution complexity
- C40 - C44 Coordination complexity

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A3 IDHEAS-DATA IDTABLE-3. The data sources for Task Complexity are organized into the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – The International HRA Empirical Study [22, 23]and the US HRA Empirical Study [16]both varied Task Complexity in the tested scenarios. The error rates were calculated for the crews' performing important human actions. Yet, the error rates were for the failure of the entire human actions which typically had more than one applicable CFM and PIF attribute. The SACADA database [26] collects UNSAT data in several complexity attributes: C1 - Detection overload with multiple competing signals, C6 - Cue or mental model for detection is ambiguous or weak, C31 - Straightforward procedure execution with many steps, and C32 - Non-straightforward procedure execution. The Korea Atomic Energy Research Institute (KAERI) operator simulator training database [34] also has data for several complexity attributes. A shortcoming in these data is that additional PIF attributes may exist in the recorded events, thus the reported UNSAT rates or error rates can be higher than the HEPs of those attributes. The German maintenance event report database [4, 5] has multiple datapoints for execution complexity. The following are several example datapoints from the analysis of the German maintenance event report database:

PIF	**CFM**	**Error rates**	**Task (and error measure)**	**PIF Measure**	**Other PIFs (and Uncertainty)**	**REF**
C31		3.3E-3 (2/651)	NPP maintenance (omitting an item of instruction)	Procedure execution with many steps		[4, 5]
C32		4.8E-3 (1/211)	NPP maintenance tasks	Long procedures, voluminous documents with checkoff provision		[4, 5]
C33		2.6E-3	Controlled actions that require monitoring action outcomes and adjusting action accordingly	Manipulating dynamically		[4, 5]

Category B – Many studies analyzed pilots and air traffic controller operational errors. Most of the studies did not relate error rates to Task Complexity. Prinzo et. al. [8, 9] analyzed pilots' errors in readback of air traffic controller clearance. The error rates were analyzed against two factors: message length in one transmission corresponding to C12- Relational complexity (Number of topics or relations in one task), and message complexity corresponding to C11 - need to decipher numerous messages (indications, alarms, spoken messages). The data are generalized in DIHEAS-DATA IDTABLE-2 as follows:

PIF	**CFM**	**Error rates**		**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
C12	U	Message relation	Error rate	Pilots listen to and read back key messages	Message relation (# of aviation topics to be related in one communication) = 1, 2, 3, and 4.	(Mixture of normal and emergent operation so other PIF attributes may exist)	[8, 9]
		1	0.038				
		2	0.061				
		3	0.085				

		4	0.26				
C11	U	# messages	Error rate	Pilots listen to and read back key messages	Message complexity - # of key messages in one transmit	(Mixture of normal and emergent operation so other PIF attributes may exist)	[8, 9]
		5	0.036				
		8	0.05				
		11	0.11				
		15	0.23				
		17	0.32				
		>20	>0.5				

Category C – There are many experimental studies about the effect of complexity on human performance. The data sources generalized from this category are primarily the controlled experiments performed with high-fidelity simulators, such as flight simulators, air traffic control simulation interfaces, driving simulators, and process control simulation. One example is the study that measured military professionals responding to compelling signals. The result showed that the error rates increased with the increasing number of annunciators to be attended. The generalized datapoints from this study are in the following:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
C1	D	0.0001 to 0.05	Respond to compelling signals	the number of annunciators from 1 to 10.		[35, 36]
C1	D	0.10 to 0.20	Respond to compelling signals	the number of annunciators 11 to 40.		[35, 36]
C1	D	0.25	Respond to compelling signals	Annunciators >40		[35, 36]

Category D – Two expert judgment studies estimated HEPs relevant to Task Complexity. The expert judgment for IDHEAS At-Power Application (NUREG-2199, Vol. 1) [6] estimated the HEPs under several attributes of Task Complexity: C22 - Alternative strategies to choose, C23 - Decision criteria are ambiguous, C24 - Advantage to the incorrect strategy, C25- Low preference for correct strategy, and C32- Execution is not straightforward. The HRA for nuclear waste facility operation [37] estimated the HEPs for C1 – Detection overload with multiple competing signals and C31 - Straightforward Procedure execution with many steps. Below are two examples of the generalized datapoints:

PIF	CFM	Error rates	Task (and error measure)	PIF Measure	Other PIFs (and Uncertainty)	REF
C31		5E-4	Nuclear facility operation - Execution procedure or script	Moderate (typical) lock out plan (4-10 lockout)	(Estimated HEP)	[37]
C31		5E-3	Nuclear facility operation - Execution procedure or script	Complex lock-out plan (11-100 lockout)	(Estimated HEP)	[37]

Category E – Data sources in this category were not generalized.

<u>Summary of Human Error Data for Task Complexity</u>

The generalized human error data are summarized according to the CFMs. The range and trends of the generalized error rates for the CFMs were examined. The numbers are directly from IDHEAS-DATA IDTABLE-3 without detaching the effects of other CFMs and PIFs.

- Failure of Detection (D) – There is at least one datapoint for every attribute of Detection complexity. There are multiple sets of datapoints for C1 "Detection overload with multiple competing signals." The error rates varied from 2.1E-3 to 5.1E-2 in SACADA data. They varied from 1E-4 to 0.25 when the number of compelling signals varied from less than 10 to greater than 40. The operational data and experimental data show consistent trends in the error rates varying with the number of compelling signals.
- Failure of Understanding (U) - The error rates varied from 3E-3 to 1. One NPP operational datapoint is that operators failed diagnosis in all the four events in which alarms signals might be triggered by maintenance work. No datapoint is generalized for two attributes: C14- Potential outcome of situation assessment consists of multiple states and contexts (not a simple yes or no) and C16- Conflicting information, cues, or symptoms.
- Failure of Decisonmaking (DM) – The error rates for NPP operators performing EOPs on a simulator is 4.5E-3 for transferring to a step in a procedure and 1.23E-2 for transferring to a different procedure. Expert judgment HEPs for NPP operators choosing wrong strategies in EOPs ranged from 9.3E-3 to 1.7E-1. Operational data are not available for three attributes: C26 - Decisionmaking involves developing strategies or action plans, C27 - Decisionmaking requires diverse expertise distributed among multiple individuals or parties, and C28 - integrating a large variety of types of cues with complex logic.
- Failure of Execution (E) – The error rates for maintenance tasks reported in the analysis of the German NPP event reporting database [4, 5] ranged from 1E-3 for simple execution (operating a pushbutton, adjusting values, connecting a cable) to 0.5 for unlearning or breaking away from automaticity of trained action scripts. The error rates from the analysis of SACADA data were 1E-2 for executing simple and distinct actions and 3.4E-2 for executing actions requiring additional mental effort.
- Failure of Interteam Coordination (T) – The only operational datapoint for this CFM is that the error rate for NPP operators notifying/requesting to personnel outside of the main control room is 1.54E-3 [38]. The expert estimated HEPs for nuclear facility operation communication ranged from 1E-3 for simple information to 5E-1 for extremely complex information communicated.

Observations from the Generalized Data

Several observations were made from the generalized data in IDHEAS-DATA IDTABLE-3.

- Task Complexity can vary human error rates from close to 0 to 1. Moreover, some continuously varying attributes alone can result in error rates from close to 0 to 1. This variation is evidence that the PIF Task Complexity is a base PIF and that it alone can drive the HEPs from very low to a very high values.
- The study on pilots' readback errors indicates that the error rates increased rapidly as the number of items to be memorized in a task was greater than 11. Similar results were also reported for detecting compelling signals. Those numbers are consistent with the 9 ~11 items of working memory span reported in many experimental studies [39, 40].
- The study on pilots' readback errors also indicates that the error rate increased significantly as the number of topics in one communication increased to 3 or 4. This is consistent with the large volume of experimental studies showing that human information processing can reliably integrate no more than 4 relations at a time [41].

Extensive data sources are available for PIF Task Complexity. The generalized data show that the PIF is a main driver for human errors. There are lots of human error data on this PIF in controlled experimental studies with isolated simple tasks. Most of those data sources were not generalized given that there were already many datapoints in Category A and B. On the other

hand, the operational data and experimental studies about the effects of complexity on Decisionmaking and Interteam Coordination mostly reported task performance measures or the number of errors made, rather than error rates. Thus, no datapoint was identified for several attributes in Decisionmaking complexity and most of the attributes in Interteam Coordination complexity. While the error rates for those attributes can probably be inferred from task performance measures, operational data for those attributes are desired.

3.1.4. IDHEAS-DATA IDTABLE-4 for Workplace Accessibility and Habitability

Introduction to the PIF Workplace Accessibility and Habitability

Workplace is where personnel perform actions. It has hardware facilities, physical structures, and travel paths to support personnel task performance. Workplace may be in an open, unprotected environment or within a building structure. Those structures should not impede personnel from entering the place needed to perform the required human actions nor impede the performance of the required tasks.

Accessibility may be limited because of adverse environmental conditions and security system operation. For example, accidents or hazards may cause workplace conditions to become less habitable or accessible for a period of time. Adverse environmental conditions include steam, high water, fire, smoke, toxic gas, radiation, electric shock risk, and roadblocks (e.g., because of extreme external hazards). Also, doors and components that are normally locked and require keys to unlock could impact accessibility (e.g., a fire or flood may cause electric security systems to fail locked).

The PIF Workplace Accessibility and Habitability has four attributes:

- WAH1 - Accessibility (travel paths, security barriers, and sustained habituation of worksite) is limited, e.g., traffic or weather impeding vehicle movement
- WAH2 - The surface of systems, structures, or objects cannot be reached or touched
- WAH3 - Habitability is reduced. Personnel cannot stay long at the worksite or experience degraded conditions for work
- WAH4 - The worksite is flooded or underwater.

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A4 IDHEAS-DATA IDTABLE-4. The data sources for Task Complexity are organized into the following categories:

A. Operational data and simulator data in nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – None of the data sources evaluated has human error data. Strom [42] reviewed and summarized the expected health impacts of radiation exposures to people delivered at high dose rates. Two major variables affecting the radiation impact on people are the amount of radiation dose and its distribution in time, that is, dose rate and fractionation. The severity of the effect is an increasing function of dose rate, with a dose threshold below which symptoms do not appear. This study does not have information about the effects of radiation exposure on cognitive abilities and human errors in task performance.

Category B – Cucinotta et al. [43] reviewed radiation risks to human central nervous systems. Possible risks include detriments in short-term memory, reduced motor function, and behavioral changes, which may affect performance and human health. This report summarized space radiobiology studies of central nervous system effects and made a critical assessment of their relevance relative to doses and dose-rates to be incurred on a Mars mission. The report does not have human error data related to radiation. Strangman et. al. [44] reviewed and summarized the cumulative results of existing studies of cognitive performance in spaceflight and analogue environments that are featured with isolation, confinement, and microgravity. The studies consistently suggest that novel environments (spaceflight or other) induce variable alterations in cognitive performance across individuals. However, the reported impairments of cognitive abilities were inconsistent across the studies. The reported data, taken together, cannot be generalized quantitively due to the inconsistency.

Category C – Barkaszi [45] studied cognitive performance of over-wintering crews in an Antarctic station and in a Space Station where the crew experienced long-term isolation, confinement, and microgravity. The results show decreased performance in cognitive tasks. The reported data were neurophysiological measures that were not directly related to human errors.

Category D – NUREG/CR-6545 [46, 47] reported expert judgment of health effects of radiation exposure. The estimated effects were about radiation damage to human health, not about behavioral performance.

Summary of Human Error Data for Workplace accessibility and habitability

No human error data on task performance were generalized for this PIF. This is because 1) the data sources relevant to the PIF attributes did not measure human error rates in behavioral task performance, 2) The reported effects on behavioral task performance were largely inconsistent due to the relatively small subject samples in the studies, and 3) the studies about workplace accessibility such as going into floods were case-specific; therefore, the results could not be generalized to other cases without explicitly knowing the detailed environmental structures. As such, the generalized datapoints for this PIF only document the qualitative effects on human performance without human error data. These datapoints cannot be used to derive PIF attribute weights. Nevertheless, they can be used as reference information for inferencing or experts' judging the PIF weights.

3.1.5. IDHEAS-DATA IDTABLE-5 for Workplace Visibility

Introduction to the PIF Workplace Visibility

Visibility of an object is a measure of easiness, fastness, and precision that the object is visually detected and recognized. It is a function of the difficulty experienced to discriminate an object visually from the background or surrounding environment. Visibility of a task in the workplace is generally determined by visibility of the most difficult element which must be detected or recognized so the task can be performed.

Personnel need to recognize objects and their surroundings to perform tasks accurately and reliably. Visibility at work is related to the illumination of the workplace. It requires a minimum level of illumination at which personnel can detect objects and discriminate spaces between objects. luminance is the most important factor for good visibility. Which is needed to reliably perform activities such as reading, writing, inspecting objects for errors, and distinguishing cues. Poor visibility impairs personnel's detection of information and execution of physical actions that require visual-motor coordination. Moreover, it also affects person's comfort and effectiveness of teamwork. In addition to luminance, visibility is also affected by light distribution such as

reflections or shadows in the workplace. Visibility is also impaired by high luminance, referred to as glare, which means that the brightness is greater than what human eyes are adapted for.

Workplace Visibility has three attributes as follows:

- VIS1 Low ambient light or luminance of the object that must be detected or recognized
- VIS2 Glare or strong reflection of the object to be detected or recognized
- VIS3 Low visibility of work environment (e.g., those caused by smoke, rain, fog, etc.)

<u>Summary of the Data Sources</u>

The data generalized for this PIF are presented in Appendix A5 IDHEAS-DATA IDTABLE-5. The data sources for Workplace Visibility are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – No data source was identified in NPP operation for this PIF. This may be due to the fact that NPP workplaces are designed using appropriate human factors engineering. Yet, poor visibility still may occur during some ex-CR actions, especially under extreme operating conditions.

Category B – Many studies on the effects of visibility have been performed in aerospace, aviation, transportation, and military workplaces. For example, strobing laser glare may present a threat to aircrews. In addition to obscuring the visibility of instruments and terrain (as continuous exposures can), strobing exposures could potentially impede visual motion processing. Beer and Gallaway [48] measured the effects of strobing vs. continuous laser exposure on performance in a visual flight task using a flight simulator. Results showed that strobing laser glare posed a legitimate threat to visual orientation control. The measured tasks were pitch control and roll control. Pilots' performance was measured as the degrees of control errors. The following is the datapoint generalized from this study to IDHEAS-DATA IDTABLE-5:

PIF	CFM	Error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
VIS2	E		Pitch control error (degree)	Roll control error (degree)	Visual flight task on a simulator (control errors)	No laser (N), Strobing (S), Continuous (C)laser exposure		[48]
		No Laser	2	5				
		C	4	9				
		S	10	20				

Category C – There have been numerous experimental studies about the effect of visibility on human performance. In particular, the numeric relationship between object luminance or luminance contrast and human perception errors has been clearly elucidated. While human error rate increases as the luminance or luminance contrast of the visual target decreases, the error rate is unchanged for "good visibility," meaning that the luminance is within the normal range for human vision. For example, Braunstein and White [15] measured human errors in

reading dials as the luminance on the dials was varied from 0.015 to 150 L/m^2. The error rate decreased with luminance. When the luminance was greater than 15 L/m^2, the error rate was low and remained the same. Many other studies reported the similar relation between luminance and error rates. The following is the datapoint generalized from this study to IDHEAS-DATA IDTABLE-5:

PIF	CFM	Error rates		Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
VIS1	D	Luminance	Reading error	dial reading error	Luminance (L/m^2)		VIS-9
		0.15	0.16				
		1.5	0.1				
		>15	0.08				

Category D – The expert judgment study on nuclear waste facility operation [37] estimated HEPs for crane/hoist striking stationary objects under different visibility conditions and the presence or absence of spotters. The following is the generalized datapoint:

PIF	CFM	Error rates		Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
VIS3	E	Spotter present	3E-5	Crane/hoist strikes stationary object	Spotter and visibility	(Expert judgment)	[37]
		No spotter, typical visibility	3E-4				
		No spotter, low visibility	3E-3				

Category E – The source data in this category are not generalized.

<u>Summary of Human Error Data for Workplace Visibility</u>

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-5 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – Most datapoints have error rates vary between 1-5 times from poor to good visibility, with a median value around 2 times. The studies that systematically varied the object luminance showed that human error rates increased about twice when the luminance was decreased two orders of magnitudes from a normal luminance value (15L/m^2).
- Failure of Understanding (U) – No data source was identified about the effect of visibility on Failure of Understanding.
- Failure of Decisionmaking (DM) – No data source was identified about the effect of visibility on failure of Decisionmaking.
- Failure of Execution (E) – Most datapoints have the error rates that vary between 2-10 times from poor to good visibility, with a median around 3 times. Several datapoints have task performance errors instead of error rates. The task performance errors increase between 1-2 times as the visibility vary from a poor to normal condition.
- Failure of Interteam Coordination (T) – Some studies reported observations that low visibility impaired team coordination. However, no quantitative data sources were identified. It is unclear that the observed impairment was due to the effects on

individual's Detection and Action Execution or it is pertinent to team coordination mechanisms.

Extensive data sources are available for the PIF Workplace Visibility. The generalized data shows that the PIF attributes moderately modify human error rates. There is a large volume of human error data on this PIF in controlled experimental studies with isolated simple tasks. Only a few datapoints from such data sources were generalized because the reported results from different studies were highly consistent. The data sources primarily studied the effect of visibility on Detection and Action Execution. It is reasonable to assume that the attributes are not applicable to the Failure of Understanding and Decisionmaking. The impairment in teamwork due to low visibility has been observed but no quantitative data sources were identified.

3.1.6. IDHEAS-DATA IDTABLE-6 for Workplace Noise

Introduction to the PIF Workplace Noise

Noise is unwanted sound disruptive to hearing. Human perceived noise is a function of the sound intensity (loudness), duration, variation of intensity, frequency of the sound waves, and the meaningfulness of the sound. Noise types include continuous sound, intermittent sound, speech, nonspeech, and mixtures of sounds. Continuous noise is constant, with no breaks in intensity. Intermittent noise changes in intensity, having gaps of relatively quiet intervals between repeated loud sounds. A major type of practical distractive noise is speech. Speech is a distracter to which humans are especially attuned.

Noise impairs human performance by interfering with cognitive processing or exerting detrimental effects on mental and physical health. It generally does not influence performance speed, but it reduces performance accuracy and short-term/working memory performance. Accuracy in cognitive and communication tasks is most vulnerable to noise effects.

Humans adapt to the environment and develop various compensatory strategies to alleviate noise effects. Humans can develop effective coping strategies for continuous noise of longer duration. Therefore, noises are typically unfamiliar disruptive sounds. Moreover, some low frequency continuous sounds such as music can increase personnel's alertness. Such sounds in workplaces are not considered as noise.

Workplace Noise has four attributes as follows:

- NOS1 Continuous loud mixture of noisy sounds
- NOS2 Intermittent non-speech noise
- NOS3 Speech noise
- NOS4 Intermittent mixture of speech/noise

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A5 IDHEAS-DATA IDTABLE-6. The data sources for Workplace Noise are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – No data source was identified in NPP operation for this PIF.

Category B – Abundant studies on the effects of noise were performed in aerospace, aviation, and military workplaces in the 1950's to 1960's. The original reports of those studies were not readily available. Later studies about Workplace Noise were primarily focused on health effects and longitudinal work performance, not human errors.

Category C – Numerous experimental studies have investigated the effects of various types of noises on human task performance. A small sample of available reports were selected from this category to represent the PIF attributes and CFMs. Some reports were selected because the noises used in the studies mimic the kind of noise in real workplaces. For example, Schlittmeier et al. [49] examined the effects of road traffic noise on cognitive performance in adults. The study tested the impact of road traffic noise at different intensity levels (50, 60,70dB) on performance in three tasks: The Stroop task, in which performance relied predominantly on attentional functions; a non-automated multistage mental arithmetic task calling for both attentional and working memory; and verbal serial recall, which placed a burden predominantly on working memory. The noise mimics 2000 cars driving by per hour, producing continuous sounds. In addition, the study also tested the noise mimicking 100 cars per hour that produced intermittent sounds. Lastly, the study tested the effect of background speech. The results show that speech has the highest detrimental performance effect, and the intermittent noise has a higher impact than the continuous noise of the same intensity level. The following are the datapoints generalized from the arithmetic task in this study.

PIF	CFM	Error rates		Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
NOS1	All	Silence	0.27	Mental arithmetic performance	NOS1 – 50 to 70DB traffic noise	(The task is for all CFMs)	[49]
		NOS1	0.3				
NOS2		Silence	0.27	Mental arithmetic performance	NOS2 – 60DB intermittent traffic	The task is for all CFMs)	[49]
		NOS2	0.3				
NOS3		Silence	0.27	Mental arithmetic performance	NOS3 – irrelevant speech	The task is for all CFMs)	[49]
		NOS3	0.4				

Category D – No expert judgment data sources were identified for this PIF. In fact, this PIF is physically measurable and adequate data are available to model the effect on human errors. There is no need for expert judgment.

Category E – Given the large amount of literature for this PIF, it is desired to get the reliable quantitative information as to how noise effects vary as a function of the characteristics of the noise itself and of the task to be performed. Szalma and Hancock [50] provided such information by means of a meta-analytic review concerning the influence of noise on human perceptual, cognitive, and psychomotor response capacities, as well as tasks requiring communication of information. The authors performed meta-analyses of noise effects as a function of task type, performance measure, noise type and schedule, and the intensity and duration of exposure. The study analyzed the data from 242 studies and calculated the standardized effect sizes (defined as the difference between the mean error rates with the presence of noise and the mean error rates of control groups divided by the standard deviation of all the error rates). The standard sizes varied as a function of each of those moderators. Collective findings identified continuous versus intermittent noise, noise type, and type of task as the major distinguishing characteristics that moderated response. The analysis results were presented as the standardized effect sizes, not human error rates. The effect size is

proportional to the difference between the measured effects at the testing condition and a baseline condition, normalized by the standard deviation of the data. Although the effect sizes cannot be used directly to infer PIF attribute weights, they provide statistically reliable information on the relative effects of the PIF attributes. The following are some datapoints generalized from this study. The first four rows' datapoints have the effect sizes of nonspeech noise on different CFMs, but they do not differentiate continuous versus intermittent noise. The last two rows have the effective sizes respectively for continuous and intermittent noise without distinguishing the CFMs.

PIF	CFM	Effect Size of Error Rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
NOS1 / NOS2	D	-0.2 [a]	Perceptual (Effect size)	Nonspeech		[50]
NOS1 / NOS2	U / DM	-0.21	Cognitive (Effect size)	Nonspeech		[50]
NOS1 / NOS2	E	-0.49	Motor (Effect size)	Nonspeech		[50]
NOS1 / NOS2	T	-0.43	Communication (Effect size)	Nonspeech		[50]
NOS1	All	-0.26	(Effect size)	Continuous noise		[50]
NOS2	All	-0.39	(Effect size)	Intermittent noise		[50]

[a] Effect size being negative means that the error rates due to the presence of the PIF attributes are reduced from the control condition of no PIF attributes.

Summary of Human Error Data for Workplace Noise

The generalized human error data are summarized according to CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-5 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – The datapoints have the error rates that vary between 1.1 to 1.5 times from no noise to high noise. Continuous low intensity noise has little effect on detection. While the effect of noise on detection errors increases with noise intensity, the changes are moderate.
- Failure of Understanding (U) – The datapoints have the error rates that vary between 1.1 to 1.4 times from no noise to noise conditions. Speech has the highest detrimental performance effect for Understanding.
- Failure of Decisionmaking (DM) – No generalized datapoint is for this CFM alone. Schlittmeier et al. [49] examined the effect of noise on three cognitive tasks that demand attention and working memory, which are the cognitive mechanisms of decisionmaking. The error rates in those tasks increased 1.1 to 1.4 times from no noise to the noise condition. The datapoints generalized from Szalma and Hancock [50] meta-analysis shows that the effect size for Understanding/Decisionmaking is -0.21, comparable to -0.2 of the effect size for Detection with nonspeech noise. However, the effect size for Understanding/Decisionmaking is -0.84 with speech noise.
- Failure of Execution (E) – The datapoints specific for failure of Execution have the error rates ~1.5 times from no noise to noise conditions. Szalma and Hancock [50] meta-analysis shows that the effect size for Execution is -0.49, about 2.5 times of the effect size for Detection.
- Failure of Interteam Coordination (T) – Szalma and Hancock [50] meta-analysis shows that the effect size for Execution is -0.43, about two times of the effect size for Detection.

Notice that the study analyzed the noise effect on communication without separating within-team or interteam communication.

Extensive data sources are available for the PIF Workplace Noise. The generalized data show that the PIF attributes only moderately modify human error rates. The highest detrimental performance effect is speech for Understanding/Decisionmaking. Overall, the effect of this PIF on human error rates is weak. Yet, notice that most studies on noise effects used normal levels of noise that would be present in most workplaces. The effect can be much more detrimental under some extreme operating conditions.

3.1.7. IDHEAS-DATA IDTABLE-7 for Workplace Temperature

<u>Introduction to the PIF Workplace Temperature</u>

Workplace Temperature includes cold, heat, and humidity. Human bodies maintain a core temperature in the vicinity of 98.6°F. Beyond a range of environmental temperature and humidity, the human's ability to regulate body temperature decreases. Cold, heat, and humidity refer to the environmental conditions that temperature or humidity have negative effects on personnel behavior and task performance.

Cold, heat, and humidity produce thermal stresses on humans. While physiological limits of endurance to temperature and humidity may be seldomly reached, personnel are subjected to thermal stresses in many work environments, such as in outdoor work under intemperate climatic conditions or loss of ventilation in control rooms. Studies on the relationship between thermal stress and accident occurrence as well as unsafe work behavior have revealed negative effects of thermal stress on task performance.

Wearing protective clothing can impose thermal stress. The effect of heat on physical work and perceptual/motor task performance may become severe in situations where personnel are required to wear heavy protective clothing in restricted or confined areas. Protective clothing worn in radiation zones may not allow adequate ventilation, which leads to heat and humidity.

Workplace Temperature has three attributes as follows:

- TMP1 Cold in workplace
- TMP2 Heat in workplace
- TMP3 High humidity in workplace

<u>Summary of the Data Sources</u>

The data generalized for this PIF are presented in Appendix A7 IDHEAS-DATA IDTABLE-7. The data sources for Workplace Temperature are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – No data source was identified in NPP operation for this PIF.

Category B – No datapoint was generalized from this category. Abundant studies on the effects of environmental temperature were performed in aerospace, aviation, and military workplaces in

the 1950’s to 1970’s. The field studies mostly focused on perceived heat or cold, the effect on body temperature, and task performance measures other than human error rates.

Category C – Numerous experimental studies have investigated the effects of heat and cold on human task performance. Many studies used operational personnel such as military solders or ship operators as the subjects of the study and/or had the subjects performed simulator tasks such as driving simulation. The studies elucidated the effects of head and cold on task performance by varying with task types, levels of heat or cold, task duration, exposure time, etc. For example, Chase et al. [51] studied the effect of heat on dual-task performance and attention allocation. The subjects performed two concurrent visual pattern match tasks for about an hour at different temperatures. Mild detrimental performance was onset at 30^{o}C while significant detrimental performance was at 35^{o}C. Moreover, the heat in the workplace narrowed the subjects’ attention allocation. While the subjects were instructed to split attention equally at the two concurrent tasks, the performance on the task at more peripheral visual fields was significantly worse than that of the task closer to the central visual field. The following is the datapoint generalized from this study.

PIF	CFM	Error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
TMP2	D / E		T1	T2	Split attention equally between two concurrent visual tasks T1 and T2	Varying temperature and splitting attention		[51]
		25^{o}C	0.3	0.23				
		30^{o}C	0.35	0.3				
		35^{o}C	0.65	0.4				

Category D – One expert judgment data source is that Basra and Kirwan [52] estimated the HEPs of offshore oil operation under extreme weather conditions. The estimated effects of extreme cold weather were about an order of magnitude higher than those measured in experimental conditions with mild cold temperatures. The following is the generalized datapoint.

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
TMP1	D, E, U, DM, T	Center and range of error factor: D (instrumentation):[1.8, 2.1, 2.7] U (cognition): [3.8, 10, 18] DM and T (management): [3., 8, 18] E (physical): [1.6, 5, 8] E (precise motor actions (connect lines to pump, remove air from lines and pumps): [13, 20, 30]	maintenance task of offshore oil and gas facility pumps (develop work orders, reconnect pump, open valve and reinstate pump)	Extremely cold	(estimation of error factors based on operational data)	[52]

Category E - Several studies reviewed and synthesized the large volume of literature on the effects of cold and heat. For example, Pilcher et al. [53] performed a comprehensive meta-analysis of 22 studies about the effects of temperature exposure on performance. The factors analyzed include the severity of temperature exposure, duration of the experimental session, duration of temperature exposure prior to task onset, type of task, and task duration. The results indicate that heat and cold in workplace negatively impact performance on a wide range of cognitive-related tasks. Statistically, Hot temperatures of 90^{o}F (32.22^{o}C) or above resulted in 14.88% decrement in performance in comparison to neutral temperature conditions and cold temperatures of 50^{o}F (10^{o}C) or less resulted in 13.91% decrement in comparison to neutral temperature conditions. Furthermore, the duration of exposure to the experimental

temperature, the duration of exposure to the experimental temperature prior to the task onset, the type of task and the duration of the task had different effects on performance. The following are some datapoints generalized from this study.

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
TEP1	D / E	%diff -7.8%	Attention/Perceptual tasks (percentage difference between neutral and experimental temperature conditions)	<65ºF	(Meta-analysis)	[53]
TEP1	D / E	%diff 1.75%	Visual tasks and control tasks requiring mathematical processing	<65ºF	(Meta-analysis)	[53]
TEP1	U	%diff -28%	Reasoning/Learning/Memory tasks	<65ºF	(Meta-analysis)	[53]
TEP1	Unsp	%diff -25%	Unspecified	<65ºF, Short task duration (<60min)	(Meta-analysis)	[53]
TEP1	Unsp	%diff -3%	Unspecified	<65ºF, long task duration (>60min)	(Meta-analysis)	[53]
TEP2	D	%diff - -14%	Attention/perceptual tasks	>80ºF	(Meta-analysis)	[53]
TEP2	U	%diff 1.75%	Reasoning/Learning/Memory tasks	>80ºF	(Meta-analysis)	[53]
TEP2	D / E	%diff -14%	Visual tasks and control tasks requiring mathematical processing	>80ºF	(Meta-analysis)	[53]

<u>Summary of Human Error Data for Workplace Temperature</u>

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-5 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – The datapoints for this CFM have the error rates that vary between 1.05 to 2.5 times from neutral to hot workplace temperatures, with the median around 1.4 times. The error rates vary between 1.01 and 1.1 times from neutral to cold workplace temperatures. The error rates can increase between 1.8 and 2.7 times in extremely cold weather for instrument reading.
- Failure of Understanding (U) – A warm to moderately hot workplace temperatures have little effect on Understanding. Meta-analysis shows that performance for reasoning and memory increases slightly as the temperature is greater than 80°F. Mildly cold temperatures decrease the performance 1.28 times. Extremely cold temperatures may increase HEPs 3 to 18 times, with the mean value of 10 times.
- Failure of Decisionmaking (DM) – No datapoint on error rates was generalized for this CFM alone. Hancock et al. [54] meta-analysis shows that the effect size for Understanding/Decisionmaking with heat is -0.27, compared to the -0.43 effect size with heat for Detection. Several studies found that completing risky tasks under elevated ambient temperatures (> 30°C) leads to a higher risk proclivity than in comfortable temperature conditions (<25°C). On the other hand, mildly cold temperatures have little effect on Decisionmaking. However, extremely cold weather may significantly increase the HEPs of task management, which involves decisionmaking and team coordination.

- Failure of Execution (E) – Compared to neutral temperatures, the datapoints for this CFM have error rates that vary between 1.05 and 2 times for hot workplace temperatures, with the median around 1.4 times. The error rates increase above those for neutral temperatures about 1.1 times for mildly cold workplace temperatures. Yet, the expert estimated HEPs for extremely cold weather increase 1.6 to 8 times for physically demanding tasks. Moreover, the estimated HEPs increase 10 to 30 times for precise motor actions that require fine finger movements, such as connecting lines to pump and removing air from lines and pumps.
- Failure of Interteam Coordination (T) – The estimated HEPs for managing tasks increases 1.6 to 8 times under extremely cold weather.

<u>Observations from the Generalized Data</u>

- Heat begins to impair performance when it exceeds 86°F, vigilance and performance of complex tasks are affected by heat.
- Performance on tasks requiring manual dexterity declines when temperature falls below 60°F. Cold exposure of the hands which is critical for manual performance affects the speed and precision of task performance.
- The range of temperatures beyond which performance is impaired depends on the kinds of tasks and exposure time. Tasks involving fine movements of the fingers and hands or manipulation of small objects are particularly sensitive to cold effects. Slow cooling is more detrimental to manual performance than rapid cooling to equivalent skin temperatures of the hands.
- Comparatively mild levels of cold, heat, and humidity exposure can increase the number of errors, speed of incorrect response, and number of false alarms. Complex reaction time slows down in heat, and more errors are made in cold.
- No datapoint is generalized for Attribute TMP3 “High humidity at workplace.” All the studies identified for this attribute used physiological measures or performance measures of tasks that cannot be related to error rates.

Extensive data sources are available for the PIF Workplace Temperature. The generalized data show that mild cold or heat only moderately increases human error rates. Overall, the effect of this PIF on human error rates is weak within the range of normal room temperature. Yet, notice that extreme cold and heat can have very strong impacts on error rates. Only one datapoint is generalized from expert judgment of HEPs for extremely cold weather. Also, cold temperature has a much stronger impact on task performance time, which would increase the time needed for completing the task and may result in higher error rates for time-critical actions. Moreover, cold and heat restrict personnel’s workplace habitability time, which can reduce the time available for personnel to complete actions and, thus, increase human errors.

3.1.8. IDHEAS-DATA IDTABLE-8 for Resistance to Personnel Movement

<u>Introduction to the PIF Resistance to Personnel Movement</u>

Resistance to Personnel Movement refers to the difficulty in making physical movement due to resisting, opposing, or withstanding of external forces such as those imposed by wind, rain, flooding, etc. Resistance to movement causes physical stress (also referred to as physical fatigue) and imposes additional physical and mental demands to complete a task. Physical stress does not lower personnel knowledge of how to get the task done, but it causes lowered physical efficiency, reduced attention, and increased susceptibility to loss of balance. Moreover, physical stress can result in unconscious lowering of performance standards. These effects can

impact task performance in ways such as making errors in timing of movement, overlooking of some important elements in the task sequence, losing accuracy and smoothness of control movement, under-controlling or over-controlling of movement, or forgetting of side tasks.

The following are example situations that could induce resistance to physical movement:

- External forces such as wind, rain, and floods.
- Postural instability may be induced by carrying heavy materials on a slippery or unstable surface while not using fall protection; or it can be induced by experiencing unexpected perturbations that cause body acceleration or deceleration. Tasks affected involve standing upright, rapid body movement, or lateral reach during lifting.
- Exposure to whole-body vibration interferes with manual tracking and visual acuity. Whole-body vibration may come from operating vehicles, walking or lying on oscillating overhead catwalks, climbing up ladders located on or over machinery, working in ventilation ducts, tending conveyors, and fixing generators, diesels, and turbines.
- Protective clothes impose a mechanical burden because body movement is limited by the clothing. That can impact manual dexterity capabilities and psychomotor performance. Wearing heavy gloves hampers performance of delicate manual tasks.

Resistance to Personnel Movement has four attributes as follows:

- PR1 Resistance to personnel movement, limited available space, postural instability
- PR2 Whole-body vibration
- PR3 Wearing heavy protective clothes, gloves, or both

<u>Summary of the Data Sources</u>

The data generalized for this PIF are presented in Appendix A8 IDHEAS-DATA IDTABLE-8. The data sources for Resistance to Personnel Movement are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – No data source was identified in NPP operation for this PIF.

Category B –Abundant studies on the effects of resistance to movement have been performed in aerospace, aviation, ground transportation, off-shore oil operation, chemical, underwater operation, and military workplaces since the 1950’s. Early field studies relevant to the PIF attributes mostly focused on physical characteristics and their impacts on human physiological reactions. Later studies have explored the effects on behavioral performance.

Category C – Numerous experimental studies have investigated the effects of various factors related to the PIF attributes on human task performance. Many studies used operational personnel such as military solders or ship operators as the subjects of the study and performed the studies in operational environments or simulation settings. The studies elucidated the effects of the PIF attributes on task performance by varying with task types, levels of intensity,

durations of tasks, and other physical characteristics such as the weights of physical loads, the depths of floodwater, the frequency of vibration, etc. Comprehensive review of the studies and synthetizations of the findings were also performed by many researchers. The accumulated research has provided a solid foundation for many engineering design standards and criteria. While most studies used physiological and task performance measures, substantial amounts of studies reported task performance accuracy or number of errors. For example, Hancock and Milner [55] examined the performance of experienced professional divers on simple mental and psychomotor tests over a range of depths in the ocean. The selected depths represented the range at which the professional diver might operate for extended periods without the associated complications of prolonged decompression. The subjects performed two tasks: the number addition task mimicking dive time calculations for safe dive profiles, and the reciprocal tapping task representing the basis of simple reaching and aiming movements while submerged. Task completion time and error rates were measured at the dryland control condition and 4.6m and 15.2m underwater diving depths. The results showed that the error rate for mental addition increased about twice while that for reciprocal tapping remained about the same at 15.2m ocean depth in comparison to those on the dryland and 4.6m ocean depth. The datapoint generalized from this study is in the following:

PIF	**CFM**	**Error rates**			**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
PR1	E		Mental addition	Tapping	Professional divers mentally added numbers or performed reciprocally tapping.	a dryland control test followed by manipulation at 4.6m and 15.2m depths in the open ocean.		[55]
		Land	0.08	0.053				
		4.6m	0.07	0.057				
		15.2m	0.15	0.056				

Category D – Two sources of expert judgment data were generalized for this PIF. The expert judgment of HRA for nuclear facility operation [37] estimated the HEPs for vehicle accidents under different weather and traffic conditions. Basra and Kirwan [52] estimated the HEPs of offshore oil operation under extreme weather conditions based on operational data. The estimated effect of strong, cold winds on operation is about an order of magnitude higher than those measured in experimental conditions. The following is the generalized datapoint:

PIF	**CFM**	**Error rates**				**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
PR1	E		T1	T2	T3	Offshore lifeboat operation T1- Incorrectly operate brake cable T2- Fail to disengage boat T3- Fail to check air support system	Controlled (C): Force 4 wind, daylight, unignited gas leak Severe (S): Force 6 wind, night, explosions/fire on platform	(Several other PIFs combined)	[52]
		C	0.02	0.02	0.028				
		S	0.04	0.07	0.158				

Category E – Many review studies and meta-analysis have well summarized the large volume of literature relevant to this PIF. For example, Conway et al. [56] performed quantitative meta-analytic examination of whole-body vibration effects on human performance. They synthesized the existing research evidence from 224 papers. Results indicate that vibration acts to degrade

goal-related activities, especially those with high demands on visual perception and fine motor control. Some studies based on statistic data also provide task performance measures related to the PIF attributes. For example, Pregnolato et al. [57] developed a depth-disruption function to emulate the impact of flooding on road transport. The function describes the relationship between depth of standing water and achievable vehicle speed. The function was constructed by fitting a curve to video analysis supplemented by a range of quantitative data that has be extracted from existing studies and road transport databases. The following is the generalized datapoint that was sampled from the continuous function.

PIF	CFM	Error rates or Task Performance indicators			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
PR4	E	Depth (W)	Small	4WD	Driving – small cars and 4WD cars (speed m/h)	Car speed with varying depth (W) of floodwater compared to 85m/h without flood	(from multiple studies and databases so other PIFs may be involved)	[57]
		100mm	10m/h	50m/h				
		150mm	0	40m/h				
		300mm	0	10m/h				

Summary of Human Error Data for Resistance to Personnel Movement

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-8 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – None of the generalized datapoints is exclusively for Failure of Detection. Although many studies reported reduced visual perception under the PIF attributes, the reduced visual perception seemed to primarily impact visuomotor tasks. Conway et al. [56] meta-analysis reported the effect size for visual perception with whole-body vibration is -1.79, as compared to the effective size of -0.89 for fine motor execution. However, the impaired visual perception reported in the meta-analysis was primary for visuomotor tasks.
- None of the generalized datapoints are exclusively for Failure of Understanding. Sherwood and Griffin [58] reported a 10% to 15% reduction in learning/memory with whole-body vibration. Yet, the study also suggested that the impairment was due to a disruption in the information input processes that are related to Detection rather than the recall process that is more related to Understanding.
- No data source was identified about the impact of this PIF on Decisionmaking.
- Failure of Execution (E) – The generalized datapoints are mostly for this CFM. Most studies relevant to the PIF attributes measured human performance of motor tasks. The datapoints from experimental studies have the error rates that vary between 1.05 to 2 times from neutral to poor attribute status. The estimated HEPs from off-shore oil shop operation vary 2 to 5 times between the controlled condition and severe weather condition.
- Failure of Interteam Coordination (T) – No data source was identified about the impact of this PIF on Failure of Interteam Coordination.

In summary, extensive data sources are available for the PIF Resistance to Personnel Movement. Overall, the effect of this PIF on human error rates is relatively weak. However, notice that most generalized datapoints are from the studies conducted in relatively mild conditions where human subjects were allowed for experimentation. Extreme PIF attribute status such as strong winds or deep floodwater can lead to devastating impacts on task performance.

3.1.9. IDHEAS-DATA IDTABLE-9 for System and Instrument & Control Transparency to Personnel

Introduction to the PIF System and Instrument & Control Transparency to Personnel

Systems and Instrument & Control (I&C) should be designed for personnel to understand their behaviors and responses in various operating conditions. This PIF models the impact of design logic and personnel's use of systems and I&C deviating from the design. When the operation of systems or I&C is not transparent to personnel, or personnel are unclear about system interdependency, they can make errors because of not understanding the systems in unusual scenarios. Also, some instrumentation, control, electrical, and fluid (water, compressed air, ventilation) systems may be aligned in alternative or unusual configurations when the initiating event occurs. For example, these configurations may apply during testing, maintenance, specific shutdown plant operating states, etc. If a system is not aligned in its normal configuration or the unusual alignment is not apparent, personnel may not correctly confirm if the system is operating properly, easily recognize the effects from equipment damage, or quickly determine how the system should be realigned to cope with the evolving scenario.

The PIF System and Instrument & Control Transparency has five attributes as follows:

- SIC1 System behaviors is complex to understand or not transparent to personnel
- SIC2 Inappropriate system functional allocation between human and automation
- SIC3 System failure modes are not transparent to personnel
- SIC4 I&C logic is not transparent
- SIC5 I&C failure modes are not transparent to personnel

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A9 IDHEAS-DATA IDTABLE-9. The data sources for System and Instrument & Control Transparency are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – No data source was identified in NPP operation for this PIF. Many reports document the cases where system or I&C transparency contributed to human failures. No quantitative operational data about the effects of the PIF attributes on human performance were identified. Some operational databases or studies reported human failures with respect to digital I&C. Yet, those studies mainly focus on design aspects of human-system interfaces, not the transparency of system or I&C design logic.

Category B – No operational data about the effects of the PIF attributes on human errors were identified.

Category C – Numerous experimental studies have investigated the effects of automation

systems on human task performance. Many studies used operational personnel such as NPP operators, pilots, or air traffic controllers as the subjects of the study in operational environments or high-fidelity simulation settings. Yet, most of those studies measured task-specific performance indicators or subjective ratings such as workload or trust to automation. There are limited studies measuring human error rates relevant to the PIF attributes. There are barely any studies quantifying human error rates that vary with I&C transparency. Thus, IDHEAS-DATA IDTABLE-9 documents some data sources that do not have human error rates. One example is a series of studies performed by the Organisation for Economic Cooperation and Development (OECD) Halden Reactor Project (HRP) on automation transparency. The report “Twenty Years of HRP Research on Human- Automation Interaction: Insights on Automation Transparency and Levels of Automation” [59] summarizes Halden’s automation studies in two decades. The studies used NPP crews performing operating procedures on high-fidelity simulators. The results showed controversial effects of automation transparency on operation performance assessment scores, i.e., automation transparency aided or hindered operator performance in different scenarios. However, those studies typically varied multiple experimental factors together, so it was difficult to elucidate the effects of transparency. On the other hand, many simulation studies with airplane pilots or air traffic controllers clearly demonstrated that lack of transparency with automation systems had detrimental impacts on task performance and increased human errors. For example, in the Trapsilawati et al. [60] study “Transparency and Conflict resolution Automation Reliability in air traffic control,” the tested air traffic controllers resolved airplane conflicts with the automation aid of the Conflict Resolution Advisor (CRA). A Vertical Situation Display (VSD) was to provide transparency of CRA to air traffic controllers. The measured error rate in resolving airplane conflicts was about double without having the VSD compared to having the VSD for transparency. Also, the air traffic controllers had higher situation awareness and spent less time resolving conflicts with the VSD. The datapoint generalized from this study is shown in the following:

PIF	CFM	Error rates				Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
SIC1	U/DM		% error	%SA	Time	Air traffic controller resolves conflicts with CRA (%incorrect)	Automation is 80% reliable VSD – Visual display providing transparency		[60]
		No VSD	0.11	59%	7.78s				
		VSD	0.06	73%	5.38s				

Category D – No data source was identified in this category.

Category E – Many studies reported the frequencies of types or causes of human errors associated with automation and digital I&C. Several datapoints from such data sources were generalized in IDHEAS-DATA IDTABLE-9 to inform the relative likelihood of CFMs and effects of PIF attributes. For example, in the report “Analysis between Aircraft Cockpit Automation and Human Error Related Accident Cases,” Kwak et al. [61] analyzed 94 cockpit automation accident cases from Flight Deck Automation Issues (FDAI). The study used a human error classification scheme to analyze and count the frequencies of error causal factors in the accidents. The study found that rule-based errors caused automation accidents most frequently. The top two causal factors to the errors are excessive automation dependency and inadequate understanding of the automation technology. The datapoint generalized from this study is shown in the following:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF

<table>
<tr><td rowspan="3">SIC1
&
SIC3</td><td rowspan="3">Unsp</td><td colspan="2">Top freq. causes in 34 accidents</td><td rowspan="3">FDAI Automation Human Error Types (frequencies of error types)</td><td rowspan="3">Accident caused by automation failure</td><td rowspan="3">(analysis did not separate system vs failure mode)</td><td rowspan="3">[61]</td></tr>
<tr><td>Lack of understanding of the system</td><td>5</td></tr>
<tr><td>Improper performance of an automation device in an abnormal situation</td><td>4</td></tr>
</table>

Summary of Human Error Data for System and Instrument & Control Transparency

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-9 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – The generalized datapoints for this CFM show that the error rates varied 1.25 to 3 times nominal due to lack of transparency.
- Failure of Understanding (U) - The limited generalized datapoints for this CFM show that the error rates varied 2 to 3 times nominal due to lack of transparency. The datapoints were from the studies using automation as a job aid. In the studies that automation is the primary system for personnel to work with, lack of transparency resulted in information unreliable or misleading, which lead to high error rates.
- Failure of Decisionmaking (DM) – Most of the datapoints for this CFM are in combination with Failure of Understanding, so it is difficult to examine the effect of the attributes on Decisionmaking alone without properly detaching the CFMs in the tasks. The only datapoint exclusively for Decisionmaking is that the pilots participating in the experiment all made the wrong decision when the decision aid gave them wrong decision advice. The failure represents a combination of several PIFs: Scenario Familiarity, Information Completeness and Reliability, and SIC2 Improper functional allocation. Overall, the generalized data are not enough to derive the range and central tendency of the effects of the PIF attributes on Failure of Decisionmaking.
- Failure of Execution (E) – No data source was identified about the impact of this PIF on Failure of Action Execution.
- Failure of Interteam Coordination (T) – No data source was identified about the impact of this PIF on Failure of Interteam Coordination.

Observations from the Data Sources Reviewed

- Studies about the effects of system transparency on human performance almost exclusively focus on automation. Most studies investigated human performance regarding trust, engagement, cooperation, and subjective opinions on automation.
- System transparency is not consistently defined in the studies. Many studies assume that transparency is presenting system information to personnel. However, personnel may not use the presented information either because of the ways the information is presented or the personnel are not available to use the information.
- Many studies on automation systems did not make distinction between job aids versus the primary system with which personnel perform their tasks. Some studies leave the system or personnel to decide when and how to use the automation. Thus, the measured results could be due to transparency, functional allocation, or both.

Overall, a limited sample of data sources were generalized for this PIF because most data sources reviewed did not have human error data. On the other hand, there are many case studies and event reports relevant to this PIF. The NRC staff at present has not analyzed those

data sources in depth to gain insights on human errors due to lack of system or I&C transparency. Compared to systems and automation, very limited studies have been done about the effects of DI&C transparency on human errors. Since many NPPs are upgrading to digital I&C control systems, operator performance data with digital I&C should be systematically collected.

3.1.10. IDHEAS-DATA IDTABLE-10 for Human-System Interface

Introduction to the PIF Human-System Interface

Human System Interface (HSI) refers to indications (e.g., displays, indicators, labels) for personnel to acquire information and controls used by personnel to execute actions on systems. HSIs are expected to support human performance. For example, advanced alarm displays in NPP control rooms organize alarms according to their urgency to help operators focus on what is most important. HSI designs of NPP control rooms generally undergo a rigorous human factors engineering design and review process; thus, HSIs should comply with human factors engineering requirements and do not impede human performance in normal and emergency operation. However, poorly designed HSIs can impede task performance in unusual event scenarios. Even a well-designed HSI may not support human performance in specific scenarios that designers or operational personnel do not anticipate. HSIs may also become unavailable or unreliable in hazardous scenarios.

The PIF Human-System Interface has 14 attributes in the following categories:

- HSI1 – HSI4: Ambiguity in sources of indications
- HSI5 – HSI7: Ambiguity in the information presentation of indications
- HSI8 – HSI9: Ambiguity in control elements
- HSI10 – HSI14: Ambiguity in the maneuvers of control elements and interaction with personnel

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A10 IDHEAS-DATA IDTABLE-10. The data sources for HSI are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – Several nuclear human performance databases have human error information related to HSI attributes. The analysis of the German NPP maintenance human event database [4, 5] shows the effects of several HSI attributes on human error rates. The analyzed error rates were reported for different types of maintenance tasks under specific PIFs. For example, the task “Operating a control element on a panel” had the error rate 1.6E-3 (7/3588) for selecting the incorrect control elements, and the contributing PIF was “Wrong control element within reach and similar in design.” This is generalized to IDHEAS-DATA IDTABLE-10 for the CFM Failure of Action Execution with PIF attribute HSI8 “Similarity in control elements”. The uncertainty in this data source is that the errors were counted for a single step “operating a control element,” while it is uncertain how many times “operating a control element” occurred in a task. Thus, when using the generalized data like this, the error rates should be calibrated with data from other sources. The following shows the datapoint generalized from this example:

PIF	CFM	Error rates	Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
HSI8	D	8.9E-4 (7/8058)	Operating a control element on a panel (Wrong element selected)	Wrong control element within reach and similar in design.	(Errors could be for a step or a task)	[4]

Category B – No operational data from other domains were generalized given that there were already many datapoints from NPP human performance databases.

Category C – Thousands of experimental studies have investigated the effects of HSIs on human task performance. Many studies used operational personnel such as pilots, ship operators, and military personnel as the subjects of the study in their operational environment or in high-fidelity simulation settings. The studies elucidated the quantitative effects of the PIF attributes on task performance. Only a limited number of data sources were selected for generalization from a large amount of available data sources in this category. For example, In the report by Eitrheim et al. [62] "Evaluation of design features in the HAMBO operator displays," NPP operators' error rates were measured with microtasks of detecting information in conventional versus innovated displays of NPP simulators. The innovate displays included features that graphically showed parameter trends and ranges. The average error rates for "check the values of multiple parameters" were 0.2 for conventional displays and 0.11 for innovate displays. The datapoint generalized from this study is shown in the following:

PIF	CFM	Error rates		Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
HSI4	D	Innovate displays	0.11	NPP operators check the values of multiple parameters (accuracy)	Innovate display – graphic features of parameters. Conventional display - numeric parameter values.		[62]
		Conventional displays	0.2				

Category D – There are several data sources in this category. Given that there are already lots of data sources in Categories A and C, only one data source of expert judgment, "An Evaluation of the Effects of Local Control Station Design Configurations on Human Performance and Nuclear Power Plant Risk" [7], was generalized, because it had estimated HEPs for Attribute HSI9 "Poor functional centralization – multiple displays/panels needed together to execute a task". In the study, an expert panel estimated the HEPs of nine NPP ex-control room actions in local control stations for low, medium, and high functional centralization and low, medium, and high quality of panel design. The datapoint generalized from this study is shown in the following:

PIF	CFM	Error rates				Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
HSI9	E		PD* low	PD Medium	PD High	Execute procedures in NPP local stations	PD – Panel ergonomic design FC – Functional centralization,	(expert judgment)	[7]
		FC* Low	8.62E-1	4.84E-1	2.64E-1				
		FC-medium	2.84E-1	1.29E-1	8.41E-2				

		FC-high	1.15E-1	6.24E-2	4.04E-2		FC Low - too many panels FC High – 1-2 panels		

* FC: functional centralization; PD: Panel ergonomic design

Category E – No datapoint was generalized from this category of data sources given that there are many data sources in other categories.

<u>Summary of Generalized Human Error Data for HSI</u>

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-10 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – The datapoints have the error rates for Failure of Detection ranging from 1.2 ~ 6 times nominal with the presence of the HSI attributes. Notice that the attributes, such as HSI5 "Poor indication salience", were examined in the normal range of human perception, i.e., the information displayed is above the perceptual thresholds such as the minimum font size or luminance contrast of text.
- Failure of Action Execution - The datapoints have the error rates for Failure of Execution ranging from 1.1 ~ 15 times nominal with the presence of the HSI attributes. The high values of error rates due to the HSI attributes are often from the data of which the number of times the task was performed was relatively low. Thus, the reported error rates may also be associated with PIF Scenario Familiarity even if this was not annotated in the data sources.
- Failure of Understanding (U) and Failure of Decisionmaking (DM) – No datapoints were generalized for these two CFMs. In fact, there are many data sources studying the effect of HSI on tasks involving Understanding and Decisionmaking. However, the factors investigated in those studies were best represented by other PIFs such as Information Availability and Reliability or Task Complexity.
- Failure of Interteam Coordination - No datapoints were generalized for this CFM. There are many studies on how HSIs enhance human performance in teamwork and coordination. No data source was identified having error rates of team coordination due to HSI attributes.

<u>Observations from the data sources reviewed</u>

- HSI is perhaps the most well studied PIF. Countless studies have investigated the effects of various HSI features on human performance. Moreover, the results in many cognitive and neuroscience research of information processing are applicable to the effects of HSI features. The research has established a solid technical basis for human factors design of HSIs. Many human factors design standards or requirements have been in place to ensure that HSIs are designed within the normal range of human perception and ergonomics. The later research on HSIs has been shifted to investigating the functional aspects of HSIs, described in many attributes of IDHEAS-G's PSFs.
- The impacts of HSI features on human error rates are generally consistent across different studies in different fields. This is because most HSI features impact human performance through challenging the capacities of human information perception and information processing commonly for personnel with normal perception and cognition abilities.

In summary, there are abundant data sources for the effects of HSI attributes on Failure of Detection and Failure of Execution. Moreover, the human error data from data sources are generally consistent with each other in the quantitative effects of the HSI attributes. On the other hand, no data sources were identified for Failure of Understanding and Failure of Decisionmaking. This should be inherited from IDHEAS-G definitions of the CFMs. The

definition of Failure of Understanding in IDHEAS-G is under the assumption that personnel correctly detected the given information, and the definition of Failure of Decisionmaking is under the assumption that personnel have a correct understanding of the situation. The HSI attributes are pertinent to detecting information and executing actions. The aspects of HSI affecting Understanding and Decisionmaking are mostly represented by task specific PIFs such as Information Availability and Reliability or Task Complexity. Lastly, there are qualitative data showing the effects of HSI on teamwork and coordination, yet no data source for human error rates has been identified to quantify the effect.

3.1.11. IDHEAS-DATA IDTABLE-11 for Portable Equipment, Tools, and Parts

Introduction to the PIF Portable Equipment, Tools, and Parts

Portable Equipment, Tools, and Parts (ETPs) assessed in an event include all those needed to support critical human actions. For example, use of a portable diesel pump would include the vehicle to tow the pump to its staging location, the water source, pipes, hoses, junctions and fittings (e.g., to connect to fire hydrants), and other things; ladders or scaffolding may be needed to access equipment that must be operated or local instrumentation that must be checked. ETPs should be available and readily usable. In event scenarios, portable equipment or special tools may be needed. Examples are portable radios, portable generators, torque devices to turn wheels or open flanges, flashlights, ladders to reach high places, and electrical breaker rack-out tools. Although ETPs should be designed for easy use, personnel may have difficulties using them. For example, personnel may not know how to calibrate a measurement tool, or the instructions for using the equipment do not indicate what to do if the equipment is operating outside of the specified range.

The ETPs in this PIF refer to portable ones that are, unlike HSIs, usually not designed with rigorous human factors engineering review and not maintained under mandatory administrative rules. Personnel may not be not trained to use them following nuclear power plants' Systematic Approach to Training (SAT). An exception may be FLEX equipment. Following the accident at Fukushima Daiichi, implementation of the Diverse and Flexible Coping Strategies (FLEX) resulted in the purchase of portable equipment (including diesel generators and diesel-driven pumps) specifically intended to support plant shutdown after extreme external events. Much of the equipment can also be used as added defense in depth to mitigate the consequences of non-FLEX-designed accident scenarios (involving anticipated internal initiating events) in which installed plant equipment fails. Many nuclear power plants have considered using FLEX equipment during non-FLEX-designed accident scenarios and are taking credit for the additional equipment and mitigation strategies in their probabilistic risk assessments (PRAs). Consequently, many NPPs may begin to include FLEX in the Maintenance Rule and Systematic Approach to Training (SAT). Thus, HRA analysts may evaluate FLEX equipment in the same way as evaluating HSIs.

This PIF has four attributes as follows:

- ETP1 ETP is complex, difficult to use, or has poor suitability for the work
- ETP2 Rarely used ETP does not work properly or is temporally not available
- ETP3 ETP labels are ambiguous or do not agree with document nomenclature
- ETP4 Personnel are unfamiliar or rarely use the ETP

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A11 IDHEAS-DATA IDTABLE-11. The data sources for the PIF are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – No NPP operational data on human failures with ETPs were identified. There have been operational experience notifications on FLEX equipment that did not work properly due to human errors. Yet, statistical data are not available for this report.

Category B – Several sources of statistical operational data from other domains were generalized. The data sources have frequencies of ETPs as the causes of the analyzed operational events or accidents. The data sources did not provide quantitative information about the impact of the PIF attributes on human error rates. They provide the likelihood of ETPs contributing to human failures.

Category C – Only one data source was identified in this category. Jacob et. al. [63] studied the effects of work-related variables on human errors in observing and noting measurements. The study isolated and quantified the effects of the variables separately. The study was designed to quantify the effects of selected work-related variables of two sets of human subjects (experienced and inexperienced technicians). Analysis of the results revealed that the variables identified and studied significantly affected measurement errors. One of the work variables tested was analog versus digital multimeters for measuring voltage and resistance. Digital tools were less complex and more intuitive to use. The result showed that technicians made 2-3 times more errors with analog tools than with digital tools. The datapoint generalized from this study is shown in the following:

PIF	CFM	Error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
ETP1	D /E		FN	AN	Experienced technicians used analog and digital multimeters to measure voltage and resistance (%measurement errors)	Tools - Digital vs analog, Time of work – Before noon (FN) and afternoon(AN)	(The errors are applicable to Detection and Execution)	[63]
		Digital	4.45	5.74				
		Analog	11.07	13.7				

Category D – One relevant data source was identified. In 2018, the NRC conducted a formal expert elicitation on FLEX HRA [3]. An expert panel estimated HEPs of a set of human actions in using portable FLEX equipment. The HEPs were estimated for a FLEX-designed scenario (seismic caused) and non-FLEX designed scenario. Even in the non-FLEX designed scenario, personnel are still challenged with scenario unfamiliarity and rare use of the equipment. The estimated HEPs for the tasks of transporting, connecting, and operating the FLEX equipment were much higher than those for operating the routine stationary equipment. Below shows the datapoint generalized from this data source:

PIF	CFM	Error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
ETP4	E		Non-FLEX	FLEX-designed	Use of portable generator or pump on a	Personnel rarely use the equipment and	Scenario unfamiliar, rarely	[3]
		Transport	0.057	0.14				

		Connect	0.088	0.16	sunny day vs. severe accident	training is infrequent,	performed actions, poor training (Expert judgment)	
		Operate	0.052	0.12				

Category E – No data source was identified from this category.

Summary of Generalized Human Error Data for Portable Equipment, Tools, and Parts

Most datapoints generalized in IDHEAS-DATA IDTABLE-11 do not meet the criterion for informing PIF attribute weights because they do not have error rates of two or more PIF attribute states. In fact, most datapoints only have the information about the association of ETPs and human events or accidents. While human error data for the PIF attributes are sparse, many operational experience notifications and accident reports have documented extensive empirical evidence that critical ETPs needed for important human actions can detrimentally impact human performance and increase human errors. Systematic data collection and experimental studies are needed to elucidate the quantitative impacts.

3.1.12. IDHEAS-DATA IDTABLE-12 for Staffing

Introduction to PIF Staffing

Staffing refers to having adequate, qualified personnel to perform the required tasks. Staffing includes the number of personnel, their skill sets, job qualifications, staffing structure (individual and team roles and responsibilities). Adequate and qualified staff is normally expected. In event scenarios, there may be a shortage of staffing, lack of staff with specific skills, or ambiguous staff roles and responsibilities. Some personnel may not be available for a period after an initiating event. For example, in an NPP external event, the offsite personnel may not be available immediately because of site inaccessibility. Staffing consideration should not be limited only to the human action being analyzed, but also it should be considered within the scope of the entire event. Staffing can be inadequate when many human actions are concurrent. Specifically, HRA analysts need to consider other activities that are not modeled explicitly in the PRA but may share the same staff. For example, personnel may be allocated to mitigate failures or damage of non-safety systems that are important for overall plant investment protection or for perceived improvement of overall plant conditions. Even in normal operation scenarios, staffing can become a concern—for example, key personnel may be temporally called away for other duties.

Fitness for duty is a requirement for staff. It refers to whether an individual is fit to perform the required actions of their job. Factors that may affect fitness for duty include fatigue, illness, drug use (legal or illegal), and personal problems. Personnel may become unfit for duty as the result of excessively long working hours or illness caused by the harsh environment.

This PIF has four attributes as follows:

- STA1 Shortage of staffing (e.g., key personnel are missing, unavailable or delayed in arrival, staff pulled away to perform other duties)
- STA2 Ambiguous or incorrect specification of staff roles, responsibilities, and configurations,
- STA3 Lack of certain knowledge, skills, and abilities needed for key personnel in unusual events
- STA4 Lack of administrative control on fitness-for-duty

<u>Summary of the Data Sources</u>

The data generalized for this PIF are presented in Appendix A12 IDHEAS-DATA IDTABLE-12. The data sources for the PIF are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – Several simulation studies examined the effect of staffing on NPP operator performance. Title 10 of the Code of Federal Regulations (10 CFR) Part 55 [64] requires a minimum crew size of three licensed operators in US NPP control rooms. However, NPPs typically have more than 3 operators in the control rom. New technologies challenge traditional staffing levels by using automation to support crew size reductions. Over the last two decades, OECD Halden Reactor Project has conducted a series of high-fidelity simulations to examine various staffing configurations in NPP control rooms. For example, the study, “Staffing Strategies in Highly Automated Future Plants - Results from the 2009 HAMMLAB Experiment,” Eitrheim et. al. [62] examined two control room staffing configurations: the traditional staffing with a crew of three operators responsible for one reactor process versus the untraditional staffing configuration in which a crew of three operators simultaneously controlled two nuclear processes with the aid of control room automation. This untraditional staffing solution was compared with a traditional staffing solution based upon current operational practices. Operator performance data were gathered from nine crews of licensed NPP operators in eight scenarios per crew. The findings from the experiment favored the traditional staffing solution. However, the operators managed to perform a considerable amount of prescribed tasks when they worked untraditionally. The results show that the untraditional solution is feasible in easy scenarios, but the performance score decreased to an unacceptable level in difficult scenarios. The results suggest that reduced staffing levels might be sufficient during normal operation, but specialized support teams and roles may be necessary to handle disturbances and upset situations. The two staffing solutions in this study involve both staff size and configuration. The study used a set of human performance measures including the task-specific operator performance assessment score, situational awareness, and several workload measures. The datapoint generalized from this study documents operator performance assessment scores, which are related to human error rates. The datapoint is shown in the following:

PIF	**CFM**	**Error rates or task performance indicator (**			**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
STA1 /STA 2	D/U/D M/E		Easy scenario	Difficult scenario	Nine 3-person NPP crews performed 8 scenarios (OPAS - Operator Performance Assessment Scale in 0-100, the higher the better)	Staffing configuration T - Traditional staffing – 3 persons for **one** reactor UT - Untraditional staffing – 3 persons for two reactors with automation	(Automation use varied)	[62]
		T	82.5	66.2				
		UT	75.5	45.7				

Category B – Adequate staffing is essential for safety-critical work domains. Proper staffing size,

configuration, and required knowledge, skills, and abilities (KSAs) have been examined in every safety-critical work domain such as health care, aviation, and emergency medical services. Several studies of staffing examination in safety-critical work domains were generalized in IDHEAS-DATA IDTABLE-12. For example, NIST [65-68] conducted a series of field studies to understand the effects of staffing size and configurations of emergency medical service crews. One of the studies showed that it took about 23 minutes for a 2-person crew and about 16 minutes for a 4-person crew to complete all the essential tasks needed on low hazard structure fires. In addition, key crew members' average heart rate was about 90% of the maximum allowed heart rate with a 2-person crew. Both measures were related to human failure events of fire rescue actions. Such data can be used to infer error rates when direct error rate data are sparse.

Category C – A few data sources were selected in this category. Because the isolated variables studied in controlled experiments are often the sub elements that affect more than one Staffing attribute, the generalized data are not specific to individual attributes. For example, many experimental studies examined the boredom effect of long personnel idle time during a long-lasting task. Cummings et al. [69] examined the effect of low personnel utilization (i.e., the percent of time on tasks) in a 4-hour session of multiple unmanned vehicle supervisory control. The study measured where personnel's attention was directed at any given time and when they switched attention. The study used three types of attention: 1) Directed, which is when participants were directing their gaze at the interface or interacting with the interface, (2) Divided, when participants were looking or glancing at the interface but also engaged in other tasks such as talking to other participants, eating while watching the screen, etc., and (3) Distracted, which was coded as a participant not in a physical position to see the interface, such as turned around in a chair while talking to other participants, at the table getting something to eat, working on a personal laptop, etc. The results showed that personnel had 32% directed attention time, 22% divided attention time, and 46% distracted attention time. Less directed attention means a higher chance of making errors. The datapoint generalized from this study is shown as follows:

<table>
<tr><th>PIF</th><th>CFM</th><th colspan="2">Error rates or task performance indicator (% of time in different attention state)</th><th>Task (and error measure)</th><th>PIF measure</th><th>Other PIFs (and Uncertainty)</th><th>REF</th></tr>
<tr><td rowspan="4">STA2</td><td rowspan="4">D /DM</td><td colspan="2">% mean attention state</td><td rowspan="4">Monitor status and replan tasks in a 4-hour UAV supervisory control session with 2-10% utilization of time (% attention state: directed on task, divided between task and other things, distracted away from the task)</td><td rowspan="4">Low task utilization time (2-10%) in long working sessions</td><td rowspan="4">(student subjects may differ from licensed crews)</td><td rowspan="4">[69]</td></tr>
<tr><td>Directed</td><td>32%</td></tr>
<tr><td>Divided</td><td>22%</td></tr>
<tr><td>Distracted</td><td>46%</td></tr>
</table>

Category D – No datapoint was generalized from this category.

Category E – No datapoint was generalized from this category.

<u>Summary of Generalized Human Error Data for Staffing</u>

Most datapoints generalized in IDHEAS-DATA IDTABLE-12 document task performance measures other than human error rates. Adequate staffing is typically based on workload measures, task completion time or other task-specific measures, so most studies do not report data about the effect of this PIF on human error rates. The human performance measures are

related to human error rates; thus, they can be used to infer PIF attribute weights. Another issue with the generalized data is that most of the data sources are from studies of whole events or full scenarios, therefore the data are not specific to the CFMs or PIF attributes. Integrating these data to develop PIF attribute weights for individual CFMs will be largely based on engineering judgment.

3.1.13. IDHEAS-DATA IDTABLE-13 for Procedure, Guidance, and Instruction

Introduction to the PIF Procedure, Guidance, and Instruction

Procedures, guidance, and instructions (PGIs) refer to availability and usefulness of operating procedures, guidance, instructions (including protocols). PGIs in safety-critical domains such as emergency operating procedures (EOPs) in NPPs are developed through a rigorous process and validated. Personnel are well trained on PGIs in various operating scenarios. Following PGIs should lead to the success of important human actions.

Nuclear power plant operation is procedure-based. PGIs direct operators to perform important human actions; operators are expected to comply with their PGIs. Normally, PGIs are available and facilitate human performance. However, there are human actions in special situations in which no procedure is available or not applicable, then personnel need to perform the actions based on their knowledge and skill-of-craft. There may even be situations in which PGIs may not apply to the scenario, thus PGIs give inadequate or incorrect guidance for important human actions. Other problems with PGIs include ambiguity of steps, lack of adequate detail, or conflict with the situation.

Nuclear power plants have many types of PGIs, including Normal Operating Procedures, Alarm Response Procedures, Abnormal Operating Procedures (AOPs), EOPs, Severe Accident Management Guidelines (SAMGs), and lately the FLEX Support Guidelines (FSGs). Some procedures can have several hundreds of steps. Various operating crews may execute the same procedure differently because a procedure has many branching points that require operators' judgment. Moreover, use of PGIs depends on administrative control and how personnel are trained to use them.

Traditionally, PGIs are written on papers, referred to as paper-based procedures. Over the last two decades, computerized procedures, referred to as computer-based procedures, have been introduced to nuclear power plant control rooms [70]. Some computer-based procedures are simply hard copies of the paper procedures, while other computer-based procedures have various levels of automation interfaces built in and can automatically perform procedure segments. Evaluation of computer-based procedures for HRA not only involves the PGI attribute, but also involves other PIFs such as HSIs and system and I&C transparency.

This PIF has seven attributes as follows:

- PGI1 Procedure design is inadequate and difficult to use
- PGI2 Procedure requires judgment
- PGI3 Procedure lacks details
- PGI4 Procedure is ambiguous, confusing
- PGI5 Mismatch - Procedure is available but does not match the situation
- PGI6 Procedure is not applicable or not available
- PGI7 Procedure is misleading

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A13 IDHEAS-DATA IDTABLE-13. Because of the diversity and complication of NPP operating PGIs, only data sources from nuclear operational or simulator data were generalized for this PIF. Most studies on PGIs were about their effectiveness without measuring human errors. Some data sources identified for generalization to IDHEAS-DATA having task performance measures or cognitive measures such as situational awareness or workload can be used to infer the range of error rates. The data sources identified include studies on normal operating procedures, EOPs, low power shutdown procedures. The studies show that PGI is a prevalent cause of human errors in NPPs. For example, Kim et. al. [71] performed a root cause analysis of 53 low power shutdown events. The result showed that procedures are the second most frequent main drivers (in 24 of the 53 events) while personnel and team are the most frequent (in 29 of the 53 events).

Several identified data sources studied computerized procedures in comparison to paper-based procedures. Overall, operators seemed to make fewer errors with computerized procedures in certain tasks. For example, In Converse's study [72] on evaluating the effectiveness of COPMA-II, a computer-based procedure system, eight teams of two reactor operators operated a PWR simulator under normal and accident conditions, using both computerized and traditional paper-based procedures. The measurement of operator performance includes error rates, times to initiate procedures, times to complete procedures, and subjective estimates of workload. Interestingly, the results showed that the crews on average made three times more errors with paper-based procedures than with computerized procedures in the LOCA scenario; however, they made about the same number of errors in the SGTR scenario. Yet, the report did not provide detailed analysis on what kind of errors the operators made, thus the CFMs applicable to the errors were unknown. It was also unclear why there was a big performance difference with the two types of procedures in the LOCA scenario but not in the SGTR scenario. The datapoint generalized from this reference is shown as follows:

PIF	CFM	Total number of errors made in a scenario			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
PGI1	Unsp.		LOCA	SGTR	Sixteen licensed operators worked in teams of SRO/RO perform LOCA and SGTR scenarios	Computerized (CP) vs paper procedures (BP)	(whole scenarios)	[72]
		Computer procedure	4	12.75				
		Paper procedure	18.75	13				

Summary of Generalized Human Error Data for Procedure, Guidance, and Instruction

About half of the datapoints generalized in IDHEAS-DATA IDTABLE-13 have unspecific CFMs, the other half have the data about PGI attributes on Failure of Execution. Thus, the data generalized so far do not have information specific to Failure of Detection, Understanding, or Decisionmaking. The generalized datapoints have error rates for failure of Execution 2~20 times nominal, varying with PIF attributes. Notice that the error rates for PGI5, "Procedure not matching the situation" and PGI6, "Procedure not available or not applicable" are extremely high from NPP maintenance human performance data. However, those error rates were calculated under the conditions that the tasks were extremely rarely performed, so the high error rates were primarily due to the base PIF, "Scenario Familiarity".

In summary, the human error data identified for this PIF were limited to NPP operational or simulator studies. Such data sources generally studied events or whole scenarios, thus the effects of the PIF attributes on individual CFMs were not isolated. Nevertheless, the studies on PGIs have demonstrated that the PIF is among the most prevalent main drivers to human failure events in NPPs, thus more sophisticated studies on PGIs are highly desirable to provide a solid data basis for HRA.

3.1.14. IDHEAS-DATA IDTABLE-14 for Training and Experience

Introduction to the PIF Training and Experience

The PIF, "Training and Experience" refers to the adequacy of the job training that personnel receive to perform their tasks and personnel's work-related experience. 10 CFR 55 [64] specifies training requirements for U.S. nuclear power plants. To comply with 10 CFR 55, U.S. nuclear power plants have adopted the Systematic Approach to Training (SAT). It is an approach that provides a logical progression from the identification of the competencies required to perform a job to the development and implementation of training to achieve it. With SAT, the competence requirements of jobs in an NPP can be established and met in an objective manner.

Without SAT, there is the risk that important elements of training will be omitted, which would adversely affect the safety and reliability of the plant. Also, training frequencies may not be adequate, which would adversely affect the safety and reliability of the plant. Yet, not all NPP job aspects are under SAT. Some training programs may be too extensive and are only needed for extremely rare events such as beyond design basis events.

Even with SAT, training may not address all possible event scenarios. For example, NPP operator training focuses on use of normal and Emergency Operating Procedures (EOPs); the training may not adequately emphasize how operators need to develop novel strategies to handle unusual accidents or hazard situations. One lesson learned from the Fukushima accident is the need for training on knowledge of system processes.

This PIF has seven attributes as follows:

- TE1 Inadequate training frequency/refreshment - Lack of or poor administrative control on training (e.g., not included in the Systematic Approach of Training)
- TE2 Inadequate amount or quality of training
- TE3 Deficient training practicality
- TE4 Poor or lack of training on procedure adaptation
- TE5 Poor of lack of knowledge-based problem-solving training
- TE6 Inadequate or ineffective training on teamwork
- TE7 Personnel are fully trained but inexperienced (compared to expert-level experienced professional)

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A14 IDHEAS-DATA IDTABLE-14. The data sources for the PIF are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain

E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – The assumption in the data sources of Category A is that NPPs have SAT for operator training. Several studies from nuclear power plant operations and simulations examined the effects of training on operator performance. For example, Preischl and Hellmich [4, 5] reported that even though control actions appeared in the wrong order in the written procedure for testing the emergency feedwater supply system during power operation, operators were able to infer the proper order from professional knowledge, thus only one error was made out of 1200 times the task was performed.

Category B – Only a few studies were selected from this category to cover the breadth of the PIF attributes. One example is the study by Goodstein [73]. In the study, chemical process plant operators received three types of training for fault diagnosis: "Theory" and "Rules" groups were given a simplified account of how the plant worked. In addition, the "Rules" group exercised in applying diagnostic rules. The baseline, "No story", group received no prior instruction of either sort. The results showed that the three groups made about the same number of incorrect diagnoses for "old" system faults that were included in the training. However, for new faults not previously seen by the operators during practice, the "Rules" group made about twice as many correct diagnoses than that of the baseline group. The result reveals the impact of training on rule-based problem solving. The datapoint generalized from this study are shown in the following:

<table>
<tr><th>PIF</th><th>CFM</th><th colspan="3">Number of mean correct diagnoses</th><th>Task (and error measure)</th><th>PIF measure</th><th>Other PIFs</th><th>REF</th></tr>
<tr><td rowspan="5">TE2.2</td><td rowspan="5">U</td><td></td><td colspan="2"># of mean correct diagnoses</td><td rowspan="5">Training for fault diagnosis in the chemical process plant area. (# of correct diagnoses). "
- OLD" for the faults previously seen by the operators during practice.
"NEW" for new faults not previously seen by the operators during practice.</td><td rowspan="5">"Theory" and the "rules" groups were given a simplified account of how the plant worked, in addition the "rules" group exercised in applying diagnostic rules, "no story" group received no prior instruction of either sort.</td><td rowspan="5"></td><td rowspan="5">[73]</td></tr>
<tr><td></td><td>OLD</td><td>NEW</td></tr>
<tr><td>No story (baseline)</td><td>7.7</td><td>2.5</td></tr>
<tr><td>Theory</td><td>7.8</td><td>3.5</td></tr>
<tr><td>Rules</td><td>7.6</td><td>5.5</td></tr>
</table>

<u>Category C –</u> Data sources from several controlled studies were selected for generalization because they explicitly isolated some PIF attributes. For example, in the study by Ha and Seong [74], 15 graduate students with five-year nuclear engineering backgrounds were trained on 14 tasks in three nuclear power plant emergency operation scenarios. They were tested immediately before and after training as well as six months later. The error rates increased twice after six months. The result indicates that the retention of trained skills and knowledge dramatically decrease after six months of not using them. The datapoint generalized from this study is shown in the following:

<table>
<tr><th>PIF</th><th>CFM</th><th colspan="5">Error rates</th><th>Task (and error measure)</th><th>PIF measure</th><th>Other PIFs</th><th>REF</th></tr>
<tr><td rowspan="3">TE1</td><td rowspan="3">D</td><td></td><td>1B</td><td>1A</td><td>2B</td><td>2A</td><td rowspan="3">15 graduate students with nuclear engineering backgrounds of 5.2</td><td rowspan="3">1B-before training
1A-after training</td><td rowspan="3">N/A</td><td rowspan="3">[74]</td></tr>
<tr><td>MMS</td><td>32</td><td>88</td><td>44</td><td>97</td></tr>
<tr><td>LOCA</td><td>0.14</td><td>0</td><td>N/A</td><td>N/A</td></tr>
</table>

		SGTR	0.45	0.14	0.28	0.04	years performed 14 tasks in three scenarios (LOCA, SGTR, SLB) (MMS – mental model score, error rates of failing detection)	2B- 6 months later before training 2A – 6 months later after training		
		SLB	0.44	0.1	0.35	0.16				

Category D – Two relevant expert judgment studies were included. One was the expert judgment of HEPs for IDEHAS-At Power method, in which the HEPs of several crew failure modes ware estimated for "good" and "poor" training along with other PIF attributes. Another is the expert judgment of the HEPs of human actions in implementing FLEX strategies, conducted by the NRC through a formal expert elicitation process in 2018. At the time, training for FLEX strategies was not under SAT. The expert panel assessed that HEPs would decrease by a factor of 10 had the FLEX training been included in the SAT programs.

Category E – No datapoint was generalized for this category.

Summary of Generalized Human Error Data

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-14 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – The error rates for Failure of Detection increased 2 to 10 times over nominal with the presence of the Training and Experience attributes. Interestingly, the data from expert judgment had HEPs about 10 times higher for poor training, while operational data and experimental studies had human error rates 2 to 4 times higher for poor training.
- Failure of Understanding (U) – The error rates increased by a factor of 1.5 to 3 over nominal with the presence of the Training and Experience attributes. Several studies show that the long-term retention of training for knowledge and skills needed in diagnosis tasks seems to be better than those needed for Action Execution.
- Failure of Decisionmaking – The error rates from Category A increased by a factor of 2 to 3 over nominal with the presence of the Training and Experience attributes. However, the expert judgment HEP for misinterpreting procedures in response planning in internal at-power events has a factor of 20 between "good" and "poor" training.
- Failure of Action Execution – The error rates increased 2 to10 times over nominal with the presence of the Training and Experience attributes. Again, the data from expert judgment had a factor of about 10 for poor training, while operational data and experimental studies had the factor around 2 to 5.
- Failure of Interteam Coordination (T) – No data source was identified for this CFM.

In summary, the data sources identified for this PIF were limited, and a big portion of the data sources had inseparable PIF attributes or the data were collected for full scenarios; therefore, the CFMs were unspecific in the study. Overall, there are relatively sparse studies about the effect of training on NPP operator errors. This might be because most NPPs have SAT programs to ensure that training is adequate.

3.1.15. IDHEAS-DATA IDTABLE-15 for Team and Organization Factors

Introduction to the PIF Team and Organization Factors

Team factors refer to everything affecting team communication, coordination, and cooperation. Teamwork activities include planning, communicating, and executing important human actions

across individuals, teams, and organizations. Examples of teamwork problems seen in event analysis are critical information not being communicated during shift turnover, loss of command and control between the operational center and personnel in the field, and coordination issues between multiple parties at different locations. Distributed locations increase the likelihood of breakdowns in communication, increase the work required to maintain shared situational awareness (common ground) and possibly diverge the team's understanding of the situation and goals to be achieved, and make it less possible to catch and correct other errors.

Safety-critical organizations foster safety culture and have mechanisms for identifying, reporting, and correcting human errors or factors that may lead to human failure events. For example, organizations should document and treat any evidence obtained during the review of an operating event indicating intergroup conflict or indecisiveness or an uncoordinated approach to safety. An organization should also maintain an effective corrective action program to address safety issues such as failure to prioritize, failure to implement, failure to respond to industry notices, or failure to perform risk analyses. The attribute of poor safety culture that impedes safety can vary greatly among organizations.

This PIF has five attributes as follows:

- TOF1 Inadequate team
- TOF2 Poor command & control with problems in coordination or cooperation
- TOF3 Poor communication infrastructure
- TOF4 Poor resource management
- TOF5 Poor safety culture

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A15 IDHEAS-DATA IDTABLE-15. Because team and organization structures vary greatly for different work domains and types of organizations, the data sources identified for IDTABLE-15 were primarily from nuclear power plant operation or simulation. Those studies investigated operator team performance with whole events or scenarios to probe team characteristics. Yet, the human performance data from those studies did not differentiate cognitive failure modes. Several non-nuclear studies were selected for the data sources because they explored the effects of specific team factors on different CFMs. For example, De Dreu and Weingart [75] performed a meta-analysis of 50+ papers studying the associations between relationship conflict within a team, task conflict, team performance, and team member satisfaction. The results revealed strong and negative correlations between relationship conflict and team performance. Specifically, there were stronger negative relations with team performance in decisionmaking tasks than in production (executing procedures or instructions) tasks.

Summary of Generalized Human Error Data

The limited data in IDTABLE-15 are not enough to derive the range or trends of human error rates for the CFMs. Moreover, most studies only reported task performance measures or correlations instead of error rates. The generalized data in IDTABLE-15 establish the initial technical basis that the PIF attributes negatively impact operator task performance and they increase human errors in task performance. More studies are needed to establish the quantitative relation between the PIF attributes and CFMs.

3.1.16. IDHEAS-DATA IDTABLE-16 for Work Process

Introduction to the PIF Work Process

Work Process refers to aspects of structuring operation and conduct of operation. Good work process in safety-critical work domains sets high standards of performance. According to the International Atomic Energy Agency guidance on conduct of operation at NPPs [76], good work processes ensure "making safety related decisions in an effective manner; conducting control room and field activities in a thorough and professional manner; and maintaining a nuclear power plant within established operational limits and conditions… To ensure safety, it is necessary that the management of a nuclear power plant recognizes that the personnel involved in operating the plant should be cognizant of the demands of safety, should respond effectively to these demands, and should continuously seek better ways to maintain and improve safety." Included in NPP work processes are functions and tasks of plant operations, shift complement and functions, operating practices, pre-job briefings, and work control and authorization.

An important aspect of work processes affecting human reliability is verification of personnel's task performance. Verification may come in forms of professional self-verification, independent verification, peer-checking, and/or close supervision. In addition, NPP control rooms also have a shift technical advisor performing independent checking and advising. Verification can capture a large portion of errors personnel made in the first place and correct them. Lack of verification greatly reduces human reliability.

This PIF has five attributes as follows:

- WP0 No impact – Professional licensed personnel with good work practices
- WP1 Lack of professional self-verification or cross-verification
- WP2 Poor attainability to task goal, individual's roles, or responsibilities
- WP3 Poor infrastructure or practice of overviewing operation information or status of event progression
- WP4 Poor work prioritization, planning, scheduling

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A16 IDHEAS-DATA IDTABLE-16. The data sources identified for the PIF are primarily from nuclear power plant operation or simulations. The studies investigated operator performance in normal or EOP scenarios. The human performance data from those studies did not differentiate cognitive failure modes. Also, most of the studies reported task performance measures or number of errors operators made in a scenario, because it was difficult to quantify the number of error opportunities in a scenario. Nevertheless, the relation between the task performance measures and error rates can be inferred from a large set of simulation studies such as those performed by Halden Reactor Project, then the generalized data can be used to estimate the changes of error rates due the changes of the PIF attributes. For example, Skraaning [77] analyzed the data in several Halden Reactor Project experiments in which NPP crews performed whole scenario simulation with fixed or free seating. In fixed seating, operators in a crew except the shift supervisor were restrained to their workstation, while free seating allowed operators to move freely in the control room. Because, in the experiment, operators had transparent displays which allowed them to see reactor process information from every workstation, operators in free seating frequently left their own workstations and grouped with other operators. As a result, the operators highly engaged in group discussion and became less attained to their own task goals, roles, and responsibilities. The data showed that the Operator Performance Assessment Score (OPAS)

was much lower for free seating than for fixed seating. The datapoint generalized from this study is shown in the following:

PIF	CFM	Operator Performance Assessment Score (0-100)			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
WP2	Unsp.		OPAS	Comm per minute	NPP crews performed 2 normal and 2 emergency scenarios (OPAS- Operator Performance Assessment Score and Comm- total communications per minute)	Two seatings: Free - moved freely Fixed - remained seated at workstation, restricted movement except RO	(HSI automation was used in the experiment)	[77]
		Free seating	57	1.05				
		Fixed seating	74	2.75				

Currently, only a few NPP studies were generalized in IDTABLE-16. Many more data sources identified for this PIF have not been generalized. Generalizing data for Work Process from the literature needs exceptional attention to the various aspects of how the study was performed. For example, some studies reported that providing overviews of reactor process information on a large screen display did not improve operator task performance compared with the situation of no overview display. However, several uncontrolled factors in the study could have contributed to the result. The study did not report to what extent that the operators needed the information displayed, to what extent they used the overview display, and how they integrated the overview information with that from their own workstation.

Several non-nuclear operational studies were included in the data sources because the studies were either filling the gaps in data sources for the PIF attributes or they reported data about the effect of a PIF attribute on a specific CFM. For example, many studies from other work domains such as the Transportation Security Administration and radiological medical diagnosis examined the effects of different types of task performance verification. The results showed that independent verification of detecting targets by a second person is more effective than one person performing the task twice.

Summary of Generalized Human Error Data

The limited data in IDTABLE-16 are not enough to derive the range or trends of human error rates of the PIF attributes on the CFMs. Moreover, most studies only reported task performance measures or correlations instead of error rates. The quantitative relationship between the task performance measures and error rates needed to be established. The NRC staff plans to work with researchers from the Halden Reactor Project to better understand the operators' work processes in many simulation studies then generalize the data to IDTABLE-16.

3.1.17. IDHEAS-DATA IDTABLE-17 for Multitasking, Interruption, and Distraction

Introduction to the PIF Multitasking, Interruption, Distraction

Multitasking refers to performing concurrent and intermingled tasks. Interruption and distraction refer to activities that interfere with personnel's performance of the primary task. Interruption means that personnel must stop the primary task momentarily to perform a different task then resume the primary task. Distraction means that a person performs the primary task and purposely or subconsciously uses his or her spare cognitive resources to attend to a distractive activity.

When personnel concurrently perform more than one task, each task demands cognitive resources such as attention, working memory, mental computation, executive controls, etc. Cognitive resources are capacity limited. Personnel need to either split resources attending to multiple tasks at once or quickly switch between the tasks. Both increase the likelihood of making errors. An example of multitasking is concurrently implementing multiple procedures; personnel may skip procedure steps when switching between procedures. An example of extreme multitasking is a situation in which decisionmakers must handle several operational systems (e.g., reactor units) that are in different critical states and the system responses are interdependent. In this example, decisionmaking may mix or transpose related information items about different systems.

Interruption means that personnel must stop the primary task momentarily an attend to the interruptive task. Personnel do not need to switch between the primary task and the interruptive task because they can resume the primary task after completion of the interruptive task. Thus, interruption mainly demands personnel's working memory, maintaining an ongoing cognitive process online and attending to it later. If the primary task does not require continuous thinking or following a sequential order, interruption may have little effect on it. Prolonged interruption refers to situations in which personnel are kept from the primary task for a prolonged period or are interrupted by cognitively demanding requests. Such interruption can severely impact the reliability of resuming the primary task.

Examples of distractions are phone calls, requests for information, and activities other than the primary task. Experienced professionals are trained to manage the cognitive demands of distractive activities. For example, NPP control room operators can manage not being distracted by many irrelevant conversations and alarms so that they can focus on the primary tasks, while they attend to non-primary tasks in their spare time or after completing a primary task. Sometimes, a distraction of low cognitive demand stimuli can enhance a person's vigilance and, therefore, enhance the reliability of performing the primary tasks.

The PIF has ten attributes in the following categories:

- MT1 Distraction by other on-going activities that demand attention
- MT2 Interruption taking away from the main task
- MT3-10 Concurrent multitasking

<u>Summary of the Data Sources</u>

The data generalized for this PIF are presented in Appendix A17 IDHEAS-DATA IDTABLE-17. Numerous studies are relevant to this PIF. However, the terminology of multitasking, interruption, and distraction has been inconsistently used in the literature. Identifying the data sources for this PIF needs to be done carefully. For example, many research papers studied dual-tasking. Yet, dual-task diagrams in the literature can be concurrently multitasking, interruption, or distraction. Having a cell phone conversation while driving was referred to as distraction or multitasking in the literature. For every data source selected for IDTABLE-17, the NRC staff carefully analyzed the experimental method in the description of the original paper and verified the method in several other papers by the same authors or research labs. This ensures that the context of the study is properly mapped to the corresponding PIF attributes.

Most data sources selected for IDTABLE-17 are not from operational data or simulation studies in nuclear power plants or other domains. Operational data and full scenario simulations usually do not distinguish the PIF attributes, and the attributes cannot be controlled throughout an event because they vary at different parts of the operation or a simulated scenario. Moreover, licensed professionals such as nuclear power plant operators or medical physicians use various

strategies managing multitasking, interruption, and distraction, while those strategies are usually not documented in the data sources.

The data sources selected for IDTABLE-17 are mostly from Category C the controlled experimental studies. Thousands of research papers relevant to this PIF are available. The ones selected mostly employed experimental settings that mimicked the tasks in safety-critical jobs such as driving, flying airplanes, attending to medical patients, operating chemical process systems, etc. Other selected data sources include several studies investigating the effects of distraction, interruption, or concurrent tasking on basic cognitive activities such as mental computation, reasoning, or selecting. Although such data are not specific to one of the CFMs, they are useful in calibrating the effects of the PIF attributes on the CFMs involving those basic activities.

Summary of Generalized Human Error Data

The generalized human error data are summarized according to PIF attribute categories. The summary is from the generalized data in IDHEAS-DATA IDTABLE-17 without detaching the effects of other PIFs and uncertainties.

- Distraction – Most datapoints are for Failure of Detection (D) or Failure of Execution (E). The error rates vary from 0.8 to 2 times nominal with the presence of the attribute. The datapoints with the error rates lower in the presence of distraction are typically for the distraction of low salience and low relevance to primary tasks. No datapoint is exclusively for the effect of distraction on failure of Understanding (U) or Decisionmaking (D). It is possible that the effects of distraction on these two macrocognitive functions are negligible.
- Interruption – Most datapoints are for Failure of Detection (D) or Failure of Execution (E). The error rates range from 2 to10 times nominal with the presence of interruption, depending on interruption duration, frequency, and the complexity of resuming the primary task. If the primary task is non-sequential, interruption has little effect on it. The datapoints for Failure of Understanding (U) have error rates between 1.2 and 3 times higher with the presence of the attribute. Yet, it is interesting that the datapoints for Failure of Decisionmaking (DM) show a positive impact on performance with the presence of the attribute. Nicholas and Cohen [78] studied how interruption affects the decisionmaking process. They found that people put forth more effort collecting information and considering alternative strategies after interruptions.
- Concurrent multitasking – Performing concurrent tasks has a profound impact on human reliability. The datapoints for concurrent multitasking attributes have error rate increased 10 to 40 times with the presence of the attributes. The changes to error rates vary dramatically depending on the macrocognitive function, the level of task intermingling, and the cognitive demands of the tasks. That is why seven distinctive attributes are used to represent the variety of concurrent multitasking. For example, a concurrent task can increase the error rate 20 times higher for detecting changes in auditory signals and 5 to 10 times for detecting changes in visual signals. Concurrently diagnosing multiple problems can increase diagnosis errors up to 37 times higher.

In summary, there are abundant data sources in controlled experimental studies for the effects of Multitasking, Interruption, and Distraction. On the other hand, operational data and full scenario simulation with professionals usually mix various attributes of this PIF, thus those data sources were not included in IDTABLE-17. More importantly, the literature showed that licensed professionals have various strategies for managing multitasking, interruption, and distraction to mitigate the impact. It is desired to develop guidance for HRA analysts to evaluate the attributes with the consideration of licensed operators' mitigating strategies.

3.1.18. IDHEAS-DATA IDTABLE-18 for Mental Fatigue

Introduction to the PIF Mental Fatigue

Mental fatigue is a condition triggered by prolonged periods of demanding cognitive activity, which temporally hampers overall cognitive functions, brain productivity, and reliability. When personnel have mental fatigue, they experience various levels of decrement of vigilance, attention span, working memory, and abilities such as reasoning relating information to performing complex cognitive tasks.

Mental fatigue results from psychological, socioeconomic, and environmental factors that affect the mind and the body. It can also result from performing high-demand cognitive tasks for an extended period. A typical situation leading to mental fatigue is sleep restriction or total sleep deprivation. Moreover, mental fatigue can result from an extended period of low mental productivity. For example, monitoring for rare abnormal signals for long hours appears to be not demanding and not productive, but staying vigilant without stimuli for extended periods demands sustained attention and leads to mental fatigue.

The effects of mental fatigue on cognitive activities have been well studied and are generally well understood. The degree to which fatigue affects human performance can range from slight to catastrophic. Personnel can manage and quickly recover from mild mental fatigue. Research had shown that mental fatigue leads to loss of vigilance, difficulty in maintaining attention, reduced working memory capacity, and use of shortcuts in diagnosing problems or making decisions. Moreover, mental fatigue also impairs physiological performance because physiological activities are controlled by mental activities and the central nervous system.

The PIF has four attributes in the follows:

- MF1 Sustained high-demanding cognitive activities requiring sustained focused attention
- MF2 Long working hours with high cognitively demanding tasks
- MF3 Sleep deprivation
- MF4 Change of cognitive state

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A18 IDHEAS-DATA IDTABLE-18. Limited data sources from nuclear power plant operation were identified. Nuclear power plants have fitness-for-duty rules that specify hours of work shifts to ensure that personnel are fit for the job. Also, there are minimum staffing requirements to ensure that operators do not experience severe mental fatigue. Most operational data and studies on mental fatigue in NPPs are from surveys of subjective ratings of fatigue levels. Yet, studies are few on NPP operators' mental fatigue in severe accidents where operators work on highly cognitive demanding tasks for long hours and experience sleep deprivation.

Mental fatigue is well studied in many safety-critical domains such as military operation, aviation, and healthcare. Operational studies examined the effects of shift work, time on task, and sleep deprivation. The effects of mental fatigue on human performance are well understood through numerous controlled experimental studies that isolated the PIF attributes. Controlled experiments typically use three ways to induce mental fatigue: time on task, high cognitive demanding tasks, and sleep deprivation. Sleep deprivation (or hours of wakefulness) is often used because it is relatively simple to achieve and straightforward to measure. With the numerous studies on sleep deprivation, several meta-analysis studies consolidated the

experimental findings and fitted the data with linear regression of human error rates varying with hours of continuous wakefulness or number of hours and days of sleep restriction.

Summary of Generalized Human Error Data

The PIF Mental fatigue has four attributes. Except for MF1, that sustained attention is needed for detection and visual-motor execution, there is no apparent distinction in effects of other attributes on different macrocognitive functions. The effects of the attributes on cognitive task performance include loss of vigilance, reduced attention span, reduced working memory capacity, reduced prospective memory, narrowly focused reasoning and relating information. These cognitive abilities are needed for all the macrocognitive functions in complex tasks.

The effects of the attributes on the CFMs vary continuously with the levels of the attributes, e.g., the time on the sustained attention task, number of wakefulness hours, etc. For example, error rates of Failure of Detection began to increase after 20 minutes of sustained attention and is roughly doubled by 40 minutes. Error rates for long term sleep restriction (e.g., having less than 5 houses of sleep) could increase error rates by four times over rates with normal sleep. Thus, the attributes should be implemented as continuous or multi-scale variables in HRA methods.

Included in the literature of the data sources are various strategies that personnel manage mental fatigue to mitigate the impact. For example, Fysh [79] studied continuously face-matching tasks for passport control. The tasks included identifying the matched faces and mismatched faces among multiple pictures. Error rates for detecting matched faces began to decrease after 15~20mins, while the error rates for detecting mismatched faces remained about the same. With the reduced vigilance, participants allocated their attention resources on the more likely targets. Also, personnel can adapt to mental fatigue. For example, although most studies found that error rates are higher for tasks performed at night compared to the day, professionals working on shifts year-round such as NPP operators or nurses are better adapted to hours of the day compared to people who occasionally work at night. HRA analysts should consider the mitigation strategies and adaptation when evaluating the mental fatigue attributes.

Although the effects of mental fatigue on human performance is generally well understood, one area lacking human error data is the effect of sudden change of cognitive alertness (from a period of low activity to high or vice versa) in nuclear power plant operation. This is particularly important for modeling operator reliability in and immediately after severe accidents.

3.1.19. IDHEAS-DATA IDTABLE-19 for Time Pressure and Stress

Introduction to the PIF Time Pressure and Stress

Time Pressure refers to the sense of time urgency to complete a task, as perceived by personnel. This sense of time urgency creates psychological pressure affecting performance. Time pressure arises when making a tradeoff between thoroughness in performing the task and completing the task in time. Because time pressure is based on perception and understanding the situation, it may not reflect the actual situation. Therefore, although time pressure is most likely to occur when marginal or inadequate time is available, it also could occur in scenarios with adequate available time, but personnel have an incorrect perception of time. For example, some training protocols emphasize the importance of making assertive, immediate decisions, and they reward personnel for rapid correct responses. This type of training can instill an inappropriate sense of urgency, reluctance to question initial impressions, and resistance to deliberative team consultation.

Mental stress, such as anxiety, frustration, threats, or fear, can increase the level of physiological stretch and affect task performance. Examples of stress are concern for families in

emergency conditions, fear of potential consequences of the event, and worrying about personal safety. Such concerns are prevalent during scenarios that involve extreme hazards such as seismic events, floods, high winds, etc. Related to mental stress is the reluctance to implement some planned actions due to concerns or fear of undesirable consequences.

The PIF Time Pressure and Stress has 4 attributes:

- TPS1 Time pressure due to perceived time urgency
- TPS2 Emotional stress (e.g., anxiety, frustration)
- TPS3 Cumulative physical stress
- TPS4 Reluctance to execute an action plan due to potential negative impacts

Summary of the Data Sources

The data generalized for this PIF are presented in Appendix A19 IDHEAS-DATA IDTABLE-19. The data sources for the PIF are organized in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – None of the current nuclear human performance databases such as SACADA or HuREX collects data for this PIF. The databases collect operator simulator training data while operator training is generally performed under normal stress, or operators are trained to attain their performance under stress. The analysis of the German NPP maintenance human event database [5] reported several error rates under moderately high or extreme high stress. For example, the error rate for not memorizing key steps in “Carrying out a sequence of task” was 1/48 given that the type of the task was rarely performed. The error rate for the same type of failure was 2/41 with moderately high levels of stress. The data source did not provide detailed information to discern what kind of stress was involved in the errors made. Thus, the corresponding PIF attributes for this datapoint were unspecified. The following shows the datapoint generalized from this example:

PIF	CFM	Error rates		Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
Unsp	E	No stress	2.45E-2 (1/48)	Carrying out a sequence of tasks (Memorized task step not remembered)	No stress - Rarely performed, no other error promoting factors	(unspecified stress)	[5]
		With stress	5.62E-2 (2/41)		Stress - Rarely performed, moderately high level of stress		

Category B – No operational data from other domains were generalized. Military organizations such as the US Coast Guard research lab have performed many studies on understanding what caused stress and the impacts on military personnel’s task performance.

Category C – Numerous experimental studies have investigated the effects of time pressure and stress on human task performance. Many studies used operational personnel such as nurses, medical physicians, athletes, and military soldiers to perform realistic tasks in operational environments or simulation settings. Controlled experimental studies have also

examined the effects of time pressure and stress on basic cognitive activities such as vigilance, attention, working memory, and reasoning. The studies elucidated the quantitative effects of the PIF attributes on task performance. For example, in Leon and Revelle' s study [80], 120 college students completed 100 geometric analogies with nine levels of complexity under relaxed and time pressure conditions. The relaxed condition had no time limit on performing the tasks. In the time pressure condition, the participants were told that they had only a short length of time to answer each analogy problem before it disappeared from the screen and the next analogy was presented, and if they failed to solve a problem before it disappeared, it would be scored as an error, while in fact only 20% of problems disappeared from the monitor screen and those problems were given adequate time before disappearing. The participants made more errors under time pressure condition. The datapoint generalized from this study is shown in the following:

PIF	CFM	Error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
TPS1	U		Low complex	High complex	120 subjects completed 100 geometric analogies with nine levels of complexity (# of Element and # of Transforms) (%incorrect)	TPS-1: relaxed (non-time-limited) or under time pressure (ego-threat, time-limited)	(time available is sufficient)	[80]
		Relaxed	0.012	0.083				
		Time pressure	0.046	0.375				

Category D – No data source was generalized from this category.

Category E – Given the largely available data sources in Category C, only one meta-analysis study of Category E was documented in IDTABLE-19. Szalma et al. [81] reviewed 281 papers about the effects of time pressure on task performance and quantified the effect sizes from the data in 125 studies. The results showed that the effect size of accuracy is -0.33 for perception (detection) tasks, -0.66 for cognition (understanding and decisionmaking) tasks, and 0.01 for execution tasks. The results suggest that time pressure impairs Understanding and Decisionmaking accuracy more than it does Detection, while it barely affects Execution tasks. The datapoint generalized from this study is shown in the following:

PIF	CFM	Effect size of error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
TPS1	D & U/DM & E	effect-size is a standardized mean difference between the experimental and control conditions.			Controlled lab settings and real-world settings in which temporal constraints impose stress and workload on operators.	time stress: (e.g., instructions to complete tasks as quickly as possible, deadlines, or stimulus presentation rate)	125 of 281 papers with 827 data for meta-analysis	[81]
			accuracy	RT				
		Perception(D)	-0.33	0.26				
		Cognition (U & DM)	-0.66	0.57				
		Motor (E)	0.1	-0.6				

<u>Summary of Generalized Human Error Data for Time Pressure and Stress</u>

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-19 without detaching the effects of other PIFs and uncertainties.

- Failure of Detection (D) – The datapoints have the error rates for Failure of Detection ranging 1.2 ~ 2 times nominal with the presence of the attributes. Among the attributes Time Pressure and Mental Stress have relatively mild effects on error rates, and Physical Fatigue barely impair Detection accuracy.
- Failure of Understanding (U) - - The datapoints have the error rates for Failure of Understanding ranging 1.5 ~ 7 times nominal with the presence of the attributes. The data also reveal speed-accuracy tradeoffs in tasks that require reasoning and relating (both are needed for situation understanding).
- Failure of Decisionmaking (DM) - The datapoints have the error rates for Failure of Understanding ranging 1.5 ~ 7 times nominal with the presence of the attributes. The data on decisionmaking performance measures showed that more decisionmaking errors under stress were due to premature closure of collecting available information and evaluating fewer alternatives.
- Failure of Action Execution – Many datapoints show that the attributes had no impact on execution tasks. Some datapoints show that time pressure even slightly reduces error rates of skill-based tasks by 10% to 20%. Physical fatigue increases skill-based task error rates 1.1 to 1.5 times.
- Failure of Interteam Coordination – No error rate data were identified for this CFM, but several studies showed that coordination and communication were impaired under Time Pressure and Stress. Moreover, personnel became less aware of other team member's work, thus further impairing team coordination.

In summary, there are abundant data sources for the effects of Time Pressure and Stress on human task performance. Existing qualitative data shows the effects of Time Pressure and Stress on teamwork and coordination, yet no data source with error rates was identified to quantify the effect.

3.1.20. IDHEAS-DATA IDTABLE-20 for Physical Demands

<u>Introduction to the PIF Physical Demands</u>

Physical Demands indicate that a task requires extraordinary physical efforts, such as handling heavy objects, performing fine motor dexterity, or operating special equipment. Physical demands challenge motor, physical, and physiological limits. There are professional standards guiding job design to ensure that the physical demands of actions are within human physical limits. High physical demands, even within the professional standards, still have the potential to impair human reliability in task performance. For example, the study "Independent Oversight Study of Hoisting and Rigging Incidents within the Department of Energy" [82] reviewed the incidents over a 30-month interval, from 1993 to 1996 and found that most incidents were caused by human errors rather than equipment failure.

The effects of high physical demands on human errors are twofold, people failing to execute the action properly and personnel injuries. Personnel safety indicates that there is the likelihood of injury when performing certain actions. In practice, personnel safety would most likely apply to scenarios with extreme operating conditions, such as those involving plant damage from internal hazards (fires, floods, etc.), external events (seismic events, floods, high winds, aircraft crashes, etc.), impending or actual core damage, large releases of radiation or toxic chemicals, etc. It accounts for the effects of personnel's concerns about their own personal safety and possible harm or known injuries to their co-workers on task performance. The effects from this PIF may be manifested by personal fear, cognitive distractions, enhanced sense of urgency, additional time delays for cognitive response and action implementation, supervisory reluctance to send personnel into specific plant locations, operator reluctance to perform local actions, etc.

The PIF has five attributes in the following categories:

- PD1 Physically strenuous Action Execution – Approaching or exceeding physical limits (e.g., lifting, handling, or carrying heavy objects, opening/closing rusted or stuck valves)
- PD2 High spatial or temporal precision of fine motor movement needed for Action Execution
- PD3 Precise coordination of joint action by multiple persons
- PD4 Unusual loading or unloading materials (e.g., unevenly balanced loads, reaching high parts, dry cask loading)
- PD5 Handling objects using crane/hoist

Summary of the Data Sources

The NRC staff has not systematically collected data sources for this PIF. IDHEAS-DATA IDTABLE-20 documents a few data sources to demonstrate the attributes. There have been accumulated operational data and studies on this PIF in work domains of high Physical Demands, such as military operation, construction, offshore oil operation, etc.

Regulatory standards and safety work practices minimize the impact of physical demands of human actions. The NRC has developed specific regulations for the handling of heavy loads within the nuclear industry. IDHEAS-DATA IDTABLE-20 documents several HRA applications involving this PIF. In the report "Savannah River Site Human Error Data Base Development for Non-reactor Nuclear Facilities (U)," [37] the HEPs of three Physical Demands actions were estimated: Dropping of load when using forklift, dropping of load when using crane/hoist, and crane/hoist strikes stationary object. Those actions were included in their PRA of site construction and installation. The report "Preliminary, Qualitative Human Reliability Analysis for Spent Fuel Handling" [83] examined how human performance of dry cask storage operations could plausibly lead to radiological consequences that impact the public and the environment. The study investigated typical cask drop scenarios and analyzed human performance vulnerabilities that impact fuel-loading activities and cause cask drops. Examples of human errors in spent-fuel handling include, "Crane operator translates cask into fuel pool wall; cask drops" and "Crane operator raises cask too high; cable breaks & cask drops." The report "Heavy Load Accidents in Nuclear Installations" [84] reviewed operating experience from 114 selected events involving the lifting of heavy loads or the operation of lifting devices. The report highlighted several types of events, such as, "collisions of fuel assemblies with different obstacles during fuel-handling operations;" "inadequate structural design of cranes and other hoisting equipment, particularly regarding seismic resistance;" and "misunderstandings among operations staff leading to loads being handled in unsafe conditions (weight of the load unknown, other operations in progress at the same location, lack of supervision, etc.)." In the NRC's recent work "Effects of environmental conditions on manual actions for flood protection and mitigation"[85] ,the analysis showed that environmental factors impair performance of manual actions especially those associated with high physical demands.

Overall, the attributes of PIF Physical demands are generally not present in NPP control room actions, but they can be present and have significant impacts on human reliability in events outside control rooms. Lots of operational data relevant to this PIF are available in the nuclear and other domains. Human error data related to this PIF have already been collected in some previous HRA efforts. The different sources of data should be consolidated and generalized to inform HRA of special applications.

3.1.21. IDHEAS-DATA IDTABLE-21 for Lowest HEPs of the Cognitive Failure Modes

Introduction to Lowest HEPs of the CFMs

In the IDHEAS-G HEP quantification model, the lowest HEPs are used as the values for the base HEPs when all the base PIF attributes are absent. The Lowest HEP IDTABLE-21 documents the datapoints of human error rates that were measured under the following two criteria:

- None of the known PIF attributes were present or there was no prevalent known PIF attribute present
- The error rates were measured from a sufficiently large number of times that the task was performed so that the measured error rate was statistically reliable.

The human error rates measured under these criteria correspond to the lowest HEP that a CFM of tasks can achieve.

Ideally, data sources for lowest HEPs should also meet the following conditions:

1) The task was performed as the time available was adequate,
2) there was professional self-verification, peer-checking, and/or supervision for task performance
3) the error rate was for a single CFM of a single task, and
4) the error rate was measured without recovery actions.

Hardly any data source can meet the two criteria and all four desired conditions. The NRC staff identified data sources for the lowest HEPs if they met the two criteria. When analyzing data sources for the lowest HEPs, it is important to annotate if any of the four conditions was not met, such as whether there was lack of peer-checking or whether the task of which the error rates were measured had multiple applicable CFMs. The data sources for the lowest HEPs were identified and generalized to IDHEAS-DATA IDTABLE-21. Each row of IDTABLE-21 is for one datapoint. One data source may have multiple datapoints. A datapoint has six dimensions of information presented in the columns: the applicable CFM, the error rate, the task of which the error rate was measured, the notes about whether the conditions are met, the uncertainties in the data, and the reference of the data source.

Summary of the Data Sources

The data generalized for the lowest HEPs are presented in IDHEAS-DATA IDTABLE-21. The data sources are organized into the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

Category A – Several NPP human performance databases and simulator data collection studies have data on lowest HEPs, such as SACADA[24, 26], HuREX/OPERA[38, 86], and the UJV HRA data collection[87]. The databases collect operator simulator training data. Operator training or simulator runs are generally performed by crews with peer-checking, with adequate time, and maybe allowing for recovery to some extent. The error rates for the task types or tasks sharing the same CFM were calculated. Those error rates met the two criteria for lowest HEPs, thus they were generalized to IDHEAS-DATA IDTABLE-21. In addition, the analysis of German

NPP maintenance human event database [4, 5] reported some error rates of which no poor PIF was present or prevalent. Notice that the data sources collected human error data at different levels of detail. For example, SACADA collects operator errors made to training objectives, which are basic tasks of multiple steps in procedures. However, HuREX collected operator errors made at individual procedure steps. The different levels of detail may affect the lowest error rates reported.

Category B – Data sources of the lowest HEPs were identified from air traffic control, NASA Command Center operation, off-shore oil drilling operation, and others. The error rates from real operational data inherits uncertainties and variations in the context under which the task was performed. For example, although most times air traffic controllers perform their tasks with adequate time, at times, they must handle situations when the time available for an action was shorter than needed. To generalize those data sources, the NRC staff reviewed relevant documents about how the jobs were performed in those work domains to understand the nature of the data.

Category C – Error rate data in controlled experiments have the advantage of informing the lowest HEPs because the context is controlled and remains the same for the number of times that the task is performed. The disadvantage is that the same task usually is not performed for many times to get reliable error rates. The data sources identified from this category typically used simple tasks such as detecting signals or performing simple manipulations. Another disadvantage of the data in this category is that the subjects of the experiments usually were not licensed professionals, thus there might be greater individual variability in the same task performed by many subjects. Also, the subjects were not as well trained as licensed professionals. Such uncertainties were documented in the datapoints generalized and should be considered when the data are integrated.

Category D – Several data sources of statistical analysis of human events were generalized in IDTABLE-21. For example, Knecht [88] analyzed flight accident rates of general aviation pilots and reported that general aviation pilot error rates causing accidents was 0.00385 per flight operation (from taking off to landing). Such datapoints do not have specific information about the CFMs applicable to the errors. Moreover, the context under which the errors were made must have had some poor PIFs, thus such error rates do not meet the criteria that no poor PIF should be present or prevalent. The presence of poor PIFs would make the error rates higher than the lowest HEPs. However, the data sources inherit great uncertainties in data collection, where a significant portion of human errors might not be documented because those errors did not lead to reportable consequences. Such uncertainties would make the observed error rates lower than the actual ones. Thus, the data cannot be used to inform the values of the lowest HEPs of the CFMs, but they can be used to calibrate the lowest HEPs estimated.

Summary of Generalized Human Error Data

The generalized human error data are summarized according to the CFMs. The summary is from the generalized data in IDHEAS-DATA IDTABLE-21 without considering the conditions under which the tasks were performed and uncertainties in data collection.

- Failure of Detection (D) and Failure of Action Execution – Many datapoints were generalized for these CFMs from data sources in all the categories. The datapoints appear to be consistent in that the lowest HEPs are in the range of E-3 to E-4, and they vary with the conditions of time adequacy, self-verification or peer-checking, and whether recovery was allowed.
- Failure of Understanding (U) – Substantial datapoints were generalized for this CFMs. Most datapoints were about diagnosis errors. The error rates were in the range of E-3 to E-2.

Notice that diagnosis tasks usually have certain levels of Understanding complexity, therefore, the data sources do not fully meet the criterion of absence of poor PIFs. When the generalized data are used to inform the lowest HEPs of Failure of Understanding, the effect of diagnosis complexity needs to be detached.

- Failure of Decisionmaking – Only a few datapoints were identified for this CFMs, given that controlled experiments usually do not run the same decisionmaking tasks for a sufficiently large number of times, while operational data often do not distinguish decisionmaking errors from other cognitive errors. An exception is the SACADA database[24, 26]. It collects operator errors in decisionmaking. The error rates are around E-2.

In summary, there are substantial data sources to inform the lowest HEPs of the CFMs. For NPP HRA applications, the most accountable data sources are operator performance data from numerous simulator runs. One weakness in IDHEAS-DATA IDTABLE-21 is that there were only a few datapoints for Failure of Decisionmaking. It is expected that the SACADA database will produce more data to better inform the lowest HEP for Failure of Decisionmaking.

3.1.22. IDHEAS-DATA IDTABLE-22 for PIF Interaction

Introduction to PIF Interaction

The PIF Interaction IDTABLE-22 documents the combined effects of multiple PIFs. A longstanding belief in the HRA community is that multiple PIFs interact to affect performance such that the combined effect of the PIFs is the multiplication of the effects of individual PIFs on HEPs. To develop the HEP quantification model in IDHEAS-G, the NRC staff identified over two hundred research papers in which human errors or task performance indicators were measured when more than one PIF varied individually and jointly. Using the definition of PIF attribute weight in IDHEAS-G, the staff examined the individual versus combined PIF weights in the reported data and had the following observations:

- For the majority of the data reviewed, there was little interaction between the PIFs such that the combined PIF weight can be predicted with the addition of the individual PIF weights; When the individual PIF weights are large, the combined weights tend to be less than the addition of the individual weights.
- The multiplication of individual PIF weights tends to over-estimate the combined effects measured in the studies;
- PIF interaction was observed in a small portion of the data as a “gating” effect: The additive effect of joint PIFs is only effective when the weight of one PIF is significantly high. For example, the combined effect of Task Complexity and mental fatigue is additive for complex tasks while mental fatigue has little effect when the Task Complexity is low. Such gating effects are more associated with the three base PIFs: Scenario familiarity, Information completeness and reliability, and Task Complexity.
- Some individual and combined effects of joint PIFs behave differently if both PIFs demand the same capacity-limited cognitive resources and the demand of a single PIF is already approach to the capacity limit. The combined effect is more than the addition of individual effects and reflect the catastrophic effect of exceeding the capacity limit. For example, in a dual-task experiment, if the complexity of the primary task demands working memory approaching to the limit, simultaneously performing a secondary task that also demands working memory would lead to a very high error rate, greater than the sum of the error rates of performing each task alone.

The NRC staff performed a pilot study with a small sample of the reviewed data (in Appendix D of [1]). The study calculated individual and combined PIF weights of the error rates in the sample data and fitted the weights to the addition rule and multiplication rule. The result

confirmed the above observations 1) and 2). Thus, the staff developed the IDHEAS-G [1] quantification model based on the observations. The quantification model adds individual PIF weights for joint PIFs, yet it allows HRA analysts to model PIF interaction with an interaction factor '***C***' in the HEP quantification model.

The NRC staff has not yet generalized and documented all the identified data sources of joint PIF effects in IDHEAS-DATA IDTABLE-21. At present, IDTABLE-21 mainly documents several studies of meta-analyses or literature reviews and analyses of joint PIFs. The main findings of those studies are consistent in that the multiplication effect of joint PIFs was not supported by the data. The following is a summary of those studies:

Van Iddekinge et. al. [89] performed a meta-Analysis of the interactive, additive, and relative effects of cognitive ability and motivation on performance. They analyzed the human performance data from 55 reports to assess the strength and consistency of the multiplicative effects of cognitive ability and motivation on performance. The results showed that the combined effects of ability and motivation on performance are additive rather than multiplicative. For example, the additive effects of ability and motivation accounted for about 91% of the explained variance in job performance, whereas the ability-motivation interaction accounted for only about 9% of the explained variance. In addition, when there was an interaction, it did not consistently reflect the predicted form (i.e., a stronger ability-performance relation when motivation is higher).

Liu & Liu [90] performed regression fitting of human error data on empirical combined effects of multiple PIFs from 31 human performance papers. They calculated the multiplicative and additive effects. The median of the multiplicative effect was greater than that of the empirical combined effect, whereas the median of the additive effect was not significantly different from that of the empirical combined effect. Thus, the multiplicative model might yield conservative estimates, whereas the additive model might produce accurate estimates. It was concluded that the additive form is more appropriate for modeling the joint effect of multiple PIFs on HEP.

Mount, Barrick, and Strauss [91] studied the joint relationship of conscientiousness and general mental ability with performance to test their hypothesis of PIF interaction. This study investigated whether conscientiousness and ability interact in the prediction of job performance. The study performed moderated hierarchical regression analyses for three independent samples of 1000+ participants. Results in the study provided no support for the interaction of general mental ability and conscientiousness. The regression analysis showed that the interaction did not account for unique variance in job performance data beyond that accounted for by general mental ability and conscientiousness alone. These findings indicate that general cognitive ability does not moderate the relationship of conscientiousness to job performance.

Hancock and Pierce [92] examines the combined effects of heat and noise upon behavioral measures of human performance. Specifically, they reviewed the capabilities on a variety of neuromuscular and mental tasks with respect to personnel's vulnerability to joint thermal and acoustic action. Most of the evidence indicates that such stressors do not interact significantly within the ranges experienced commonly in the industrial setting. Yet, the authors warned that various experimental and methodological inadequacies in the meager data base cautioned against a simple interpretation of this apparent insensitivity.

Murray and McCally [93] reviewed human performance and physiological effects of combined stress interaction. They grouped the possible effects into four major types.

I. No effect. Combinations produce no effects greater than those of any of the included stressors alone.

II. Additive effect. Combinations produce effects greater than any single stressors, but not greater than the addition of effects from single stressors.
III. Greater than additive effect. Combinations produce effects greater than mere addition of single stress effects. This possible result is sometimes referred to as "synergistic."
IV. Subtractive effect. Combinations produce effects lower than effects produced by single stressors. This result may be referred to as "antagonistic."

These four types of outcomes seem to be likely on a theoretical basis of possible interactions among PIFs. Type I seemed most likely when the stressors included in the combination are unequal in their effects. Then the more severe stress would dominate the results, and variables with less effect would make no detectable addition to the overall result. Type II seemed to be the most likely when the stressors are about equal in their effects, and their mechanisms of action are independent. Type III and Type IV, synergistic and antagonistic effects were rarely observed in reported experiments.

Grether [94] reviewed the studies about the effect of combined environmental factors on human errors. The reviewed environmental factors included noise, temperature, sleep deprivation, and others. The results showed that the combined effect was no more than the added single effects and could be predicted from single effects. The study suggests that the combined environmental stresses do not present a special hazard in flying that could not be anticipated from the results of single factor studies. The findings are consistent to those in Broadbent's study [95] that reviewed many experiments applying different stresses to comparable subjects performing similar tasks. The study found that the experiments on the simultaneous application of two stresses show that the effects of heat appear to be independent of those of noise and sleeplessness, while the latter two conditions partially cancel each other.

Given that the above listed meta-analysis and review studies are, in general consistent, the additive effect of joint PIFs seems to be applicable for the majority of PIF weight ranges, it may not add much value to generalize the large amount of identified data sources into IDHEAS-DATA IDTABLE-22. Rather, in-depth studies are desirable to understand the nature of PIF interactions and elucidate the situations that the joint effects become synergistic rather than additive, because such situations represent great hazards to safety-critical operation.

3.1.23. IDHEAS-DATA IDTABLE-23 for Distribution of Time Needed

<u>Assess the Time Needed</u>

Using empirical data (e.g., training data or actual event data) is the recommended method to estimate the time needed (TN). In many cases, plant-specific empirical data may not be available. When the data of similar plants are available, the analyst may use the data to support the TN assessment. The relevant data may show a significant difference in TN. This section discusses the factors that should be considered in assessing TN using data of similar plants or similar scenarios. The purpose of this section is to raise awareness about factors that could significantly affect TN. The discussion does not intend to provide a comprehensive list of factors nor provide guidance on assessing TN. That requires a study of its own.

<u>Acuteness Disturbance on Symptom</u>
Table 3-1 shows the operators' response time in 8 steam generator tube ruptures (SGTRs) [96]. It shows that the operator response time can be divided into two groups based on the steam generator (SG) rupture flow. The Point Beach 1 and Fort Calhoun, with the ruptured flow rates of less than 130 gallons per minute (gpm), had a significantly longer time for the diagnosis and isolation of the ruptured SG. The times are counted from the beginning of the SGTR.

Table 3-2 shows the means and standard deviations of the two groups. In both groups, the standard deviations of the time to reach a diagnosis are less than 5 minutes. The similar standard deviation (3 vs 4.4) and significant difference in mean (29 vs 4.8) is an indication that the SGTR rupture flow rate affects operator diagnosis time. The time to isolate the ruptured SG (from the beginning of an SGTR) of the two groups show a significant difference in mean values. The difference is an inherent effect of the difference in diagnosis times. The authors of had the same observation and concluded that the Point Beach 1 and Fort Calhoun events were more complex than the others because of ambiguous conditions [97]. Another explanation could be simply that the operators did not have the urgency to quickly respond to the events without acute disturbances to the system. Based on the conventional nuclear power plant design, an SG rupture flow rate between 130 and 300 gpm minimizes the acuteness of system disturbance.

Table 3-1 The operator response times in SGTR events [96].

Plant	Capacity (MWe)	Vendor (# of loop)	Event Year	Tube Rupture Flow Rate (gpm)	Time to SGTR Perception (Minute)[a]	Time to SG Isolation (Minute) [a]	Plant State
Point Beach 1	500	WEC(2)	1975	125	24 ~ 28	58	Full Power
Surry 2	823	WEC(3)	1976	330	< 5	18	Full Power
Prairie Is. 1	545	WEC(2)	1979	336	5 ~ 18.5	27	Full Power
Ginna	490	WEC(2)	1982	760	< 1	15	Full Power
North Anna 1	947	WEC(3)	1987	637	< 5	18	Full Power
McGuire 1	1100	WEC(3)	1989	500	< 1	11	Full Power
Mihama 2	470	WEC(2)	1991	700	< 5	22	Full Power
Fort Calhoun	476	CE(2)	1984	112	< 32	40	Startup

[a]The time after the SGTR started.
WEC: Westinghouse Electric Company. CE: Combustion Engineering.

Table 3-2 Time needed analysis based on the example Table 3-1 data

Tube Rupture Flow Rate	Time to Reach an SGTR Diagnosis [a] (Minutes)		Time to Isolate the ruptured SGs [a] (Minutes)	
	Mean	Standard Deviation	Mean	Standard Deviation
< 150 gpm	29	3	49	13
> 300 gpm	4.8	4.4	18.5	5.5

[a]The time after the SGTR started.

Simulated Events vs. Actual Events

Based on the experience of the authors of this report, most nuclear power plant operator instructors believe that operators behave similarly in simulated and actual events. One instructor indicated that his plant had an actual event similar to a simulated event, and the operators' responses were the same in the actual and simulated events. Table 3-3 provides supporting evidence. Table 3-3 shows the times to isolate the ruptured SG in actual and simulated events, including:

- Actual US SGTR events shown in Table 3-1 above [96] with SGTR rupture flow rates greater than 300 gpm.
- Korean crews in a Korea standard nuclear power plant (KSNP) simulator [96], which is a 1000MWe CE type pressurized water reactor (PWR) with conventional control interfaces.
- Korean crews in a KSNP simulator [98], which is a 950MWe Westinghouse 3-loop PWR with conventional control interfaces.

- US HRA Benchmark Study [16], an SGTR event with a 500 gpm rupture flow rate.
- International HRA Benchmark Study [22]. The study was conducted in an experimental facility. The main control room was fully digitalized.

Table 3-3 shows that, in SGTR events, the time to the isolate the ruptured SG in actual events and simulated events and in Westinghouse and Combustion Engineering pressurized water reactors, are very consistent. The 2 to 3 minutes shorter response time in the International HRA Benchmark Study [22] could be because the study was conducted in a fully digitalized main control room. In the other studies and actual events, the operators are in conventional main control rooms. Even though the reports [96] and [22] does not document the SG rupture flow rates, it is expected the SGTR symptoms in the two studies are comparable to a greater than 300 gpm SGTR event. All the simulated SGTR events in Table 3-3 are basically (straightforward) SGTR events.

The KSNP-Westinghouse data shown in Table 3-3 were not documented in [98] but through an information exchange with the authors of [98] (see "Verify the Outlier Data" discussion of this section. The authors of [98] attributed the short response time of the KSNP-Westinghouse crews in Table 3-3 to their early detection of SGTR symptoms before reactor trip and promptly responded to the event.

Table 3-3 The time to isolate the ruptured steam generator in actual events and simulated events.

SGTR Studies	Mean Time to Isolate the Ruptured SG (s) (Minutes)[a]	Standard deviation to Isolate the Ruptured SG(s) (Minutes)
Actual events (6 events, > 300 gpm)	18.5	5.5
KSNP-CE (23 crews)	19.8	3.0
KSNP-Westinghouse (6 crews)	13.8	3.6
US HRA Benchmark (3 crews, SGTR)	19.0	3.5
International HRA Benchmark (14 crews, basic SGTR)[b]	15.9	3.6

[a]The time is from the SGTR occurrence to the ruptured SG isolation.
[b]The study was conducted in an experimental facility with a digitalized main control room.

<u>Basic vs. Complicated Scenarios</u>
Both the US HRA Benchmark Study [16] and the International HRA Benchmark Study [22] performed basic and complicated SGTR events. In the US HRA Benchmark Study, the complicated SGTR event started with a loss of feedwater event that required establishing feed-and-bleed (F&B) to maintain cooling of the reactor coolant system.

> After F&B has been established, the crew will be able to establish auxiliary feedwater (AFW) flow to one or several SGs by either closing the recirculation valve and/or cross-connecting the flow from the running AFW pump to the other SGs.
> As soon as the crew has established AFW flow, the trainers will initiate a tube rupture in the first SG that is fed. The crew will want to fill an SG to be able to exit FR-H1, and the tube rupture may be masked by AFW flow to the SG, as long as it is being fed. The leak size of the ruptured tube is about 500 gallons per minute (gpm) at 100% power, but the flow will depend on the differential pressure between the reactor coolant system (RCS) and the ruptured SG. There is initially no secondary radiation because there is only a minimum steam flow. The blowdown (BD) and sampling are secured because of the SI.

By the time the crews fill the SG(s) enough to exit FR-H1, they may have problems with the RCS integrity status tree and be forced to enter procedure FR-P1, which will delay the possibility of transitioning to the SGTR procedure E-30. [22]

The international HRA Benchmark Study [22] studied basic and complicated SGTR events. The main scenario differences between the complicated and basic events were:

a) the event starts off with a major steamline break with a nearly coincident SGTR in SG #1 that will cause an immediate automatic scram and expectations that the crew will enter the EOP-0 procedure; and
b) auto closure (as expected) of the Main Steam Isolation Valves (MSIVs) in response to the steamline break along with the failure of any remaining secondary radiation indications (not immediately known nor expected by the crew) as part of the simulation design.

Table 3-4 shows the comparison of the response times. It shows that, compared to the basic events, in complicated events, the operators take a longer time to isolate the ruptured SG and have a larger standard deviation. In the complicated SGTR event of the US HRA Benchmark Study, four data points are available: 11.2, 33.0, 24.5, and 94.5 minutes. The last data point (94.5 minutes) is considered as an outlier. It is considered to be caused by cognitive failures that should not be included in IDHEAS-ECA's TN assessment. The analysis shown in Table 3-4 has excluded the outlier data point.

Table 3-4 Comparing the response time of simple and complicated SGTR events

SGTR Scenarios	Mean Time to Isolate the Ruptured SG (s) (Minutes)[a]	Standard deviation to Isolate the Ruptured SG(s) (Minutes)
US HRA Benchmark (3 crews, basic)[b]	19.0	3.5
US HRA Benchmark (3 crews, complicated)[bd]	22.9	11.0
International HRA Benchmark (14 crews, basic)[c]	15.9	3.6
International HRA Benchmark (14 crews, complicated)[c]	26.9	6.4

[a]The time is from the SGTR occurrence to the ruptured SG isolation.
[b]The study was conducted in a conventional main control room.
[c]The study was conducted in an experimental facility with a fully digitalized main control room.
[d]Excluded a data point (from a total of four data points) that was considered an outlier.

<u>Verify the Outlier Data</u>

The response times of different studies performed in similar settings could vary significantly. The analysts should perform a "sanity check" to identify the outlier data points and, if feasible, to verify the data to prevent misinterpretation. An example is that a journal paper [98] documents operator response time to a basic SGTR event as shown in Table 3-5. The test facility is a Westinghouse 3-loops PWR (950MWe and conventional interfaces). The Tasks 1 to 8 in Table 5 cover the procedural step to respond to an SGTR event to the point that the ruptured SG is isolated. The sum of the average time spent on the tasks is about 5.5 minutes. That is significantly shorter than the other data (ranging from 16 to 20 minutes as shown in Table 3-1). Upon discussion with the authors of the journal paper [98], the task times in Table 3-5 are only the time spent on that task (the egress time minus ingress time of the task). The time spent between tasks is not counted. The authors of the journal paper [98] checked the original data records and provided the mean and standard deviation of 13.8 and 3.6 minutes, respectively. Those values are relatively close to the values of the other data points. The data are shown in

Table 3-5. A lesson learned is that when suspecting a data point is an outlier, the analysts should verify with the data providers to ensure correct data interpretation.

Table 3-5 The crew performance time in a basic SGTR event of a Westinghouse 3-loop PWR [98]

Task ID	Task Description	Time[a]	SD[b]
1	Confirming immediate response after reactor trip	41.9	25.5
2	Confirming the isolation of essential valves	12.0	2.9
3	Confirming the operation of essential pumps	17.9	5.6
4	Verifying containment status	33.9	22.3
5	Verifying the delivery of SI and AFW flow	55.4	27.8
6	Verifying the status of RCS heat removal	38.9	16.0
7	Entering E-3 procedure according to the status of SGs	34.7	10.3
8	Identifying and isolating faulty SGs	97.0	25.6

[a] Averaged task performance time in second
[b] Standard deviation in second

3.1.24. IDHEAS-DATA IDTABLE-24 for Modification of Time Needed

Introduction to Modification to task completion time

Many factors modify task completion time. These factors contribute to the uncertainty in time distribution. The time uncertainty model in IDHEAS-G requires HRA analysts to estimate the distribution of time needed for a human action. The center, range, and shape of time distribution can be modified by many time uncertainty factors such as weather or environmental conditions. IDTABLE-24 documents the modifications of task completion time under various time uncertainty factors.

Summary of the Data Sources

The NRC staff has not generalized the data sources identified for this PIF. IDHEAS-DATA IDTABLE-24 documents a few data sources for demonstration. There have been accumulated operational data and experimental studies for modifications of task completion time. In fact, most data sources identified for IDHEAS-DATA IDTABLE-1 through IDTABLE-20 also have data about the effect of the studied PIFs on task completion time.

A data source for IDTABLE-24 should have task completion times under at least two different states of time uncertainty factors to inform the effect of the factor on task completion time. The most useful data for IDTABLE-24 would be operational data from tasks performed by licensed, professional personnel. However, operational data typically do not systematically record action performance times under different factors. On the other hand, controlled experimental studies have data on task completion times with varying time uncertainty factors.

The NRC staff identified data sources from three categories. Category A is nuclear power plant operation or simulation. KAERI has systematically collected operator task performance times in control room operation. The data were recorded as operators performed training or requalification examinations, thus the factors contributing to task performance time were known. Operator simulation studies by many NPP organizations reported operator task performance times in different scenarios and conditions. For example, Park et.al. [99] investigated the relationship between performance influencing factors and operator performances in the digital main control rooms. In the study, crews performed scenarios that varied in complexity and urgency. The study involved the participation of licensed NPP operators and the use of an APR1400 simulator. Half of the participants had some experience with the APR1400 simulator. The other half had not worked with it before. During the simulation, operator performance such

as completion time, errors, and situational awareness were measured and collected. The results indicated that task completion time, measured as seconds per procedure instruction, varied with the factors tested. The operators' experience with the APR1400 simulator was most impactful on task completion time, with the mean varying from 9 to 16 seconds per instruction. On the other hand, the mean value of task completion time did not change with scenario urgency, but the range of task completion time among the crews was more broadly distribute for less urgent scenarios than for urgent scenarios. The datapoint generalized from this study is shown in the following:

CFM	PIF	Task completion time (mean and SD)		Task	PIF measure	Other PIFs (and Uncertainty)	REF
		PIF-Lo	PIF-Hi				
Unsp	TE	9(1.5)s per instruction	16(2)	4 NPP crews perform EOP scenarios	Lo – Experienced with AP1400 Hi – No experience with AP1400	(4 crews)	[99]
Unsp	TPS	13(2.5)	12(4)	EOP scenarios	Lo- urgent Hi- less urgent	(4 crews)	[99]
Unsp	SF / INF	12(5)	14(2)	EOP scenarios	Lo – Design basis event Hi - Design basis event + masking	(4 crews)	[99]

The Category B data sources are from operation or simulation of job performance in non-nuclear domains. The data sources from nuclear power plants are primarily from control room operation and they do not have data about the effects of many factors outside control rooms, such as environmental factors on manual actions. The data sources identified from other work domains are used to fill the gaps. For example, Kelly [100] examined the effect of military soldiers wearing MOPP IV gear on cognitive task performance. The results showed that performance time on simple response tasks increased 10~20% after one hour wearing the gear, and the increased performance time was accompanied with decrements in performance accuracy. Thus, the modification to task completion time represents the overall performance decrement. Taylor and Orlansky [101] studied the effects of wearing protective chemical warfare combat clothing on human performance of different types of jobs such as combined arms in nuclear and chemical environments, military manual actions, fire rescue operation, etc. For example, one of the studies showed that the average time for crews to perform a maintenance task "Remove and Replace M60A3 Transmission" was 73.5 minutes in battle uniform dress and 125.9 minutes wearing MOPP protective clothing.

The data sources in Category C are from controlled experimental studies, and most data sources selected from this category for IDTABLE-24 involved tasks and experimental settings that mimicked tasks in real operation domains. Although the studies were low-fidelity simulations, the individual factors were isolated to elucidate the effects of individual factors on task completion time. For example, Speier et. al. [102] studied the influence of interruption on individual decision making. In the experiment, the number of information items to be integrated for decisionmaking was manipulated as simple versus complex tasks. Interruption was manipulated at different frequencies of interruption and the content similarity between interruption and the decisionmaking tasks. The results showed that interruptions improved decisionmaking performance on simple tasks and lowered performance on complex tasks. For complex tasks, the frequency of interruptions and the dissimilarity of content between the primary and interruption tasks was found to exacerbate this effect. The decrement in performance was represented with increased task completion time and decreased accuracy. The datapoint in IDTABLE-24 from this study is shown in the following:

CFM	PIF	Task completion time: mean (SD) in second		Task	PIF measure	Other PIFs (and Uncerta-inty)	REF
		PIF-Lo	PIF-Hi				
DM	MT2	110.3 (27.6)	90.8 (30.8)	Simple decisionmaking	Lo – No interruption Hi – With interruption		[102]
DM	MT2	608.3 (284.4)	760.8 (293.8)	Complex decisionmaking	Lo – No interruption Hi – With interruption		[102]
DM	MT2	831.3 (238.7)	1702.5 (526.8)	Complex decisionmaking	Lo- low interruption freq. Hi- High interruption freq.		[102]
DM	MT2	1317.4 (613.9)	1842.0 (741.6)	Complex decisionmaking	Lo- Different content Hi- Similar content		[102]

Status of IDTABLE-24

At present, only a few datapoints are documented in IDTABLE-24 for demonstration. Documenting all the data sources identified on Modification of Task Completion Time is time-consuming. Moreover, the identified data sources by the NRC staff are only a very small proportion of the data available in public domain. Before generalizing data sources to IDTABLE-24, a screening study should be performed first to identify the factors that modify time significantly. Based on the data generalized in IDTABLE-23, IDTABLE-24, and other documents, the NRC staff intends to develop guidance on estimating uncertainty distributions of time needed to assist the use of IDHEAS in HRA applications.

3.1.25. IDHEAS-DATA IDTABLE-25 for Dependency of Human Actions

IDHEAS-DATA Dependency examples

This section provides examples of the three types of dependency: consequential dependency, resource sharing dependency, and cognitive dependency. Consequential dependency is the outcome of one task directly affects the performance of the other tasks. Resource sharing dependency occurs when two tasks share the same resources (e.g., containment spray and reactor coolant system (RCS) cooling share the same water source, or there is limited manpower to perform multiple tasks). Cognitive dependency is the same cognitive mechanism that failed a task failed the subsequent tasks. The examples are from operations experience. Each example starts with a brief explanation of the dependency then followed with the detailed narrative of the operation experience.

Consequential dependency

Example 1: Failure to control RCS inventory, that resulted in a liquid-solid pressurizer, consequently affecting the performance of terminating safety injection.

> On April 17, 2005, at 8:29 a.m., Millstone Power Station, Unit 3, a four-loop pressurized-water reactor, experienced a reactor trip from 100-percent power [103]. The trip was caused by an unexpected “A” train safety injection (SI) actuation signal and main steamline isolation caused by a spurious “Steam Line Pressure Low Isolation SI” signal. As a result of the main steam isolation signal, the main steam isolation valves and two of the four main steamline atmospheric dump valves automatically closed. With the closure of the main steam isolation valves, the main steamline safety valves opened

to relieve secondary plant pressure. Control room operators entered Emergency Operating Procedure (EOP) E-0, “Reactor Trip or Safety Injection,” and manually actuated the “B” train of SI and actuated the “B” main steam isolation train in accordance with station procedures. Both motor-driven auxiliary feedwater (AFW) pumps started to maintain steam generator (SG) levels. The turbine-driven AFW pump attempted to start but immediately tripped on overspeed. Operators were dispatched to investigate the cause of the turbine-driven AFW pump trip.

At approximately 8:42 a.m., the shift manager noted that a “B” main steam safety valve had remained open for an extended time. In consultation with the unit supervisor and shift technical advisor, the shift manager declared an alert based on a stuck open main steam safety valve. The crew determined that the stuck open main steam safety valve represented a non-isolable steamline break outside containment. The main steam safety valves were in fact functioning as designed to relieve post-reactor-trip decay heat with a main steamline isolation signal present. In this event, the main steam safety valves closed once the operators took positive control of decay heat removal by remotely opening the atmospheric dump bypass valves.

At 8:45 a.m., because of the addition of the inventory from the SI, the pressurizer reached water solid conditions and the pressurizer power-operated relief valves cycled many times to relieve RCS pressure and divert the additional RCS inventory to the pressurizer relief tank. No pressurizer safety valve actuations occurred, and the pressurizer relief tank rupture diaphragm remained intact. At approximately 8:59 a.m., the operating crew transitioned from EOP E-0 to ES-1.1, “Safety Injection Termination.” The SI was reset, the crew terminated SI at 9:12 a.m., and normal RCS letdown was reestablished at 9:20 a.m.

Example 2: Failure to complete the isolation valve leakage test that resulted in the system being in a wrong configuration to perform the valve stroke test, caused the failure of the valve stroke test.

On October 4, 1990, at 1:24 a.m., Braidwood Unit 1 experienced a loss of approximately 600 gallons of water from the reactor coolant system (RCS) while in cold shutdown [104]. Braidwood 1 technical staff was conducting two residual heat removal (RHR) system surveillances concurrently, an isolation valve leakage test and a valve stroke test. After completing a leakage measurement per one surveillance procedure, a technical staff engineer (TSE) in the control room directed an equipment attendant to close an RHR system vent valve. However, before those instructions could be carried out, another TSE in the control room directed that an RHR isolation valve be opened per another surveillance procedure. While the equipment attendant was still closing the vent valve, RCS coolant at 360 psig and 180 °F exited the vent valve, ruptured a Tygon tube line and sprayed two engineers and the equipment attendant in the vicinity of the vent valve. This loss of coolant was reported to the control room and the control room personnel quickly identified the cause and isolated the leak.

Resource-sharing dependency

Example: Performing the atmospheric dump valve (ADV) Partial Stroke Test (that caused excessive letdown) and the boron injection flow test (that limited charging flow) simultaneously caused a loss of letdown.

On May 7, 2004, Palo Verde [19] simultaneously testing the atmospheric dump valve and boron injection systems resulted in a loss of letdown event on high regenerative

heat exchanger temperature. The procedures of the two surveillances were " atmospheric dump valve (ADV) 30% Partial Stroke Test" and "Boron Injection Flow Test." The simultaneous performance of these evolutions caused a loss of letdown due to the high regenerative heat exchanger outlet temperature. This condition occurred due to a single charging pump operation per "Boron Injection Flow Test" procedure, combining excessive letdown flow to accommodate the RCS heat up following ADV partial stroke testing.

Cognitive dependency

Example: Failure to deisolate two wide range indicators (0-3000 psig) and one low range indicator (0-800 psig) because of failure of the same cognitive mechanism.

On March 20, 1990, at about 09:30, Catawba Station Unit I experienced an overpressurization of the Residual Heat Removal System (RHR) and Reactor Coolant System (RCS) during the procedure to initially pressurize the RCS to 100 psig following a refueling outage [105]. The operators had three indicators for monitoring RCS pressure (two wide range indicators, 0-3000 psig, and one low range indicator, 0-800 psig) which were being closely monitored for a detectable rise in RCS pressure. However, unknown to the control room operators on duty, all three RCS pressure instrument transmitters were still isolated after the welding of the tubing fittings during the refueling outage.

3.1.26. IDHEAS-DATA IDTABLE-26 for Recovery of Human Actions

The primary sources of information for IDHEAS-DATA IDTABLE-26 are the event reports, ASP/SDP analysis reports, operational experience reviews, and reports on operator performance in simulators. Several examples were included in IDTABLE-26 to demonstrate recovery actions and different kinds of data sources. The examples are summarized as follows:

- In OECD/NEA report, "Human Factor Related Common Cause Failure - Part 1, Report from the Expanded Task Force on Human Factors," [20], many human failure events in NPPs were analyzed for common cause failure and recovery actions. Among 17 maintenance human failure events analyzed, eleven events occurred in the outage phase, and 5 of these during start up. Another might be during power operation. Scheduled periodical tests detected nine of the events. This reference provides a datapoint of error recovery rate in maintenance surveillance tests as 0.53 (9/17).
- In the study, "A HAMMLAB HRA Data Collection with U.S. Operators," by Massaiu and Holmgren [106] of Halden Reactor Project, five US crews performed three challenging emergency scenarios: Multiple SGTRs, ISLOCA, and Loss of all feedwater. The crews made totally 65 errors and only 13 of them were recovered. Detection and Execution errors had much higher recovery rates (2/5 and 5/18) than those of Understanding and Decisionmaking errors(1/17 and 4/25).
- In the report, "An empirical study on the human error recovery failure probability when using soft controls in NPP advanced MCRs," by Jang et al. [107], 48 subjects performed tasks from emergency scenarios. The study recorded the error recovery rates for eight types of error modes in Failure of Execution as the following:

  Recover rate (operation selection omission) = 0.052
  Recover rate (operation execution omission) =0.71
  Recover rate (wrong screen selection) =0.93
  Recover rate (wrong device selection) =0.5

Recover rate (wrong operation) =0.6
Recover rate (mode confusion) =0.8
Recover rate (inadequate operation) =0.5
Recover rate (delayed operation) =0.02

The results show that, even for the same CFM, recovery rates can vary greatly.

These studies show that human error recovery probability, just like HEPs, vary with CFMs and the context of recovery actions. Thus, it is possible that recovery actions can be modeled the same way as important human actions in HRA, with specific attention to the dependency between the recovery action and the failure of the important human action.

In summary, modeling recovery actions is still an underdeveloped area in HRA. While the PRA standard and some HRA methods have guidance for assessing the feasibility of recovery actions, none of the HRA methods have explicitly modeled the quantification of failure probabilities of recovery actions. IDTABLE-26 made an initial effort to systematically collect qualitative and quantitative datapoints of recovery action. As more datapoints are populated in IDTABLE-26, the information will provide the basis for modeling recovery actions in HRA.

3.1.27. IDHEAS-DATA IDTABLE-27 for Main Drivers to Human Failure Events

IDHEAS-DATA IDTABLE-27 generalizes situations or contexts that are the main drivers to human failure events in operational or simulated events. The data sources in IDTABLE-27 are primarily from the nuclear domain. The NRC staff has investigated data sources but has not systematically collected and analyzed them. This section summarizes viable data sources. IDHEAS-DATA IDTABLE-27 presents several examples to demonstrate the generalization of data sources For Main Drivers to Human Failure Events.

Event or accident analysis

Analysis of major or significant nuclear events has been performed by the NRC, industries, and research organizations. For example, there have been many studies of human error or human factors analysis for major NPP events such as the Fukushima accident, Three-Mile Island accident, or Robinson fire event. In-depth analyses document the event context and identify human errors in the event along with the main drivers to the errors. This kind of data source allows the NRC staff to represent the main drivers in IDHEAS-G CFMs and PIFs. Such data sources document information about single events, thus they do not inform HEPs or the frequencies of the main drivers. However, as more data sources are documented in IDHEAS-DATA IDTABLE-27, events or accidents with similar contexts or main drivers can be grouped together to provide HRA analysts a holistic understanding of what can happen to human performance for similar situations.

Operator performance simulation studies

Simulation studies of NPP operator performance are usually conducted with licensed operators on high-fidelity training or research simulators. The studies use hypothetical, yet realistic scenarios and real procedures. Such studies observe operators' behaviors and measure operators' performance such that human failures and the main drivers in such simulated events can be elucidated. Simulation studies also have the advantage that the same scenario is typically performed by multiple crews, thus the studies have human error data although with large uncertainties due to the small numbers of the participants.

Analysis of data in operator performance databases

Operator performance databases collect data from many operators or operating crews performing the same tasks multiple times. For example, SACADA [24, 25] collects human performance data from operator simulator training, in which operators perform the same training objective tasks in the same and different scenarios. SACADA documents operators' success or satisfaction of task performance along with the types of failures made and the situational factors under which a task is performed. Analysis of the large amount of data collected in the database can reveal the types of failures with high unsatisfactory rates and aggregate the situational factors associated with satisfactory performance. The aggregated situational factors are likely the main drivers of the unsatisfactory performance. Another example is the analysis of German NPP maintenance performance database [4, 5]. The analysis shows that most of the very high error rates are associated with rarely performed tasks. Thus, scenario or task familiarity appears as one of the main drivers to human errors in NPP maintenance tasks. Such analysis aggregates data of the same task under a variety of situational factors, thus the analysis may not reveal all the main drivers and some rarely presented main drivers could be missed.

Human error analysis

Human error analysis, sometimes also referred to as root causal analysis, uses a taxonomy or classification scheme to analyze a set of human events. The analysis classifies human errors in an event to predefined error types and the associated context to causal factor categories. The studies then calculated the frequencies of the types of errors or causal factors appearing in all the events analyzed. The frequency is often represented as the percent of an error type or causal factor that occurred in an analyzed sample of human events. Because each event can be associated with multiple error types and causal factors, the sum of the percentages of all the error types or causal factors are usually greater than 100%. For example, Gertman et. al.[108] studied the contributions of human performance to risk in operating events at commercial nuclear power plants. They reviewed 48 events described in licensee event reports (LERs) and augmented inspection team reports. Human performance did not play a role in 11 of the events so they were excluded from the sample. In the remaining 37 events, 270 human errors were identified, and multiple human errors were involved in every event. The results show maintenance practices was highest (54%), followed by design deficiencies (49%), and procedures (38%). Errors in communication and errors in configuration management were each present in 27% of the events. The numbers or percentages of error occurrences inform the prevalent types of human errors in the event sample analyzed. Yet, they do not necessarily relate to main drivers that occurred less frequent but had significant impacts on the likelihood of human errors.

In summary, compared to most other IDHEAS-DATA Tables, IDTABLE-27 for Main Drivers to Human Failure Events is still in its exploratory stage. The NRC staff has not yet demonstrated how the information documented in this IDTABLE will be integrated and used for HRA. One potential approach is to aggregate the datapoints in IDTABLE-27 and then link the aggregated information to the corresponding CFMs and PIF attributes in the IDHEAS-ECA tool.

3.2. Integration of the generalized data for IDHEAS-ECA

This section describes an example of integrating the data in IDHEAS-DATA to provide the basic numbers for calculating HEPs in the IDHEAS-ECA method [2]. The integration process was described in Section 2.5.

The following is the recapture of what was described in Section 2.5 about the general process of integrating human error data for lowest HEPs, base HEPs, and PIF attribute weights:

- Assess and organize the datapoints according to the data source categories and datapoint types.
- Use single-component, Category A, B, and C datapoints to make initial estimates of a base HEP or PIF weight;
- Use the initial estimation to detach multi-component data into single-component data.
- Integrate all the data available from the single-component and detached multi-component datapoints to estimate the range and mean of a base HEP or PIF weight.
- Use Category D and E and range datapoints to calibrate the estimated HEPs and PIF weights and adjust the mean values accordingly to represent the breath of the available data.
- Iterate the process and calibrate the estimated HEP to represent the breath of the available data.
- If there are no single-component or multi-component detachable datapoints available, then use multi-component undetachable or range datapoints for HEP estimation.

The biggest challenge in using the human error data is that most datapoints are not exclusively for one PIF attribute and one CFM. The most essential step in integrating the data is detaching the effects of other PIFs in the human error rates to make the data exclusively represent the effect of the PIF attribute being analyzed. Detaching makes the integration process iterative. Initial estimates of some frequently involved base HEPs and PIF weights must be made for the use of detaching; the detached error rates are used to make estimates of the base HEPs and PIF attribute weights.

The following section shows an example of integrating the datapoints in IDHEAS-DATA IDTABLE-21 to obtain the lowest HEP of the CFM Failure of Detection. The NRC staff followed the general process and made engineering judgment as necessary. The example demonstrates the integration process without excluding or rejecting reasonable alternative lowest HEP values in other HRA methods.

3.2.1. Assessing and organizing the datapoints

The first step is to assess and organize the datapoints for the CFM Failure of Detection in IDTABLE-21. Datapoints in IDHEAS-DATA Tables are referred to as the following types:

Single-component datapoint – A datapoint has the error rates for a single CFM with the presence of a single PIF attribute.

Multi-component detachable datapoint – A datapoint has the error data with the presence of multiple PIF attributes. The PIF attributes are clearly defined in the data source and the combined effects can be detached into the effects of individual attributes.

Bounding datapoint – Those datapoints have the range or trend of the human error rate for the CFM and PIF attribute being studied. For example, a datapoint has human error rates of certain error modes and the error rates were calculated from statistical data that involved different scenarios or contexts. Also included in this category are datapoints with error rates of human actions or whole events in which multiple CFMs are involved and the data are inseparable. These datapoints cannot be directly used for calculating the base HEPs and PIF weights, but they can be used for reasonableness checks and calibration of the estimated HEPs or PIF weights.

The data sources in IDHEAS-DATA were in the following categories:

A. Operational data and simulator data in the nuclear domain
B. Operational data of human performance from non-nuclear domains
C. Experimental data in the literature
D. Expert judgment of HEPs in the nuclear domain
E. Unspecific-context data (e.g., statistic data, ranking, frequencies of errors or causal factors)

The datapoints for a given CFM and PIF attribute are assessed and organized according to the types and source categories. Table 3-6 shows the 12 datapoints in IDHEAS-DATA IDTABLE-21 for Failure of Detection. The first column has the IDs assigned to the datapoints. The rest of the columns are the same as those in IDTABLE-21: Error rate, task, criteria for lowest HEPs, uncertainties, and source reference. The four criteria for lowest HEPs are: adequate time performing the task, having self-verification (trained as licensed operators), having team-verification (peer-checking and/or close supervision), and not having creditable recovery.

Table 3-6: IDHEAS-DATA IDTABLE-21 Lowest HEPs for Failure of Detection

CFM	Error rate	Task	Criteria for lowest HEPs: TA - Time adequacy SelfV - Self verification TeamV – Team verification Rec - Recovery O - other factors (Y-Yes, N – No, M-Mixed Un-Unknown)	Uncertainty	REF
1	2.1E-3 (4/1872)	NPP operators alarm detection in simulator training. Alarms are self-revealing	TA-Yes, SelfV-Y, TeamV-Y, R-Unknown O – Y (unspecified)	(Other PIFs may exist)	[26]
2	3.4E-3 (3/870)	NPP operators check indicators in simulator training, procedure directed checking.	TA-Yes, SelfV-Yes, TeamV-yes, Rec – Unknown O - Y (unspecified)	(Other PIFs may exist)	[26]
3	5E-4	Military operators read meters, Alphanumeric reading, Detection straight-forward	TA-Y, SelfV-Y, TeamV-No, Rec-No	(Maybe time constraint, 10K+ source data trials)	[109]
4	E-4	Estimated lowest probity of human failure events	TA-Yes, SelfV-Yes, TeamV-yes, Rec - Unknown	(Engineering judgment)	[110]
5	E-4	Simplest possible tasks	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec - Unknown	(Engineering judgment)	[111]
6	E-3	Routine simple tasks	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec – Unknown O – Maybe weak complexity	(Engineering judgment)	[111]
7	5E-3	Line-oriented text editor. Error rate per word	TA-Yes, SelfV-Yes, TeamV-No, Rec - No	No apparent uncertainty	[112]
8	5E-3	Reading a gauge incorrectly. Per read	TA-Yes, SelfV-Yes, TeamV-No, Rec – Unknown O – HSI	No apparent uncertainty	[113]
9	E-3	Interpreting indicator on an indicator lamp. Per interpretation	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec – Unknown	(Engineering judgment)	[109]

			O- complexity in interpreting indicator		
10	9E-4	NPP operator simulator runs	TA – Y, Selv-V – Y TeamV – Y, R – Unknown O – Mixed complexity	No apparent uncertainty	[114, 115]
11	5.3E-4	Gather information and evaluate parameters	TA – Y, Selv-V – Y TeamV – Y, R – Yes	No apparent uncertainty	[116]
12	9E-3	Collision avoidance and target monitoring in simulated ship control, Fixed situation, routine response	TA – Y, Selv-V – Yes TeamV – No, R – Yes O – Dual task, and maybe mixed complexity, mental fatigue, time pressure	Dual task	[27]

The datapoints are organized according to the types and data source categories, as shown in Table 3-7. The rows of Table 3-7 are for data source categories and the columns are for datapoint types. The numbers in the IDTABLE are datapoint identifiers in the first column of Table 3-6.

Table 3-7. The organized identifiers of the datapoints for the lowest HEP of Failure of Detection

	Single-component	Multiple component detachable	Range or trend
A - Nuclear operation		1, 2, 10	
B - Other operation	11	3, 7, 8	
C – Controlled experiment		5, 6, 12	
D – Expert judgment	4	9	
E – Unspecific			

3.2.2. Detaching multi-component human error data

The critical step in the process is detaching multi-component datapoints. The following rules are derived from initial estimates of base HEPs of task complexity and PIF attribute weights. They are used for detaching:

1) If SelfV=NO or TeamV=NO, the detached error rate is the original error rate divided by a factor of 5; If both are NO, the detached error rate is the original error rate divided by a factor of 10.
2) If Recovery = YES, the detached error rate is the original error rate multiplied by a factor range of 2 to 10.
3) If there are other PIFs, the detached error rate is the original error rate divided by multiplication of a factor range of (5 to 10 for complexity) and the sum of the weights of other PIF attributes. The weights of the PIF attributes are from the initiation estimation of the single-component data in IDHEAS-DATA.

Table 3-8 shows the detached error rates. The first column is the datapoint identifier, the second column and third column are the original error rates and lowest HEP criteria, the fourth column is the detached error rate, and the last column contains the notes about the basis of detaching.

Table 3-8: Detached human error rates for the lowest HEP of Failure of Detection

CFM	Error rate	Criteria for lowest HEPs	Detached error rate	Notes
1	2.1E-3 (4/1872)	TA-Yes, SelfV-Y, TeamV-Y, R-Unknown O – Y (unspecified)	2.1E-3 / (5 to 10) = 2.1E-4 to 4E-4	A factor of 5 to 10 represents the combined effect of possible other PIFs
2	3.4E-3 (3/870)	TA-Yes, SelfV-Yes, TeamV-yes, Rec – Unknown O - Y (unspecified)	3.4E-3 / (5 to 10) = 3.4E-4 to 7E-4	A factor of 5 to 10 represents the combined effect of possible other PIFs
3	5E-4	TA-Y, SelfV-Y, TeamV-No, Rec-No	5E-4 / 5 = 1E-4	Divided by 5 for no team verification
4	E-4	TA-Yes, SelfV-Yes, TeamV-yes, Rec - Unknown	E-4	No change
5	E-4	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec - Unknown	E-4	No change
6	E-3	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec – Unknown O – Maybe weak complexity	E-3 / 5 = 2E-4	Divided by 5 for weak complexity
7	5E-3	TA-Yes, SelfV-Yes, TeamV-No, Rec - No	5E-3 / 10 = 2E-4	Divided by (5+5) for lack of self and team verification
8	5E-3	TA-Yes, SelfV-Yes, TeamV-No, Rec – Unknown O – Maybe HSI	5E-3 / (5+2) = 7E-4	Divided by (5+2) for lack of self verification and possible HSI attributes
9	E-3	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec – Unknown	E-3 / 5 = 2E-4	Divided by 5 for no team verification.
10	9E-4	TA – Y, Selv-V – Y TeamV – Y, R – Unknown O – Mixed complexity	9E-4 / (5 t o10) = 9E-5 to 4.8E-4	Divided by (5 to 10) for mixed complexity
11	5.3E-4	TA – Y, Selv-V – Y TeamV – Y, R – Yes O – Mixed complexity	5.3E-4 x 2 / (5-10) = 1.06E-4 to 2.12E-4	Multiplied by 2 for existence of recovery
12	9E-3	TA – Y, Selv-V – Yes TeamV – No, R – Yes O – Dual task, and maybe mixed complexity	9E-3 / (5 to 10) x (5-10) = 9E-5 to 3.6E-4	Divided by (5 to 10) for mixed complexity and divided by (5 to 10) for dual task.

3.2.3. Estimating the lowest HEP

The error rates are organized according to the types and data source categories as shown in Table 3-9.

Table 3-9. Single-component and detached multi-component human error rates for the lowest HEP of Failure of Detection

	Single-component	Multi-component detachable	Bounding
A - Nuclear operation		2.1E-4 to 4E-4, 3.4E-4 to 7E-4, 9E-5 to 4.8E-4	
B - Other operation	1.06E-4 to 2.12E-4	1E-4, 2E-4 7E-4	
C – Controlled experiment		E-4, 2E-4 9E-5 to 3.6E-4	
D – Expert judgment	E-4	2E-4	
E - Unspecific			

Figure 3-1 plots these data points. The vertical axis represents error rates. The datapoints are arranged along the horizontal axis in the order of Category A, B, C, D, E from left to right, and within each category the datapoints are arranged with single-component, detached multi-component, and range or trend. A single error rate value is shown as a filled circle, and the detached ranges of error rates are shown as vertical line segments. The graph shows that the lower end of the data distribution, i.e., the lowest HEP for Failure of Detection, is around 1E-4.

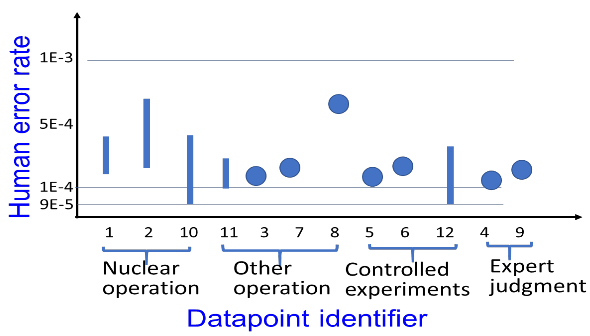

Figure 3-1. The human error rates for the lowest HEP of Failure of Detection

The mean and range of the error rates are calculated for Category A, B, C datapoints separately and for the datapoints in all the three categories. The mean is calculated as the average of the midpoints of the error rate ranges and the single error rate values. The lower bound is calculated as the average of all the lower ends of the error rate ranges, and the upper bound is calculated as the average of all the upper ends of the error rate ranges. The calculated numbers are as follows:

Category A datapoints: [1.8, 3.6, 5.3]E-4 for lower bound, mean, and upper bound;

Category B datapoints: [1.06, 2.8, 2.1]E-4

Category C datapoints: [0.9, 1.7, 3.6]E-4

Category A, B, C datapoints: [1.4, 1.8, 4.4]E-4

Overall, the differences in the mean and range of the error rates of individual categories are less than a factor of 2, thus the error rates from different categories are convergent in the main body and range of their distributions. While the mean value is more representative for the overall datapoints, the lower bound is more appropriate for estimating the lowest HEPs. Based on the data, the value **1E-4** is taken as the lowest HEP for Failure of Detection. This value is lower than the average lower bound 1.4E-4 and slightly larger than the lower bound of 0.9E-5 of two datapoints.

3.2.4. Reasonableness checking and Calibration of the estimated HEP

Category D and E datapoints and range datapoints are used to verify and calibrate the estimated HEP. The two error rates from expert judgment are 1E-4 and 2E-4, thus the estimated lowest HEP of **1E-4** represents those error rates from expert judgment.

Table 3-6 did not include data from "Unspecific" category. The "Unspecific" datapoints in IDHEAS-DATA IDTABLE-21 follow in Table 3-10. There are six unspecific datapoints for lowest HEPs. Two of them are pilot error rates in aviation accidents, two are the rates of air traffic controller (ATC) operational errors, and two are NPP operator error rates in simulator runs for requalification examination. All the reported error rates are from human failure events that may consist of multiple CFMs.

Table 3-10. The "Unspecific" datapoints in IDHEAS-DATA IDTABLE-21 for lowest HEPs of the CFMs.

Unsp	2E-5 (800/4E7)	ATC OE per operation	SelfV – Y TeamV – Y Recov - Y	Recovery is high	[117]
Unsp	2E-4 (290/1.4E6)	ATC OE per shift	SelfV – Y TeamV – Y Recov - Y	Recovery is high	[118]
Unsp	1.47E-2	NPP Requal simulate data – Perform procedures	SelfV – Y TeamV – N Recov - Unknown		[87]
Unsp	7.3E-3	NPP Requal simulate data – Perform procedures	SelfV – Y TeamV – Y Recov - Unknown		[87]
Unsp	3.85E-3	Pilot errors causing accidents	TA – Mixed SelfV – Y TeamV – Y Recov - Mixed		[88]
Unsp	5.5E-6 (686/(1.25×E8))	Pilot error rate x ATC error rate = NTSB reported human error accident rate TABLE 1. The Event Classifications of the 686 Events Reviewed in the NTSB Database from about 1.25×10^{8} Total Flights.	TA – Mixed SelfV – Y TeamV – Y Recov - Y		[119]

Among these datapoints, the first row has the error rate of 2E-5 for air traffic control operational error per operation. This number was obtained including recovery. Using a recovery factor of 5 to 10, the detached error rate would be 2E-4 to 4E-4, and it is larger than the estimated lowest HEP of 1E-4.

The last row of the unspecified datapoints has an error date of 5.5E-6 for the human error rate in aviation accidents from the National Transportation Safety Board (NTSB) Database. Note that the reported pilot errors were actually the combined errors of air traffic controllers and pilots. A rough estimation is that the NTSB reported human error accident rate equals the pilot error rate multiplied by the ATC error rate, without considering the dependency between air traffic

controller and pilot actions. There are two ways to estimate pilot errors. The first one is to equally split the errors between air traffic control and pilots then the pilot error rate would be 1E-3. The second way is using the detached air traffic controller operational error rate of 2E-4 to 4E-4 then the pilot error rate would be 1.3E-2 to 2.7E-2. In either case the error rate is larger than the estimated lowest HEP of 1E-4. Therefore, the estimated lowest HEP of 1E-4 for Failure of Detection is reasonable for the datapoints generalized so far.

The rest of the Unspecific datapoints all have the error rates larger than the estimated lowest HEP of 1E-4. Overall, the reasonableness check verified the estimated lowest HEP of 1E-4 for Failure of Detection.

Using a similar process as described in this section, the lowest HEPs for other CFMs were estimated as 1E-3 for Failure of Understanding, 1E-3 for Failure of Decisionmaking, 1E-4 for Failure of Execution, and 1E-3 for Failure of Interteam Coordination. These were the lowest HEPs used for IDHEAS-ECA [2].

4 DISCUSSION AND CONCLUDING REMARKS

4.1. Generalization of human error data from various sources

IDHEAS-DATA uses the IDHEAS-G [1] framework to organize characteristics of human error data. IDHEAS-DATA is capable of generalizing human error data of various sources to the formats that can be used for HEP quantification.

4.2. Integration of the generalized data to inform IDHEAS-ECA

The NRC staff integrated the human error data in IDHEAS-DATA to infer the base HEPs and PIF weights for IDHEAS-ECA[2]. This integration advances HRA method development in that the calculated HEPs have traceable and updateable data sources. Moreover, the data sources provide HRA analysts the technical basis in their quantitative HRA analyses. The limitation in the current status is that the data integration required different approaches and engineering judgment due to the limited availability of the generalized data and gaps in data sources.

4.3. Limitations in the current status of IDHEAS-DATA

1) Every IDHEAS-DATA TABLE has gaps in data sources.
2) Only a very small sample of data sources was generalized for IDTABLE-23 thorough IDTABLE-27.
3) IDHEAS-DATA is intended to capture available human performance data and empirical evidence to support HRA applications. It is not intended to cover everything in HRA. For this reason, some under-developed areas in HRA, such as error of commission and dynamic HRA, are not included in the current IDHEAS-DATA structure.

4.4. Perspectives of HRA data sources

1) Only a small portion of available nuclear operation and simulation data were generalized. As of 2019, only the effects of several base PIF attributes were analyzed in the SACADA database [26]. The effects of more PIF attributes are being analyzed. Only a few datapoints were generalized from HuREX [34]. The NRC staff is working with HuREX developers to understand the context in HuREX data and the relation between HuREX [120] and the SACADA [25]taxonomy. Moreover, the Halden Reactor Project has conducted NPP simulation experiments over the last three decades. Most of the experimental results are not generalized to IDHEAS-DATA because the studies reported operator task performance indicators other than error rates. However, it is feasible to establish the relation between the performance indicators and error rates based on empirical evidence in the experiments. The NRC staff expects that these efforts would greatly enrich IDHEAS-DATA.
2) The structure of IDHEAS-DATA is generic because it is based on the IDHEAS-G CFMs and PIFs that model human cognition and behavior. IDHEAS-DATA is also flexible because its 27 IDTABLEs operate independently and the datapoints in each IDTABLE can be at different levels of detail. These two features make IDHEAS-DATA a candidate for serving as a hub for HRA data exchange. Different NPP human performance databases can be generalized to IDHEAS-DATA, and the generalized data can be used for different HRA applications.

4.5. Concluding Remarks

1) Data generalization is generic for the IDHEAS CFMs and PIF attributes. Data integration is specific for the HRA method.

2) To close the gaps in existing HRA methods, data generalization should be an on-going, continuous effort. As such, the NRC intends to continue to update its data sources, generalize the information, and integrate the data into its methods.

Appendix A. IDHEAS-DATA Tables

Introduction to Appendix A

Appendix A presents human error data generalized in the 27 IDHEAS-DATA IDTABLEs. Note that the datapoints in the IDTABLEs have not been independently verified for their accuracy and appropriateness. They are being made available to the public in this Research Information Letter only for the purposes of communicating information and demonstrating the data basis of IDHEAS-ECA. It is not recommended that these DRAFT data IDTABLEs be used by HRA practitioners without first verifying the data validity.

Appendix A-1 through Appendix A-3 are for human error data of the three base PIFs. Appendix A-4 through Appendix A-20 are for the rest of the 17 PIFs. Each of these appendices has two sub-tables; the first one presents the PIF attributes and their identifiers; the second one presents the generalized datapoints, with each row usually for one datapoint (except some rows combining several datapoints from the same data source) and each column for a dimension of information. Appendix A-21 through Appendix A-27 present the IDHEAS-DATA IDTABLEs for Lowest HEPs, PIF interaction, Distribution of task completion time, Time factor effects on task completion time, Dependency between human actions, Recovery actions to human failures, and Main drivers to human failures.

The detailed structures of IDHEAS-DATA TABLEs are described in Chapter 2 of this report. Below briefly list the symbols for frequently used terminology that describes the datapoints of the TABLEs.

Column "CFM"

This column is for the cognitive failure modes. The labels D, U, DM, E, and T are for Failure of Detection, Failure of Understanding, Failure of Decisionmaking, Failure of Action Execution, and Failure of Interteam Coordination. The symbols used in this column are the following:

"/" – The symbol "/" separating two CFMs means that the reported error data could be for one of the CFMs or applicable to both CFMs.

"&" – The symbol "&" separating two CFMs means that the reported error data is the sum of the two CFMs.

"," - The symbol "," separating two CFMs means that the datapoint in a row contains error data from the same data source for each CFM in the sub-rows or sub-columns of the "Error rate" column.

"Unsp." – This means that the CFMs of the reported error data were unspecific in the data source. Because much of the error data is event data, it could involve all the CFMs, thus the reported data are unspecific to any CFM.

Column "PIF attribute"

This column is for the PIF attribute applicable to the error data. The labels in the column are the PIF attribute identifiers shown in the Appendix. The symbols used in this column are the following:

“/” – The symbol “/” separating two identifiers means that the reported error data is applicable to both PIF attributes.

“&” – The symbol “&” separating two identifiers means that the reported error data is due to the combined effects of the two PIF attributes.

“,” - The symbol “,” separating two identified means that the datapoint in a row contains error data from the same data source for each PIF attribute in the sub-rows or sub-columns of the “Error rate” column.

“Unsp.” – This means that the PIF attributes applicable to the reported error data were unspecific in the data source. For example, a data source may only report the error rates under “Good” versus “Poor” human-system-interfaces without providing specific information to infer what HSI attributes correspond to “Poor” HSI.

The column “Error rates and Task Performance Indicators”

This column presents the human error data in selected data sources. Unless otherwise specified, the numbers in the column are human error rates. They could be measured as the number of errors made divided by the number of tasks performed, and they could also be from engineering estimates or expert judgment. While most datapoints have error rates in this column, some datapoints only have task performance indicators instead of error rates. The task performance indicators are annotated briefly in this column. Below are some frequently used task performance indicators:

- No. of errors made – The indicator is the total or average number of errors made in the tasks. The data sources did not report the number of times the same task was performed. Some data sources of full scenario simulation only reported the numbers of errors made in the simulation without reporting the number of error opportunities in the scenario.
- Effect size – Effect size is a quantitative measure of the magnitude of a phenomenon in meta-analysis. It quantifies the difference between two groups as the following:

  $$\text{Effect Size} = \frac{[\text{Mean of experimental group}] - [\text{Mean of control group}]}{\text{Standard Deviation}}$$

  If the effect size is calculated for human error rate difference between the presence of a PIF attribute and the control condition, then the positive value means that the error rates with the presence of the PIF attribute is higher than those without the attribute. The higher the effect size, the larger the difference is.
- Correlation coefficient - The coefficient measures the correlation of a PIF attribute and the human error rate or the task performance indicator.
- Frequency (freq.) of occurrence – The frequency of occurrence is typically used in studies of human error analysis or root cause analysis. It calculates the percent of different types of human failure modes or error factors occurring in the analyzed sample of human events, incidents, or accidents.

Column “Task (and error measure)”

This column has a brief description of the task performed. The definition of the error measure is in the parentheses. The default definition is the error rate of incorrect task performance.

Column “PIF measure”

This column has a brief description about the context or experimental manipulation of the context under which the task was performed. The context is represented by the PIF attributes.

Column "Other PIFs (and uncertainty)"

This column annotates other PIFs that were present but not manipulated in the context under which the task was performed. For example, a data source studied the effect of heat by manipulating the work environment temperature, while the tasks were performed in the presence of noise. Thus, the PIF attribute being studied was heat, and noise is annotated as "Other PIFs." This column also annotates uncertainties in the data source as well as uncertainties in representing the data source with the CFM and PIF attributes. The uncertainties are presented in the parentheses. Below are several frequently used annotations in this column:

- "No apparent uncertainty" – This typically applies to well-controlled experiments. There could be uncertainties in the error data that were not described in the data source.
- "Not analyzed" – The data source did not provide detailed information to assess whether other PIFs were present and what the uncertainties were in the data.
- "Meta-analysis" – The datapoints were generalized from meta-analysis of many research papers on the topic.
- "Expert judgment" – The error data were obtained through a formal expert elicitation process.
- "Engineering judgment" – The reported error data were based on experts' analysis of available information and estimates of the HEPs instead of a formal expert elicitation process.

Appendix A1 PIF Attributes and Base HEPs for Scenario Familiarity

Table A1-1 Attribute Identifiers and Descriptions for PIF Scenario Familiarity

ID	PIF Attribute
SF0	**No-impact** • frequently performed tasks in well-trained scenarios, • routine tasks
SF1	**Unpredictable dynamics in known scenarios**
SF1.1	Shifting objectives
SF1.2	Unpredictable dynamics
SF2	**Unfamiliar elements in the scenario**
SF2.1	Non-routine, infrequently performed tasks,
SF2.2	Unlearn a technique and apply one that requires the application of an opposing philosophy
SF2.3	Personnel are unfamiliar with system failure modes.
SF2.4	Personnel are unfamiliar with worksites for manual actions.
SF3	**Scenario is unfamiliar**
SF3.1	Scenarios trained on but infrequently performed
SF3.2	Scenario is unfamiliar, rarely performed, e.g., • Notice adverse indicators that are not part of the task at hand • Notice incorrect status that is not a part of the routine tasks
SF3.3	Scenario is extremely rarely performed, e.g., • Lack of plans, policies and procedures to address the situation • No existing mental model for the situation • Rare events such as the Fukushima accident
SF4	**Bias, preference for wrong strategies, or mismatched mental models**
SF4.1	Wrong expectation or bias
SF4.2	Mismatched mental models
SF4.3	Preference for wrong strategies in decisionmaking

Table A1-2 IDHEAS-DATA IDTABLE-1 – Base HEPs for PIF Scenario Familiarity

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates**	**Task (and error measure)**	**PIF Measure**	**Other PIFs (and Uncertainty)**	**REF**
SF0	D	9E-3	Collision avoidance and target monitoring in simulated ship control	Fixed situation, routine response	Dual task	[27]
SF1.1	D	1.4E-2	Collision avoidance and target monitoring in simulated ship control	Alerting target, normal response	Dual task	[27]
SF1.1	D	1.3E-2	Collision avoidance and target monitoring in simulated ship control	Alerting target, routine response	Dual task	[27]
SF1.1 & SF2.1	D	1.06E-1	Collision avoidance and target monitoring in simulated ship control	Alerting target, emergency response	Dual task, (Time urgent)	[27]
SF2.1	D	6.7E-2	Collision avoidance and target monitoring in simulated ship control	Fixed situation, emergency response	Dual task, (Time urgent)	[27]

SF0	D	2E-4	NPP crews attend to source of information in EOP (estimated HEP)	Good familiarity with the Source	(Expert judgment)	[6]
SF2	D	4E-3	NPP crews attend to source of information in EOP (estimated HEP)	Poor familiarity with the Source	(Expert judgment)	[6]
SF0	D & U	1E-4	Air traffic control (Operational error)	100+min on shift	(with recovery)	[118]
SF1.2	D & U	4.1E-4	Air traffic control (Operational error)	first 30min on shift, unpredictable dynamics	(with recovery)	[118]
SF0	U	7.6E-3 (13/1718)	NPP operators diagnose in simulator training	Standard scenario	(Other PIFs may exist)	[26]
SF2.1	U	8.8E-3 (7/800)	NPP operators diagnose in simulator training	Novel scenario	(Other PIFs may exist)	[26]
SF3.1	U	1.2E-1 (8/69)	NPP operators diagnose in simulator training	Anomaly scenario	(Other PIFs may exist)	[26]
SF0	E	0.04	Go / No-go based on pattern match	Simple “X” for Go and “O” for No-go	No verification	[28]
SF0	D	0.018	Go / No-go based on object recognition	Female vs male faces or one-story vs. two-story houses	No verification	[28]
SF2.2	E	0.177	Diagnosing a pattern; personnel use structured information to guide diagnosis	Rare Stop-trails need to unlearn Go-trials	Task complexity	[28]
SF0	U & DM	3.8E-3	Pilot flight (error rates)	Flight hour > 5000	(Other PIFs may exist)	[88]
SF2.3	U & DM	6E-2	Pilot flight (error rates)	Flight hour < 500	(Other PIFs may exist)	[88]
SF0	DM	5.1E-3 (24/4691)	NPP operators decisionmaking in simulator training	Standard scenario	(Other PIFs may exist)	[26]
SF3.1	DM	1.1E-2 (1/92)	NPP operators decisionmaking in simulator training	Anomaly scenario	(Other PIFs may exist)	[26]
SF0	E	6.8E-4 (1/1470)	NPP maintenance Carrying out a sequence of tasks from memory	Frequently performed	(Ex-CR actions)	[4]
SF3.1	E	2.1E-2 (1/48)	NPP maintenance Carrying out a sequence of tasks from memory	Rarely performed	(Ex-CR actions)	[4]
SF1.1 & SF3.1	E	2.8E-2 (2/70)	NPP maintenance Carrying out a sequence of tasks from memory	Rarely performed	Dynamic environment	[4]
SF3.2	E	1.43E-1 (1/7)	NPP maintenance Carrying out a sequence of tasks from memory	Rarely performed	Dynamic environment	[4]
SF0	E	7.42E-4 (1/1347)	NPP maintenance; Operation of a manual control	Frequently performed task, part of professional knowledge	No apparent uncertainty	[4]
SF3.2	E	7.77E-2 (1/13)	NPP maintenance; Operation of a manual control	Rarely performed test procedure consisting of many sub-steps	high task load, procedure consisting of many sub-steps	[5]
SF0	E	9.78E-4 (3/3067)	Sequence of tasks	Frequently performed,	No apparent uncertainty	[5]
SF3.2	E	2.1E-2 (1/48)	Sequence of tasks	Rarely performed	No apparent uncertainty	[5]

SF3.3	E	3.33E-1 (1/3)	Sequence of tasks	Extremely rarely performed	No apparent uncertainty	[5]
SF0	DM	1.13E-3 (1/888)	Identifying or defining the task	Frequently performed	No apparent uncertainty	[4]
SF3.1	DM	2.33E-2 (4/172)	Identifying or defining the task	Rarely performed	No apparent uncertainty	[4]
SF3.1	DM	1.36E-1 (3/22)	Identifying or defining the task	Rarely performed	Other PIFs	[4]
SF 0	E	9.58E-4 (2/2088)	Procedure execution with professional knowledge (incorrectly remembered professional knowledge)	Part of frequently performed procedure	No apparent uncertainty	[5]
SF3.1	E	1.42E-2 (6/423)	Procedure execution with professional knowledge (incorrectly remembered professional knowledge)	Part of rarely performed procedure	No apparent uncertainty	[5]
SF0	E	9E-4	Maintenance and repair in cable production process	Familiarity with the task in-hand	(data and engineering judgment)	[121]
SF3.2	E	7.64E-2	Maintenance and repair in cable production process	Unfamiliar	(data and engineering judgment)	[121]
SF3.2	D	0.1	Notice adverse indicators when reaching for wrong switch or items	Not part of the task at hand	(Other PIF may exist)	[111]
SF3.2	D	0.1	Roving inspection (Fail to notice incorrect status)	Not part of the routine tasks	(Other PIF may exist)	[111]
SF3.3	DM	0.5	Medicine dispensing	Lack of plans, policies and procedures to address the situation	Inadequate time, Training, procedure	[122]
SF4	D	0.2	Railroad operators start new workshift (fail to check hardware unless specified)	New workshift, task not specified so no mental model for checking	(Other PIF may exist)	[123]
SF0	U	1.6E-3	Situation assessment in EOP (HEP of Critical Data Dismissed/Discounted)	Inappropriate Bias not formed, No Confirmatory Information	(Expert judgment)	[6]
SF4	U	2.5E-1	Situation assessment in EOP (HEP of Critical Data Dismissed/Discounted)	Inappropriate Bias formed, No Confirmatory Information	(Expert judgment)	[6]
SF0	U	3.5E-4	Critical Data Collection (Premature Termination of Critical Data Collection)	Expectations or Biases not formed	(Expert judgment)	[6]
SF4	D / U	8.2E-3	Critical Data Collection (Premature Termination of Critical Data Collection)	Expectations or Biases formed	The failure mode could be either D or U. (Expert judgment)	[6]
SF0	E	2.3E-3	Execution of EOPs (Critical Data Not Checked with Appropriate Frequency)	Good Match with Expectations	(Expert judgment)	[6]
SF4	E	1.3E-2	Execution of EOPs (Critical Data Not Checked with Appropriate Frequency)	Poor Match with Expectations	(Expert judgment)	[6]

Appendix A2 PIF Attributes and Base HEPs for Information Availability and Reliability

Table A2-1 Attribute Identifiers and Descriptions for PIF Information Availability and Reliability

	PIF Attribute
INF0	**No impact – Key information is reliable and complete**
INF1	**Key information is incomplete**
INF1.1	Information is temporarily incomplete or not readily available • Updates of information are inadequate (e.g., information perceived by one party who fails to inform another party). • Feedback information is not available in time to correct a wrong decision or adjust the strategy implementation.
INF1.2	Information of different sources is poorly organized and/or is not specific.
INF1.3	Primary sources of information are not available, while secondary sources of information are not reliable or readily perceived.
INF1.4	Information is moderately incomplete (e.g., a small portion of key information is missing.)
INF1.5	Information is largely incomplete – • Key information is masked, • Key indication is missing.
INF2	**Information is unreliable**
INF2.1	Personnel are aware that source of information could be temporally unreliable.
INF2.2	Overriding information - Pieces of information change over time at different paces; they may not all be current by the time personnel use them together.
INF2.3	Source of information is moderately unreliable, and personnel likely recognize it.
INF2.4	Ambiguity, uncertainty, incoherence, or conflicts in information.
INF2.5	Information is unreliable, and personnel are not aware of it.
INF2.6	Information is misleading or wrong.

Table A2-2 IDHEAS-DATA IDTABLE-2 – Base HEPs for PIF Information Availability and Reliability

1 PIF	2 CFM	3 Error rates	4 Task (and error measure)	5 PIF Measure	6 Other PIFs (and Uncertainty)	7 REF
INF0	U	9.5E-3 (24/2524)	NPP operators diagnose in simulator training	Poor Information Timing does NOT exist	Other PIFs exists	[26]

INF1.1	U	4.5E-2 (4/89)	NPP operators diagnose in simulator training	Poor Information Timing exists	Other PIFs exists, infrequent task	[26]
INF0	U	4E-2	Student controllers performed air traffic control (near miss rate)	Full information displayed	Task complexity and training	[124]
INF1.1	U	8E-2	Student controllers performed air traffic control (near miss rate)	Partially information displayed, full information available upon request	Task complexity and training	[124]
INF0	U & DM	7.9E-2	Pilots in flight deicing (Percent of early buffet, i.e., about to stall)	Accurate information timely with status displays	Inadequate time	[30]
INF1.1	U & DM	20.6E-2	Pilots in flight deicing (Percent of early buffet)	Accurate information not timely without status displays	Inadequate time	[30]
INF0	U & DM	1.8E-1	Pilots in flight deicing (Percent of stall)	Accurate information timely	Complexity, inadequate time	[30]
INF1.1	U & DM	3E-1	Pilots in flight deicing (Percent of stall)	Accurate information not timely	Complexity, inadequate time	[30]
INF0	U	7.7E-3 (10/1293)	NPP operators diagnose in simulator training	Information specificity - specific	Other PIFs exists	[26]
INF1.2	U	1.5E-2 (16/1077)	NPP operators diagnose in simulator training	Information NOT specific	Other PIFs exists	[26]
INF1.2	U	5E-2	Medicine dispensing (Wrong conclusion drawn)	Competing/unclear information	(Distraction	[122]
INF0	DM	5E-3	Maintenance in cable production process (wrong task plan)	Good quality of information		[121]
INF1.2 / INF1.4	DM	4.5E-2	Maintenance in cable production process (wrong task plan)	Poor/impoverished quality information	(Information not organized or missing)	[121]
INF0	DM	3E-2	Licensed driver simulator (%collision)	Fast driving, early real-end information	Time inadequate	[125]
INF1.4	DM	1.1E-1	Licensed driver simulator (%collision)	Fast driving, moderate real-end information	Time inadequate	[125]
INF1.4 / INF1.5	DM	2.2E-1	Licensed driver simulator (%collision)	Slow driving, late real-end information	Time inadequate	[125]
INF0	U	7.7E-3 (20/2582)	NPP operators diagnose in simulator training	No missing information	Other PIFs exist	[26]
INF1.5	U	2.6E-1 (8/31)	NPP operators diagnose in simulator training	Missing information	Other PIFs exist	[26]
INF1.5	U	9/10	NPP crew diagnose SG tube leak and tube rupture in simulator	information of a tube lake was masked in a tube rupture	Licensed crew with peer-checking	[106]
INF1.5	U	4/5	NPP crew diagnose LOCA in simulator	Key information of a small LOCA was masked in a big LOCA	Licensed crew with peer-checking	[16]
INF0	U	0 (9/9)	Physician diagnosis	High-context with all information	(Experiment study)	[126]
INF1.5	U	0.46 (5/9)	Physician diagnosis	Low-context with limited information	(Experiment study)	[126]
INF1.5	DM	3.9E-1	Licensed driver simulator (%collision)	Fast driving late real-end information	Time inadequate	[125]
INF0	DM	1.3E-2	Licensed driver simulator (%collision)	Fast driving early real-end information	Time inadequate	[125]

INF1.5	DM	3.1E-1	Licensed driver simulator (%collision)	Fast driving late real-end information	Time inadequate	[125]
INF1.5	DM	0.01	Students match patterns	No masking	Time inadequate	[127]
INF1.5	DM	0.25	Students match patterns	Visual masking	Time inadequate	[127]
INF1.5	DM	0.33	Students match patterns	Strong visual masking	Time inadequate	[127]
INF0	U	1.6E-3	MCR critical tasks with EOPs (Critical Data Dismissed/Discounted)	Indications Reliable	(Expert judgment)	[6]
INF2.1	U	3.3E-3	MCR critical tasks with EOPs (Critical Data Dismissed/Discounted)	Indications NOT Reliable and no Inappropriate Bias	(Expert judgment)	[6]
INF0	DM	9E-5	Maintenance of the disc brake assembly (decided to omit part of the task)	No over-riding information	(Expert judgment)	[123]
INF2.2	DM	1.1E-2	Maintenance of the disc brake assembly (decided to omit part of the task)	Over-riding information	(Expert judgment)	[123]
INF2.3	U & DM	3.6E-1	Pilots in flight deicing (Percent of stall)	(30%) inaccurate information and pilots were informed the inaccuracy trend	Complexity, inadequate time	[123]
INF2.3	U	1.2E02	MCR critical tasks with EOPs (failed to use alternative source of information)	Primary source of information obviously incorrect	Licensed crew with peer-checking	[6]
INF0	U	9.8E-3 (25/2552)	NPP operators diagnose in simulator training	No misleading information	Other PIFs exists	[26]
INF2.3 / INF2.5 / INF2.6	U	4.9E-2 (3/61)	NPP operators diagnose in simulator training	Misleading information	Other PIFs exists	[26]
INF0	U	8.1E-3 (19/2350)	NPP operators diagnose in simulator training	Ambiguous information does NOT exist	Other PIFs exists	[26]
INF2.4	U	3.4E-2 (9/263)	NPP operators diagnose in simulator training	Ambiguous Information exists	Other PIFs exists	[26]
INF2.4	DM	0 to 0.4 (Sigmoid function)	Students make 2-alternative choices	100% to 10% of information coherence	No apparent uncertainty	[31]
INF2.4	DM	0-0.6 (Sigmoid function)	Students make 4-alternative choices	100% to 10% of information coherence	No apparent uncertainty	[31]
INF2.4	DM	0.3	Pattern matching	70% coherence of information	(experimental study, simple decision)	[128]
INF2.5	U & DM	6.4E-1	Pilots in flight deicing (Percent of stall)	(30%) inaccurate information and pilots did not know of the inaccuracy	Complexity, inadequate time	[123]
INF2.5	U	2.5E-1	MCR critical tasks with EOPs (Critical Data Dismissed/Discounted)	Indications NOT Reliable and Inappropriate Bias formed	Bias exists (Expert judgment)	[6]
INF2.6	U & DM	73.6E-2	Pilots in flight deicing (Percentage of early buffet)	(30%) inaccurate information on status displays	Inadequate time	[30]
INF2.6	U & DM	8.9E-1	Pilots in flight deicing (Percent of stall)	(30%) inaccurate information timely	Complexity, inadequate time	[30]

INF2.6	U & DM	6.4E-1	Pilots in flight deicing (Percent of stall)	(30%) inaccurate information	Complexity, inadequate time	[123]
INF2.6	DM	0.37	Physician decisionmaking for drugs	Information is inaccurate or misleading	(Maybe other PIFs)	[129]
INF2.6	U	3.2E-1	MCR critical tasks with EOPs (failed to use alternative source of information)	Primary source of information NOT obviously Incorrect	(Expert judgment)	[6]

Appendix A3 PIF Attributes and Base HEPs for Task Complexity

Table A3-1 Attribute Identifiers and Descriptions for PIF Task Complexity

	PIF Attribute
	C0-C7 are for Detection complexity
C0	No impact on Detection HEP, simple straightforward attending to alarms, monitoring, or checking information, directed by procedures, routinely, or well-known knowledge
C1	Detection overload with multiple competing signals • Track the states of multiple systems • Monitor many parameters • Memorize many pieces of information detected • Detect many types or categories of information
C2	Detection is moderately complex • Criteria are not straightforward, • Information of interest involves complicated mental computation • Comparing for abnormality
C3	Detection demands for high attention • Need split attention • Need sustained attention over a period of time • Need intermittent attention For example, determining a parameter trend during unstable system status or monitoring a slow-response-system behavior without a clear time window to conclude that monitoring requires attention for a prolonged period.
C4	Detection criteria are highly complex • Multiple criteria to be met in complex logic • Information of interest must be determined based on other pieces of information • Detection criteria are ambiguous and need subjective judgment
C5	Cues for detection are not obvious – e.g., detection is not directly cued by alarms or instructions and personnel need to actively search for the information
C6	Weak or no cue or mental model for detection
C6.1	Cue or mental model for detection is ambiguous or weak • Time gap between the cue for initiating detection to the time detection is performed • Incoherent, uncertain, or inconsistent cues for initiating the detection
C6.2	No rules / procedures / alarms to cue the detection; Detection of the critical information is entirely based on personnel's experience and knowledge
C7	Low signal probability for detection
	C10-C16 are for Understanding complexity
C10	No impact – straightforward diagnosis with clear procedures or rules
C11	Working memory overload • Need to decipher numerous messages (indications, alarms, spoken messages) • Multiple causes for situation assessment: Multiple independent 'influences' affect the system and system behavior cannot be explained by a single influence alone
C12	Relational complexity (Number of unchunkable topics or relations in one understanding task) • Relations involved in a human action are very complicated for understanding • Need to integrate (use together) multiple relations
C13	Understanding complexity - Requiring high level of comprehension • Multiple causes for situation assessment: Multiple influences affect the system, and system behavior cannot be explained by a single influence
C14	Potential outcome of situation assessment consists of multiple states and contexts (not a simple yes or no)
C15	Ambiguity associated with assessing the situation • Key information for understanding has hidden coupling • Pieces of key information are intermingled or with complex logic • The source of a problem is difficult to diagnose because of cascading secondary effects that make it difficult to connect the observed symptoms to the originating source
C16	Conflicting cues or symptoms
	C20-C29 are for Decisionmaking complexity
C20	No impact – simple, straightforward choice

	PIF Attribute
C21	Transfer step in procedure – integrating a few cues
C22	Transfer procedure (Multiple alternative strategies to choose) – integrating multiple cues
C23	Decision criteria are intermingled, ambiguous, or difficult to assess
C24	Multiple goals difficult to prioritize, e.g., advantage for incorrect strategies
C25	Conflicting goals (e.g., choosing one goal will block achieving another goal, low preference for correct strategy, reluctance & viable alternatives)
C26	Decision-making involves developing strategies or action plans
C27	Decisionmaking requires diverse expertise distributed among multiple individuals or parties who may not share the same information or have the same understanding of the situation
C28	Integrating a large variety of types of cues with complex logic
	C30-39 are for Execution complexity
C30	No impact - Simple execution with a few steps
C31	Straightforward procedure execution with many steps
C32	Non-straightforward procedure execution • Out-of-sequence steps • Very long procedures, voluminous documents with checkoff provision • Multiple procedures needed, action sequences are parallel and intermingled
C33	Simple continuous control that requires monitoring parameters and adjusting action accordingly
C34	Continuous control that requires manipulating dynamically and sustained attention
C35	Long-lasting action, repeated discontinuous manual control (need to monitor parameters from time to time)
C36	No immediacy to initiate execution - Time span between annunciation (decision for execution made) and operation
C37	Complicated or ambiguous execution criteria • Multiple, coupled criteria • Restrictive, irreversible order of multiple steps • Open to misinterpret
C38	Action execution requires close coordination of multiple personnel at different locations – Transport fuel assemblies with fuel machines
C39	Unlearn or break away from automaticity of trained action scripts
	C40-C44 are for Interteam Coordination complexity
C40	No impact – Clear, streamlined, crew-like communication and coordination
C41	Information to be communicated is complex
C42	Complex or ambiguous command-and-control
C43	Complex or ambiguous authorization chain
C44	Coordinate activities of multiple diverse teams or organizations

Table A3-2 IDHEAS-DATA IDTABLE-3 – Base HEPs for PIF Task Complexity

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates**	**Task (and error measure)**	**PIF Measure**	**Other PIFs (and Uncertainty)**	**REF**
C1	D	2.1E-3 (2/953)	NPP operators alarm detection in simulator training	Alarm board dark	(Other PIFs may exist)	[26]
C1	D	5.0E-3 (5/991)	NPP operators alarm detection in simulator training	Alarm board busy	(Other PIFs may exist)	[26]
C1	D	3.9E-2 (6/155)	NPP operators alarm detection in simulator training	Alarm board overloaded	(Other PIFs may exist)	[26]
C1	D	2.8E-3 (2/711)	Indicator checking	No Concurrent demands	Possible multitasking	[26]
C1	D	7.8E-3 (10/1289)	Indicator checking	Concurrent demands	Possible multitasking	[26]
C1	D	2.5E-2 (5/198)	Indicator checking	Multiple concurrent demands	Possible multitasking	[26]

C1	D	3E-3	Detecting signals in nuclear facility operation	Few competing signals	(expert elicitation)	[37]
C1	D	1E-2	Detecting signals in nuclear facility operation	Several competing signals	(expert elicitation)	[37]
C1	D	1E-1	Detecting signals in nuclear facility operation	Many competing signals	(expert elicitation)	[37]
C1	*D*	0.0001 to 0.05	Respond to compelling signals	The number of annunciators from 1 to 10	Not analyzed	[35, 36]
C1	D	0.10 to 0.20	Respond to compelling signals	The number of annunciators 11 to 40.	Not analyzed	[35, 36]
C1	D	0.25	Respond to compelling signals	Annunciators >40	Not analyzed	[35, 36]
C1	*D*	2E-3	Reading meters	One meter	Not analyzed	[35, 130, 131]
C1	*D*	1.3E-2	Reading meters	Multiple meters	Not analyzed	[35, 130, 131]
C1	*D*	0.24	Students acquires information from air traffic control timelines	3-9 categories of information to be detected	High time constraint and dual task	[132]
C1	*D*	0.2	Students detects abnormal signals (omitted signals)	3-6 categories of information to be detected	High time constraint and dual task	[132]
C1	*D*	0.3	Students detects abnormal signals (omitted signals)	9 categories of information to be detected	High time constraint and dual task	[132]
C1	*D*	L 0.14 M 0.24 H 0.29	Highly experienced drivers simulate driving (miss rate on peripheral detection task)	Driving environment: L-low complexity M-medium complexity H-high complexity	Time constraint, dual-task	[10]
C1	*D*	0.01	Drivers recognize names while simulating driving	Few names	Time constraint, Dual task	[133]
C2	D	0.05	Military professionals read meters	Analog meter reading with limit marks	(Maybe time constraint)	[134-137]
C2	D	8.4E-4	NPP crews perform EOPs on simulator (failure of verifying information)	Synthetically verifying information	No apparent uncertainty	[138]
C2	D	2.2E-3	NPP operators perform EOPs on simulator	Comparing for abnormality	No apparent uncertainty	[138]
C3	D	3.14E-3	NPP operators perform EOPs on simulator	Detection requires sustained attention	No apparent uncertainty	[138]
C0	D	5E-4	Military operators read meters	Alphanumeric reading, Detection straightforward	(Maybe time constraint)	[134-137]
C4	D	0.1	Military operators read meters	Analog meter reading without limit marks	(Maybe time constraint)	[134-137]
C4	D	0.2	Military operators check information	Geometric symbols - Detection criteria need interpretation	(Maybe time constraint)	[134-137]
C5	D	6.4E-3(5/782)	NPP operators check indicators in simulator training	Not procedure directed, awareness/ inspection needed	(Other PIFs may exist)	[26]
C0	D	2.1E-3 (4/1872)	NPP operators alarm detection in simulator training	Alarms self-revealing	(Other PIFs may exist)	[26]
C6.1	D	5.1E-2(9/177)	NPP operators alarm detection in simulator training	Alarms not self-revealing and need operators' awareness/inspection	(Other PIFs may exist)	[26]
C7	D	4.2E-2	Students detect signals	Signal probability =0.1	No apparent uncertainty	[139]

C7	D	3.7E-2	Students detect signals	Signal probability =0.35	No apparent uncertainty	[139]
C11	U	M Error rate 5 0.036 8 0.05 11 0.11 15 0.23 17 0.32 >20 0.5	Pilot read-back communication ((error rate of incorrect read-back messages) 5 - 3.6% 8 - 5% 11 - 11% 15 - 23% 17 - 32% 20> -50%	M= Message complexity (# of messages and relations)	(other PIFs may exist)	[8]
C11	U	0.04	Navy controllers perform ATC simulation (near miss separation)	Low task load	Poor training	[124]
C11	U	0.09	Navy controllers perform ATC simulation (near miss separation)	High task load	Poor training	[124]
C11	U	0.05	Students test on relational working memory	4 simultaneously presented items	No apparent uncertainty	[124]
C12	U	R Error rate 1 0.038 2 0.061 3 0.085 4 0.26	Pilot read-back communication (error rate of messages incorrectly communicated)	R= Message relation (# of aviation topics in one communication)	(other PIFs may exist)	[8]
C12	U	0.3	Students test on relational working memory	4 sequentially presented items	No apparent uncertainty	[140]
C13	U	0.028	Understand requirements (Misinterpret NPP procedure)	Procedure complexity	No apparent uncertainty	[141]
C13	U	0.03	Pharmacists dispense medicine	Typical understanding	(Other PIFs)	[122]
C13	U	0.15	Pharmacists dispense medicine	Requiring high level of comprehension	(Other PIFs)	[122]
C13	U	0.035	Interpret cues in flight simulator	Domain (location) cues require little comprehension	(Other PIFs)	[142]
C13	U	0.136	Interpret cues in flight simulator	Importance cues require comprehending info	(Other PIFs)	[142]
C13	U	0.169	Interpret cues in flight simulator	Importance cues require comprehending and matching info	(Other PIFs)	[142]
C14	U	1/17	NPP maintenance - Orally give work permit (Incorrect plant state interpretation)	Interpretation of plant state consists of multiple states and context (not a simple yes or no)	Rarely performed tasks	[4]
C12 & C15	U	1.8E-2 ~ 3E-1	Diagnosis that needs to decipher numerous indications and alarms, and the ambiguity associated with assessing the situation	Difficulty as the level of the ambiguity associated with assessing the situation	Stress and team dynamics	[143]
C12 & C15	U	3E-3 ~ 1.8E-1	Diagnosis that needs to decipher numerous indications and alarms, and the ambiguity	Easy to somewhat difficult	Stress and team dynamics	[143]

			associated with assessing the situation			
C12 & C15	U	5E-2 to 1E-4	Diagnosis that needs to decipher numerous indications and alarms, and the ambiguity associated with assessing the situation	Very easy	stress and team dynamics	[143]
C15	U	0.27	Simulated process control (Prospective memory failures)	System failure was accompanied by a simultaneous disabling of the relevant control panel	Multitasking	[144]
C15	U	9.5E-1. (4 /4)	NPP events	Alarms signal may be triggered by maintenance work and difficult to identify initiation criteria	No apparent uncertainty	[4]
C21	DM	0.08	Go/no-go switching task in flight simulator (incorrectly choosing no-switching)	Integrating two cues	(Other PIFs)	[142]
C21	DM	0.08	Go/no-go switching task in flight simulator (incorrectly choosing not to switch)	Integrating two cues	(Other PIFs)	[142]
C21	DM	4.5E-3	NPP operators perform EOPs on simulator	Transfer step in procedure	(Other PIFs)	[116]
C22	DM	1.23E-2	NPP operators perform EOPs on simulator	Transfer procedures	(Other PIFs)	[116]
C22	DM	9.3E-3	Choose wrong strategy	Alternative strategies to choose	(Expert judgment)	[6]
C23	DM	3.4E-3	Delayed implementation (incorrect Assessment of Margin)	Decision criteria are ambiguous	(Expert judgment)	[6]
C24	DM	3.3E-2	Choose wrong strategy	Advantage in using the incorrect strategy	(Expert judgment)	[6]
C25	DM	1.4E-1	Choose wrong strategy	Low preference for correct strategy	(Expert judgment)	[6]
C25	DM	1.7E-1	Choose wrong strategy	Competing strategies, reluctance & viable alternatives exist	(Expert judgment)	[6]
C28	DM	0.274	Students perform DM tasks	Integration of simple spatial cues	No apparent uncertainty	[145]
C28	DM	0.451	Students perform DM tasks	Integration of complex spatial cues	No apparent uncertainty	[145]
C30		E-3	NPP maintenance	Simple execution (operating a pushbutton, adjust values, connect a cable)	No apparent uncertainty	[4]
C30	E	1E-4	Nuclear facility operation - Execution procedure or script	Nominal (simple) lock out plan (1-4 lock out)	(Estimated HEP)	[37]
C31		5E-4	Nuclear facility operation - Execution procedure or script	Moderate (typical) lock out plan (4-10 lockout)	(Estimated HEP)	[37]
C31		5E-3	Nuclear facility operation - Execution procedure or script	Complex lock-out plan (11-100 lockout)	(Estimated HEP)	[37]
C31		3.3E-3 (2/651)	NPP maintenance (omitting an item of instruction)	Procedure execution with many steps	(Other PIFs may exist)	[4]
C31	E	1E-2	NPP operators execute actions on simulator	Simple and distinct	(Other PIFs may exist)	[26]
C32	E	3.4E-2	NPP operators execute actions on simulator	Additional mental effort required	(Other PIFs may exist)	[26]

C32	E	3.8E-3	NPP crew performs EOPs	Execution is not straightforward	(expert judgment)	[6]
C32		4.7E-3 (1/211)	NPP maintenance tasks	Long procedures, voluminous documents with checkoff provision	Not analyzed	[5]
C33		3.4E-4	Controlled actions that require monitoring action outcomes	Simple continuous control	Not analyzed	[4]
C33		2.6E-3	Controlled actions that require monitoring action outcomes and adjusting action accordingly	Manipulating dynamically	Not analyzed	[4]
C33		0.0015 ~ 0.0086	Operating controls while monitoring dynamic displays	Discrete controls	Not analyzed	[35, 131, 137, 146, 147]
C34		0.0029 ~ 0.0356	Operating controls while monitoring dynamic displays	Continuous controls	Not analyzed	[35, 131, 137, 146, 147]
C35		0.02 (1/50)	Maintenance	Repeated discontinuous manual control – demand for working memory	Not analyzed	[4]
C36		0.3E-3 (2 /608)	NPP maintenance (operated too late)	Short time span between annunciation and operation	Not analyzed	[4]
C36		8.2E-3	NPP crews execute procedures	No immediacy to initiate execution	(expert judgment)	[6]
C37		0.036 (1/28)	NPP maintenance	Complex execution criteria - Fast response/ correction in case of deviation	With moderately high level of stress	[4]
C37		0.028 (1/36)	NPP maintenance	Ambiguous execution criteria - High similarity between right or wrong position	No procedure (this may be an HSI attribute)	[4]
C37		0.012 (1/84)	NPP maintenance	Ambiguous execution criteria	No procedure (this may be an HSI attribute)	[4]
C37		0.024 (1/40)	NPP maintenance tasks	Ambiguous task description in procedures	Not analyzed	[4]
C38		0.14 (1/7)	Transport fuel assemblies with fuel machines	Action execution requires close coordination of multiple personnel at different locations	Time pressure, visualization aid not available, unfavorable ergonomic design	[4]
C39		0.5 (2/4)	NPP maintenance tasks	Unlearn or break away from automaticity of trained action scripts – Switch off automatic fuses	(Similar fuses within reach, unfavorable labeling design and working document design)	[4]

C31-C39	D / U/E	3E-3 ~ 1.8E-1	NPP crews perform SGTR events: HFE-2A, cool down the reactor coolant system	Base case	(Estimated HEP bounds)	[22]
C31-C39	D / U/E	9E-5 ~5E-2	NPP crews perform SGTR events: HFE-2B, cool down the reactor coolant system	Complex case	(Estimated HEP bounds)	[22]
C31-C39	D / U/E	3E-3 ~ 1.8E-1	NPP crews perform SGTR events: HFE-3A, depressurize the reactor coolant system	Base case	(Estimated HEP bounds)	[22]
C31-C39	D / U/E	1.8E-2 ~ 3E-1	NPP crews perform SGTR events: HFE-3B, depressurize the reactor coolant system	Complex case	(Estimated HEP bounds)	[22]
C31-C39	E	9E-5 ~ 5E-2	NPP crews perform SGTR events: HFE-4A, terminate safety injection	Base case	(Estimated HEP bounds)	[22]
C41	T	1E-3	Nuclear facility operation Communication	Simple information	(Estimated HEP)	[37]
C41	T	5E-2	Nuclear facility operation Communication	Moderate complex	(Estimated HEP)	[37]
C41	T	5E-1	Nuclear facility operation Communication	Extremely high complex	(Estimated HEP)	[37]
C41	T	1.54E-3	Notifying/requesting to ex-MCR	Ex-CR communication		[116]

Appendix A4 PIF Attributes and Weights for Workplace Accessibility and Habitability

Table A4-1 Attribute Identifiers and Descriptions for PIF Workplace Accessibility and Habitability

ID	Attribute
WAH1	Accessibility (travel paths, security barriers, and sustained habituation of worksite) is limited because of physical threats to life in the environment (e.g., Traffic or weather impeding vehicle movement)
WAH2	Habitability is reduced; Personnel cannot stay long at the worksite or they experience degraded conditions for work, • challenges to living conditions (e.g., isolation, confinement, microgravity) • environmental hazards like radiation or earthquake aftershocks
WAH3	The worksite is flooded or underwater
WAH4	The surface of systems, structures, or objects to be worked on cannot be reached or touched (e.g., because the surface is too hot to touch or the object is too high to reach).

Table A4-2 IDHEAS-DATA IDTABLE-4 – PIF Weights for Workplace Accessibility and Habitability

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates or task performance indicators**	**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncer-tainty)**	**REF**
WAH 1	E	Accidents increase 75% in adverse weather	Drive	All weather included	(Statistical data)	[148]
WAH 1	E	Heavy rain: 4-6% increase in travel time Heavy snow: 7.4 – 11% increase in travel time	Macroscopic travel times in UK the Greater London area (% increase in travel time)	Light, moderate, heavy rain and snow on travel time	(Statistical data)	[149]
WAH 1	E	Light rain: 9% (4-10%) increase in travel time Heavy rain: 20% (8-30%) increase in travel time	Drive in rain (% increase in travel time)	Rain intensity: Light rain – 0.25-6.4mm/h Heavy rain > 6.4mm/h	(From multiple studies)	[57]
WAH 1	E	Depth of floodwater / Small car / 4WD vehicle 100mm / 10m/h / 50m/h 150mm / 0 / 40m/h 300mm / 0 / 10m/h	Driving – small cars and 4WD cars (speed m/h)	Car speed with varying depths of floodwater compared to at 85m/h without flood	(From many data sources)	[57]
WAH 1	E	Precipitation is associated with a 75% increase in traffic collisions and a 45% increase in related injuries, as compared to `normal 'seasonal conditions	Travel risk in mid-sized Canadian cities	Risk levels vary depending on the characteristics of the weather event	(Statistic data)	[150]
WAH 1	E	Nominal – No congestion no weather: 1E-7/mi Moderate - Typical highway environment: 1E-6/mi	Vehicle collision/ accident (probability of collision or accident per mile)	Highway congestion and weather	(Expert judgment)	[37]

		congestion, many objects close to road, and bad weather 1E-5/mi				
WAH 2	D	Continuously decrease with radiation doses	Perception, attention tests	Varying radiation doses	(Not analyzed)	[42, 43, 46, 151]
WAH 2	U	Continuously decrease with radiation doses	Memory and reasoning tests	Varying radiation doses	(Not analyzed)	[42, 43, 46, 151]
WAH 2	D M	Continuously decrease with radiation doses	Judgment and decisionmaking, Space Shuttle operation	Varying radiation doses	(Not analyzed)	[42, 43, 46, 151]
WAH 2	E	Continuously decrease with radiation doses	Visual-motor tasks, tracking, spatial transformation, Space Shuttle operation	Varying radiation doses	(Not analyzed)	[42, 43, 46, 151]
WAH 2	D	Continuously increase perception time, difficulty in concentrating or focusing attention, and divided attention	Cognitive tests and Space Shuttle operation	Novel environments (spaceflight or other), confinement, CO2 level increases	(Not analyzed)	[44]
WAH 2	U	Deficits only with personally relevant stimuli	Reasoning tests	Novel environments (spaceflight or other), confinement, CO2 level increases	(Not analyzed)	[44]
WAH 2	D M	No observed changes	Cabin Air Management System (a simulation task); problem-solving test, Iowa Gambling	Novel environments (spaceflight or other), confinement, CO2 level increases	(Not analyzed)	[44]
WAH 2	E	Motor slowing, and increased motor variability	Visual-motor tests and space shuttle operation	Novel environments (spaceflight or other), confinement, CO2 level increases	(Not analyzed)	[44]
WAH 2	T	Degradations in social functioning, e.g., social integration, team cohesion	Process social cues or social decision making	Novel environments (spaceflight or other), confinement, CO2 level increases	(Not analyzed)	[44]
WAH 2	E	Statistically significant deterioration of intellectual efficiency as isolation time increased	Various cognitive tests	Astronauts in space station or in simulated lab	(Not analyzed)	[152]
WAH 2	D	Lack of detectable impairment over time	Various cognitive tests	Over-wintering crew over 6-month winter in Antarctic station - Hypoxia, isolation, confinement	(Not analyzed)	[45]
WAH 2	D/ U/ E	Attention function impairments and reduced P3a wave	Attention tests	Astronauts in Space Station - isolation, confinement	(Not analyzed)	[45]
WAH 2	E	Impairments in psychomotor function, arithmetical skills, working memory, and multitasking	Various cognitive tests	Hypoxia, isolation, confinement	(Not analyzed)	[45]
WAH 3	E	9% flood-death reaching a destination	Go into flood to reach a destination or rescue (Poor risk	Flood	(Statistic al data)	[153]

			perception, underestimated the risk)			

Appendix A5 PIF Attributes and Weights for Workplace Visibility

Table A5-1 Attribute Identifiers and Descriptions for PIF Workplace Visibility

ID	Attribute
VIS1	Low ambient light or luminance of the object that must be detected or recognized
VIS2	Glare or strong reflection of the object to be detected or recognized
VIS3	Low visibility of work environment (e.g., those caused by smoke, rain, and fog)

Table A5-2 IDHEAS-DATA IDTABLE-5 – PIF Weights for Workplace Visibility

1	2	3		4	5	6	7
PIF	**CFM**	**Error rates or task performance indicators**		**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
VIS1	D	Luminance	Reading error	Dial reading error	Luminance (L/m^2)	Not analyzed	[154]
		0.15	0.16				
		1.5	0.1				
		>15	0.08				
VIS1	D	Contrast	Error rate	Visual discrimination	Contrast (%) of the target to be discriminated	No apparent uncertainty	[155]
		5%	0.1				
		6.9%	0.034				
VIS1	D	Good Visibility	6E-4	Read meters	Visibility	(Uncertainty in good vs. poor visibility)	[35, 130, 131]
		Poor Visibility	3.5E-3				
VIS1	D	Good Visibility	5E-4	Read computer display	Visibility	(Uncertainty in good vs. poor visibility)	[134 - 137, 156]
		Poor Visibility	2.4E-3				
VIS1	E	Good Visibility	2.8E-3	Operate continuous control while monitoring dynamic display	Visibility	Dual task (Uncertainty in good vs. poor visibility)	[131, 137, 146]
		Poor Visibility	3.5E-2				
VIS1	E	Good Visibility	3.6E-3	Adjust control while tracking a dynamic target signal	Visibility	Dual task (Uncertainty in good vs. poor visibility)	[131, 137, 146]
		Poor Visibility	3.5E-2				
VIS2	D	No glare	5.3%	Reading from computer LCD (reading errors)	Ambient light to LCD	Subjects adjusted chair to mitigate glare	[157]
		Glare	4.6%				
VIS2	D	No glare	0.1	Reading from paper strip mimicking inspection tasks (reading errors)	Glare source angles	Subjects adjusted positions to mitigate glare	[158]
		Glare 15^o	0.09				
		Glare 40^o	0.126				
VIS2	D	Just imperceptible	1.17E-2	Text reading and categorizing from computer display	Glare source luminance and subjective	(Subjective definition of glare levels)	[159]
		Just acceptable	1.69 E-2				
		Just uncomfortable	1.82 E-2				

<table>
<tr><td></td><td></td><td colspan="2">Just intolerable</td><td>4.3 E-2</td><td></td><td>evaluation of glare</td><td></td><td></td></tr>
<tr><td rowspan="4">VIS2</td><td rowspan="4">E</td><td></td><td>Pitch control error (degree)</td><td>Roll control error (degree)</td><td rowspan="4">Visual flight task on a simulator (control errors)</td><td rowspan="4">No laser (N) Strobing (S) vs. continuous (C) laser exposure</td><td rowspan="4">Small sample size</td><td rowspan="4">[48]</td></tr>
<tr><td>No Laser</td><td>2</td><td>5</td></tr>
<tr><td>C</td><td>4</td><td>9</td></tr>
<tr><td>S</td><td>10</td><td>20</td></tr>
<tr><td rowspan="4">VIS3</td><td rowspan="4">D & E</td><td>Fog level</td><td>Distance headway</td><td>Velocity error</td><td rowspan="4">Simulated driving (mean distance headway and velocity error)</td><td rowspan="4">Fog level (luminance contrast)</td><td rowspan="4">No apparent uncertainty</td><td rowspan="4">[160]</td></tr>
<tr><td>0.05</td><td>19.5</td><td>5.5</td></tr>
<tr><td>0.1</td><td>19.6</td><td>6</td></tr>
<tr><td>0.2</td><td>17</td><td>7</td></tr>
<tr><td rowspan="3">VIS3</td><td rowspan="3">D & E</td><td>Fog level</td><td>Lane deviation</td><td>Velocity deviation</td><td rowspan="3">Simulated driving (Lane deviation and velocity error)</td><td rowspan="3">Fog level</td><td rowspan="3">No apparent uncertainty</td><td rowspan="3">[161]</td></tr>
<tr><td>Low</td><td>0.5</td><td>2.3</td></tr>
<tr><td>High</td><td>0.55</td><td>2.6</td></tr>
<tr><td rowspan="2">VIS3</td><td rowspan="2">E</td><td colspan="2">Low Visibility</td><td>5 errors</td><td rowspan="2">Using mono and stereo TV to position a manipulator (# of errors)</td><td rowspan="2">Environmental visibility (V) in undersea vehicles</td><td rowspan="2">No apparent uncertainty</td><td rowspan="2">[162]</td></tr>
<tr><td colspan="2">High Visibility</td><td>12 errors</td></tr>
<tr><td rowspan="3">VIS3</td><td rowspan="3">E</td><td colspan="2">Spotter present</td><td>3E-5</td><td rowspan="3">Crane/hoist strikes stationary object</td><td rowspan="3">Spotter and visibility(V)</td><td rowspan="3">(Expert judgment)</td><td rowspan="3">[37]</td></tr>
<tr><td colspan="2">No spotter, typical Visibility</td><td>3E-4</td></tr>
<tr><td colspan="2">No spotter, low Visibility</td><td>3E-3</td></tr>
</table>

Appendix A6 PIF Attributes and Weights for Workplace Noise

Table A6-1 Attribute Identifiers and Descriptions for PIF Workplace Noise

ID	Attribute
NOS1	Continuous loud mixture of noisy sounds
NOS2	Intermittent non-speech noise
NOS3	Speech noise
NOS4	Intermittent mixture of speech/noise

Table A6-2 IDHEAS-DATA IDTABLE-6 – PIF Weights for Workplace Noise

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates or task performance indicators**	**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
NOS1	Unsp	-0.26 (Effect size)	Unspecified	Continuous noise	(statistic)	[50]
NOS2	Unsp	-0.39 (Effect size)	Unspecified	Intermittent noise	(statistic)	[50]
NOS3	Unsp	-0.84 (Effect size)	Unspecified	Speech	(statistic)	[50]
NOS4	Unsp	-0.46 (Effect size)	Unspecified	Mixture of speech and noise		[50]
NOS3	D	-0.06 (Effect size)	Visual Detection	Speech	(statistic)	[50]
NOS3	D	-0.74 (Effect size)	Aural Detection	Speech	(statistic)	[50]
NOS3	U or DM	-0.84 (Effect size)	Cognitive tasks	Speech	(statistic)	[50]
NOS1 / NOS2	D	-0.2 (Effect size)	Perceptual	Nonspeech	(statistic)	[50]
NOS1 / NOS2	U/ DM	-0.21 (Effect size)	Cognitive	Nonspeech	(statistic)	[50]
NOS1 / NOS2	E	-0.49 (Effect size)	Motor	Nonspeech	(statistic)	[50]
NOS1 / NOS2	T	-0.43 (Effect size)	Communication	Nonspeech	(statistic)	[50]
NOS1, NOS2, NOS3	D	Quiet 11.03 Noise 9.41	View word lists and recall them (# of correct recalls) Stroop task requiring attention	55-dB(A) background noise or white noise amplified through wall speakers to 95 dB(A) NOS1 – 50-70dB traffic noise NOS2 – 60dB intermittent traffic NOS3 – irrelevant speech	No apparent uncertainty (Attention is for all CFMs)	[163]
NOS1, NOS2,	All*	NOS1 0.032 – 0.048		NOS1 – 50-70dB traffic noise	(Attention is for all CFMs)	[49]

NOS3		NOS2 NOS3	0.038 0.034	Stroop task requiring attention verbal serial recall that requires working memory	NOS2 – 60dB intermittent traffic NOS3 – irrelevant speech	(Working memory is for all CFMs)	
NOS1, NOS2, NOS3	All	Silence NOS1 NOS2 NOS3	0.27 0.18-0.227 0.24 0.314	Verbal serial recall that requires working memory Mental arithmetic performance	NOS1 – 50 to 70dB continuous traffic noise NOS2 – 60dB intermittent traffic NOS3 – irrelevant speech	(Working memory is for all CFMs) (The task is for all CFMs)	[49]
NOS1, NOS2, NOS3	All	Silence NOS1 NOS2 NOS3	0.27 0.3 0.3 0.40	Mental arithmetic performance Five-choice control task	NOS1 – 50 to 70dB traffic noise NOS2 – 60dB intermittent traffic NOS3 – irrelevant speech Low frequency continuous noise	(The task is for all CFMs) (low frequency noise improves vigilance)	[49]
NOS0 NOS0	E, D	Control Noise	0.021 0.014	Five-choice control task Detect signals in vigilance task	Low frequency continuous noise Low frequency continuous noise	(Low frequency noise improves vigilance) (low frequency noise improves vigilance)	[164]
NOS0 NOS2	D D, U, DM, E, T	Control Noise	0.43 0.33	Detect signals in vigilance task Arithmetic - calculate the answer	Low frequency continuous noise Noise bursts	(Low frequency noise improves vigilance) (Arithmetic calculation can be in all macrocognitive functions)	[164]
NOS2 NOS2	D, U, DM, E, T D	No noise Noise	0.18 0.32	Arithmetic - calculate the answer Read a number	Noise bursts Noise bursts	(Arithmetic calculation can be in all macrocognitive functions) No apparent uncertainty	[165, 166]
NOS2 NOS1	D E	No noise Noise	0.27 0.25	Read a number 5-choice control task (# of errors)	Noise bursts 95dB continuous noise	No apparent uncertainty No apparent uncertainty	[165, 166]
NOS1 NOS2	E D	No noise Noise	7 10.5	5-choice control task (# of errors) Perception	95dB continuous noise Noise burst	No apparent uncertainty Not analyzed	[165, 166]
NOS2 NOS2	D U	No noise Noise	0.2 0.34	Perception N-back working memory test	Noise burst Noise burst	Not analyzed Not analyzed	[167-171]
NOS2	U	No noise Noise	0.36 0.38	N-back working memory test	Noise burst	Not analyzed	[167-171]

All* - The generic task, such as mental arithmetic performance, can be involved in every macrocognitive function.

Appendix A7 PIF Attributes and Weights for Cold/Heat/Humidity

Table A7-1 Attribute Identifiers and Descriptions for PIF Cold/Heat/Humidity

ID	Attribute
TEP1	Cold in workplace
TEP2	Heat in workplace
TEP3	High humidity in workplace

Table A7-2 IDHEAS-DATA IDTABLE-7 – PIF Weights for Cold/Heat/Humidity

1 **PIF**	2 **CFM**	3 **Error rates**	4 **Task (and error measure)**	5 **PIF measure**	6 **Other PIFs (and Uncertainty)**	7 **REF**
TEP1	Unsp	(Effect size on accuracy) 0.05	Unspecified	Cold	(Meta-analysis)	[172]
TEP1	Unsp	(Effect size on reaction time) -0.11	Unspecified	Cold	(Meta-analysis)	[172]
TEP1	D	(Effect size on accuracy) -1.07	Unspecified	Cold	(Meta-analysis)	[173]
TEP1	U / DM	(Effect size on accuracy) 0.05	Unspecified	Cold	(Meta-analysis)	[173]
TEP1	E	(Effect size on accuracy) 0.58	Unspecified	Cold	(Meta-analysis)	[173]
TEP1	E	(Effect size on reaction time) -1.1	Unspecified	Cold	(Meta-analysis)	[173]
TEP1	D / E	%difference -7.8%	Attention/Perceptual tasks	<65ºF	(Meta-analysis)	[53]
TEP1	D /E	% difference (+) 1.75%	Mathematical processing tasks	<65ºF	(Meta-analysis)	[53]
TEP1	U	% difference -28%	Reasoning/Learning/ Memory tasks	<65ºF	(Meta-analysis)	[53]
TEP1	Unsp	% difference -25%	Unspecified	<65ºF, Short task duration (<60min)	(Meta-analysis)	[53]
TEP1	Unsp	% difference -3%	Unspecified	<65ºF, long task duration (>60min)	(Meta-analysis)	[53]
TMP1	E	T_{car} T_{stop} Normal 4s 12s Cold 3s 8s	Simulate driving (T_{car} time hitting brake from car, T_{stop} time hitting brake from STOP sign)	Cold temperature for 40mins	No apparent uncertainty	[174]
TMP1	D, E, U, DM, T	Center and range of error factor (i.e., PIF weight): D (instrumentation): [1.8, 2.1, 2.7] U (cognition): [3.8, 10, 18] DM and T (management): [3., 8, 18] E (physical): [1.6, 5, 8] E (precise motor actions (connect lines to pump, remove air from lines and pumps): [13, 20, 30]	Maintenance task of offshore oil and gas facility pumps (develop work orders, reconnect pump, open valve and reinstate pump)	Extremely cold	(Estimation of error factors based on operational data)	[175]
TEP2	Unsp	(Effect size on accuracy) -0.33	Unspecified	Heat	(Meta-analysis)	[172]
TEP2	Unsp	(Effect size on reaction time) -0.11	Unspecified	Heat	(Meta-analysis)	[172]

<table>
<tr><td>TEP2</td><td>D</td><td>(Effect size) -0.78</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[172]</td></tr>
<tr><td>TEP2</td><td>U / DM</td><td>(Effect size) -0.23</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[172]</td></tr>
<tr><td>TEP2</td><td>E</td><td>(Effect size) -0.31</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[172]</td></tr>
<tr><td>TEP2</td><td>D</td><td>(Effect size on accuracy) -0.41</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[173]</td></tr>
<tr><td>TEP2</td><td>U / DM</td><td>(Effect size on accuracy) -0.27</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[173]</td></tr>
<tr><td>TEP2</td><td>E</td><td>(Effect size on accuracy) -0.59</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[173]</td></tr>
<tr><td>TEP2</td><td>E</td><td>(Effect size on reaction time) -1.1</td><td>Unspecified</td><td>Heat</td><td>(Meta-analysis)</td><td>[173]</td></tr>
<tr><td>TEP2</td><td>D</td><td>%diff - (percentage difference between neutral and experimental temperature conditions) -14%</td><td>Attention/Perceptual tasks</td><td>>80°F</td><td>(Meta-analysis)</td><td>[53]</td></tr>
<tr><td>TEP2</td><td>U</td><td>%diff 1.75%</td><td>Reasoning/Learning/ Memory tasks</td><td>>80°F</td><td>(Meta-analysis)</td><td>[53]</td></tr>
<tr><td>TEP2</td><td>D / E</td><td>%diff -14%</td><td>Mathematical processing tasks</td><td>>80°F</td><td>(Meta-analysis)</td><td>[53]</td></tr>
<tr><td>TEP2</td><td>Unsp</td><td>%diff -17.8%</td><td>Unspecified</td><td>>80°F, Short experimental session (<120min)</td><td>(Meta-analysis)</td><td>[53]</td></tr>
<tr><td>TEP2</td><td>Unsp</td><td>%diff -5%</td><td>Unspecified</td><td>>80°F, long experimental session (>120min)</td><td>(Meta-analysis)</td><td>[53]</td></tr>
<tr><td>TEP2</td><td>D</td><td><table><tr><td></td><td>20°C</td><td>50°C</td></tr><tr><td>RVP</td><td>0.03</td><td>0.04</td></tr><tr><td>PRM</td><td>0.04</td><td>0.08</td></tr><tr><td>SSP</td><td>0.16</td><td>0.22</td></tr></table></td><td>RVP- rapid visual processing
PRM-pattern recognition memory
SSP-spatial span</td><td>Normal: 20°C
Hot: 50°C</td><td>No apparent uncertainty</td><td>[176]</td></tr>
<tr><td>TMP2</td><td>D /E</td><td><table><tr><td>70°F</td><td>22 (# errors)</td></tr><tr><td>92°F</td><td>46 (# errors)</td></tr></table></td><td>Monitor displays (# of errors)</td><td>Vigilance error of omission varying temperature</td><td>No apparent uncertainty</td><td>[177]</td></tr>
<tr><td>TMP2</td><td>D /E</td><td><table><tr><td>37°C</td><td>0.1</td></tr><tr><td>38°C</td><td>0.14</td></tr></table></td><td>Visual vigilance task (% of missed signals)</td><td>Varying temperature</td><td>No apparent uncertainty</td><td>[177]</td></tr>
<tr><td>TMP2</td><td>D /E</td><td><table><tr><td>37°C</td><td>0.35</td></tr><tr><td>38°C</td><td>0.47</td></tr></table></td><td>Auditory vigilance task (% of missed signals)</td><td>Varying body temperature</td><td>No apparent uncertainty</td><td>[177]</td></tr>
<tr><td>TMP2</td><td>D /E</td><td><table><tr><td>82°F</td><td>0.52</td></tr><tr><td>92°F</td><td>0.56</td></tr></table></td><td>Visual vigilance task (% of missed signals)</td><td>Varying temperature</td><td>No apparent uncertainty</td><td>[177]</td></tr>
<tr><td>TMP2</td><td>D /E</td><td><table><tr><td>Min</td><td>20</td><td>40</td><td>60</td></tr><tr><td>74° F</td><td>0.02</td><td>0.02</td><td>0.02</td></tr><tr><td>82° F</td><td>0.06</td><td>0.06</td><td>0.06</td></tr><tr><td>90° F</td><td>0.06</td><td>0.10</td><td>0.15</td></tr></table></td><td>Visual vigilance task (% of missed signals)</td><td>Varying temperature and duration</td><td>No apparent uncertainty</td><td>[177]</td></tr>
<tr><td>TMP2</td><td>D /E</td><td><table><tr><td>19°C</td><td>0.32</td></tr></table></td><td></td><td></td><td></td><td>[177]</td></tr>
</table>

		33°C	0.49		Vigilance task (% of missed signals)	Varying temperature	Time constraint and other PIFs	
TMP2	D /E	19°C	0.35		Vigilance task (% of missed signals)	Varying temperature	Time constraint and other PIFs	[177]
		33°C	0.45					
TMP2	D / E		T1	T2	Split attention 50/50 percent between two concurrent visual tasks T1 and T2	Varying temperature and splitting attention	Time constraint and other PIFs	[51]
		25°C	0.3	0.22				
		30°C	0.35	0.3				
		35°C	0.65	0.4				
TMP2	DM	No significantly difference in the switching point in the lottery task			Lottery Game	Neutral 25°C vs warm 32°C	Time constraint	[178]
TMP2	DM		CDQ score	RSQ score	CDQ (Choice Dilemma Questionnaire) and RSQ (Risk Scenario Questionnaire)	Neutral 25°C vs warm 32°C	Time constraint	[178]
		25°C	5.92	3.68				
		30°C	5.01	4.96				
TMP2	DM		#click	Sum	BART (Balloon Analogue Risk Task, # of average clicks and sum of burst balloons)	Neutral 25°C vs warm 32°C	No apparent uncertainty	[178]
		25°C	6.77	32				
		30°C	9.73	40				

Appendix A8 PIF Attributes and Weights for Resistance to Physical Movement

Table A8-1 Attribute Identifiers and Descriptions for PIF Resistance to Physical Movement

ID	Attribute
PR1	Resistance to personnel movement, limited available space, postural instability
PR2	Whole-body vibration
PR3	Wearing heavy protective clothes or gloves or both

Table A8-2 IDHEAS-DATA IDTABLE-8 – PIF Weights for Resistance to Physical Movement

<table>
<tr><td>1</td><td>2</td><td>3</td><td>4</td><td>5</td><td>6</td><td>7</td></tr>
<tr><td>PIF</td><td>CFM</td><td>Error rates or task performance indicators</td><td>Task (and error measure)</td><td>PIF measure</td><td>Other PIFs (and Uncer-tainty)</td><td>REF</td></tr>
<tr><td>PR1</td><td>E</td><td><table><tr><td>Size for access</td><td>Right side</td><td>Real location</td></tr><tr><td>35mm</td><td>100s</td><td>50s</td></tr><tr><td>30mm</td><td>100s</td><td>70s</td></tr><tr><td>20mm</td><td>200s</td><td>380s</td></tr></table></td><td>Removing two nuts (task completion time in seconds)</td><td>Sizing and configuration for access - aperture size (in mm) and task location (right side and real location)</td><td>(No error data)</td><td>[179]</td></tr>
<tr><td>PR1</td><td>E</td><td>35% increase in task completion time with suited compared to unsuited.</td><td>Mobility moving through hatchways, tunnels</td><td>Size and configuration of hatchways, tunnels – unsuited and suited</td><td>Accuracy is more sensitive, but no data reported</td><td>[179]</td></tr>
<tr><td>PR1</td><td>E</td><td><table><tr><td></td><td>Mental addition</td><td>Tapping</td></tr><tr><td>Land</td><td>0.08</td><td>0.053</td></tr><tr><td>4.6m</td><td>0.07</td><td>0.057</td></tr><tr><td>15.2m</td><td>0.15</td><td>0.056</td></tr></table></td><td>Professional divers mentally added numbers or performed reciprocally tapping</td><td>A dryland control test followed by manipulation at 4.6m and 15.2m depths in the open ocean</td><td>No apparent uncertainty</td><td>[55]</td></tr>
<tr><td>PR1</td><td>E</td><td><table><tr><td></td><td>T1</td><td>T2</td><td>T3</td></tr><tr><td>C</td><td>0.02</td><td>0.02</td><td>0.028</td></tr><tr><td>S</td><td>0.04</td><td>0.07</td><td>0.158</td></tr></table></td><td>Offshore lifeboat operation
T1- Incorrectly operate brake cable
T2- Fail to disengage boat
T3- Fail to check air support system</td><td>Controlled (C): Force 4 wind, daylight, unignited gas leak
Severe (S): Force 6 wind, night, explorations/fire on platform</td><td>(Data-based estimation)</td><td>[52]</td></tr>
<tr><td>PR1</td><td></td><td><table><tr><td>S-SM</td><td>12s</td></tr><tr><td>D-SM</td><td>22s</td></tr><tr><td>S-TM</td><td>12s</td></tr><tr><td>D-TM</td><td>36s</td></tr></table></td><td>Use space mitten (SM) and tool mitten (TM) to screw bolts (task completion time in secs)</td><td>Space tool mitten cylinder mode - static (S) vs (D) dynamic</td><td></td><td>[180]</td></tr>
<tr><td>PR1</td><td></td><td><table><tr><td>Measures</td><td>% changes</td></tr><tr><td>FVC</td><td>6-15%</td></tr></table></td><td>Male infantry soldiers marched on six occasions wearing</td><td>Six occasions wearing either: no load,</td><td>(No error data)</td><td>[181]</td></tr>
</table>

		Expiratory: 17% Breathing frequency: 3 to 26 breaths per min 72% of participants experienced expiratory flow limitation whilst	loads (% changes in physiological measures)	15 kg, 30 kg, 40 kg or 50 kg. Each loaded configuration included body armor which was worn as battle-fit or loose-fit (40 kg only).		
PR2	E	Up to 40% more errors than occurring when tracking under static conditions	Completion of tracking tasks	Low frequency vertical vibration between 0.20g and 0.80g	No apparent uncertainty	[58, 182, 183]
PR2	D / E	10% to 15% reduction in error rates of information processing tasks (the impairment was due to a disruption in the information input processes)	Processing of information in short-term memory	Exposure to 16 Hz WBV at a magnitude of 2.0m/s2 rms .	No apparent uncertainty	[58, 182, 183]
PR2	D / E	Effect size -1.79	Perception in task performance	Vibration duration, intensity, frequency	(Meta-analysis)	[56]
PR2	U / E	Effect size -0.52	Cognition in task performance	Vibration duration, intensity, frequency	(Meta-analysis)	[56]
PR2	E	Fine motor continuous: Effect size -0.89 Fine motor discrete: Effect size -0.84	Motor execution in task performance	Vibration duration, intensity, frequency	(Meta-analysis)	[56]
PR3	E	Percent error increased 17%-23%; map plotting diminished by approximately 40%	Military tasks - investigator-paced tasks and map plotting	7-h periods on 4 successive days with or without protective cloths	No apparent uncertainty	[184]
PR3	E	Normal: 3.3CM Arctic cloth and gloves: 4.0CM	Turning bolt with common screwdriver (Maximum space needed in centimeter (CM))	Wearing arctic leather jacket and gloves	No apparent uncertainty	[179]
PR3	E	T1A (# of errors) / T1B (# of errors) / T2A (# of errors) No suit: 2.7 / 4.6 / 0.71 suit: 18 / 25 / 3.15 Task completion time increased 109%	Members of the National Guard's Civil Support Team (CST) performed T1A, B - Minnesota Dexterity test - placing, turning, and displacing small objects (# of errors) T2A- Mirror Tracer Test (# of errors)	Level A suits - fully encapsulating, bulky, and heat retentive	No apparent uncertainty	[185]

Appendix A9 PIF Attributes and Weights for System and I&C Transparency to Personnel

Table A9-1 Attribute Identifiers and Descriptions for PIF System and I&C Transparency to Personnel

ID	Attribute
SIC0	No impact
SIC1	System behavior is complex to understand or not transparent to personnel • Decision bias – Personnel use cues as heuristics for making decision without fully understanding the context of the cues • Feedback about system state, action, and intention is not provided
SIC2	Inappropriate system functional allocation between human and automation • Over-reliance – System is highly autonomous and personnel are not alerted for actions to take
SIC3	System failure modes are not transparent to personnel • System behavior is not consistent • System failures are not obvious to personnel • System failures are coupled or interdependent
SIC4	I&C logic is not transparent, e.g., complex logic for personnel to understand, I&C reset unclear to personnel
SIC5	I&C failure modes are not transparent to personnel

Table A9-2 IDHEAS-DATA IDTABLE-9 – PIF Weights for System and I&C Transparency to Personnel

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates or task performance indicators**	**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
SIC1 (Inf2.6)	D / U	No Automation: 0.03 Unreliable automation: 0.41	Monitor status with or without automation aid in triple tasks (missing targets)	Unreliable A – Automation aid for monitoring is 91% reliable No automation- no automation aid	(The task was understanding when automation failed)	[186]
SIC1	U / DM	Rate of pilots who made errors Routing tasks: <0.3 Mode awareness and understanding automation: 0.7 Answer consequence of automation failure: 0.46	20 pilots fly 1-hour scenario with automation system failure (rate of pilots who made errors)	Pilots' mental model and knowledge about automation system	(Small sample size)	[187]
SIC1	U / DM	(A-C / A-F) NoVSD: 0.05 / 0.3 VSD: 0.025 / 0.1	ATC resolves conflicts with automation assistance, i.e., Conflict Resolution Advisor (incorrect rate)	Automation is 80% reliable A-C – Automation correct A-F – Automation failure VSD – Visual display for transparency	(Other PIFs may exist)	[60]
SIC1	U / DM	(% error / %SA / Time) No VSD: 0.11 / 59% / 7.78s	ATC resolves conflicts with automation assistance,	Automation is 80% reliable A-C – Automation correct	(Other PIFs may exist)	[60]

		VSD	0.06	73%	5.38s	SA – Situation awareness	A-F – Automation failure VSD – Visual display for transparency		
SIC1	Unsp.	Traditional	75.9 (0-100)			NPP crew performs normal procedures (operator performance assessment score 0-100)	Traditional vs. transparent automation interface	(Many other factors involved)	[59]
		Transparent	67.5 (0-100)						
SIC1	Unsp.	No difference between transparency and non-transparency				NPP crews performs normal procedures	Traditional vs. transparent automation interface	(Many other factors involved)	[59]
SIC2	Unsp.		Paper	CP		NPP crews perform normal procedures	Level of automation – paper procedures vs. computerized procedures (CPs)		[59]
		Performance score and response time	No difference						
		Situation awareness score (1-10)	4.5	5.5					
Inf2.6	U	0.65 (0.59 corresponding to the belief that automation will lead to high accuracy)				Commission error in simulated flight	Conflicting info, automation misleading	(Other PIFs may exist)	[186]
SIC2	D		Triple tasks	Single task		Monitoring status with automation aid (% of failing to detect automation failure)	Automation reliability – constant vs. variable Simultaneous triple vs. single task	Triple tasks	[188]
		Variable	0.18	0.03					
		Constant	0.67	0.03					
SIC2	D	Time on monitoring in the triple task				Monitoring status with automation aid (time on monitoring task)	Automation reliability – constant vs. variable.	Triple tasks	[188]
		Variable	4.0s						
		Constant	2.9s						
SIC2 (Inf2.6)	U	0.55				25 pilots simulated 4 flight events (failing to detect automation failure)	Automation failure in the scenarios. There was other correct information available	(More experience leads to higher error rates)	[189]
SIC2 (Inf2.6)	DM	1				25 pilots simulated 4 flight events (commission rate)	Decision aid was wrong. Pilots should use other information	(Level of experience varied)	[189]
SIC2	D / U	Frequency of error classification: 35% failure to monitor 23% related to task distraction 5% related to over-reliance on automation				Flight automation failure accident (frequency of error classification)	Automation-induced complacency	(Error classification)	[190]

SIC2	D	15s (RMS error) 72 30s (RMS error) 85 60s (RMS error) 90	6 pilots simulated flight tracking task (RMS errors)	Task allocation - Duration of tracking automatic cycle	(Small sample size)	[191]
SIC1 & SIC3	Unsp.	Freq. in 34 accidents Understanding of automation may be inadequate: 6/34 Pilots may over-rely on automation: 5/34 Automation may not work well under unusual conditions: 4/34	Aviation FDAI Automation Human Error Types	Accident caused by automation failure	(Error classification)	[61]
SIC1 & SIC3	Unsp	Top freq. in 34 accidents Lack of understanding of the system: 5/34 Improper performance of an automation device in an abnormal situation: 4/34	FDAI Automation Human Error Types (frequencies of error types)	Accident caused by automation failure	(Analysis did not separate system vs failure mode)	[61]
SIC4	DM	(NoT / T) % caution with decision: 30% / 57% # attempts: 1.43 / 1.73 # identified targets: 13.5 / 19.5 Accuracy: 0.83 / 0.87	Identify threating targets under uncertainties	T (transparent) - visual display of target uncertainty NoT– no transparency	(Measures are not error rates)	[192]
SIC5	Unsp.	HSI: 34% Software: 32 % Hardware: 34%	Relative percent of errors reported in LERs	Digital I&C failures in LERs between 1994-1998	Not analyzed	[193]
Unsp	Unsp	The results (Figure A9-1) indicated that instrumentation is more prone to human error than the rest of maintenance. Instrumentation & control equipment and software (IC), electrical equipment (EL), process valves, ventilation dampers or channel hatches (VAL), mechanical equipment (other than valves, MEC), block or primary valves in instrument lines (IVAL).	Human errors recognizable in connection with maintenance were looked for by reviewing about 4400 failure and repair reports and some special reports which cover two nuclear power plant units on the same site from 1992–94	Equipment types involved in single human errors 1992–1994, together 206 cases.	(Root causal analysis)	[194, 195]

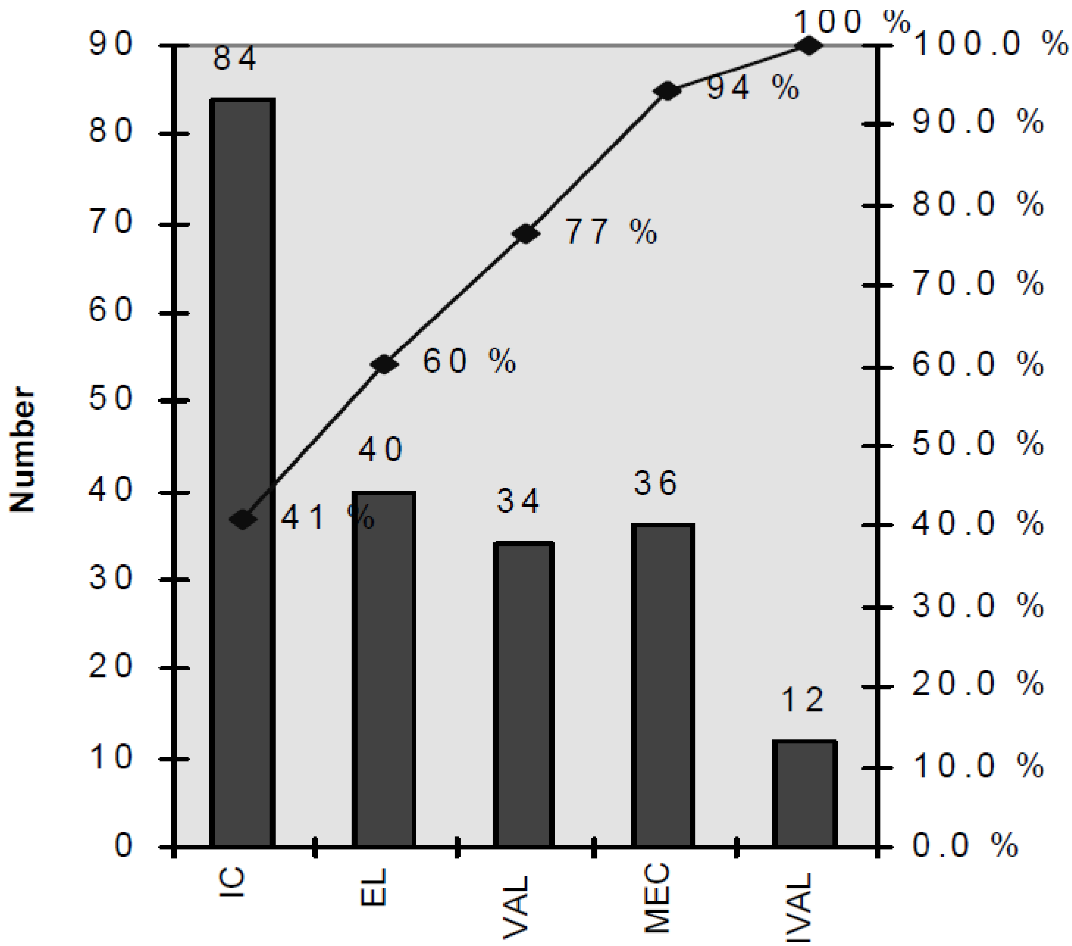

Figure A9-1 Instrumentation is more prone to human error than the rest of maintenance

Appendix A10 PIF Attributes and Weights for Human-System Interfaces

Table A10-1 Attribute Identifiers and Descriptions for PIF Human-System Interfaces

ID	Attribute
HSI0	No impact – well designed HSI supporting the task
HSI1	Indicator is similar to other sources of information nearby
HSI2	No sign or indication of technical difference from adjacent sources (meters, indicators)
HSI3	Related information for a task is spatially distributed, not organized, or cannot be accessed at the same time
HSI4	Un-intuitive or un-conventionnel indications
HSI5	Poor salience, eccentric location, or low text readability of the target (indicators, alarms, alerts) out of the crowded background
HSI6	Inconsistance – Physical représentation of information, mesurément units, symbols, or tables
HSI7	Inconsistent interpretation of displays
HSI8	Similarity in control elements - Wrong element selected in operating a control element on a panel within reach and similar in design
HSI9	Poor functional centralization –multiple displays/panels needed together to execute a task
HSI10	Ergonomic deficits • Controls are difficult to maneuver • Labels and signs of controls are not salient or low readability • Labels are confusing (e.g., using unconventional measurement units) • Inadequate indications of states of controls - Small unclear labels, difficult reading scales • Maneuvers of controls are un-intuitive or unconventional
HSI11	Labels of the controls do not agree with document nomenclature, confusing labels
HSI12	Controls do not have labels or indications
HSI13	Controls provide inadequate or ambiguous feedback, i.e., lack of or inadequate confirmation of the action executed (incorrect, no information provided, measurement inaccuracies, delays)
HSI14	Confusion in action maneuver states (e.g., automatic resetting without clear indication)

Table A10-2 IDHEAS-DATA IDTABLE-10 – PIF Weights for Human-System Interfaces

1	2	3			4	5	6	7
PIF	**CFM**	**Error rates**			**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
HSI1	D	Perceived contrast	Central	Eccen.	Perceive target visual contrast (% contrast perceived)	Target location – Fovea and 12º eccentric Surrounding – similar visual stimuli surround the target	No apparent uncertainty	[196]
		No surrounding	40%	15%				
		With surrounding	26%	3%				
HSI1	D	Random	0.004		Read numbers from screen	Nearby similar text – Ordered from small to large numbers, randomly in line, randomly in cloud	(Small subject sample)	[197]
		Ordered	0.004					
		Cloud	0.015					
HSI1	D	Random	0.0995		Search targets and count the total number (incorrect counting)	Target numbers are arranged orderly, randomly with similar	(Ordered has the maximum similarity	[197]
		Ordered	0.224					
		Cloud	0.194					

					distractors in line, or embedded among similar distractors (cloud)	between target and distractors)	
HSI2	D	2.1E-2 (1/56)		Verifying the state of indicator lights on the front side of a control cabinet (Erroneous operation of a push button)	No indication of technical differences between two adjacent plant units provided	Rarely performed	[4]
HSI3	U	With integration	0.23	96 male students diagnosed leak location using NPP simulator displays (Diagnosis accuracy)	Integration - the process information was integrated into the alarm display and presented as alarm bars	Time constraint (students not proficient with the tasks)	[198]
		Without integration	0.29				
HSI3	D	Without HUD	0.27	Pilots detect off-normal event out-of-window (missing events)	HUD (head-up-Display) and target in different spatial location for view	Multitasking (collective data from many studies)	[199]
		With HUD	0.36				
HSI3	D	Without HITS	0.22	Pilots detect off-normal event out-of-window (missing events)	HITS(Highway-in-the-sky) A HITS display integrates 3-D information of the flight path into a perspective path through the air	Multitasking (collective data from many studies)	[199]
		With HITS	0.45				
HSI4	D	Innovate	0.13	NPP operators identify parameter trends on NPP simulators (% incorrect identification)	Innovate display – graphically show trends Conventional display - show numeric parameter values	No context, no peer-checking, time constraint (small sample)	[200]
		Conventional	0.33				
HSI4	D	Innovate	0.11	NPP operators check the values of multiple parameter (% incorrect identification)	Innovative display – graphic features of parameters. Conventional display - numeric parameter values	No context, no peer-checking, time constraint (small sample)	[200]
		Conventional	0.2				
HSI5	D	Salient	0.008	Detect visual notification of a pending interrupting task while performing an arithmetic task	Non-salient: Exclamation marks appeared over a clock icon in the controller display Salient – pop-out color or blinking visual icon that captured attention	Dual-task in non-salient display	[142]
		Non-salient	0.167				
HSI5	D	Central	0.04	Students detect visual targets (missing rate)	Location of the target in the central/eccentric visual field	No apparent uncertainty	[201]
		Eccentric	0.11				
HSI5	D	Font size = 7.434 * EXP(-contrast/0.6297) + 5.028		Read text from displays (error-free font size)	Error-free angular (arc min) font size is a function of text contrast	(Error-free: error rate < 0.01)	[202]

		Formula is fitted from experimental data				
HSI5	D	Minimum salience (luminance contrast and color contrast) for reliable perception — Lumin. / Color Attention: >20cd/m^2 / >0.24 in CIE Idnetification: <20cd/m^2 / > 0.04 in CIE Separation: >15~20% / >0.004 Text reading: >30% /	Four basic tasks: Salient target to capture attention; Use colors to identify information categories; Separate information; Read text	Luminance context and color contrast of target or text from the background or surround distractors, Apparent luminance, number of colors	(Numbers from many experimental papers)	[203, 204]
HSI5	D	View distance / Error rate 1.21meter / 0.06 2,13m / 0.18 3.05m / 0.42	Read green text on black background in daylight	View distance (meters) from the CRT screen, viewing time was 0.5s.	(0.06 is the lowest error rate with strong ambient light)	[202]
HSI5	D	— Match / Mismatch Salient: 0.03 / 0.1 Non-salient: 0.16 / N/A	Prospective memory-based decision-making with cue/task match	Cue (alert) saliency – flicking vs. static Cue-task match vs. mismatch	No apparent uncertainty	[205]
HSI6	D	Standard: 0.15 Physical: 0.04 Conventional: 0.29 Conceptual: 0.2	IT Professionals learned and answered questions with e-learning systems (error rate of answering questions)	Information displayed inconsistently across displays	Subjects were in training and not proficient yet	[206]
HSI7	D	W=5.7	Information gathering tasks	Inconsistent interpretation of displays	(Engineering judgment)	[121]
HSI8		7.29E-3 (1/162)	Pulling an isolating terminal in a control cabinet (Wrong terminal pulled)	Similar terminals nearby, terminals arranged in regular patterns, similar terminal identification codes	(Errors could be for a step or a task)	[5]
HSI8		8.9E-4 (7/8058)	Operating a control element on a panel (Wrong element selected)	Wrong control element within reach and similar in design	(Errors could be for a step or a task)	[5]
HSI8		1.3E-3 (1/888)	Reassembly of component elements (Wrong element)	Similar design and close spatial proximity between correct and wrong element	(Errors could be for a step or a task)	[4]
HSI8		1.2E-3 (1/948)	Operating a pushbutton control (Wrong button selected)	Similar buttons nearby, ergonomically well-designed panel	(Errors could be for a step or a task)	[4]

<table>
<tr><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>HSI8</td><td></td><td>9.2E-4
1/1146</td><td>Operating a push button control (Wrong button selected)</td><td>Similar buttons within reach, text labeling only</td><td>(Errors could be for a step or a task)</td><td>[4]</td></tr>
<tr><td>HSI8</td><td></td><td>8.9E-4
(1/1332)</td><td>Operating a rotary control (Wrong switch selected)</td><td>Switch within reach,
similar switches nearby, text labeling only</td><td>(Errors could be for a step or a task)</td><td>[4]</td></tr>
<tr><td>HSI8</td><td></td><td>7.8E-4
(1/1512)</td><td>Connecting a cable between an external test facility and an electronic module (Connected to wrong module)</td><td>Module access ports within reach, similar access ports nearby, frequently performed task, color coding of ports</td><td>(Errors could be for a step or a task)</td><td>[4]</td></tr>
<tr><td>HSI8</td><td></td><td>1.2E-3
(3/2630)</td><td>Operating a control element on a panel (Wrong control element selected)</td><td>Plain text labeling, similar controls within reach</td><td>(Errors could be for a step or a task)</td><td>[5]</td></tr>
<tr><td>HSI8</td><td></td><td>2.1E-3
(4/1958)</td><td>Operating a control element
on a panel
(Wrong control element selected)</td><td>Mimic diagrams, color coding, similar
controls within reach</td><td>(Errors could be for a step or a task)</td><td>[5]</td></tr>
<tr><td>HSI8</td><td></td><td>1.6E-3
(7/4588)</td><td>Operating a control element on a panel (Wrong control element selected)</td><td>Wrong control element within reach and similar in design</td><td>(Errors could be for a step or a task)</td><td>[5]</td></tr>
<tr><td>HSI9</td><td>E</td><td><table>
<tr><td></td><td>PD low</td><td>PD M</td><td>PD High</td></tr>
<tr><td>FC Low</td><td>8.62E-1</td><td>4.84E-1</td><td>2.64E-1</td></tr>
<tr><td>FC-medium</td><td>2.84E-1</td><td>1.29E-1</td><td>8.41E-2</td></tr>
<tr><td>FC-high</td><td>1.15E-1</td><td>6.24E-2</td><td>4.04E-2</td></tr>
</table></td><td>Execute procedures in NPP local stations</td><td>PD – Panel ergonomic design

FC – Functional centralization, low for too many panels</td><td>(Expert judgment)</td><td>[7]</td></tr>
<tr><td>HSI10</td><td>E</td><td>8.78E-4
(1/1347)</td><td>Operation of a manual control at an MCR control panel (Task not remembered)</td><td>Position of indicator lamps ergonomically unfavorably designed</td><td>(Errors could be for a step or a task)</td><td>[5]</td></tr>
<tr><td>HSI12</td><td>E</td><td>1.93E-3
(1/612)</td><td>Operating a continuously adjustable rotary handle (Handle rotated too far)</td><td>No markings and no end stop present</td><td>(Errors could be for a step or a task)</td><td>[5]</td></tr>
<tr><td>HSI12</td><td></td><td>9.83E-3
(1/120)</td><td>Reinstallation of control rod drive motors (Drive motor mounted to wrong control rod, false</td><td>No position labels on control rods, position inferred indirectly from</td><td>(Errors could be for a step or a task)</td><td>[5]</td></tr>
</table>

			identification of position)	secondary information		
HSI12		7.57E-3 (1/156)	Connecting transducers to pressure sensing lines (Connections swapped, professional knowledge remembered incorrectly)	Frequently performed task, no labeling	(Errors could be for a step or a task)	[5]
HSI13	E	W=5.5	Unspecified manipulations	Controls provide inadequate or ambiguous feedback, i.e., lack of adequate confirmation of the action executed	(Engineering judgment)	[121]

Appendix A11 PIF Attributes and Weights for Equipment and Tools

Table A11-1 Attribute Identifiers and Descriptions for PIF Equipment and Tools

ID	Attribute
ETP0	No impact –ETPs are easy to use and well maintained under proper administrative control
ETP1	ETP is complex, difficult to use, or has poor suitability for the work, e.g., • Using ETPs require calculations • ETP has ambiguous or unintuitive interfaces • ETP is difficult to maneuver • Labels on ETPs are not salient
ETP2	Rarely used ETP does not work properly or is temporally not available (due to lack of proper administrative control, lack of accessories, incompatibility, improper calibration, etc.)
ETP3	ETP labels are ambiguous or do not agree with document nomenclature
ETP4	Personnel are unfamiliar or rarely use the ETP, e.g., • Failure modes or operational conditions of the ETP are not clearly presented to personnel • Personnel are not familiar with the ranges, limitations, and requirements of ETP

Table A11-2 IDHEAS-DATA IDTABLE-11 – PIF Weights for Equipment and Tools

1 PIF	2 CFM	3 Error rates or task performance indicators	4 Task (and error measure)	5 PIF measure	6 Other PIFs (and Uncertainty)	7 REF
Unsp.	Unsp.	| X^2 correlation coefficient Unsafe Conditions | 29.7 Machinery and equipment | 34.1 Tools | 3.9 Organizational factors | 3.9	Human errors in 225 automotive manufacture accidents (X^2 correlation coefficient)	Associations between types of human error and contributing factors evaluated by cross-Pearson's Chi-test	(CFMs and PIFs unspecified)	[207]
Unsp.	Unsp.	| Cronbach 's Alpha Coefficient Human error in maintenance | 0.818 Machine availability | 0.761 Machine performance | 0.823 Product quality | 0.776	Cronbach 's Alpha Coefficient of the factors studied and overall equipment effectiveness	Relationship between human errors in maintenance and overall equipment effectiveness in food industries	(CFMs and PIFs unspecified)	[208]
ETP1	E	| Before noon | Afternoon Digital | 4.45 | 5.74 Analog | 11.07 | 13.7	Experienced technicians used digital and analog multimeters to measure voltage and resistance (%measurement errors)	Tools - Digital vs analog, Time of work – before noon (FN) and afternoon (AN)	(The errors are applicable to Detection and execution)	[63]
EAP1, EAP2	E	Freq. of EAP as the cause in 100 accidents EAP1-suitability | 44	Construction work (freq. of EAP as the	Suitability, usability, and conditions of	(Statistical data, no error rates)	[209]

<table>
<tr><td rowspan="2"></td><td rowspan="2"></td><td>EAP1- usability</td><td colspan="2">19</td><td rowspan="2">cause of an accident)</td><td rowspan="2">equipment and tools</td><td rowspan="2"></td><td rowspan="2"></td></tr>
<tr><td>EAP2- conditions</td><td colspan="2">12</td></tr>
<tr><td rowspan="6">EAP1</td><td rowspan="6">E</td><td># of calculation</td><td colspan="2">%error</td><td rowspan="6">Calculation needed in construction work</td><td rowspan="6">Number of calculations in construction work</td><td rowspan="6">(Unverified original data source)</td><td rowspan="6">[210-212]</td></tr>
<tr><td>2</td><td colspan="2">0.01</td></tr>
<tr><td>3</td><td colspan="2">0.04</td></tr>
<tr><td>4</td><td colspan="2">0.05</td></tr>
<tr><td>5</td><td colspan="2">0.07</td></tr>
<tr><td>8</td><td colspan="2">0.1</td></tr>
<tr><td rowspan="7">ETP1</td><td rowspan="7">E</td><td>Code interpretation</td><td colspan="2">0.015</td><td rowspan="7">Calculation needed in construction work (error rate in calculation)</td><td rowspan="7">Types of calculation needed in construction work</td><td rowspan="7">(Types pf calculation are applicable to use ETPs)</td><td rowspan="7">[210-212]</td></tr>
<tr><td>Ranking</td><td colspan="2">0.014</td></tr>
<tr><td>Table look-Up</td><td colspan="2">0.013</td></tr>
<tr><td>Loading coefficients</td><td colspan="2">0.133</td></tr>
<tr><td>Loading directions</td><td colspan="2">0.10</td></tr>
<tr><td>Reduction factors</td><td colspan="2">0.80</td></tr>
<tr><td>Loading combinations</td><td colspan="2">0.42</td></tr>
<tr><td rowspan="8">Unsp</td><td rowspan="8">E</td><td>Equipment</td><td colspan="2">Freq.</td><td rowspan="8">Proportion of accidents caused by humans vs specific equipment/systems in chemical process industry</td><td rowspan="8">Human error contributor freq. in 364 chemical process plant accidents, each accident has ~2 contributors on average</td><td rowspan="8">(Unspecified human error contributors)</td><td rowspan="8">[213]</td></tr>
<tr><td>Piping system</td><td colspan="2">25%</td></tr>
<tr><td>Storage tank</td><td colspan="2">14%</td></tr>
<tr><td>Reactor</td><td colspan="2">14&</td></tr>
<tr><td>Heat transfer Eq.</td><td colspan="2">10%</td></tr>
<tr><td>Process vessel</td><td colspan="2">8%</td></tr>
<tr><td>Separation Eq.</td><td colspan="2">7%</td></tr>
<tr><td>Machineries</td><td colspan="2">5%</td></tr>
<tr><td rowspan="4">ETP4</td><td rowspan="4">E</td><td></td><td>Non-FLEX</td><td>FLEX</td><td rowspan="4">Use of portable generator or pump in a Non-FLEX-designed scenario (sunny day) vs. a FLEX-designed scenario (severe accident)</td><td rowspan="4">Personnel rarely use the equipment and training is infrequent. Non-FLEX scenario– no complication FLEX scenario – Post seismic and rain</td><td rowspan="4">Scenario unfamiliar, rarely performed actions, poor training (Expert judgment)</td><td rowspan="4">[3]</td></tr>
<tr><td>Transport</td><td>0.057</td><td>0.14</td></tr>
<tr><td>Connect</td><td>0.088</td><td>0.16</td></tr>
<tr><td>Operate</td><td>0.052</td><td>0.12</td></tr>
</table>

Appendix A12 PIF Attributes and Weights for Staffing

Table A12-1 Attribute Identifiers and Descriptions for PIF Staffing

ID	Attribute
STA0	No impact – adequate staffing
STA1	Shortage of staffing • key personnel are missing • unavailable or delayed in arrival • staff pulled away to perform other duties
STA2	Ambiguous or incorrect specification of staff roles, responsibilities, and configurations • Inappropriate staff assignment • Personnel utilization (percent of time on task)
STA3	Lack of certain knowledge, skills, or abilities needed for key personnel in unusual events, e.g., Key decision maker's knowledge and ability are inadequate to make the decision (e.g., lack of required qualifications or experience)
STA4	Lack of administrative control on fitness-for-duty

Table A12-2 IDHEAS-DATA IDTABLE-12 – PIF Weights for Staffing

<table>
<tr><th>1</th><th>2</th><th colspan="3">3</th><th>4</th><th>5</th><th>6</th><th>7</th></tr>
<tr><th>PIF</th><th>CFM</th><th colspan="3">Error rates</th><th>Task (and error measure)</th><th>PIF measure</th><th>Other PIFs (and Uncertainty)</th><th>REF</th></tr>
<tr><td rowspan="4">STA 1</td><td rowspan="4">Unsp.</td><td></td><td>M- minimal staffing</td><td>N- normal staffing</td><td rowspan="4">Crew performed five EOP scenarios:
T1- Primary tasks
T2 – Announcement and notifications
T3- Cooldown and stabilization
(Performance rating scale 1-10)</td><td rowspan="4">Crew size
Reduced I- SRO &RO
Minimum (M)- CRS, RO, BOP
Normal (N)- CRS, RO, BOP, control room technician</td><td rowspan="4">(Scenario differences)</td><td rowspan="4">[214]</td></tr>
<tr><td>T1</td><td>2.9</td><td>2.9</td></tr>
<tr><td>T2</td><td>3.1</td><td>3.3</td></tr>
<tr><td>T3</td><td>2.65</td><td>3.25</td></tr>
<tr><td rowspan="5">STA 1</td><td rowspan="5">Unsp</td><td colspan="3">Operator workload (6-60)</td><td rowspan="5">Crew performed five EOP scenarios
S- Supervisor
RO, BOP
(Operator workload level rated from 6 to 60)</td><td rowspan="5">Crew size
Reduced
Minimum (M)- CRS, RO, BOP
Normal (N)- CRS, RO, BOP, control room technician</td><td rowspan="5">(Scenario differences)</td><td rowspan="5">[214]</td></tr>
<tr><td></td><td>N</td><td>M</td></tr>
<tr><td>S</td><td>38</td><td>49</td></tr>
<tr><td>RO</td><td>39</td><td>41</td></tr>
<tr><td>BOP</td><td>38</td><td>38</td></tr>
<tr><td rowspan="2">STA 1</td><td rowspan="2">T</td><td colspan="2">M</td><td>N</td><td rowspan="2">Crew performed five EOP scenarios:
S- Supervisor
RO, BOP
(Team interaction score 1-5)</td><td rowspan="2">Minimum (M)- CRS, RO, BOP
Normal (N)- CRS, RO, BOP, control room technician (CT)</td><td rowspan="2">(Scenario differences)</td><td rowspan="2">[214]</td></tr>
<tr><td colspan="2">4.3</td><td>4.9</td></tr>
<tr><td rowspan="3">STA 1</td><td rowspan="3">D</td><td colspan="2">4 AVOs</td><td>0.25</td><td rowspan="3">Monitor and detect targets (% of missing)</td><td rowspan="3">Alternate Crew Configurations- # of Airforce Vehicle Operators (AVOs)</td><td rowspan="3"></td><td rowspan="3">[215]</td></tr>
<tr><td colspan="2">6 AVOs</td><td>0.05</td></tr>
<tr><td colspan="2">8 AVOs</td><td>0</td></tr>
<tr><td rowspan="3">STA 1</td><td rowspan="3">Unsp.</td><td colspan="3">Task completion time (min)</td><td rowspan="3">Firefighters complete twenty-two essential tasks that must be</td><td rowspan="3">Firefighter crew size</td><td rowspan="3"></td><td rowspan="3">[68]</td></tr>
<tr><td colspan="2">2-person</td><td>22:30</td></tr>
<tr><td colspan="2">3-person</td><td>20:37</td></tr>
</table>

		4-person: 15:46 5-person: 15:52	performed on low hazard structure fires (task completion time)			
STA 1	Uns p.	% of maximum heart rate (Driver / Officer) 2-person: 89% / 93% 3-person: 72% / 70% 4- and 5-person: 75%, 71% / 70%. 68%	Firefighters complete twenty-two essential tasks that must be performed on low hazard structure fires (% of maximum heart rate for the age)	Firefighter crew size		[65]
STA 1	Uns p	(PRT / OST) 2E&A: 4:23 / 13:46 3E&A: 3:13 / 12:06 4E&A: 2:52 / 10:23 2&A: 6:59 /	EMS (Emergency Medical Service) crews complete all EMS tasks for Trauma Patient (PRT- patient removal time, OST- overall time on the scene)	Crew size and configuration 2E &A 2-person Engine + Ambulance 3E&A 4E&A 2&A- 2-person Ambulance only		[67]
STA 1 & STA 2	D/U/ DM/ E	(Easy scenario / Diff. scenario) T: 0.825 / 0.662 U: 0.755 / 0.457	3-person NPP crews performed 8 scenarios (OPAS - Operator Performance Assessment Scale 0 to 1)	Staffing configuration T - Traditional staffing – 3 persons for one reactor U - Untraditional staffing – 3 persons for two reactors with automation	(Automation use varied)	[62]
STA 2	Uns p.	(Trauma / Cardiac) A: 10:50 / 11:00 B: 13:06 / 12:00 C: 12:38 / 10:30 D: 11:45 / 13:00	EMS crews complete all EMS tasks for Trauma Patient (OST- overall time on the scene)	Crew configuration A - 1 ALS on Engine & 1 ALS on Amb (Ambulance) B - 2 ALS on Amb C - 1 ALS on Engine/ BLSAmb D - BLSEngine/ 1 ALS on Amb		[67]
STA 2	D /DM	% mean attention state Directed: 32% Divided: 22% Distracted: 46%	Monitor status and replan tasks in a 4-hour session with 2-10% utilization of time (% attention state: directed on task, divided between task and other things, distracted away from the task)	Low task utilization time (2-10%) in long working sessions	(Student subjects may differ from licensed crews)	[69, 216]
STA 2	Uns p/	“Full-crew” in 97.6% railroad accidents	Railroad operation	Crew sizes in accidents	(Opinion article)	[217]
STA 2	D / E	(Ins / Out) 2-man: 0.04 / 0.25 1-man: 0.04 / 0.45	Simulated Tactical Air Command Pilots detect and respond to threat targets	1-man crew - all controls and displays were located in a single cockpit, 2-man crew – the controls and displays were divided between two cockpits, Ins - Inside-cockpit threat targets	Time constrained	[218]

				Out - Outside-cockpit threat targets		
STA 2	D	The crews on airplanes flown with three pilots did see more aircraft. Interestingly, the 2-person crews saw significantly more aircraft than the two pilots on the 3-person airplanes	Simulated flying and detecting targets outside cockpit	2- vs. 3-pilot crew configuration	(Original data not public)	[219]
STA 3	Uns p.	Average F-JAS score <table><tr><td></td><td>RO</td><td>SS</td><td>SE</td></tr><tr><td>C</td><td>4.49</td><td>4.91</td><td>4.83</td></tr><tr><td>I</td><td>5.1</td><td>5.47</td><td>5.31</td></tr><tr><td>P</td><td>3.05</td><td>3.19</td><td>2.34</td></tr><tr><td>S</td><td>4.28</td><td>4.46</td><td>3.6</td></tr></table>	Normal and incident NPP CR operation: SS- shift supervisor, RO-reactor operator, SE-Safety Engineer or technician (average F-JAS scores for each category)	F-JAS 51 items of ability evaluation in four categories C – Cognitive reasoning I – Interpersonal P – Psychomotor S – Sensory perception	(Subjective assessment)	[220]
SAT 2	Uns p.	SMR staffing approach requires a comprehensive analysis of all the tasks, jobs, and workload which may be required of an operator while on the job	Monitor and respond to SMRs	Staffing configuration	(No error data)	[221]
SAT 2	Uns p.	<table><tr><td>Utilization</td><td># errors/ task</td></tr><tr><td>48%</td><td>1.5</td></tr><tr><td>51%</td><td>2.2</td></tr><tr><td>59%</td><td>12</td></tr></table>	Dispatching in managing networks of railroads and flights (# errors/task)	Utilization - % time on task	(Different types of tasks)	[222]
STA 3	Uns p.	Ability requirements in two work conditions for reactor operators (Figure 12-1). The F-JAS scales ranged from 1 to 7 with larger numbers reflecting higher ability requirements * $p < .05$, ** $p < .01$.	F-JAS ability requirements for NPP CR crew members: SS- shift supervisor, RO-reactor operator, SE-Safety Engineer or technician (average F-JAS scores for each category)	Normal and incident NPP CR operation N= 87 reactor operators, 60 shift supervisors, and 40 safety engineers	(Subjective assessment)	[220]

Ability	Normal operation		Incident		*d*
	M	*SD*	*M*	*SD*	
Time Sharing (C)	5.31	1.09	5.88	0.87	-0.57*
Behavior Flexibility (SI)	4.47	1.11	5.31	0.90	-0.84**
Self-Control (SI)	5.44	1.27	6.34	0.94	-0.81**
Oral Defense (SI)	4.06	1.34	4.88	1.16	-0.65*
Auditory Attention (SP)	4.56	1.52	5.31	0.90	-0.60*
Speech Recognition (SP)	4.34	1.41	5.28	0.92	-0.79**

Figure 12-1 Ability requirements in two work conditions for reactor operators

Appendix A13 PIF Attributes and Weights for Procedures, Guidelines and Instructions

Table A13-1 Attribute Identifiers and Descriptions for PIF Procedures, Guidelines, and Instructions

ID	Attribute
PG0	No impact – well validated procedures like most EOPs
PG1	Procedure design is inadequate and difficult to use • Difficult layout, lack of placeholders • Graphics or symbols not intuitive • Fold-out page not salient • Complicated logic and mental calculation required (e.g., unit conversion) • Poor standardization in use of terminology • Multiple versions not clearly labeled • Inconsistency between procedures and displays
PG2	Procedure requires judgment • Assessment of trends • Foldout use • Mental representation of the given situation
PG3	Procedure lacks details, e.g., • Lack of verification in procedure for verifying key parameters for detection or execution • Lack of guidance to seek confirmatory data when data may mislead for diagnosis or decisionmaking • Lack of detailed steps for non-skill-of-craft actions
PG4	Procedure is ambiguous, confusing, e.g, • Wrong or incomplete descriptions in certain key steps • Conflict between step's literal meaning and step intention
PG5	Mismatch - Procedure is available but does not match the situation (e.g., needs deviation or adaptation)
PG6	Procedure is not applicable or not available
PG7	Procedure is misleading

Table A13-2 IDHEAS-DATA IDTABLE-13 – PIF Weights for Procedures, Guidelines, and Instructions

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates**	**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
Unsp	Unsp	Freq. (%) of causes Personnel (team): 29% Procedure: 24% Planning: 11% Training: 10% Communication: 9%	Identification and classification of root causes in 53 NPP LPSD human or human-related events (Freq. of procedure as the cause)	Pre-defined categories of root causes	(Root causal analysis)	[71]
Unsp	E	No PIF: 0.033 Inadequate procedure: 0.5	Elevator installation (All kinds of human errors made in installation)	"Inadequate" procedure is for unspecified PIF attributes	(Statistical data and model fitting)	[223]
Unsp	E	(Good workload / Poor Workload) Good P: 4.53E-5 / 1.56E-5	NPP operators manipulating simple (discrete) control in Low Power Shutdown (error rate in	Good vs. Poor procedure (P) (Unspecific definition of good or poor procedure)	With recovery	[138]

<table>
<tr><td></td><td></td><td><table><tr><td>Poor P</td><td>3.53E-3</td><td>1.58E-5</td></tr></table></td><td>executing procedure steps)</td><td></td><td></td><td></td></tr>
<tr><td>Unsp</td><td>E</td><td><table><tr><td>Good</td><td>3.20.E-3</td></tr><tr><td>Poor procedure + poor training</td><td>1.63.E-1</td></tr></table></td><td>Verifying state of indicator</td><td>Nominal vs. poor procedure + poor training</td><td>Scenario familiarity may be in the high HEP</td><td>[138]</td></tr>
<tr><td>PGI1</td><td>D</td><td><table><tr><td colspan="2"># of ROs failed</td></tr><tr><td>T1-R</td><td>0</td></tr><tr><td>T1-G</td><td>3</td></tr><tr><td>T2-ES1.3</td><td>3</td></tr><tr><td>T2-E0</td><td>1</td></tr><tr><td>T3</td><td>1</td></tr></table></td><td>3 RO use CPS
(# of ROs failed)
T1-Detecting failures of the automatic evaluation function.
T1-R (red x)
T1-G (Green x)
T2- Detecting failures of the place-keeping function
T3- Total loss of CPS and transition to paper</td><td>CP indicators: red and green x for automation failure, Place-keeper, transition to paper procedure</td><td>(3 subjects tested)</td><td>[224]</td></tr>
<tr><td>PGI1</td><td>Unsp</td><td><table><tr><td colspan="3">Total # of errors made</td></tr><tr><td></td><td>LOCA</td><td>SGTR</td></tr><tr><td>CP</td><td>4</td><td>12.75</td></tr><tr><td>BP</td><td>18.75</td><td>13</td></tr></table></td><td>Sixteen licensed operators worked in teams of SRO/RO perform LOCA and SGTR scenarios</td><td>Computerized (CP) vs paper procedures (BP)</td><td>(Errors in whole scenarios)</td><td>[72]</td></tr>
<tr><td>PGI1</td><td>D/E</td><td><table><tr><td colspan="2"># of operation errors</td></tr><tr><td>CP</td><td>0.53</td></tr><tr><td>BP</td><td>1.08</td></tr></table></td><td>45 OPERATORS executed decision and action tasks to deal with alarm signals, while detecting occasional system errors in the interface (# of operation errors)</td><td>CP vs BP,
Team size (1,2,3-person)</td><td>(Whole scenarios)</td><td>[225]</td></tr>
<tr><td>PG1</td><td>D/E</td><td>See Figure A13-1</td><td>45 OPERATORS executed decision and action tasks to deal with alarm signals, while detecting occasional system errors in the interface (Subjective scores)</td><td>CP vs BP,
Team size (1,2,3-person)</td><td>(Whole scenarios)</td><td>[225]</td></tr>
<tr><td>PG1</td><td>E</td><td>3.3E-3
(2/651)</td><td>NPP maintenance tasks using a procedure (Omitting an item of instruction)</td><td>Long list, checkoff provisions</td><td>Not analyzed</td><td>[4]</td></tr>
<tr><td>PG1</td><td>E</td><td>3.38E-3 (1/350)
(In comparison, 0/2010 for reading instructions in a written procedure, long procedure, checkoff provisions, task also part of professional knowledge)</td><td>Performing a manual control action at an MCR panel (Task omitted)</td><td>Long procedure, no checkoff provisions</td><td>Not analyzed</td><td>[5]</td></tr>
<tr><td>PGI2</td><td>Unsp</td><td>Y=difficulty, x=% of the level descriptions
(Figure A13-2)</td><td>Experts rated difficulty score for procedures</td><td>Percent of intermedium procedure description</td><td>(Subjective rating)</td><td>[226]</td></tr>
</table>

				(requiring judgment)		
PGI3	Unsp	Y=difficulity, x=% of the level descriptions (Figure A13-3)	Experts rated difficulty score for procedures	Percent of detailed procedure description	(Subjective rating)	[226]
PGI3	E	0.5 (1/2)	Testing electronic modules in the reactor protection system. (Signal plugs erroneously removed in all redundancies)	Lack of detailed instructions in procedures, rarely performed task, unfavorable ergonomic design of alarm indication	Rarely performed task, poor HSI	[5]
PGI4	Unsp	Y= difficulty, x=% of the level descriptions (Figure A13-4)	Experts rated difficulty score for procedures	Percent of problematic (confusing, ambiguous) procedure descriptions	(Subjective rating)	[226]
PGI2 / PGI4	Unsp	VPP values for EOPs (VPP value is proportional to time needed) LOCA: 47 SBO: 26 SGTR: 26 Loop: 17 General task: 1.77	Measuring variability of EOP progression (VPP)	VPP features: Task complexity, same task covered by contiguous alternative steps, conflict between steps, literal meaning and step intention, foldout use, assessment of trends, mental representation of the given situation, control modes of EOPs	(No error rate)	[227]
PGI4	E	2.9E-2 (1/40) (For comparison: 2.7E-3 for step in a procedure not read, task omitted in securing a valve in open position, long procedure with checkoff provisions)	Activation of both mid loop level measurement devices (One channel not activated, task description in procedure misinterpreted)	Ambiguous task description in procedure	Moderately high level of stress, infrequently performed (1/40)	[4]
PGI4	E	2.41E-3 (1/490)	Manually opening a locally operated valve (Opened too early, false interpretation of oral instruction)	Ambiguous oral instruction	Not analyzed	[5]
PGI5	E	1.05E-3 (1/112)	Plugging connectors to jacks in control cabinets (Connected to wrong jack, incorrect task generation)	Very error prone written instructions, recall of rarely used professional knowledge necessary	Infrequently performed (1/112)	[5]
PGI5	E	7.97E-3 (1/148)	Returning a power switch to operational condition at a local switchgear cabinet	Imprecise written procedure, rarely performed,	Infrequently performed	[5]

				Performing an inadmissible switching operation (false interpretation of written procedure	professional knowledge necessary for proper interpretation		
PGI6	E	No procedure	3.3E-2 (1/36)	Reassembly of component elements (Element position remembered incorrectly)	No written procedures available, similarity between correct and wrong position	(Infrequently performed (1/36))	[4]
		With procedure	1.3E-3 (1/888)				
PGI6	E	1/1		Closing pegging steam control valves after SCRAM (Not fully closed, error in task generation)	Special operating mode, no written procedure available, complex thermo-hydraulic context	Rarely performed, system not transparent	[5]
PGI6	E	1/1		Testing the 24 V DC power supply (Failed to check the presence of essential test system prerequisites)	No indication in written procedure, rarely used professional knowledge	Rarely performed, no mental model	[5]
PGI6	E	1/1		Start-up of reactor (Further increase of thermal power despite a lacking prerequisite)	Special operating mode, no written procedures available	Rarely performed	[5]

Question items	Mean values*		
	Computerized procedures	Paper procedures	Results of K-W test
Operation skills	0.49	0.53	N.S.
Situation awareness	0.29	0.13	N.S.
Workload	0.05	0.40	N.S.
Communication	0.52	0.72	N.S.
Coordination skills	0.53	0.90	$p<0.05$*

Figure A13-1 Operator performance statistics

Plant experience	Regression analysis (R^2)
less than 5 years	y = 0.014x + 2.204 (0.765)
5 to 10 years	y = 0.0176x + 2.079 (0.775)
10 to 15 years	y = 0.009x + 1.950 (0.603)
greater than 15 years	y = 0.013x + 2.046 (0.657)

Figure A13-2 Operator performance statistics

Plant experience	Regression analysis (R^2)
less than 5 years	y = -0.026x + 4.375 (0.687)
5 to 10 years	y= -0.027x + 4.438 (0.743)
10 to 15 years	y = -0.018x + 3.494 (0.610)
greater than 15 years	y= -0.021x + 3.821 (0.875)

Figure A13-3 Operator performance statistics

Plant experience	Regression analysis (R^2)
less than 5 years	y = 0.026x + 2.423 (0.957)
5 to 10 years	y = 0.025x + 2.549 (0.891)
10 to 15 years	y = 0.021x + 2.177 (0.992)
greater than 15 years	y = 0.017x + 2.380 (0.875)

Figure A13-4 Operator performance statistics

Appendix A14 PIF Attributes and Weights for Training

Table A14-1 Attribute Identifiers and Descriptions for PIF Training

ID	Attribute
TE0	No impact - professional staff have adequate training required
TE1	Inadequate training frequency/refreshment • Lack of or poor administrate control on training (e.g., not included in the Systematic Approach to Training Program) • Training frequency is longer than needed for retention of proficient knowledge/skills
TE2	Inadequate amount or quality of training
TE2.1	Inadequate training on skills and basic knowledge, deficient mental model of the systems
TE2.2	Inadequate training specification/requirement, deficient knowledge on rules and action control
TE2.3	Inadequate training in system processes for knowledge-based human actions
TE3	Deficient training practicality • No hands-on training • Not drilled together • Training on parts, not whole scenario together
TE4	Poor or lack of training on procedure adaptation: Training focuses on procedure-following without adequately training personnel to seek alternative interpretations, evaluate the pros and cons of alternatives, and adapt the procedure for the situation
TE5	Poor of lack of knowledge-based problem-solving training, e.g., • Inadequate training or experience with sources of information (such as applicability and limitations of data or the failure modes of the information sources) • Inadequate specificity on urgency and the criticality of key information such as key alarms • Not trained to seek confirmatory information when dismissing critical data • Premature termination of critical data collection in diagnosis due to inadequate training on system failure modes • Poor training on assessing action margin in deciding implementation delay • Poor training on interpreting procedure in the context of the scenario for decisionmaking • Poor training on the importance of data in frequently checking data for execution
TE6	Inadequate or ineffective training on teamwork
TE7	Personnel are fully trained but inexperienced (compared to expert-level experienced professionals)

Table A14-2 IDHEAS-DATA IDTABLE-14 – PIF Weights for Training

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates or task performance indicators**	**Task (and performance measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
TE1	E	Use FLEX generator With SAT: 0.036 NO SAT: 0.36	Use of FLEX generator: -Transport and stage -Connect and start -Operate (Estimated HEPs)	FLEX-designed scenarios, training is under SAT vs. not under SAT	(Expert judgment)	FLEX-EE
TE2	D	# of missed detection One type at a time / One number at a time Baseline: 1.4 / 0.8	16 non-operators - check numbers on an information heavy display screen and make sure that the numbers were inside specified safety ranges, otherwise the numbers need to be marked	HD - “highlight and disappear” increased the efficiency of task completion. HM- “highlight missed” increased the participants’	(May not be a training PIF)	[228]

		HD	1.6	0.8		confidence during the task "Heat map" is a training feedback tool by highlighting areas that need more attention.		
		HM	1.2	0.6				
		Heat map	0.9	0.35				
TE2.3	D	Perceived urgency high	9.6E-4		Key Alarm Not Attended To	Training on knowledge-based response – Perceived urgency of alarms high vs. low	HSI poor (expert judgment)	[6]
		Perceived urgency low	7.3E-3					
TE2	D	Training good	1.3E-5		Critical Data Misperceived	Training good vs. poor	(Expert judgment)	[6]
		Training poor	1.3E-4					
TE2.3	DM	Training good	3.2E-3		Misinterpret procedure in response planning	Training good vs. less than adequate	(Expert judgment)	[6]
		Less than adequate training	7.3E-2					
TE2.1	E	Good	1.3E-2		Critical data not checked with appropriate frequency for initiating execution	Training - Importance of data understood good vs. poor when monitoring is not optimized	(Expert judgment)	[6]
		Poor	3.2E-2					
TE2	E	Good	3.8E-03		Failure to correctly execute response (Complex task)	Training good vs. poor	(Expert judgment)	[6]
		Poor	5.1E-2					
TE2.1	E	1.13E-4 (0/2010)			Adjusting a process parameter by push-button controls (Operated too long)	Training has no negative impact - frequently performed task, part of professional knowledge	(Error rate for a single step)	[5]
TE2.3	E	9.86 E-4 (1/1200)			Testing the emergency feedwater supply system during power operation (Wrong order of steps)	Control actions appear in wrong order in written procedure, proper ordering was to be inferred from professional knowledge, frequently performed task	(Error rate for a single step)	[5]
TE2 / TE6	DM / E	# of errors			Crews performed two scenarios. The difficult LCOA transient was "partial breakdown of plate fixing bolts of the primary manifold of the steam generator" (# of errors all the crews made in each scenario)	Deficiencies in knowledge and action control, problems related to procedures, collective operational strategy	Whole event scenario)	[229]
		Leak in the live steam manifold	57 errors (by 8 crews)					
		Difficult LOCA-transient	155 (by 12 crews)					
TE2.1	D	Without KR	0.2		Students detect rare targets	With CKB – with feedback of composite knowledge of results Without KR – without knowledge of results	(Student subjects)	[230]
		With CKR	0.1					
TE2.1	D	Day1 start	0.58			Beginning and end of day1 and day2 training	(Inadequate	[231]
		Day1 end	0.41					

<table>
<tr><td></td><td></td><td><table><tr><td>Day2 start</td><td>0.48</td></tr><tr><td>Day2 end</td><td>0.32</td></tr></table></td><td>Nurses intensively trained on discriminating sounds of alarms</td><td></td><td>amount of training)</td><td></td></tr>
<tr><td>TE1 & TE2.1</td><td>D</td><td><table><tr><td></td><td>1B</td><td>1A</td><td>2B</td><td>2A</td></tr><tr><td>MMS</td><td>32</td><td>88</td><td>44</td><td>97</td></tr><tr><td>LOCA</td><td>0.14</td><td>0</td><td>NA</td><td>NA</td></tr><tr><td>SGTR</td><td>0.45</td><td>0.14</td><td>0.28</td><td>0.04</td></tr><tr><td>SLB</td><td>0.44</td><td>0.1</td><td>0.35</td><td>0.16</td></tr></table></td><td>15 graduate students with nuclear engineering backgrounds of 5.2 years performed 14 tasks in three scenarios (MMS – mental model score, error rates of failing detection)</td><td>1B-before training
1A-after training
2B- 6 months later before training
2A – 6 months later after training</td><td>(Not licensed operators)</td><td>[74]</td></tr>
<tr><td>TE1 & TE2 & TE3</td><td>E</td><td>See Figure A14-1</td><td>Engineering students trained to perform process system control (% control failures)</td><td>T_0 – Test immediately after training
T2w – Test 2 weeks after training
T6w– Test 6 weeks after training
Three training methods:
EST – emphasize knowledge
EST/SA – EST + situational awareness
P&D – Practice and drills</td><td>(Not licensed operators)</td><td>[232]</td></tr>
<tr><td>TE1 & TE2 & TE3</td><td>U</td><td>See figure A14-2</td><td>Engineering students trained to diagnose system faults (%Diagnostic errors)</td><td>T0 – Test immediately after training
T2w – Test 2 weeks after training
T6w– Test 6 weeks after training
Three training methods:
EST – Emphasize knowledge
EST/SA – EST + situational awareness
P&D – Practice and drills</td><td>(Not licensed operators)</td><td>[232]</td></tr>
<tr><td>TE5 /TE6</td><td>DM / T</td><td><table><tr><td></td><td>VP</td><td>PM</td></tr><tr><td>Low</td><td>0.26</td><td>0.23</td></tr><tr><td>Nominal</td><td>0.15</td><td>0.13</td></tr><tr><td>High</td><td>0.12</td><td>0.1</td></tr></table></td><td>This study used SPAR-H to evaluate HEPs of action and diagnosis in the Fukushima Daiichi accident management model (HEP evaluated with SPAR-H)
VP- Vice president
PM-plant manager</td><td>Training level
Low: inadequate practice in tasks with abnormal conditions.
Normal: more than 6 months of relevant training in tasks with abnormal conditions.
High: Training with extensive knowledge and practice in a wide range of potential scenarios.</td><td>(SPAR-H based assessment)</td><td>[233]</td></tr>
</table>

TE5	D / E	Training: # of errors N: 3.31 V: 2.84 S: 2.59	144 students performed 4 simulator flight tasks of tracking and monitoring (# of errors out of 6 omission and 6 commission opportunities)	Training types: N - Normal training V – Trained on verifying automation S – Specific explicitly trained about automation bias	Multitasking	
TE5	U	Solo pilots: 0.52 Crews: 0.43 • Training types had no significant effect on error rates • Crew only reduced automation bias slightly compared to solo pilots	20 2-person pilot crews and 8 solo pilots performed flight simulation in which 6 automation failures would result in omission errors if pilots did not verify automation function	Training types: N - Normal training V – Trained on verifying automation S – Specifically trained on automation bias	Multitasking	[234]
TE5	Unsp.	Unexperienced: 0.002 (errors/ trails) Experienced: 0.008 (errors/ trails)	Licensed operators use APR1400 simulator to perform 6 scenarios varying as DBA, DBA+masking, and BDBA (Errors/trials, not error rate in percent.))	Experienced operators: APR1400 and other types of PWRs Unexperienced - No APR1400, but other types of PWRs	(Whole scenario)	[99]
TE7	Unsp.	Years of experience: Subjective error rate 0-2: 0.8 2-6: 0.37 6-20: 0.2 >20: 0.07	Operation error in electric utilities	Years of experience	(Survey results)	[235]
TE7	D	GNP proportion (Expert / Novelty) Normal: 0.48 / 0.52 One malfunction: 0.32 / 0.18 Multiple malfunctions: 0.14 / 0.05	Astronaut experts and novelty perform flight simulator Space Shuttle (GNP proportion is the proportion of eye fixation time on navigation display vs on systems)	Expert vs novelty	(Small sample)	[159]
TE2.3	U	# of correct diagnosis (OLD / NEW) No story: 7.7 / 2.5 Theory: 7.8 / 3.5 Rules: 7.6 / 5.5	Training for fault diagnosis in the chemical process plant area (# of correct diagnosis) "NEW" for new faults not previously seen by the operators during practice	"Theory" and the "rules" groups were given a simplified account of how the plant worked in addition, the "rules" group exercised in applying diagnostic rules, "No story" group received no prior instruction of either sort	(Other PIFs may exist)	[73]
TE1	U	Diagnosis test score (0-100) after training 1st test: 89%	Operator trainees performed static diagnostic tests (Diagnosis test score 0-100)	Four tests immediately after training and one 5 months later	(Not dynamic	[73]

		2nd test	99%			scenarios)	
		3rd test	99%				
		4th test	99%				
		After 5 months	73%				

System control failures (in percentages) as a function of training and fault type (SD in parentheses).

Fault type	Drill and practice			EST			EST/SA		
	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}
Practiced	11.6 (9.0)	3.5 (2.3)	8.6 (5.5)	9.5 (8.0)	4.0 (6.0)	6.4 (4.0)	17.6 (14.1)	5.1 (5.7)	9.8 (6.1)
Novel	8.8 (6.9)	7.4 (6.3)	17.5 (3.6)	6.6 (2.3)	4.3 (3.1)	15.5 (5.3)	8.4 (4.6)	6.2 (4.9)	15.7 (6.9)

Figure A14-1 System failures as a function of training and fault type.

Diagnostic performance as a function of training and fault type (SD in parentheses).

Fault type	D&P			EST			EST/SA		
	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}	T_0	T_{2w}	T_{6w}
Diagnostic errors (%)									
Practiced	21.5 (12.8)	51.3 (17.3)	33.3 (19.2)	40.0 (29.4)	59.0 (30.9)	41.0 (30.9)	50.0 (35.7)	57.1 (27.5)	52.4 (31.3)
Novel	46.2 (24.7)	51.3 (35.0)	46.2 (32.0)	38.5 (30.0)	51.3 (25.9)	38.5 (18.5)	53.6 (23.7)	52.4 (38.6)	35.7 (33.2)

Figure 14 -2 Diagnosis performance as a function of training and fault type

Appendix A15 PIF Attributes and Weights for Team and Organization Factors

Table A15-1 Attribute Identifiers and Descriptions for PIF Team and Organization Factors

ID	Attribute
TF0	No impact – adequate, crew-like teams
TOF1	Inadequate team • Deficient teamwork structure, e.g., knowledge gaps in the team, deficient reconciling viewpoints, deficient team monitoring, lack of adaptability • Distributed or dynamic teams • Poor team cohesion (e.g., newly formed teams, lack of drills/experience together)
TOF2	Poor command & control with problems in coordination or cooperation • Ambiguous specifications of function, responsibilities, and authorization for personnel in the command & control • Inadequate coordination between site personnel and decision-makers (e.g., adapt or modify planned actions based on site situation) • Inadequate verifying the plan with decision-makers • Inadequate overseeing action execution and questioning current mission
TOF3	Poor communication infrastructure
TOF4	Poor resource management, e.g., managing competing resources among multiple entities involved in an event
TOF5	Poor safety culture
TOF5.1	Deficient practice (e.g., pre-job briefing) for personnel to be aware of potential pitfalls in performing the tasks
TOF5.2	Deficient practice for safety issue monitoring and identification, e.g., no regular inspection
TOF5.3	Deficient practice for safety reporting
TOF5.4	Hostile work environment

Table A15-2 IDHEAS-DATA IDTABLE-15 – PIF Weights for Team and Organization Factors

1 PIF	2 CFM	3 Error rates		4 Task (and error measure)	5 PIF measure	6 Other PIFs (and Uncertainty)	7 REF
TOF 1	Unsp.	Agreeableness (1-10) 4 7 9	Team performance Poor Average Excellent	16 NPP crews performed the same scenario "*Failure of one turbine unit*" (Teamwork measures)	Agreeableness – reconciling different viewpoints	(No error data)	[236]
TOF-1	Unsp.	Team performance Poor Average Unbalance Excellent	Freq. of OIQ 22 14 9 6	16 NPP crews performed the same scenario "*Failure of one turbine unit*"	Open Information Question (OIQ) – having less information, knowledge about the cues in the scenario	(No error data)	[236]
TOF-1	Unsp.	Team performance Poor	% Coherent communication 65%	16 NPP crews performed the same scenario	Coherent communication means that the	(No error data)	[236]

		Excellent 90%	"*Failure of one turbine unit*".	team members are aware of the information distributed by others, and react to the received information, creating a semantic connection in the information sharing activity		
TOF 1	Unsp	Team (S/F) / OPAS / LSC / CI #1 (F) / 20 / 0 / n/a #2 (F) / 46 / 0.26 / 0.43 #3 (F) / 42 / 0.19 / 0.37 #4 (F) / 47 / 0.26 / 0.32 #5 (F) / 51 / 0.30 / 0.37 #6 (F) / 40 / 0.13 / 0.18 #7 (S) / 45 / 0.32 / 0.36 #8 (S) / 55 / 0.28 / 0.50 #9 (S) / 63 / 0.43 / 0.48	9 crews performed LOCA scenario (OPAS-Operator Performance Assessment Score and success/failure (S/F) of the scenario)	LSC - each member's SA weighted by numbers of team members who share confidence, Consensus Index (CI) – degree of consensus in team decisionmaking	(No error data)	[237]
TOF 1	Unsp.	See Figure A15-1 OPAS score= 37.044x (team cohesion) +14.221 OPAS score= 47.826x(amount of communication) +30.553	9 crews performed ISLOCA scenario (OPAS-Operator performance assessment score)	Team cohesion and amount of communication within the team	(Not analyzed)	[238]
TOF 1	Unsp	Correlation of TC or RC with performance / TC / RC Decision-making / -.16 / -.33 Project / -.22 / -.15 Production / .03 / -.07 Multiple types / -.35 / -.31	50+ studies: production tasks -overt task execution while striving to meet standards, decisionmaking tasks - require reaching consensus on issues with no right answer, project tasks include a variety of group tasks (The most uncertain, most complex, or least routine.)	Team relationship conflict (RC) Task conflict (TC)	(Meta-analysis of 50+ papers)	[75]
TOF-2	D	Correlation with detection errors Coordination problems / 0.17 Cooperation problems / 0.49	NPP operator expert evaluation of effect on operator performance in outage and normal operation	Coordination problems with planners Cooperation problems with work permit managers	(Subjective evaluation)	[239]

<table>
<tr><td>TOF-2</td><td>U</td><td>Correlation with misinterpretation errors<table><tr><td>Coordination problems</td><td>0.48</td></tr><tr><td>Cooperation problems</td><td>0.56</td></tr></table></td><td>Operator expert evaluation of effect on operator performance in outage and normal operation</td><td>Coordination problems with planners
Cooperation problems with work permit managers</td><td>(Subjective evaluation)</td><td>[239]</td></tr>
<tr><td>TOF 2</td><td>Unsp</td><td><table><tr><td># of errors made</td><td>Detection</td><td>Diagnosis</td><td>DM</td><td>Execution</td></tr><tr><td>A</td><td></td><td></td><td></td><td>1</td></tr><tr><td>B</td><td></td><td>3</td><td>2</td><td>1</td></tr><tr><td>C</td><td></td><td></td><td>3</td><td></td></tr></table></td><td>Experiment 1: NPP crews performed LOCA and SGTR scenarios (# of errors made)</td><td>A-insufficient concentration
B-insufficient communication
C-unclear division of tasks</td><td>(Not error rates)</td><td>[73]</td></tr>
<tr><td>TOF 2</td><td>Unsp</td><td><table><tr><td># of errors made</td><td>Detection</td><td>Diagnosis</td><td>DM</td><td>Execution</td></tr><tr><td>A</td><td></td><td></td><td>1</td><td></td></tr><tr><td>B</td><td></td><td></td><td>7</td><td>4</td></tr><tr><td>C</td><td></td><td></td><td></td><td></td></tr><tr><td>D</td><td></td><td></td><td>11</td><td></td></tr></table></td><td>Experiment 2: NPP crews performed LOCA and SGTR scenarios in training. Experiment 1 transient is more straightforward and physically transparent than experiment 2 transient. (# of errors made)</td><td>A-insufficient concentration
B-insufficient communication
C-unclear division of tasks
D-lack of operational strategy</td><td>(Not error rates)</td><td>[73]</td></tr>
<tr><td>TOF 2</td><td>U & DM</td><td>Error distribution (%) regard to causes<table><tr><td></td><td>Exp1</td><td>Exp2</td></tr><tr><td>HSI</td><td>5.5%</td><td>3.2%</td></tr><tr><td>PGI</td><td>31%</td><td>17%</td></tr><tr><td>TOF</td><td>13.7%</td><td>16%</td></tr><tr><td>TRI</td><td>34.5%</td><td>55%</td></tr><tr><td>Simulator effect</td><td>11%</td><td>7.7%</td></tr></table></td><td>NPP crews performed LOCA and SGTR scenarios in training. Experiment 1 transient is more straightforward and physically transparent than Experiment 2 transient. (Error distribution regard to causes)</td><td>HSI - Control room layout
PGI - Procedure
TOF - Cooperation
TRI - Knowledge and action control</td><td>(Operators in training)</td><td>[73]</td></tr>
<tr><td>TOF 5</td><td>Unsp.</td><td>See Figure 15-2</td><td>Correlation of 4 NPP crews simulator performance (five abnormalities in the scenario) and State of team safety culture index</td><td>Safety culture elements:
IA - Operation Information Acquisition
PA - Personal Accountability
RC - Respectful Cooperation
NU- Recognition of Nuclear as Unique Technology</td><td>(Subjective assessment of safety culture)</td><td>[240]</td></tr>
</table>

						CD - Conservative Decision Making QA - Questioning Attitude RI - Regular Inspection CL - Continuous Learning		
TOF-5	Unsp.		*Intercept of Poisson Regression with # of errors*	Correlation with # of errors	Safety culture assessment and treatment errors of 123 residents from 25 medical wards	1.Year of residency 2.Level of fatigue 3. Active learning climate 4. Priority of safety 5. Managerial safety practices	(Subjective assessment of PIF attribute)	[241]
		1	0.10	0.10				
		2	0.49	0.07				
		3	-0.90	0.17				
		4	1.32	-0.06				
		5	2.10	0.14				

Two kinds of communication characteristics and the associated crew performance scores collected under ISLOCA 1 scenario.

ID	Success	Amount of communications	Crew cohesion	Crew performance
Crew 1	No	0.294	0.857	48
Crew 2	No	0.268	0.714	42
Crew 3	No	0.346	1.095	48
Crew 4	Yes	0.464	0.905	50
Crew 5	No	0.196	0.857	47
Crew 6	No	0.137	0.524	31
Crew 7	Yes	0.543	1.000	57
Crew 8	No	0.294	0.857	43

Figure A15-1 Two kinds of communication characteristics and the associated crew performance scores collected under ISLOCA 1 scenario

Table 4. Probability of Being Success

	Normal Probabilities of Each Category						Probability of Being Safe State of Team Safety Culture	Operational Performance Assessment System (OPAS) Score
	IA	PA	RC	NU	CD	QA		
Crew1	0.635	0.311	0.459	1.000	1.000	0.800	0.097	39
Crew2	1.000	0.611	0.556	1.000	1.000	0.875	0.340	68
Crew3	0.882	0.818	0.800	1.000	0.818	0.857	0.472	74
Crew4	1.000	1.000	0.846	1.000	0.612	0.923	0.518	83

Figure 15-2 Probability of being successful

Appendix A16 PIF Attributes and Weights for Work Processes

Table A16-1 Attribute Identifiers and Descriptions for PIF Work Processes

ID	Attribute
WP0	No impact – Professional licensed personnel with good work practices
WP1	Lack of professional self-verification or cross-verification (e.g., 3-way communication), peer-checking, independent checking or advising, or close supervision
WP2	Poor attending to task goals, individual's roles, or responsibilities, e.g., • Poor practice of attending to the task goals (so personnel disengages from the goal too early) • Poor practice of keeping personnel in assigned roles and responsibilities • Excessive disturbance to planned work and assigned responsibilities • Bad shift handovers
WP3	Poor infrastructure or practice of overviewing operation information or status of event progression
WP4	Poor work prioritization, planning, scheduling, e.g., • Poor planning of work permits • Many extra instructions regarding task prioritization and scheduling • The purpose and object of the work permit was not specified • Work permits were not handed in on time and, therefore, delayed other activities • Indistinct information concerning the prioritization of different work activities • Insufficient information in operational order concerning performance of tasks

Table A16-2 IDHEAS-DATA IDTABLE-16 – PIF Weights for Work Processes

1	2	3	4	5	6	7
PIF	**CFM**	**Error rates**	**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
WP-1	D	Diagnostic error in radiology reported to be as high as 20% Double reading increases sensitivity around 10% (10–14%, 9%, 15% and 8.1%, 9%, 9.5% in different literature) Single reporting: 0.3 Double reporting: 0.17	Diagnosis in radiology images (diagnosis error rate)	Independent double reporting, where personnel have no knowledge of each other's report	Meta-analysis	[242]
WP 1	D	(LSOD / WSD) Individual: 0.15 / 0.07 Team: 0.06 / 0.04	NPP operators performed microtask detecting information from GPWR simulator	Individual vs team – team has independent checking LSOD – Large overview display WSD – Workstation display	(Microtasks were mixtures of complexity, speed-accuracy biased toward speed, no recovery)	[243, 244]
WP 1 & WP 4	D, E	PIF weights calculated E - Training and supervision: 14.21 (Dynamic) D/E - Supervisi: 6.78 (Value) 32.44	NPP operator performance data in low power and shut down (LPSD) Synthetical - Synthetically verifying information	PIF weights calculated: WP1 - Supervision WP4 - Task planning	(PIF definition and weight calculation may not be the same as those in	[138]

		on and task planning – (Synthetical) E - Task planning – 4.57 (Dynamic) 12.00 (Synthetical) D/E - Supervision – 5.53 (Dynamic) 2.70 (Synthetical)	Value – Reading simple value Dynamic - Manipulating dynamically		IDHEAS-DATA)	
WP 2	E	5.04E-3 (1/324)	Refilling nitrogen to SCRAM accumulator (Inadmissible control action performed)	Time consuming procedure, operator intended to save time by departing from procedure	Not analyzed	[5]
WP 2	DM	Advantage to the correct strategy - Yes – 9.3E-03 Advantage to the correct strategy – No – 3.3E-02	Choose inappropriate strategy in procedure-based decisionmaking	Advantage to the correct strategy – operators more likely attend to rules	(Expert judgment)	[6]
WP 1	E	Monitoring Optimized – 2.3E-3 Monitoring NOT Optimized – 1.3E02	Critical data not checked with appropriate frequency in action execution	Monitoring optimized (verification or peer-checking) vs. not optimized	(Expert judgment)	[6]
WP 1 / WP 2	E	Good work practice – 8.0E-05 Poor work practice – 8.0E-04	Failure to correctly execute response (Complex task)	Good vs poor work practice	(Expert judgment)	[6]
WP 2	Unsp.	Vertical axis shows the times the personnel was seated vs unseated in the scenario (Figure A16-1)	Behavior observation of 5 NPP crews running startup and shutdown scenarios	Operation positions seated vs unseated (Away from his working panel)	(No error data)	[245]
WP 2	Unsp.	– / OPAS / Comm / minute Free seating / 57 / 1.05 Fixed seating / 74 / 2.75	NPP crews performed 2 normal and 2 emergency scenarios (OPAS-Operator Performance Assessment Score and Comm- total communication per minute)	Two seatings: Free - moved freely Fixed - remained seated at workstation, restricted movement except RO	(HSI automation was used in experiment)	[77]
WP 3	Unsp.	– / Without IPad / With IPad SS effort / 0.75 / -0.75 SS readiness / 0.5 / -0.5 SS-PUS / 5.8 / 6.25 FO-PUS / 2.6 / 5.5	NPP crews performed 2 normal scenarios (effort and readiness, process understanding score (PUS 1-10))	SS (Shift supervisor) and FO (Field operator) with vs without IPad for overview *of* process information	(HSI automation was used in experiment)	[77]

<table>
<tr>
<td>WP 2 & WP 3</td>
<td>Unsp.</td>
<td>Average # of errors per scenario
(For each scenario: SGTR=3.3, ISLOCA=3.6, LOFW=5.9)
<table>
<tr><td></td><td>No Tool</td><td>Tool</td></tr>
<tr><td>No STA</td><td>3.2</td><td>5.6</td></tr>
<tr><td>STA</td><td>4.9</td><td>4.2</td></tr>
</table>
</td>
<td>9 NPP crews run SGTR, LOCA, and LOFW scenarios (Average # of errors on required safety-important actions)</td>
<td>WP2: With or without STA,
WP3: With or without decision-support tools (Displaying important plant information)</td>
<td>(Interaction between STA independence and use of tools)</td>
<td>[246]</td>
</tr>
<tr>
<td>WP 2 & WP 3</td>
<td>Unsp.</td>
<td>
<table>
<tr><td></td><td>No Tool</td><td>Tool</td></tr>
<tr><td>No STA</td><td>2.8</td><td>4.3</td></tr>
<tr><td>STA</td><td>4.7</td><td>3.8</td></tr>
</table>
</td>
<td>NPP crews run EOP scenarios (average # of errors)</td>
<td>WP2: With or without STA,
WP3: With or without decision-support tools</td>
<td>(Interaction between STA independence and use of tools)</td>
<td>[246]</td>
</tr>
<tr>
<td>WP 2 & WP 3</td>
<td>Unsp</td>
<td>
<table>
<tr><td></td><td># errors</td><td>TESA-Op score</td><td>ScPerf score</td></tr>
<tr><td>No-S No-T</td><td>2.8</td><td>5.3</td><td>7.8</td></tr>
<tr><td>No-S Yes-T</td><td>4.3</td><td>4.9</td><td>6.5</td></tr>
<tr><td>Yes-S No-T</td><td>4.7</td><td>4.9</td><td>6.5</td></tr>
<tr><td>Yes-S Yes-T</td><td>3.8</td><td>5.3</td><td>7.3</td></tr>
</table>
</td>
<td>NPP crews performed EOP scenarios
(# of errors,
TESA-Op - Emergency Operation Handling Score,
ScPerf - # of important actions completed)</td>
<td>WP2: With or without STA,
WP3: With or without decision-support tools</td>
<td>(Interaction between STA independence and use of tools)</td>
<td>[247]</td>
</tr>
<tr>
<td>WP 2 & WP 3</td>
<td>Unsp.</td>
<td>65 unsafe acts observed in 5 crews running 3 emergency scenarios for about 2-3 hours after the initiating event.
13 unsafe acts were recovered, but in 7 cases the recovery did not avoid negative consequences to the plant or operational problems (e.g., delay). This means an average of about 4 unrecovered unsafe acts per scenario</td>
<td>5 crews performed four EOP scenarios representing the emergency response phase in which the control room team is expected to manage the accident without external technical support</td>
<td>Observation study, no independent variables, error narratives described with poor work process (WP2 and WP3)</td>
<td>(Errors reported individually)</td>
<td>[106]</td>
</tr>
<tr>
<td>WP 2</td>
<td>D</td>
<td>
<table>
<tr><td></td><td>Type1</td><td>Type4</td></tr>
<tr><td>Promotion</td><td>0.124</td><td>0.67</td></tr>
<tr><td>Prevention</td><td>0.27</td><td>0.35</td></tr>
</table>
</td>
<td>Proof reading (Missing targets)
Type 1 – easy targets
Type 4 -difficult targets</td>
<td>Promotion frame -

Regulatory prevention frame -</td>
<td>Time constrained (time-accuracy trade-off)</td>
<td>[248]</td>
</tr>
<tr>
<td>WP 2</td>
<td>D</td>
<td>
<table>
<tr><td># of errors</td><td>Beginning</td><td>Toward end of work</td></tr>
<tr><td>Promotion frame</td><td>7</td><td>8.5</td></tr>
<tr><td>NO frame</td><td>5.8</td><td>5</td></tr>
<tr><td>Regulatory prevention frame</td><td>4</td><td>1</td></tr>
</table>
</td>
<td>Proof reading (# of errors found) at the beginning and toward the end of the work</td>
<td>Promotion frame-

Regulatory prevention frame</td>
<td></td>
<td>[248]</td>
</tr>
<tr>
<td>WP 2</td>
<td>D</td>
<td>Correlation with detection errors
<table>
<tr><td></td><td>Outage</td><td>Normal</td></tr>
</table>
</td>
<td>NPP crews performed full scenario simulations</td>
<td>Item 12: Handover some of own work</td>
<td>(Subjective rating)</td>
<td>[239]</td>
</tr>
</table>

		Item 12	0.24	0.02	in outage (O) and normal (N) operation	to colleagues on shift Item 13: Decreased aspiration level Item 14: Leave work tasks to the next shift		
		Item 13	0.5	0.34				
		Item 14	0.15	0.10				
WP 2	U	Correlation with detection errors			NPP crews performed full scenario simulations in outage (O) and normal (N) operation	Item 12: Hand over some of own tasks to colleagues on shift Item 13: Decreased aspiration level Item 14: Leave work tasks to the next shift	(Subjective rating)	[239]
			Outage	Normal				
		Item 12	0.52	0.30				
		Item 13	0.47	0.44				
		Item 14	0.45	0.39				
WP 4	D	Correlation with detection errors			NPP crews performed full scenario simulations in outage (O) and normal (N) operation	Item 3: Planning problems Item 4: Work distributions	(Subjective rating)	[239]
			Outage	Normal				
		Item 3	0.35	NA				
		Item 4	0.34	0.22				
WP 4	U	Correlation with detection errors			NPP crews performed full scenario simulations in outage (O) and normal (N) operation	Item 3: Planning problems Item 4: Work distributions	(Subjective rating by observing simulations, N=90)	[239]
			Outage	Normal				
		Item 3	0.44	NA				
		Item 4	0.40	0.46				

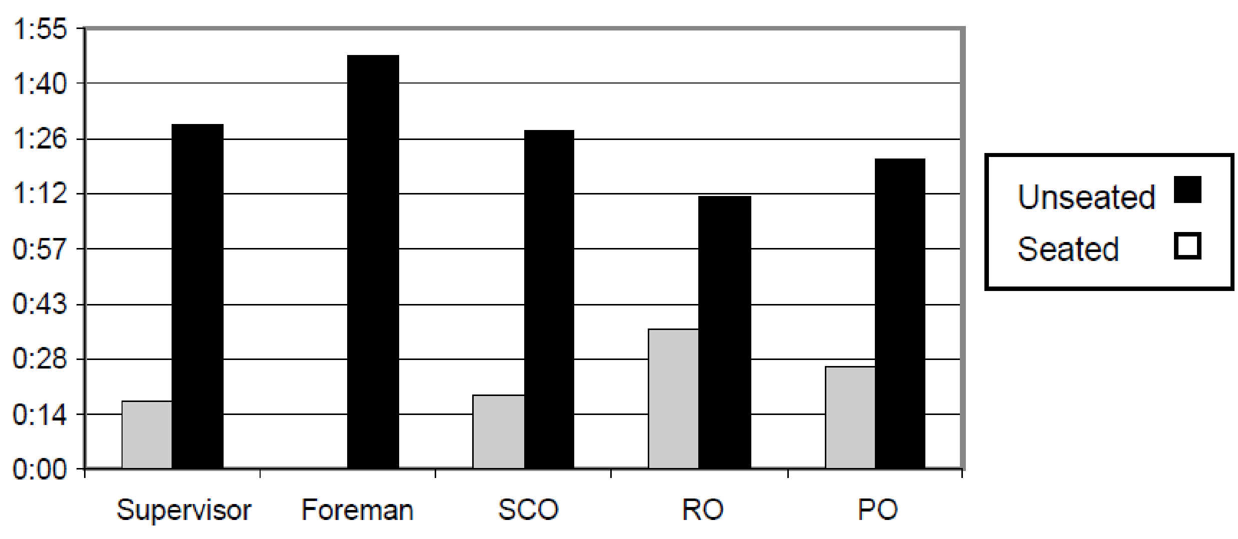

Figure A16-1 Variation in the operator positions during a change of the plant state

Appendix A17 PIF Attributes and Weights for Multitasking, Interruptions, and Distractions

Table A17-1 Attribute Identifiers and Descriptions for PIF Multitasking, Interruptions, and Distractions

ID	Attribute
MT0	No impact
MT1	Distraction by other on-going activities that demand attention
MT2	Interruption taking away from the main task
MT3	Concurrent visual detection and other tasks
MT4	Concurrent auditory detection and other tasks
MT5	Concurrent diagnosis and other tasks
MT6	Concurrently making two or more simple decisions/plans
MT7	Concurrently making intermingled complex decisions/plans
MT8	Concurrently executing action sequence and performing another attention/working memory task
MT9	Concurrently executing intermingled or inter-dependent action plans
MT10	Concurrently communicating or coordinating multiple distributed individuals or teams

Table A17-2 IDHEAS-DATA IDTABLE-17 – PIF Weights for Multitasking, Interruptions, and Distractions

1	2	3			4	5	6	7
PIF	CFM	Error rates			Task (and error measure)	PIF measure	Other PIFs (and Uncertainty)	REF
MT1	D	No distraction	0.025		Target detection in driving with cellphone conversation (Missing dangerous targets)	Distraction by other on-going activities (e.g., cell phone conversation)	No apparent uncertainty	[249]
		With distraction	0.07					
MT1	D	No distraction	0.07		Young adults performed detection of low meaningfulness stimuli	With or without distraction	No apparent uncertainty	[142]
		With distraction	0.14					
MT1	E		Cell phone	Radio control	Pursuit tracking task in which participants used a joystick to maneuver the cursor on a computer display to keep it aligned as closely as possible to a moving target and press the “brake” button when an alert appeared	Without or with cell phone conversation through a microphone, without and with radio control while listening to the radio	No apparent uncertainty	[249]
		Without	0.028	0.035				
		With	0.07	0.04				
MT1	D/E	No distraction	0.025		Driving and target detection	Distraction - auditory detection	No apparent uncertainty	[250]
		Distraction	0.05					
MT1	D/E	Accuracy ratio with distraction = 0.8 ~ 1 (distraction reduces error rates)			Driving, navigating, cognitive tasks - low complexity	Auditory - tactile	(Meta-analysis)	[251]

<table>
<tr><td></td><td></td><td colspan="3"></td><td>and/or low urgency tasks</td><td></td><td></td><td></td></tr>
<tr><td>MT1</td><td>D/E</td><td colspan="3">Accuracy ratio with distraction = 1~1.5</td><td>Driving, navigating, cognitive tasks - high complexity and/or high urgency tasks</td><td>Auditory - tactile</td><td>(Meta-analysis)</td><td>[251]</td></tr>
<tr><td>MT1</td><td>D/E</td><td colspan="3">Accuracy ratio with distraction = 0.8 ~ 1.2</td><td>Driving, navigating, cognitive tasks</td><td>Auditory - visual</td><td>(Meta-analysis)</td><td>[251]</td></tr>
<tr><td>MT1</td><td>D/E</td><td colspan="3">Accuracy ratio with distraction = 2~3</td><td>Driving, navigating, cognitive tasks</td><td>Auditory and visual redundant high complexity and/or high urgency distraction</td><td>(Meta-analysis)</td><td>[251]</td></tr>
<tr><td rowspan="2">MT1</td><td rowspan="2">D/U</td><td colspan="2">Without datalink</td><td>0.33</td><td rowspan="2">Flying simulator (% missing voice clearance)</td><td rowspan="2">Datalink is a distraction to pilots</td><td rowspan="2">(Average from four studies)</td><td rowspan="2">[199]</td></tr>
<tr><td colspan="2">With datalink</td><td>0.69</td></tr>
<tr><td rowspan="4">MT1 / MT2</td><td rowspan="4">DM</td><td colspan="3">Decision accuracy - Z-score deviation from the optimal decision, lower score means higher accuracy)</td><td rowspan="4">Production management, simple task - scheduling workloads on multiple machines, complex task - involved interrelated outcomes where the processing of one part of the task influences processing of another part of the task (Decision deviation)</td><td rowspan="4">Interruption – answering a question by acquiring information</td><td rowspan="4">(Short interruption to complex task may be distraction)</td><td rowspan="4">[252]</td></tr>
<tr><td></td><td>Simple task</td><td>Complex task</td></tr>
<tr><td>No interruption</td><td>0.18</td><td>0.13</td></tr>
<tr><td>Interruption</td><td>0.29</td><td>0.08</td></tr>
<tr><td rowspan="6">MT2</td><td rowspan="6">DM</td><td colspan="3">Z-score of decision accuracy</td><td rowspan="6">Production management complex task</td><td rowspan="6">Interruption frequency and content similarity with the primary task</td><td rowspan="6">(Short interruption to complex task may be distraction)</td><td rowspan="6">[252]</td></tr>
<tr><td colspan="2">No interruption</td><td>0.13</td></tr>
<tr><td colspan="2">Lo. frequency</td><td>0.22</td></tr>
<tr><td colspan="2">Hi. frequency</td><td>0.05</td></tr>
<tr><td colspan="2">Similar content</td><td>0.12</td></tr>
<tr><td colspan="2">Diff content</td><td>0.05</td></tr>
<tr><td rowspan="3">MT2</td><td rowspan="3">D/U/E</td><td></td><td>With notes</td><td>No notes</td><td rowspan="3">Professionals watched an interview video then tested on 25 questions of the interview (%incorrect)</td><td rowspan="3">Without or with interruption during watch, interview is interrupted from 3 min 50 sec to 4 min 30 sec by a secretary giving the interviewer a letter to sign and then leaving the room</td><td rowspan="3">(Brief irrelevant interruption on non-sequential task)</td><td rowspan="3">[253]</td></tr>
<tr><td>No interrup</td><td>0.21</td><td>0.26</td></tr>
<tr><td>interrup</td><td>0.26</td><td>0.28</td></tr>
<tr><td rowspan="3">MT2</td><td rowspan="3">D/U/E</td><td></td><td>Add</td><td>Count</td><td rowspan="3">Primary tasks are adding numbers or counting (% errors made)</td><td rowspan="3">Without and with interruption of reading comprehension or reasoning in the middle of the primary task</td><td rowspan="3">(Maybe distraction)</td><td rowspan="3">[254]</td></tr>
<tr><td>No interrup.</td><td>0.15</td><td>0.1</td></tr>
<tr><td>Interrup</td><td>0.35</td><td>0.19</td></tr>
<tr><td>MT2</td><td>U</td><td colspan="2">No interrup.</td><td>0.04</td><td>Primary tasks are reading</td><td>Without and with interruption of</td><td>(Maybe distraction)</td><td>[254]</td></tr>
</table>

		Interrup	0.12		comprehension (%errors made)	reading comprehension or reasoning		
MT2	E	No interrup.	0.08		Primary tasks are selecting items from a list (%errors made)	Without and with interruption of reading comprehension or reasoning	(Maybe distraction)	[254]
		Interrup	0.16					
MT2	E	No interruption	0.05		Physicians continuous performance of critical tasks	Excessively frequent or long interruption	(Other PIFs may exist)	[122]
		Interruption	0.2					
MT2	E	No Interrup.	0.15		Performing sequence of action steps	Interruption duration – no, 2.8s, 4.4s	No apparent uncertainty	[255]
		2.8s	0.3					
		4.4s	0.45					
MT2	E		Seq. error	Non-seq. error	Performing sequence of action steps, Sequence errors defined as the proportion of trials on which the performed step was not the immediate successor to the step performed on the previous trial, Nonsequence errors defined as the proportion of trials on which the correct step was selected but the incorrect choice was made given the stimulus	Position after interruption:1, 2, or 3 sequence steps after interruption.	No apparent uncertainty	[255]
		Interrup.	0.06	0.03				
		1 Step after	0.02	0.03				
		2 steps after	0.02	0.03				
		3 steps after	0.02	0.03				
MT2	E	Without interrp.	0.04		Military actions involving computer file operation and other procedural tasks	With vs without interruption		[256]
		With interruption	0.08					
MT2	DM		Simple	Comp	Simple vs complex decisionmaking	Without and with interruption on simple and complex decisionmaking tasks	(Other PIFs may constantly exist)	[102]
		Without Interrup.	0.08	0.18				
		With Interrup.	0.13	0.29				
MT2	D		No interrup.	With Interrup.	Recognizing simple and complex visual patterns	Weak (very short) interruption	(Maybe distraction)	[145]
		Simple Symbolic	0.26	0.23				
		Simple Spatial	0.27	0.2				
		Complex Symbolic	0.24	0.3				
		Complex Spatial	0.45	0.56				
MT2	E	No interrup.	0.02			Disruption duration		[14]

		2.8s: 0.04 13s: 0.10 22s: 0.15 32s: 0.17	Procedure execution sequentially with very short steps		No apparent uncertainty	
MT0	E	No interrup: 0.02 interrup: 0.02	Procedure execution non-sequentially with very short steps	Disruption type – inserting a disruption task to non-sequential steps is not a disruption	No apparent uncertainty	[14]
MT2	E	(PCE / SEQ / INI) No-I: 0.086 / 0.06 / 0.04 Yes-I: 0.30 / 0.23 / 0.032	Execute long procedures in which no association between subtasks PCE-post completion error SEQ- sub-task sequence error - wrong subtask selected INI- subtask initialization error – skip a procedure step	Interruption: Long duration (75s), cognitive demanding, and similar content interruption	No apparent uncertainty	[257]
MT2	DM	(#cells / #strategies / Out-come) Pre: 8 / 3.2 / 1 Post NoI: 11 / 2.8 / 1 Post YesI: 15 / 3.5 / 1	Risk-taking gamble games # cells – items viewed for information collection after interruption # strategies – alternatives considered outcomes – total wins/loss	Interruption: 8 second mental computation Pre- before interruption Post-NoI: after an interruption warning without interruption task Post-YesI: after interruption	(Very brief interruption leads to more information collection)	[78]
MT-2	E	No interruption: 0.196 With interruption: 0.276	17 participants perform medicine administration tasks while interrupted by alarms (% of active errors)	Alarms came in the middle of the primary task performance	No apparent uncertainty	[258]
MT-2	D / U, E	Wrong Answer Rate Ratio (One-time / Three-time) Cog/Cog: 1.32 / 1.43 Cog/Phy: 1.27 / 1.43 Phy/Cog: 1.48 / 1.74 Phy/Phy: 1.63 / 1.95	College students performed primary tasks with brief interruption, resumed to previous screen after interruption	Interrupted one time or three times in 10min-blocks, Primary task/ interruptive task: Cognitive/Cognitive Task Set Physical/Physical Task Set	(Very brief interruption)	[259]
MT-3	E	Single: 0.008 Dual: 0.062	Arithmetic task while monitoring (Arithmetic errors%)	Added salient cues to monitoring notification - Irrelevant but attention-demanding parallel task	No apparent uncertainty	[142]
MT-3	E	Single: 0.008				[142]

<table>
<tr><td></td><td></td><td>Dual</td><td colspan="2">0.031</td><td>Arithmetic task while monitoring (arithmetic errors%)</td><td>Irrelevant parallel task</td><td>No apparent uncertainty</td><td></td></tr>
<tr><td rowspan="2">MT-3</td><td rowspan="2">D</td><td>Single-task with salient notification</td><td colspan="2">0.008</td><td rowspan="2">Detect visual notification of a pending interrupting task while performing an arithmetic task</td><td rowspan="2">Dual-task with Non-salient notification: of an exclamation marks t appeared over a clock icon in the controller display,
single task with salient notification–pop-out color or blinking visual icon that captures attention</td><td rowspan="2">No apparent uncertainty</td><td rowspan="2">[142]</td></tr>
<tr><td>Dual-task with non-salient notification</td><td colspan="2">0.176</td></tr>
<tr><td rowspan="3">MT3</td><td rowspan="3">D</td><td></td><td>Missing changes</td><td>Missin g cues</td><td rowspan="3">Airplane pilots performing de-icing cue detection and responding to air traffic control information, concurrently detecting (monitoring or searching) multiple sets of parameters</td><td rowspan="3">Parameters in different sets may be related (missing changes or missing cues)</td><td rowspan="3">Time pressure and task complexity</td><td rowspan="3">[260]</td></tr>
<tr><td>Single task</td><td>0.028</td><td>0.05</td></tr>
<tr><td>Dual-task</td><td>0.21</td><td>0.2</td></tr>
<tr><td rowspan="2">MT3</td><td rowspan="2">D</td><td>Single</td><td colspan="2">0.15</td><td rowspan="2">Concurrently detection of dynamic system failure</td><td rowspan="2">Single vs concurrent tasks</td><td rowspan="2">Not analyzed</td><td rowspan="2">[261-263]</td></tr>
<tr><td>Concurrent</td><td colspan="2">0.35</td></tr>
<tr><td rowspan="2">MT3</td><td rowspan="2">D</td><td>Single</td><td colspan="2">0.05</td><td rowspan="2">Concurrent visual detection</td><td rowspan="2">Single vs concurrent tasks</td><td rowspan="2">Not analyzed</td><td rowspan="2">[261-263]</td></tr>
<tr><td>Concurrent</td><td colspan="2">0.3</td></tr>
<tr><td rowspan="2">MT4</td><td rowspan="2">D</td><td>Single</td><td colspan="2">0.05</td><td rowspan="2">Concurrent auditory detection</td><td rowspan="2">Single vs concurrent detection</td><td rowspan="2">Not analyzed</td><td rowspan="2">[261-263]</td></tr>
<tr><td>Concurrent</td><td colspan="2">0.5</td></tr>
<tr><td rowspan="5">MT4</td><td rowspan="5">D</td><td>Auditory alone</td><td colspan="2">0.012</td><td rowspan="5">Auditory detection of change and algebra task</td><td rowspan="5">Task performed alone vs concurrent</td><td rowspan="5">No apparent uncertainty</td><td rowspan="5">[264]</td></tr>
<tr><td>Auditory concurrent</td><td colspan="2">0.23</td></tr>
<tr><td>Algebra alone</td><td colspan="2">0.4</td></tr>
<tr><td>Algebra concurrent</td><td colspan="2">0.52</td></tr>
<tr><td>Single diagnosis</td><td colspan="2">0.01</td></tr>
<tr><td rowspan="2">MT5</td><td rowspan="2">u</td><td>Concurrent diagnosis</td><td colspan="2">0.37</td><td rowspan="2">Pilots concurrently diagnosed more than one complex event that required continuously seeking additional data to understand the events</td><td rowspan="2">Participants were asked
to report the location and severity of ice accretion, and they had to indicate whether the most recent icing cues represented a change from the previous condition. Another secondary task involved monitoring for the</td><td rowspan="2">Time urgent</td><td rowspan="2">[260]</td></tr>
<tr><td>Single</td><td colspan="2">0.04</td></tr>
</table>

					occurrence of an out-of-range value on one of two oil pressure gauges		
MT5	U / E	Baseline	0.04	Concurrently Text composition (Composition errors) and spatial visual detection task	Secondary task is visual detection, spatial location, and aural detection		[265]
		Concurrent visual	0.12				
		Concurrent spatial	0.07				
		Concurrent aural	0.13				
MT6	DM	Single	0.07	Concurrently making go vs no-go decisions	Single- or Dual-task, With or without specific training on dual-task		
		Concurrent	0.3				
MT8	E	Simulator fly lateral errors		Executing sequence and mental computation (% error in execution)	Concurrently executing action sequence and performing an attention/working memory task		
		Accuracy ratio = 10					
MT10	Unsp	See Figure A17-1.		Simulator flying (Lateral errors)	Communicating to comprehend air traffic control instructions		[266]

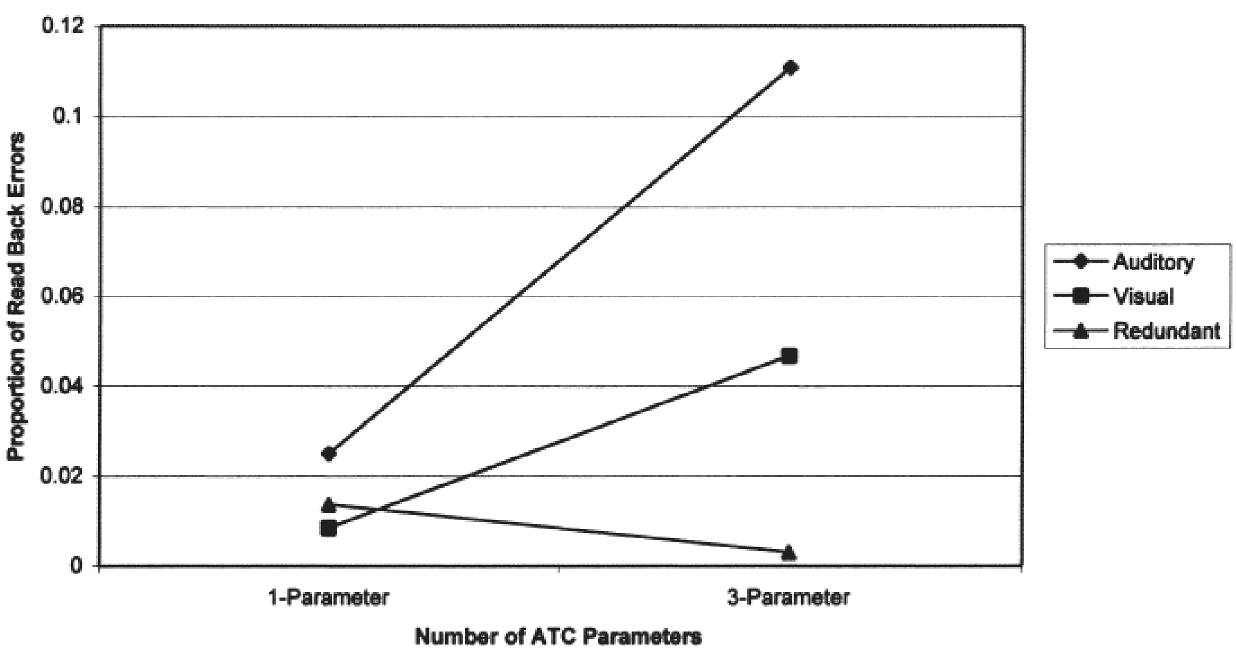

Figure A17-1 Proportion of read-back communication errors as a function of display and message length

Appendix A18 PIF Attributes and Weights for Mental Fatigue

Table A18-1 Attribute Identifiers and Descriptions for PIF Mental Fatigue

ID	Attribute
MF0	No impact
MF1	Sustained (> ~20mins) high-demand cognitive activities requiring continuously focused attention
MF2	Long working hours with high cognitively demanding tasks or hours of intensive work (e.g., taking a comprehensive examination, solving an emergency event) • Time on work, afternoon or evening working hours • Day vs night shifts, long work shift
MF3	Sleep deprivation
MF3.1	Sleep restriction (fewer sleep hours for days)
MF3.2	Total sleep deprivation (long hours of continuous wakefulness)
MF4	Change of cognitive state – • sudden increase in workload from a long period of low to high • sudden decrease in workload from high to low

Table A18-2 IDHEAS-DATA IDTABLE-18 – PIF Weights for Mental Fatigue

<table>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th></tr>
<tr><th>PIF</th><th>CFM</th><th>Error rates or task performance indicators</th><th>Task (and error measure)</th><th>PIF measure</th><th>Other PIFs (and Uncertainty)</th><th>REF</th></tr>
<tr><td>MF-1</td><td>D</td><td><table>
<tr><td colspan="3">Effective size</td></tr>
<tr><td></td><td>Sim.</td><td>Succ.</td></tr>
<tr><td>Low / Sensory</td><td>0.91</td><td>0.39</td></tr>
<tr><td>Low cognitive</td><td>0.00</td><td>0.78</td></tr>
<tr><td>High / Sensory</td><td>0.74</td><td>0.72</td></tr>
<tr><td>High / cognitive</td><td>0.47</td><td>0.76</td></tr>
<tr><td>Total</td><td colspan="2">0.71</td></tr>
</table>
<table>
<tr><td></td><td>Non-degraded</td><td>Degraded</td></tr>
<tr><td>1st 9min</td><td>0.007</td><td>0.07</td></tr>
<tr><td>18-27min</td><td>0.022</td><td>0.14</td></tr>
</table></td><td>Meta-analysis of 42 studies and 138 experimental conditions, signal detection and discrimination that needs vigilance and sustained attention, tasks last 30-60mins, (Effect size of Detection sensitivity), effect size was computed as the difference between perceptual sensitivity scores during the first and last periods of a vigil, divided by the square root of the mean square error term for the time effect</td><td>Low/High – Low/High event rates in visual detection tasks, Sensory/Cognitive – visual detection requires perception only or perception and recognition, Sim/Succ – visual targets were presented simultaneously (Sim) or successively (Succ) in visual discrimination tasks</td><td>(Meta-analysis)</td><td>[267]</td></tr>
<tr><td>MF1</td><td>D</td><td><table>
<tr><td></td><td>Traditional</td><td>Modified</td></tr>
<tr><td>First 10min</td><td>0.2</td><td>0.22</td></tr>
<tr><td>40-50min</td><td>0.42</td><td>0.26</td></tr>
</table></td><td>Traditional - button-press responses to signify detection of rarely occurring critical signals Modified - button-press responses acknowledged frequently occurring neutral stimulus events and response withholding signified critical signal detection</td><td>Five blocks of 10mins, traditional task requires constant attention, modified task promotes mindlessness via routinization</td><td>No apparent uncertainty</td><td>[268]</td></tr>
<tr><td>MF1</td><td>D</td><td><table>
<tr><td></td><td>First 10min</td><td>20-30min</td></tr>
<tr><td>High-Sim</td><td>0.12</td><td>0.13</td></tr>
</table></td><td>Discrimination of differences in line lengths (%incorrect)</td><td>Three 10min blocks Task difficulty:</td><td>No apparent</td><td>[269]</td></tr>
</table>

<table>
<tr><td></td><td></td><td><table><tr><td>High-Suc</td><td>0.13</td><td>0.31</td></tr><tr><td>Low-Sim</td><td>0.21</td><td>0.47</td></tr><tr><td>Low-Suc</td><td>0.32</td><td>0.54</td></tr></table></td><td></td><td>High vs low discriminability simultaneous- vs successive-discrimination</td><td>uncertainty</td><td></td></tr>
<tr><td>MF1</td><td>D</td><td>Probability of detecting a signal decreased dramatically over time-on-watch. This decrement was greatest when:
• The signal duration was short
• The probability of a signal was low
• The signal intensity was low
• The signal was simple rather than complex</td><td>Radar operators detect signals</td><td>Time-on-watch</td><td>No apparent uncertainty</td><td>[270]</td></tr>
<tr><td>MF1</td><td>D</td><td><table><tr><td></td><td>Match</td><td>mismatch</td></tr><tr><td>1st block</td><td>0.32</td><td>0.34</td></tr><tr><td>2nd block</td><td>0.30</td><td>0.4</td></tr><tr><td>3rd block</td><td>0.27</td><td>0.41</td></tr><tr><td>4th block</td><td>0.25</td><td>0.46</td></tr></table></td><td>Passport control face-matching task: identify 184 matched pairs and 16 mismatched pairs</td><td>Four blocks of 50 pairs of face pictures, each pair has average 5-6sec, so one block is about 5min</td><td>No apparent uncertainty</td><td>[79]</td></tr>
<tr><td>MF1 & MF2</td><td></td><td><table><tr><td></td><td>Morning</td><td>Afternoon</td></tr><tr><td>Beginning</td><td>0.03</td><td>0.08</td></tr><tr><td>45min</td><td>0.16</td><td>0.20</td></tr></table></td><td>Subjects listened to a stream of digits and were required to detect three successive odd digits that were all different; for example, 3-5-9 or 1-7-5 (& missed)</td><td>Performance at beginning vs 45min</td><td>No apparent uncertainty</td><td>[271]</td></tr>
<tr><td>MF2</td><td>DM</td><td>Subjects with a considerable fatigue induced by a lengthy college examination demonstrated greater primacy effects in their impressions than did the less fatigued ones</td><td>Read summary information about a job candidate, evaluated the candidate's qualifications and justified their impressions</td><td>3 levels of manipulated mental fatigue conditions: Before and after a regular class period and after a 2-hr final examination</td><td>(No error data)</td><td>[272]</td></tr>
<tr><td>MF2</td><td>U</td><td>Correlation coefficient<table><tr><td>Shift</td><td>Outage</td><td>Afternoon</td></tr><tr><td>Morning</td><td>0.04</td><td>0.11</td></tr><tr><td>Afternoon</td><td>.004</td><td>0.32</td></tr><tr><td>Night</td><td>0.24</td><td>0.17</td></tr></table></td><td>Correlation of NPP operators' diagnosis errors with work shift</td><td>Shifts of operator working schedule</td><td>(Other factors may exist)</td><td>[239]</td></tr>
<tr><td>MF2</td><td>D</td><td>Correlation coefficient<table><tr><td>Shift</td><td>Outage</td><td>Afternoon</td></tr><tr><td>Morning</td><td>-0.004</td><td>0.19</td></tr><tr><td>Afternoon</td><td>.06</td><td>0.33</td></tr><tr><td>Night</td><td>0.25</td><td>0.35</td></tr></table></td><td>Correlation of NPP operators' detection errors (minor errors) with work shift</td><td>Shifts of operator working schedule</td><td>Other factors may exist)</td><td>[239]</td></tr>
<tr><td>MF2</td><td></td><td><table><tr><td></td><td>Day</td><td>Night</td></tr></table></td><td>Participants performed simulated spacecraft life-</td><td></td><td>Multitasking</td><td>[144]</td></tr>
</table>

	U & DM & E	System control errors (%) PracF NovF CtrlPanF	4.01 2.18 3.70 6.16	4.17 2.18 3.75 6.58 _	support tasks: Monitor automatic subsystems, take manual control of the systems, engage in a process of fault diagnosis to identify and rectify the fault, acknowledge alarms, and remember to carry out an action at a specified time in the future (perspective memory)	Day vs night – Occasional night work Time on work – three periods F-free - fault-free condition, the automatic controller functioned perfectly well, requiring no operator intervention, In the practiced faults PracF - participants had to manage faults they were familiar with through extensive practice during the training sessions, (NovF) - novel faults were of the same general type as PracF but had not been experienced before, (CtrlPanF) - or pcontrol panel failures in which a system failure was accompanied by a simultaneous disabling of the relevant control panel		
		Diagnostic accuracy (# of errors) PracF NovF	.44 .22 .65	0.52 0.28 0.76				
		Prospective memory failures (%) F-free PracF NovF CtrlPanF	15.81 6.72 10.32 18.91 27.30	14.05 7.53 12.03 18.51 18.12				
MF2 & MF1	U/ DM	Information sampling			Participants performed simulated spacecraft life-support tasks (Information sampling, number per minute)	Day vs night – Occasional night work Time on work – three periods System fault type - Routine vs. emergency	Multitasking	[144]
		DAy / Period	Routine	Emergency				
		DAy 1	1.4	1.1				
		DAy 2	1.4	0.8				
		DAy 3	1.2	0.8				
		Night 1	1.5	0.95				
		Night 2	1.15	0.7				
		Night 3	1.0	0.8				
MF2	D		Low freq	High freq	Subjects listened to a stream of digits and were required to detect three successive odd digits that were all different; for example, 3-5-9 or 1-7-5. (& missed)	Task performed in morning vs late afternoon, Low vs High stimulus freq	No apparent uncertainty	[271]
		Morning	0.09	0.12				
		afternoon	0.12	0.20				
MF2	Unsp.		Analog	Digital	Experienced technicians used equipment to make measurement (%measurement errors)	Analog vs digital equipment (less mentally demanding) FN- Forenoon AN - Afternoon	No apparent uncertainty	[63]
		FN	0.09	0.05				
		AN	0.35	0.11				

<table>
<tr><td>MF3.1</td><td>Unsp.</td><td>Performance ratio (PR), e.g., PR = −0.05 translates to a 5% decrement in performance relative to control performance for each hour of continuous wakefulness
Circadian day Accuracy
PR = -0.004 x Hours +1
Circadian night Accuracy
PR = -0.009 x Hours +1</td><td>Complex cognitive tasks including diagnosis, decisionmaking, teamwork</td><td>Total sleep deprivation – hours of continuous wakefulness</td><td>(247 papers)</td><td>[273]</td></tr>
<tr><td>MF3.2</td><td>Unsp.</td><td>Performance ratio (PR), e.g., PR = −0.05 translates to a 5% decrement in performance relative to control performance for each consecutive day:
Mild SR Accuracy
PR = -0.008 x Hours +1
Severe SR Accuracy
PR = -0.067 x Hours +1</td><td>Complex cognitive tasks similar to real-world tasks including diagnosis, decisionmaking, teamwork</td><td>Sleep restriction (SR) in consecutive days:
Mild – 4-6 hours of sleep per 20 hours
Severe: < 4 hours</td><td>(247 papers)</td><td>[273]</td></tr>
<tr><td>MF3.2</td><td></td><td>Y is performnce ratio and X is # of average horus of sleep (Figure A18-1)</td><td>Psych-motor tasks similar to astronauts performing in long flight Space Shuttle</td><td>Short to long term sleep deprivation</td><td>(Meta-analysis)</td><td>[274]</td></tr>
<tr><td>MF3.2 & MF3.1</td><td></td><td>PIF weight derived from meta-data<table><tr><td>Well rested</td><td>0.6</td></tr><tr><td>Adequate rest</td><td>1</td></tr><tr><td>Short-term high sleep deprivation</td><td>1.7</td></tr><tr><td>Long-term moderate sleep deprivation</td><td>4.0</td></tr><tr><td>Long-term high sleep deprivation</td><td>8.7</td></tr></table></td><td>Psych-motor tasks similar to astronauts performing in long flight Space Shuttle</td><td>Short to long term sleep deprivation</td><td>(Meta-analysis)</td><td>[274]</td></tr>
<tr><td>MF3.1</td><td>U/E</td><td>Blood alcohol content (BAC%) of various tasks for the hours awake (Figure A18-2).</td><td>Tasks in the data sources:
Simulated driving task
Tracking task
Simple reaction time
Mackworth clock
Simulated driving task
Tracking task Simulated driving task
Grammatical reasoning—latency
Vigilance—latency
Vigilance—accuracy
Tracking task
(Compared % blood alcohol level, BAC)</td><td>Sleep deprivation – hours of wakefulness</td><td>(No error data)</td><td>[275]</td></tr>
<tr><td>MF3.1</td><td>E</td><td>PIF weight is between 1.2 to 2.5 for 20-80 hours of wakefulness (Figure A18-3)</td><td>34 studies, most visual-motor tasks</td><td>Sleep deprivation – hours of wakefulness</td><td>(Other PIFs may exist)</td><td>[276]</td></tr>
<tr><td>MF3.1</td><td>DM</td><td>The critical reasoning task was unaffected by sleep loss, whereas performance at the game significantly deteriorated after 32-36 h of sleep loss, when sleep</td><td>Performed dynamic and realistic marketing decision making "game" requiring flexible thinking and the updating of plans</td><td>Total sleep deprivation</td><td>(No error data)</td><td>[277]</td></tr>
</table>

<table>
<tr><td></td><td></td><td colspan="3">deprivation led to more rigid thinking, increased perseverative errors, and marked difficulty in appreciating an updated situation</td><td>in the light of new information</td><td></td><td></td><td></td></tr>
<tr><td>MF3.1</td><td>DM</td><td colspan="3">Impairment on DM is as much as on other cognitive functions</td><td>Review of decisionmaking impairment due to total sleep deprivation</td><td>Total sleep deprivation</td><td>(No error data)</td><td>[278]</td></tr>
<tr><td rowspan="4">MF-4</td><td rowspan="4">E</td><td>Alarm onset time</td><td>Sterile</td><td>Non-sterile</td><td rowspan="4">Trained students monitored NPP CR alarm onset and performed alarm response procedure in 30mins (%uncompleted by 30mins)</td><td rowspan="4">Alarm onset time: sterile condition - not allowed access to any activity that was not directly related to the task
Non-sterile: Allowed to access the Internet and read or use their own electronic devices</td><td rowspan="4">Scenario (small subject sample)</td><td rowspan="4">[279]</td></tr>
<tr><td>1:30</td><td>0.08</td><td>0.08</td></tr>
<tr><td>2:30</td><td>0.17</td><td>0.5</td></tr>
<tr><td>3:30</td><td>0.67</td><td>0.83</td></tr>
<tr><td>MF4</td><td></td><td colspan="3">Figure A18-4</td><td>Annual number of OEs distributed by the amount of time on position that had lapsed before the OE occurred, most OEs occurred in the first 30minutes on-shift</td><td>Minutes on position</td><td>Scenario familiarity (Statistical)</td><td>[118]</td></tr>
</table>

- 3 Hours of Sleep on Average: $y = 2.1625x - 0.0806$, $R^2 = 0.976$
- 4 Hours of Sleep on Average: $y = 1.0418x + 1.177$, $R^2 = 0.967$
- 5 Hours of Sleep on Average: $y = 0.7622x + 1.5545$, $R^2 = 0.972$
- 6 Hours of Sleep on Average: $y = 0.7622x + 0.1095$, $R^2 = 0.961$
- 7 Hours of Sleep on Average: $y = 0.3568x + 1.0391$, $R^2 = 0.929$
- 8 Hours of Sleep on Average: $y = 0.1836x + 1.0902$, $R^2 = 0.464$
- 9 Hours of Sleep on Average: $y = -0.0293x + 1.1474$, $R^2 = 0.139$
- 10 Hours of Sleep on Average: $y = -0.1181x + 2.6544$, $R^2 = 0.534$

Figure A18-1 Performance decrement (y) corresponding to the number of hours of sleep

Hours of wakefulness compared to BAC% levels.

Study reference	Hours awake	BAC %	Task
[63]	18.5	0.05	Simulated driving task
[64]	17	0.05	Tracking task
[66]	24	0.05	Simple reaction time
[51]	17	0.05	Mackworth clock
[63]	21	0.08	Simulated driving task
[64]	24	0.08	Tracking task
[62]	20	0.08	Simulated driving task
[65]	20.3	0.10	Grammatical reasoning—latency
[65]	24.9	0.10	Vigilance—latency
[65]	22.3	0.10	Vigilance—accuracy
[65]	25.1	0.10	Tracking task

Figure A18-2 Equivalent of the blood alcohol content (BAC%) corresponding to sleep deprivation (hours awake) in various studies.

Probability ratios for number of lapses means (standard deviations) over 20 h time blocks.

	< 20 h	20–40 h	40–60 h	60–80 h	> 80 h
Approach 1					
k=1	0.937 (0.25)	1.120 (0.47)	0.897 (0.05)	0.937 (0.12)	0.998 (0.24)
k=3	0.979 (0.58)	1.713 (1.25)	1.090 (0.07)	1.095 (0.07)	1.015 (0.06)
k=6	1.583 (1.70)	1.769 (0.69)	2.692 (0.21)	2.666 (0.23)	2.339 (0.17)
k=9	2.317 (2.63)	2.978 (3.79)			
Approach 2					
$P(T>C)/0.5$	1.122 (0.48)	1.400 (0.38)	1.528 (0.116)	1.591 (0.19)	1.657 (0.34)

Note: standard deviations are shown in parentheses.

Figure A18-3 Probability ratios for number of lapses means

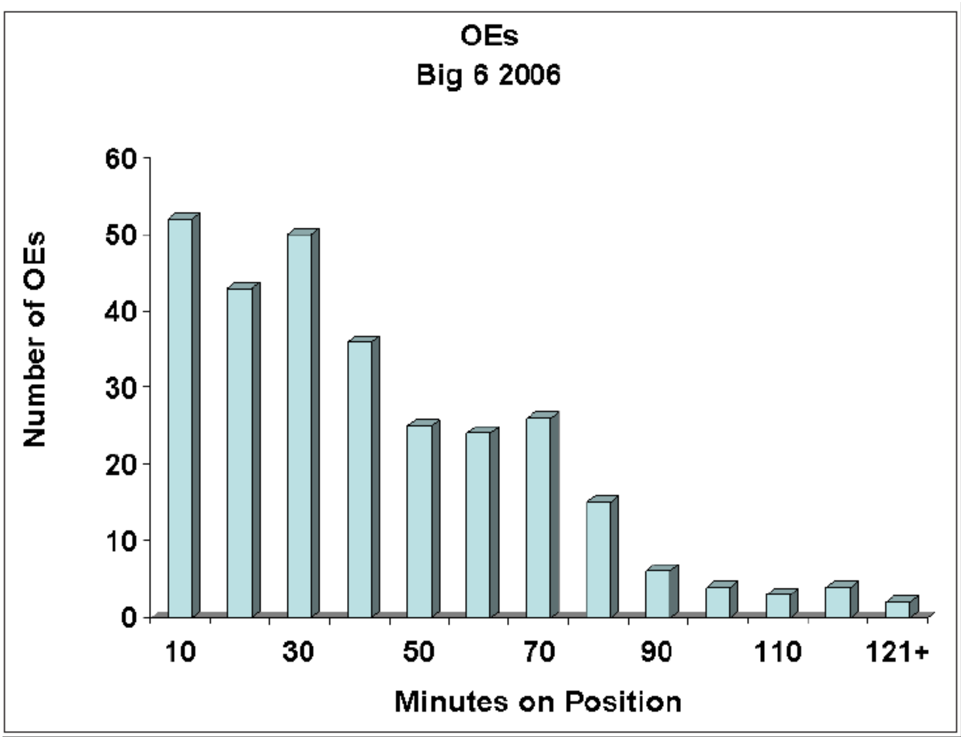

Figure A18-4 Annual number of OEs distributed by the amount of time on the position that had lapsed before the OE occurred.

Appendix A19 PIF Attributes and Weights for Time Pressure and Stress

Table A19-1 Attribute Identifiers and Descriptions for PIF Time Pressure and Stress

ID	Attribute
MF0	No impact
TPS1	Time pressure due to perceived time urgency • Receiving instructions to complete tasks as quickly as possible, deadlines, or stimulus presentation rate • Skipping self-verification due to rush the task completion (speed-accuracy trade-off)
TPS2	Emotional stress (e.g., anxiety, frustration)
TPS3	Cumulative physical stress (e.g., long hours exposure to ambient noise, disturbed dark and light rhythms, air pollution, disruption of normal work-sleep cycles, illness)
TPS4	Reluctance to execute an action plan due to potential negative impacts (e.g., adverse economic impact, or personal injury)

Table A19-2 IDHEAS-DATA IDTABLE-19 – PIF Weights for Time Pressure and Stress

<table>
<tr><th>1</th><th>2</th><th>3</th><th>4</th><th>5</th><th>6</th><th>7</th></tr>
<tr><th>PIF</th><th>CFM</th><th>Error rates</th><th>Task (and error measure)</th><th>PIF measure</th><th>Other PIFs (and Uncertainty)</th><th>REF</th></tr>
<tr><td>TPS1</td><td>D, U&DM, E</td><td>Effect-size is a standardized mean difference between the experimental and control conditions.
<table>
<tr><td></td><td>accuracy</td><td>Response time</td></tr>
<tr><td>Perception (D)</td><td>-0.33</td><td>0.26</td></tr>
<tr><td>Cognition (U & DM)</td><td>-0.66</td><td>0.57</td></tr>
<tr><td>Motor (E)</td><td>0.1</td><td>-0.6</td></tr>
</table></td><td>Controlled lab setting and real-world settings in which temporal constraints impose stress and workload on operators, as anyone who is pressed to meet proposal deadlines can attest.</td><td>time stress: (e.g., instructions to complete tasks as quickly as possible, deadlines, or stimulus presentation rate)</td><td>125 of 281 papers with 827 data for meta-analysis</td><td>[81]</td></tr>
<tr><td>TPS1</td><td>U & T</td><td>Members lose awareness of each other as time pressure increases, but far less so in terms of task-relevant than task-irrelevant information. Time-pressure has a direct effect on awareness of group members in addition to the indirect effect that would be expected with the reduced social interaction observed by Karau and Kelly (1992). This effect could be especially problematic for group coordination if group members do not consider coordination related information to be important.</td><td>3-person groups performed anagram-solving task independently but simultaneously with and in the presence of their group members</td><td>Excess time – 75% of work assignment
Moderate time Time pressure – 100% assignment
High time pressure – 150% work assignment</td><td>(Anagram-solving task is relating and reasoning)</td><td>[280]</td></tr>
</table>

TP S1	Uns p.	See Figures A19-1 and A19-2	Aircraft maintenance tasks: Skill-based errors, decisionmaking error, and procedure routine violation	678 human errors in 992 ASRS maintenance reports. Time pressure is a pressure to hastily complete a task as indicated by an approaching deadline.	(Statistic al analysis)	[281]
TP S1	U, DM, E	See Figure A19-3.	Aircraft maintenance tasks: Skill-based errors, decisionmaking error, and procedure routine violation	678 human errors in 992 ASRS maintenance reports. Time pressure is the pressure to hastily complete a task as indicated by an approaching deadline.	(Statistic al analysis)	[281]
TP S1	DM	See Figures A19-4, A19-5, and A19-6.	Dynamic decisionmaking - monitor the fitness of an athlete wi-3 is running a race and avoid athlete to collapse (i.e. to reach a fitness level of zero). To attain this goal the subject can request information and apply treatment.	Three time-pressure conditions expressed by the slopes of the functions Y =aX+ b: low time pressure (a=-0.5), moderate time pressure (a=-1}, and high time pressure (a=-2).		[282]
TP S1	E	See Figure A19-7.	Three tasks with increasing levels of execution complexity in the simple response task, participants responded with their left hand in half of each block and with their right hand in the other half. In the choice-by- location task, participants had to respond at the side where the letter was displayed. In the Simon task, participants had to press the left button when an “A” was presented, and the right button when a “B” was presented.	Participants were told that filling time varied randomly during the session. In the condition without time pressure filling time was held constant at 600ms. The starting value of the filling time for the condition with time pressure was 450ms.		[283]

TP S1	U	Without time pressure: 0.49 With time pressure: 0.67	Senior internal medicine residents diagnosed eight written clinical cases presented on computers (diagnosis accuracy)	In the time-pressure condition, after completing each case, participants received information that they were behind schedule.		[284]
TP S1	U	See Figure A19-8.	Solve syllogism through reasoning. Simple problems require a few steps to determine the logical validity. Complex problems require a larger number of steps and more difficult logical operations (e.g., reduction at absurdum) in their proofs. (Accuracy of reasoning)	Time limited vs. unlimited. Reasoning complexity - Syllogism complexity was manipulated by presenting people with simple or complex syllogisms.		[285]
TP S1	E	Error rate of response (M +/- SD): difference not significant 1s: 0.8 ± 1.0 1.5s: 0.9 ± 1.2 2s: 1.4 ± 2.6	Visual-motor response requiring motor precision	Time allowed to make response: 1s, 1.5s and 2s		[286]
TP S1	D	(Match / Mismatch) 10s: 0.22 / 0.3 8s: 0.22 / 0.45 6s: 0.22 / 0.42 4s: 0.23 / 0.45 2s: 0.24 / 0.42	Experiment 3: Students performed passport-control face picture match (%error in match and mismatch)	Time pressure - number of tasks assigned within fixed timeframe		[79]
TP S1	U & E	(Trigger event / Skill-based) W-TP: 1.7 / 1 W-NoTP: 1.1 / 1.2 FF-TP: 0.8 / 0.3 FF-NpTP: 0.6 / 0.8	Students enrolled in aviation maintenance technician program recognized 3 trigger events and performed aircraft maintenance tasks (Trigger event errors and skill-based errors)	Time pressure (TP). Shift turnover strategy: Written (W) vs. Face-toface (FF)		[287]
TP S1	DM	See Figures A19-9 and A19-10.	210 male undergraduates) were presented five pieces of information to assimilate in evaluating cars as purchase options. (# of factors had been systematically used by the processor to make the final judgment)	*High time pressure condition* - "proceed as rapidly as possible without sacrificing accuracy." Subjects were asked to record the elapsed time on their booklet when they finished. *low time pressure* – "accurately judge the cars." Each was told he would have 40 seconds to consider the information		[288]

				available and should use the entire period. The length of a 40-second interval offered plenty of processing time. *Undefined time* - no mandatory deliberation period was imposed.		
TP S1	DM	(# of factors used to make the final judgment) Undefined 2.08 Low time pressure 2.33 High time pressure 1.5	210 male undergraduates were presented five pieces of information to assimilate in evaluating cars as purchase options (# of factors used to make the final judgment)	High, low, undefined (unconstrainted) time pressure		[288]
TP S1	U	Low complex (1E0T) / High complex (3E2T) Relaxed 0.012 / 0.083 Stressed 0.046 / 0.375	120 subjects completed 100 geometric analogies with nine levels of complexity (# of Elements and # of Transforms) (%incorrect)	TPS-1: relaxed (reassurance, non-time-limited) or stressed (ego-threat, time-limited)	(Time available is sufficient)	[80]
TP S1 & TP S2	U	Low complex (1E0T) / High complex (3E2T) Relaxed & less A 0.007 / 0.061 Relaxed & more A 0.023 / 0.133 Stressed & less A 0.047 / 0.352 Stressed& more A 0.046 / 0.386	120 subjects completed 100 geometric analogies with nine levels of complexity defined as # of Elements and # of Transforms) (%incorrect)	TPS-1: relaxed (reassurance, non-time-limited) or stressed (ego-threat, time-limited) TPS2- Individual differences in trait and state anxiety: Less state anxious (Less A) and more state anxious (More A)	(Time available is sufficient)	[80]
TP S2	D	HS-ST / LS-ST / LSDT Normal 5.24 (4.75) / 48.04 (21.3) / 45.00 (17.7) Threat 6.19 (5.15) / 41.48 (19.7) / 53.10 (24.2)	The threat-of-shock Detect target in normal condition and anticipatory anxiety: Participants were informed that during these blocks, they could randomly	HS-ST - High Salience,Single Target LS-ST - Low Salience, Single Target LSDT - low Salience, Dual Target		[289]

			receive a wrist shock that was not related to performance. (% miss)			
TPS2	D	Prepicture / 8min with picture Negative: 0.11 / 0.20 Neutral: 0.9 / 0.8 Control: 0.8 / 0.12	51 participants (15 men and 36 women) performed target detection vigilance tasks while viewing a task-irrelevant picture (% miss)	Three vigilance conditions: negative-arousing pictures, neutral pictures, or a no-picture visual vigil control.		[290]
TPS2	E	Heart rate / Dart score Low anxiety: 162 / 5.2 High anxiety: 167 / 4.6	Psycho-motor performance (average heart rate and dart score per dart)	Low and high anxiety		[291, 292]
TPS2	DM	Stress showed a significantly stronger tendency to offer solutions before all available alternatives had been considered (Figure A19-11)	They were requested to solve decision problems, while being exposed to controllable stress, uncontrollable stress, or no stress at all.	No time constraint for the performance of the task. Uncontrollable stress - the computer had been programmed with the number and timing of the shocks in such a way that the subject had no control over them whatsoever. Controllable stress - Receiving shocks was presented to the subject as contingent on his or her performance.		[293]
TPS2	Unsp	Both threat of shock and anxiety disorders promote mechanisms associated with harm avoidance across multiple levels of cognition (from perception to attention to learning and executive function. This mechanism comes at a cost to other functions such as working memory, but leaves some functions, such as planning, unperturbed. We also highlight a number of cognitive effects that differ across anxiety disorders and threat of shock. These discrepant effects are largely seen in “cold” cognitive functions involving control mechanisms	Review threat of shock on cognition		(No error data)	[294]

<table>
<tr><td rowspan="3">TPS2, TPS3</td><td rowspan="3">E</td><td></td><td>Lo-A</td><td>High-A</td><td rowspan="3">Military solder shooting accuracy task (%miss)</td><td rowspan="3">TSP-2 - LoA and Hi-A: Low and high anxiety
TSP-3 - LF an HF – low and high physical fatigue</td><td rowspan="3"></td><td rowspan="3">[295]</td></tr>
<tr><td>LF</td><td>67.4 (24.9)</td><td>32.6 (26.2)</td></tr>
<tr><td>HF</td><td>66.7 (22.7)</td><td>37.1 (23.7)</td></tr>
<tr><td rowspan="3">TPS2, TPS3</td><td rowspan="3">U / E</td><td></td><td>Lo-A</td><td>High-A</td><td rowspan="3">Military solders math task (%incorrect)</td><td rowspan="3">TSP-2 - LoA and Hi-A: Low and high anxiety
TSP-3 - LF an HF – low and high physical fatigue</td><td rowspan="3"></td><td rowspan="3">[295]</td></tr>
<tr><td>LF</td><td>88</td><td>82</td></tr>
<tr><td>HF</td><td>86</td><td>76</td></tr>
<tr><td rowspan="3">TPS2, TPS3</td><td rowspan="3">U / E</td><td></td><td>Lo-A</td><td>High-A</td><td rowspan="3">Military solders memory task (%incorrect)</td><td rowspan="3">TSP-2 - LoA and Hi-A: Low and high anxiety
TSP-3 - LF an HF – low and high physical fatigue</td><td rowspan="3"></td><td rowspan="3">[295]</td></tr>
<tr><td>LF</td><td>52</td><td>61</td></tr>
<tr><td>HF</td><td>60</td><td>49</td></tr>
<tr><td rowspan="3">TPS2, TPS3</td><td rowspan="3">D</td><td></td><td>Lo-A</td><td>High-A</td><td rowspan="3">Military solders vigilance task - detecting target (0-5)</td><td rowspan="3">TSP-2 - LoA and Hi-A: Low and high anxiety
TSP-3 - LF an HF – low and high physical fatigue</td><td rowspan="3"></td><td rowspan="3">[295]</td></tr>
<tr><td>LF</td><td>0.6</td><td>0.5</td></tr>
<tr><td>HF</td><td>0.7</td><td>0.7</td></tr>
<tr><td rowspan="3">TPS2, TPS3</td><td rowspan="3">DM</td><td></td><td>Lo-A</td><td>High-A</td><td rowspan="3">Military task – decide to or not to shoot (incorrect-decisions-to-shoot ratio)</td><td rowspan="3">TSP-2 - LoA and Hi-A: Low and high anxiety
TSP-3 - LF an HF – low and high physical fatigue</td><td rowspan="3"></td><td rowspan="3">[292]</td></tr>
<tr><td>LF</td><td>0.03</td><td>0.04</td></tr>
<tr><td>HF</td><td>0.03</td><td>0.06</td></tr>
<tr><td rowspan="3">TPS2, TPS3</td><td rowspan="3">E</td><td></td><td>Lo-A</td><td>High-A</td><td rowspan="3">Military task - shoot accuracy (%miss)</td><td rowspan="3">TSP-2 - LoA and Hi-A: Low and high anxiety
TSP-3 - LF an HF – low and high physical fatigue</td><td rowspan="3"></td><td rowspan="3">[292]</td></tr>
<tr><td>LF</td><td>0.52</td><td>0.69</td></tr>
<tr><td>HF</td><td>0.60</td><td>0.58</td></tr>
<tr><td>Unsp</td><td>E</td><td colspan="3">5.8E-2 (1/20)</td><td>34 Opening a valve by MCR panel controls
Failed to open, memorized task step is not remembered</td><td>Rarely performed task sequence, moderately high level of stress</td><td>(Infrequently performed tasks)</td><td>[4]</td></tr>
<tr><td rowspan="2">Unsp</td><td rowspan="2">E</td><td>No stress</td><td colspan="2">2.45E-2 (1/48)</td><td rowspan="2">Carrying out a sequence of tasks
Memorized task step not remembered</td><td rowspan="2">No stress - Rarely performed, no error promoting factors
Stress - Rarely performed, moderately high level of stress</td><td rowspan="2">(Infrequently performed tasks, unspecified stress)</td><td rowspan="2">[5]</td></tr>
<tr><td>Stress</td><td colspan="2">5.62E-2 (2/41)</td></tr>
<tr><td rowspan="2">TPS4</td><td rowspan="2">E</td><td>Exist</td><td colspan="2">1.1E-2</td><td rowspan="2">Delay implementation of a decision/plan</td><td rowspan="2">Exist vs absence of reluctance & viable alternative. Incorrect assessment of margin and with additional cues</td><td rowspan="2">(Expert judgment)</td><td rowspan="2">[6]</td></tr>
<tr><td>Absence</td><td colspan="2">2.2E-4</td></tr>
</table>

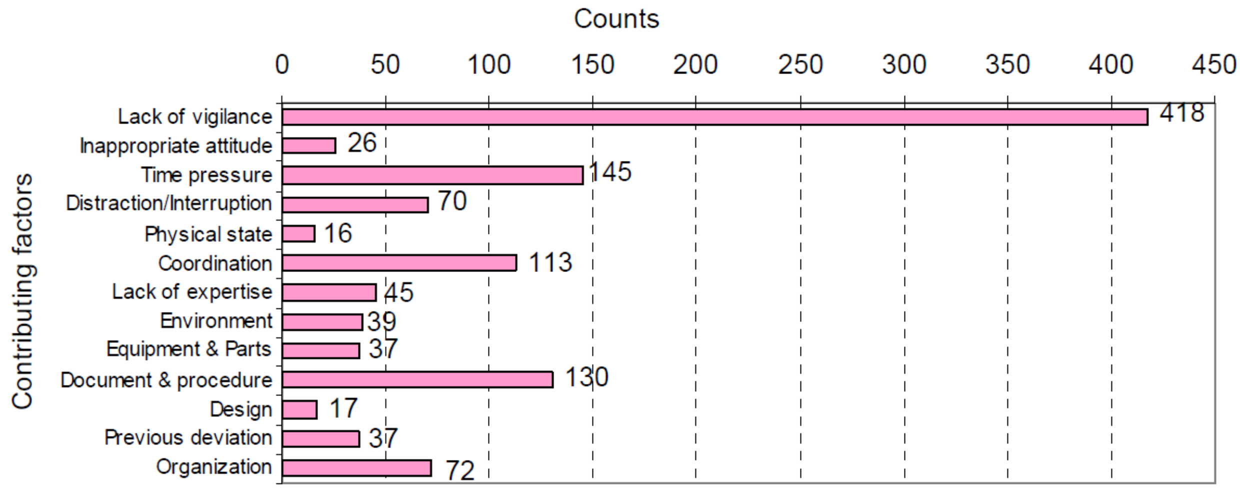

Figure A19-1 Identified contributing factors.

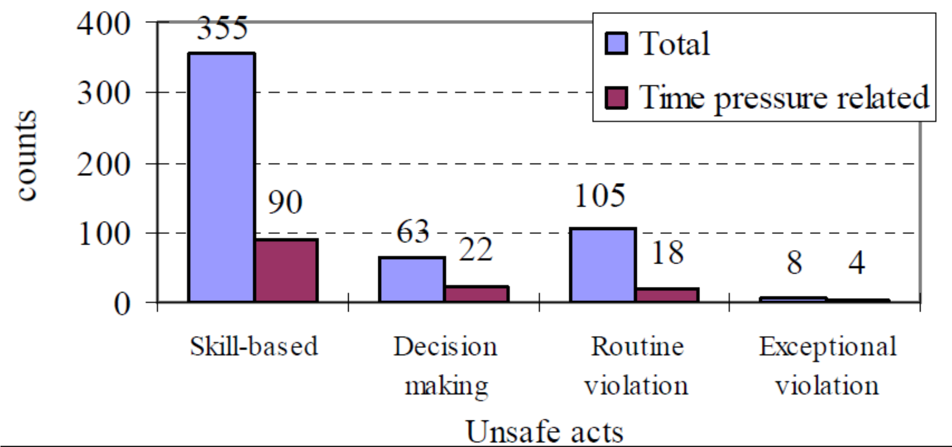

Figure A19-2 Frequency of unsafe acts.

Association between unsafe acts and contributing factors by Multinominal Logistic Regression Analysis

	Skill-based errors				Decision making errors				Routine violation			
	Beta	SE	P	OR	Beta	SE	P	OR	Beta	SE	P	OR
Lack of vigilance	1.46	0.30	0.00 **	4.30	-1.03	0.36	0.00 **	0.36	-1.10	0.34	0.00 **	0.33
Inappropriate attitude	-1.75	0.52	0.00 **	0.18	0.52	0.66	0.43	1.69	1.58	0.50	0.00 **	4.83
Time pressure	0.03	0.28	0.91	1.03	0.80	0.35	0.02 *	2.23	-0.70	0.34	0.04 *	0.50
Distraction/Interruption	0.95	0.41	0.02 *	2.57	-0.68	0.56	0.23	0.51	-0.75	0.49	0.12	0.47
Physical state	0.57	0.82	0.49	1.76	-0.36	1.07	0.73	0.70	-0.66	1.10	0.55	0.52
Coordination	-0.87	0.27	0.00 **	0.42	1.32	0.34	0.00 **	3.75	0.04	0.32	0.90	1.04
Lack of expertise	-0.34	0.38	0.37	0.71	1.70	0.41	0.00 **	5.45	-1.52	0.64	0.02 *	0.22
Environment	2.65	1.04	0.01 *	14.12					-1.85	1.04	0.08	0.16
Equipment & Parts	-0.29	0.43	0.50	0.75	0.35	0.53	0.51	1.42	0.17	0.47	0.71	1.19
Document & procedure	-0.35	0.26	0.18	0.71	0.37	0.35	0.29	1.45	0.27	0.30	0.37	1.30
Design	1.03	0.84	0.22	2.81	-0.17	1.09	0.88	0.84	-1.28	1.09	0.24	0.28
Previous deviation	-2.06	0.42	0.00 **	0.13	-0.28	0.67	0.68	0.76	2.26	0.40	0.00 **	9.57
Organization	-0.46	0.36	0.20	0.63	0.04	0.45	0.93	1.04	0.48	0.40	0.23	1.61

1. ** 99% level of significance, * 95% level of significance
2. SE = Standard Error, OR = Odds Ratio

Figure A19-3 Association between unsafe acts and contributing factors by Multinominal Logistic Regression Analysis

Table I Mean number of trials accurately dealt with in each time pressure condition (maximum=4), and the proportions correct treatments.

time pressure	low	moderate	high
mean number of trials	3.75	3.40	1.70
proportion correct treatments	0.91	0.96	0.91

Figure A19-4 Time pressure effects

Table II Mean number of information requests (total number of information requests divided by the total number of dummy trials) in each time pressure condition.

time pressure	low	moderate	high
mean number of requests	3.09	2.07	2.30

Figure A19-5 Time pressure effects

Table V Relative amount of treatment given to the athlete for each time pressure condition.

time pressure	low	moderate	high
level of restore	0.95	0.89	0.73

Figure A19-6 Time pressure effects

Table 1 RTs, proportion of correct responses (PCs), and response force for the simple response task, choice-by-location task, and the Simon task (averaged across correspondence and non-correspondence trials, and separately for correspondence and non-correspondence trials) in the case of high and low time pressure.

Task	Time pressure	RT (in ms)	PC (in%)	Response force (in N)
Simple response	High	271	95.3	14.9
	Low	335	98.2	7.6
Choice-by-location	High	373	87.5	15.9
	Low	412	96.1	9.5
Simon (averaged)	High	439	83.5	13.8
	Low	469	90.3	9.4
Correspondence	High	422	90.1	13.7
	Low	454	93.1	9.2
Non-correspondence	High	460	77.1	13.8
	Low	486	87.4	9.6

Figure A19-7 Response time and proportion of responses.

Experiment 2: Endorsement Rates

Group	Valid		Invalid	
	Believable	Unbelievable	Believable	Unbelievable
High cognitive capacity				
Time limit	.73 (.05)	.46 (.07)	.48 (.05)	.42 (.06)
No limit	.83 (.03)	.73 (.06)	.57 (.06)	.32 (.05)
Low cognitive capacity				
Time limit	.70 (.04)	.49 (.06)	.57 (.05)	.51 (.05)
No limit	.70 (.04)	.67 (.06)	.49 (.05)	.39 (.04)

Note. High cognitive capacity participants scored above the median Part I AH4 Group Test of General Intelligence score. Low cognitive capacity participants scored at or below the median. Standard errors for the per condition analyses are in parentheses.

Figure A19-8 Endorsement rates

Data usage model	Time pressure		
	High	Low	Undefined
Unbiased	5	14	15
Negative bias (1)	26	11	13
Positive bias (1)	9	15	12

Figure A19-9 Frequency of best-firs for data usage model in time pressure.

Data usage model	Time pressure			
	High	Low	Undefined	
Unbiased	.56	.70	.69	*.65*
Negative bias (1)	.62	.72	.70	*.68*
Positive bias (1)	.54	.68	.72	*.65*
	.57	*.70*	*.70*	

Figure A19-10 Mean multiple correlations for time pressure

Scanning and Quality-of-Performance Scores in the Three Experimental Conditions

Variable	Uncontrollable stress	Controllable stress	No stress	$F(2, 98)$	p
Premature closure					
M	5.21	4.36	1.29	10.13	.0001
SD	4.67	4.13	2.24		
Nonsystematic scanning					
M	8.30	8.64	3.20	7.85	.0007
SD	6.69	7.92	3.97		
Temporal narrowing					
M	3.03	3.32	3.12	0.52	.59
SD	1.13	1.32	1.18		
Quality of performance					
M	4.85	6.06	8.91	12.56	.0001
SD	2.60	3.80	3.78		

Figure A19-11 Scanning and quality of performance scores in the experiments.

Appendix A20 PIF Attributes and Weights for Physical Demands

Table A20-1 Attribute Identifiers and Descriptions for PIF Physical Demands

ID	Attribute
PD0	No impact
PD1	Physically strenuous action execution – Approaching or exceeding physical limits, e.g., lifting, handling, or carrying heavy objects, opening/closing rusted or stuck valves (Note: Heavy loads is defined in NUREG-0612: "Any load, carried in a given area after a plant becomes operational, that weighs more than the combined weight of a single spent fuel assembly and its associated handling tool for the specific plant in question.")
PD2	High spatial or temporal precision of fine motor movement needed for action execution
PD3	Precise coordination of joint action by multiple persons
PD4	Unusual loading or unloading materials (e.g., unevenly balanced loads, reaching high parts, dry cask loading)
PD5	Handling objects using crane⁄hoist

Table A20-2 IDHEAS-DATA IDTABLE-20 – PIF Weights for Physical Demands

1	2	3			4	5	6	7
PIF	**CFM**	**Error rates or task performance indicators**			**Task (and error measure)**	**PIF measure**	**Other PIFs (and Uncertainty)**	**REF**
PD1	E	Figure A20-1			Scope of load lifting & carrying task demands for US soldiers.	Weights of lifting or carrying loads	(No error data)	[296]
PD1	E	Several published regression equations can be used to predict team performance of manual materials handling. Dependent variables included measures of muscle strength, anthropometric characteristics, and gender of team members. These equations were able to account for between 35% – 98% of the variance in team performance, but most reported a relatively large standard error of the estimate, making them of limited practical use.			Team performance of manual materials handling	Personnel factors affecting team performance	(Literature review)	[296]
PD1	E	U.S. Military Standard 1472 F provides recommendations to team lifting. For two-person teams lifting from floor level to 91 cm, the standard recommends doubling the one-person load (79 kg for two men, 40 kg for two women), and a maximum of 75% of the one-person value can be added for each additional lifter beyond two.			Team lifting load or carrying tasks	Task demanding – weight, height of lifting, distance of carrying	(No error rate)	[296]
PD-1	E	Performance Demands	Operate a transport vehicle (% Contribution)	Manually lift and move designated heavy materials (% Contribution)	(Table 6.5) Notional performance demand profiles of hypothetical generalized actions in flood hazard	Contribution of generic tasks to the performance demands of a	(Engineering judgment based on task analysis)	[85]

		Detecting and noticing	40%	20%		manual action		
		Action – fine motor	30%	20%				
		Action – gross motor	30%	60%				
PD2	E		Same hand correction	Hand-switch correction	Repetition - repeat the same task Switch – randomly switch several tasks	Correction of execution errors	(Simple psych-motor tasks)	[297]
		Repetition	0.15	0.13				
		Switch	0.23	0.19				
PD3	E	See Figure A20-2			An overview of the major cognitive, sensorimotor, affective, and cultural processes supporting joint action – the variety of coordination mechanisms underlying joint action	No error data provided but many references of the paper have error data		[298]
PD4	E	HFE group 1: Before & during fuel loading Scenarios: 1. Failure in fuel-movement planning results in misload of ≤ 13 spent fuel assemblies with wrong fuel 2. Failures of multiple personnel in fuel movement results in misload of ≤ 4 spent fuel assemblies 3. Failures of one person during fuel movement results in misload of ≤ 4 spent fuel assemblies 4. Omission of in-pool staging results in misload of ≤ 4 spent fuel assemblies 5. Failures during fuel movement lead to misload with wrong fuel 6. Fuel-handling failures damage fuel during placement				No error data, 8 types of nuclear waste handling scenarios were described		[83]
PD4	E	**Distribution of events by type of load (% of events)**			114 NPP heavy load handling events were analyzed	Types of load	(Causal analysis)	[84]
		Nuclear fuel	30%					
		No load	19%					
		Control rods or parts	5%					
		Container with radiological waste	19%					
		Test load	3%					
		RPV head or internals	5%					
		Other loads	19%					
PD4 / PD5		**Distribution of events by failure mode (%)**			114 NPP heavy load handling events were analyzed, eight different main failure modes have been identified, covering more than 90% of the events	Failure modes	Causal analysis)	[84]
		Lifting interface failure	21%					
		Crane or lifting device failure	17%					
		Collision during handling	14%					
		Unauthorized crane operation	13%					
		Slings/wire/rope/chain breakdown	10%					
		Crane controls/device failure	8%					
		Hoist emergency breaks failure	6%					
		Other	9%					
PD4		Low	E-4/operation		25. Dropping of load when using forklift	NA	(Expert judgment)	[37]
		Nominal	E-3					
		High	E-2					
PD5		Low	E-5/operation		27. Dropping of load when using crane/hoist	NA	(Expert judgment)	[37]
		Nominal	E-4					
		High	E-3					
PD5		Low	E-5/operation		28. Crane/hoist strikes stationary object	NA		[37]
		Nominal	E-4					

		High	E-3			(Expert judgment)	
PD5		See Figure A20-3		"Independent Oversight Special Study of Hoisting and Rigging Incidents within the Department of Energy" covers a 30-month interval, from October 1, 1993 to March 31, 1996	Root causes as shown in the data table.	Causal analysis)	[82]
PD5	E	The number of incidents associated with operator failure is an astonishing 90 to 95% (Figure A20-4)		Navy crane incidents: Failure of the Trudock crane system at the waste isolation pilot plant (WIPP)	Incidents due to equipment failure vs due to operator failure	Causal analysis	[299]

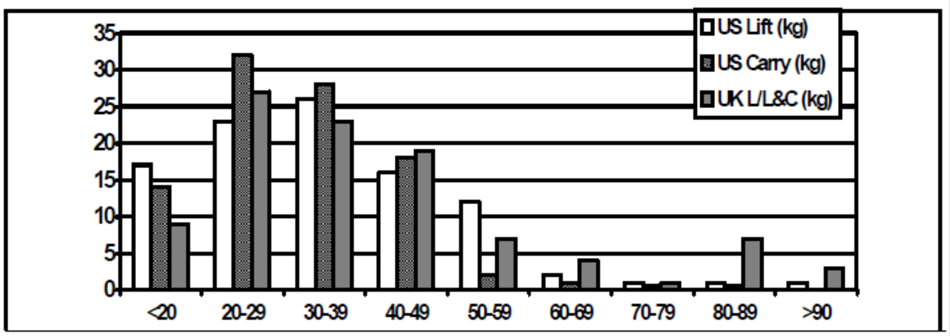

Figure 5-2: Frequency Distribution of Loads Handled by US and UK Soldiers.

Figure A20-1 A frequency diagram of the loads lifted and carried by US and UK Army Soldiers.

Coordination mechanism	Example
Mental representations in joint action	
Joint action goal	Relocating a sofa by lifting and moving it together
Task (co-)representation	Carrying a sofa forward or backward
Monitoring	Noticing errors in a co-actor's performance
Sharing sensorimotor information	
Joint attention and shared gaze	Being mutually aware of an obstacle in the way
Sensorimotor prediction	Predicting a co-actor's movement direction
Sensorimotor communication	Pushing a co-actor into a certain direction
Haptic coupling	Feeling a co-actor pushing the sofa
Multisensory processing	Integrating information from different senses
Emotion understanding and expression	Realizing how exhausted a co-actor is
General mechanisms supporting coordination	
Coordination smoothers	Distributing the task of moving forward or backward
Affordances	Being constrained by available space and a co-actor's physical strength
Conventions and culture	Appreciating rules about who carries more weight

Figure A20-2 Overview of different coordination mechanisms supporting joint action, along with a set of examples.

Root Cause	Crane	Forklift	Other
Inattention to Detail	20%	23%	8%
Work Organization and Planning	18%	3%	27%
Procedure Not Used or Used Incorrectly	9%	15%	0%
Policy Not Adequately Defined, Disseminated, or Enforced	9%	10%	4%
Inadequate or Defective Design	5%	5%	19%
Defective or Inadequate Procedure	9%	5%	0%
Inadequate Administrative Control	9%	0%	4%
Defective or Failed Part	5%	5%	8%
Other Management Problem	3%	3%	12%
Other Human Error	3%	3%	0%
Inadequate Work Environment	0%	10%	0%
Lack of Procedure	2%	3%	4%
Insufficient Refresher Training	3%	3%	0%
Insufficient Practice or Hands-On Experience	5%	0%	0%
Communication Problem	2%	3%	4%
Inadequate Supervision	0%	3%	4%
Error in Equipment or Material Selection	0%	3%	4%
Weather	0%	3%	0%
No Training Provided	0%	0%	4%

*Rounded to the nearest whole number.

Figure A20-3 Root Cause of Hoisting and Rigging Incidents by Equipment Type

Frequencies of Navy Crane Incidents

Year	1998	1997	1996
Total no. of incidents	196	167	154
No. due to equipment failure/percentage	11/5.6%	16/9.6%	7/4.6%
No. due to operator failure/percentage	185/94.4%	151/90.4%	147/95.4%

Figure A20-4 Frequencies of Navy crane incidents.

Appendix A21 Lowest HEPs

Table A21-1 IDHEAS-DATA IDTABLE-21 – Lowest HEP

1	2	3	4	5	6
CFM	**Error rate**	**Task and context**	**Criteria for lowest HEPs:** **TA - Time adequacy** **SelfV - Self verification** **TeamV – Team verification** **Rec - Recovery** **O - other factors** **(Y-Yes, N – No, M-Mixed Un-Unknown)**	**Uncertainty**	**REF**
D	2.1E-3 (4/1872)	NPP operators alarm detection in simulator training - Alarms are self-revealing	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	(Other PIFs may exist)	[26]
D	3.4E-3 (3/870)	NPP operators check indicators in simulator training - procedure directed checking.	TA-Yes, SelfV-Yes, TeamV-yes, Rec – Unknown O - Y (unspecified)	(Other PIFs may exist)	[26]
D	5E-4	Military operators read meters, Alphanumerics reading, - Detection straightforward	TA-Y, SelfV-Y, TeamV-No, Rec-No	(Maybe time constraint, 10K+ source data trials)	[109]
D	E-4	Estimated lowest probability of human failure events	TA-Yes, SelfV-Yes, TeamV-yes, Rec - Unknown	(Engineering judgment)	[110]
D	E-4	Simplest possible tasks	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec - Unknown	(Engineering judgment)	[111]
D	E-3	Routine simple tasks	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec – Unknown O – Maybe weak complexity	(Engineering judgment)	[111]
D	5E-3	Line-oriented text editor (Error rate per word)	TA-Yes, SelfV-Yes, TeamV-No, Rec - No	Not analyzed	[112]
D	5E-3	Reading a gauge incorrectly (Error rate per read)	TA-Yes, SelfV-Yes, TeamV-No, Rec – Unknown O – HSI	Not analyzed	[113]
D	E-3	Interpreting indicator on an indicator lamp (Error rate per interpretation)	TA-Yes, SelfV-Yes, TeamV-Unknown, Rec – Unknown O- complexity in interpreting indicator	(Engineering judgment)	[109]
D	9E-4	NPP operator simulator runs	TA – Y, Selv-V – Y TeamV – Y, Rec – Unknown O – Mixed complexity	No apparent uncertainty	[114]
D	5.3E-4	Gather information and evaluate parameters	TA – Y, Selv-V – Y TeamV – Y, R – Yes	No apparent uncertainty	[300]

D	9E-3	Collision avoidance and target monitoring in simulated ship control - Fixed situation, routine response	TA – Y, Selv-V – Yes TeamV – No, R – Yes O – Dual task, and maybe mixed complexity, mental fatigue, time pressure	Dual task	[27]
U	8.1E-3 (19/2350)	NPP operators diagnose in simulator training - Ambiguous Information NOT existing	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	Other PIFs exists	[26]
U	7.7E-3 (10/1293)	NPP operators diagnose in simulator training - Information specificity: specific	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	Other PIFs exists	[26]
U	7.7E-3 (20/2582)	NPP operators diagnose in simulator training - No missing information	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	Other PIFs exists	[26]
U	9.8E-3 (25/2552)	NPP operators diagnose in simulator training - No misleading information	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	Other PIFs exists	[26]
U	0.0143 (9/360)	NPP crew simulation with soft control in CR (Diagnosis error). See Figure A21-1.	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	No apparent uncertainty	[301]
U	4E-2	Student controllers performed air traffic control (near miss rate)	TA-Yes, SelfV- Unknown, TeamV- No, Rec -Unknown O – Y (Task complexity and poor training)	Task complexity and poor training	[124]
U	3.9E-3	NPP operator simulator runs	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	No apparent uncertainty	[114]
U	1.9E-3	Identify procedure	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	No apparent uncertainty	[300]
U	1E-4	Plan and decide command strictly following procedures	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	No apparent uncertainty	[300]
DM	4.6E-3	NPP operator simulator runs - Follow procedure	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	No apparent uncertainty	[114]
U	0.04	Diagnosing a pattern; personnel uses structured information to guide diagnosis - Predictive situation	TA-Yes, SelfV-Y, TeamV-Y, Rec -Unknown O – Y (unspecified)	Task complexity	[28]
U	1E-4	Air traffic control (Operational error) - 100+min on shift,	TA-Yes, SelfV-Y, TeamV-Y, Rec - Unknown O – Unknown	With teamwork, recovery, and pilot redundancy	[118]
U	0(9/9)	Physician diagnosis - High-context with all information		(Experiment study)	[126]

U &DM	3.8E-3	Pilots' flight (error rates) - Flight hour > 5000			[88]
DM	9E-5	Maintenance of the disc brake assembly (decided to omit part of the task) - No over-riding information			[123]
DM	5E-3	Maintenance in cable production process (wrong task plan) - Good quality of information		(Estimation)	[121]
DM	6.2E-2	NPP operator simulator runs - Plan for manipulation	TA – Y Selv-V – Y TeamV – Y Rec - Unknown	(Error definition may be different)	[114]
DM	1.3E-2	Licensed driver simulator (%collision) - fast driving early real-end information	TA-No, SelfV-No, TeamV – No Rec - No O – Y (unspecified)	Time inadequate	[125]
U & DM	7.9E-2	Pilots in-flight deicing (Percentage of early buffet, i.e., a low stool or hassoc) - Accurate information timely with status displays	TA-No, SelfV-Y, TeamV – No Rec - Mixed O –Multitasking	Inadequate time	[30]
E	4E-3 (5/1281)	NPP crew simulation with soft control in CR – Operation omission (Figure A21-1)	TA-Yes, SelfV-Y, TeamV-Y, Rec - Y O – Y (unspecified)	(Error definition may be different)	[301]
E	7.9E-3	NPP operator simulator runs - execute procedures	TA – Y Selv-V – Y TeamV – Y Rec - Unknown	(Error definition may be different)	[114]
E	9E-4	Maintenance in processing plant soldering	TA – Y Selv-V – Y TeamV – Unknown Rec - Unknown	Data-based estimation	[302]
E	4.8E-3	Component selection	TA – Y Selv-V – Y TeamV – Unknown Rec - Unknown	Data-based estimation	[302]
E	5E-3	Not available	TA – Yes V – SelfV and teamV Rec – Yes	Not analyzed	[303]
E	3E-4	Bank machine operators, errors per check	TA- Y V – SelfV Rec – Un	Not analyzed	[304]
E	E-4	Simplest possible tasks	Not available	Not analyzed	[111]
E	E-3	Routine simple	Not available	Not analyzed	[111]

E	8 E-4 (1/1470)	Manually operating a local valve, frequently performed task, valve not operated, step in a sequence of different steps not remembered - No known PIF exists	TA – Y, SelfV- Y, TeamV - Unknown Rec - Unknown	Error rates were for steps of a task, most tasks performed may not have peer-checking, some errors made may have been recovered so they did not get into the reporting system.	[4, 5]
E	8.9E-4 (7/8058)	Operating a control element on a panel, wrong control element selected - Similar controls within reach			
	8.78E-4 (1/1347)	59 Operation of a manual control at MCR control (Task not remembered) - Frequently performed task, part of professional knowledge, position of indicator lamps ergonomically unfavorably designed			
E	7.78E-5 (1/15,200)	Pulling and replugging a simulation pin on an electronic module front cover in a control cabinet; Errors were replugging omitted, highly trained task, not part of a written procedure but part of professional knowledge, favorable ergonomic design - No known PIF exists			
E	1.13E-4 (0/2010)	Reading instructions in a written procedure; Errors were Omitting to read one instruction			
E	1.13E-4 (0/2010)	Adjusting a process parameter by push- button controls, Frequently performed task, part of professional knowledge - Long procedure, checkoff provisions			
E	1.04E-3 (2/2088)	Remembering professional knowledge, remembered incorrectly, part of frequently performed procedure			
E	1.03E-3 (3/3067)	Carrying out a sequence of tasks, errors were skipped steps, frequently performed			
E	1.2E-3 (1/948)	Operating a pushbutton control Wrong button - selected button within reach, similar buttons nearby, ergonomically well designed panel			
E	1.3 E-3 (1/ 913)	Adjusting actuation value of a pressure limiting valve (Deviation out of tolerance) - High accuracy necessary			
E	8.9E-4 (1/1332)	5 Operating a rotary control Wrong switch - selected switch within reach, similar switches nearby, text labeling only			

E	7.8E-4 (1/1512)	7 Connecting a cable between an external test facility and an electronic module. Connected to wrong module panel, mimic layout - Module access ports within reach, similar access ports nearby, frequently performed task, color coding of ports			
E	1E-3 (1/1146)	9 Operating a push button control (Wrong button selected) - Similar buttons within reach, text labeling only			
E	1.2E-3 (3/2630)	Plain text labeling, similar controls within reach			
E	2.1E-3 (4/1958)	Operating a control element on a panel (Wrong control element selected) Mimic diagrams, color coding, similar controls within reach			
E	1.6E-3 (7/4588)	Operating a control element on a panel (Wrong control element selected) - Wrong control element within reach and similar in design			
E	E-4	Lowest HEP of an event or task (performing off-shore oil operation)	TA – Y, Selv-V – Y TeamV – Y, Rec – Y O - No	(Engineering judgment)	[110]
E	E-5	Lowest HEP of an event or task (performing off-shore oil operation)	TA – Y, Selv-V – Y TeamV – Y, Rec – Y O - No	(Engineering judgment)	[110]
E	2.7E-3	Nuclear hard-copy data - During a shift the transport department brought a chemical load to the compound after permission had been arranged between two supervisors, but the correct paperwork did not arrive with the chemicals. Consequently this led to two cans of highly enriched chemical solution being processed instead of six cans of low enriched chemical	TA – Y, Selv-V – Y TeamV – Y, Rec – Unknown O – Y (unspecified)	(Engineering judgment)	[305]
E	3.9E-4	Manufacturing (Confidential) real data - A component has a different profile machined on each end. The operator inadvertently machines the aft end profile on the forward end.	TA – Y, Selv-V – Y TeamV – Y, Rec – Unknown O – Y (unspecified)	(Engineering judgment)	[305]
E	48 students majoring in nuclear engineering - NPP simulator procedure execution (Figure A21-2)		TA – Y SelfV – Y TeamV – Y Recov - No		[107]
E	Failure of recovery - 48 students majoring in nuclear engineering - NPP simulator procedure execution (Figure A21-3)		TA – Y SelfV – Y TeamV – Y		[107]
E	9E-4	Maintenance and repair in cable production process			[121]

		- familiarity with the task in-hand			
D / E	0.007	Omission errors - Operator crew simulator re-training	TA-Y SelfV-Y TeamV-Y Recov-Y		[306]
D /E	0.01	Unrecovered omission errors - Operator crew simulator re-training	TA-Y SelfV-Y TeamV-Y Recov- Y		[306]
D /E	4E-3	Commission errors - Operator crew simulator re-training	TA-Y SelfV-Y TeamV-Y Recov- Y		[306]
D /E	2E-3	Unrecovered commission errors - Operator crew simulator re-training	TA-Y SelfV-Y TeamV-Y Recov- Y		[306]
T	2E-3	Speech sample (speech errors) per word	TA-Y SelfV-Y TeamV-No Recov- No		[307]
T	2E-3	Aviation communication errors	TA-Y SelfV-Y TeamV-No Recov- No		[305]
Un sp	2E-5 (800/4E7)	ATC OE per operation	SelfV – Y TeamV – Y Recov - Y	Recovery is high	[117]
Un sp	2E-4 (290/1.4 E6)	ATC OE per shift	SelfV – Y TeamV – Y Recov - Y	Recovery is high	[118]
Un sp	1.47E-2	NPP Requal simulation data – Perform procedures	SelfV – Y TeamV – N Recov - Unknown		[87]
Un sp	7.3E-3	NPP Requal simulation data – Perform procedures	SelfV – Y TeamV – Y Recov - Unknown		[87]
Un sp	3.85E-3	Pilot errors causing accidents	TA – Mixed SelfV – Y TeamV – Mixed Recov - Mixed		[88]
Un sp	5.5E-6 (686/(1.25×E8))	Pilot error rate x ATC error rate = NTSB reported human error accident rate TABLE A21-2. The event classifications of the 686 Events Reviewed in the NTSB database from about 1.25×10^{8} Total Flights.	TA – Mixed SelfV – Y TeamV – Y Recov - Y		[119]

Number of human errors and probabilities of human errors according to error modes.

	Errors/opportunities	Probabilities of human errors	95% confidence limits	Error factors
E1 (operation omission)	5/1281	0.0039	0.0005–0.0073	3.87
E2 (wrong object)	11/756	0.01455	0.0060–0.0231	1.96
E3 (wrong operation)	5/441	0.01134	0.0015–0.0213	3.82
E4 (mode confusion)	1/42	0.02381		
E5 (inadequate operation)	12/504	0.02381	0.0105–0.0371	1.88
E6 (delayed operation)	71/504	0.01389	0.0037–0.0241	2.56
Diagnosis error	9/360	0.0143	0.0050–0.0236	2.17

Figure A21-1 Number of human errors and probabilities of human errors according to error modes.

Human error modes	Errors/opportunities	HEP	5–95% Confidence interval
E_0 (operation selection omission)	5/1296	4.00E-3	0.0018–0.0076
E_1 (operation execution omission)	15/5232	2.90E-3	0.0018–0.0043
E_{2SS} (wrong screen selection)	44/2208	2.00E-2	0.0155–0.0253
E_{2DS} (wrong device selection)	19/2544	7.50E-3	0.0051–0.0107
E_3 (wrong operation)	11/1488	7.50E-3	0.0044–0.0118
E_4 (mode confusion)	22/672	3.30E-2	0.0229–0.0456
E_5 (inadequate operation)	11/720	1.55E-3	0.0091–0.0243
E_6 (delayed operation)	15/3024	5.00E-3	0.0032–0.0074

Figure A21-2 Human error probabilities with 5–95% confidence interval.

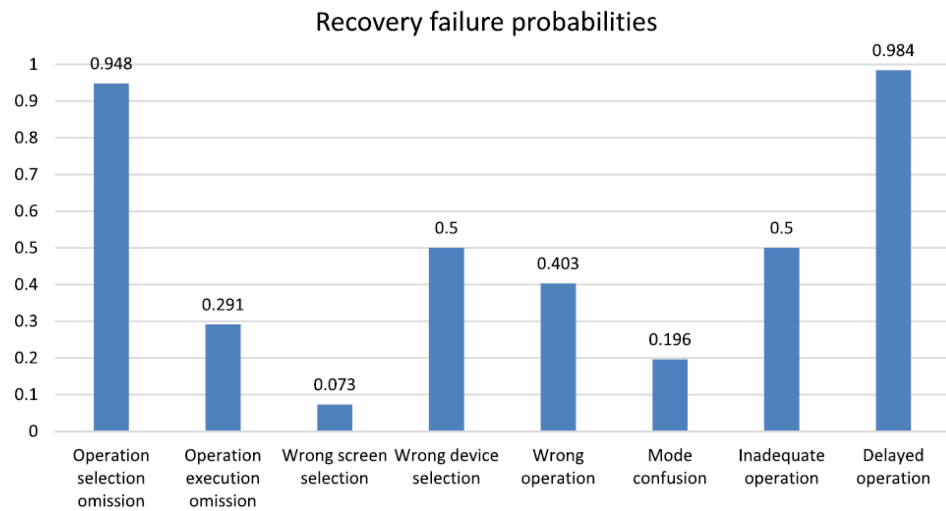

Figure A21-3 Recovery failure probabilities according to human error modes obtained from the experiments.

Table A21-2 Pilot error event classification

Classification	# of events
HFEs attributed to pilots, ATC and GTC	179
HFEs attributes to ground service (e.g., snowplowing and deicing)	71
Human-in-operation successfully avoided an undesired consequence	270
The situation is beyond the control of the human-in-operation	3
Insufficient information to determine	27
Passenger or flight attendant injury not attributed to pilots' fault	136
Total	686

Appendix A22 PIF Interaction

Table A22-1 IDHEAS-DATA IDTABLE-22 – PIF Interaction

<table>
<tr><th rowspan="2">CFM</th><th rowspan="2">Task and error measure</th><th rowspan="2">PIF measures</th><th colspan="2">PIF1 – Lo</th><th colspan="2">PIF1 -High</th><th rowspan="2">Other PIFs (and uncertainty)</th><th rowspan="2">Ref</th></tr>
<tr><th>PIF2-Lo</th><th>PIF2-High</th><th>PIF2-Lo</th><th>PIF2-High</th></tr>
<tr><td>D</td><td>Pilots read aircraft instrument dials as the luminance (c/m^2) of dials and degree of acceleration (+Gx) vary, errors are percent of misreading dials</td><td>PIF1 – VIS: Luminance Lo=15 c/m^{2,} Hi=0.15 c/m^2
PIF2 – PR: Acceleration
Lo = 2G, Hi=4G</td><td>0.07</td><td>0.15</td><td>0.2</td><td>0.45</td><td>Maybe time constraint</td><td>[14]</td></tr>
<tr><td>D</td><td>Pilots read aircraft instrument dials as the luminance (c/m^2) of dials and degree of acceleration (+Gx) vary, errors are percent of misreading dials</td><td>PIF1 – VIS: Luminance 0.015, 0.15, 1.5, 15, 150 c/m^{2,}
PIF2 – PR: Acceleration
Lo = 2G, Hi=4G</td><td colspan="4"><table><tr><td>PIF1 \ PIF2</td><td>2G</td><td>4G</td></tr><tr><td>150 c/m^2</td><td>0.07</td><td>0.07</td></tr><tr><td>15</td><td>0.07</td><td>0.15</td></tr><tr><td>1.5</td><td>0.10</td><td>0.20</td></tr><tr><td>0.15</td><td>0.20</td><td>0.45</td></tr><tr><td>0.015</td><td>0.50</td><td>0.63</td></tr></table></td><td>Maybe time constraint</td><td>[14]</td></tr>
<tr><td>Unsp</td><td>Meta-analysis of 55 reports to assess the strength and consistency of the multiplicative effects of cognitive ability and motivation on performance</td><td>PIF1 – Cognitive abilities
PIF2 - Motivation</td><td colspan="5">The effects of ability and motivation on performance are additive rather than multiplicative. For example, the additive effects of ability and motivation accounted for about 91% of the explained variance in job performance, whereas the ability-motivation interaction accounted for only about 9% of the explained variance. In addition, when there was an interaction, it did not consistently reflect the predicted form (i.e., a stronger ability-performance relation when motivation is higher).</td><td>[89]</td></tr>
<tr><td>Unsp</td><td>Regression fitting of human error data on empirical combined effects of multiple PSFs from 31 human performance papers and calculated their multiplicative and additive effects</td><td>Unspecified, all kinds of PIFs</td><td colspan="5">The median of the multiplicative effect was greater than that of the empirical combined effect, whereas the median of the additive effect was not significantly different from that of the empirical combined effect. Thus, the multiplicative model might yield conservative estimates, whereas the additive model might produce accurate estimates. The additive form is more appropriate for modeling the joint effect of multiple PSFs on HEP.</td><td>[308]</td></tr>
<tr><td>Unsp</td><td>This study investigated whether conscientiousness and ability interact in the prediction of job performance - Moderated hierarchical regression analyses for three independent samples of 1000+ participants</td><td>PIF1 – general mental ability (GMA)
PIF2 - conscientiousness</td><td colspan="5">Results in the present study provided no support for the interaction of GMA and conscientiousness. It showed that the interaction did not account for unique variance in the prediction of supervisory ratings of job performance beyond that accounted for by GMA and conscientiousness. These findings indicate that ability does not moderate the relationship of conscientiousness to job performance. (See Figure A22-1)</td><td>[91]</td></tr>
<tr><td>Unsp.</td><td>Analyzed 23 datapoints of human error rates varying</td><td>Different PIFs, e.g., shown in Figure A22-2</td><td colspan="5">1. The multiplicative rule tends to over-estimate the combined effect of PIF</td><td>[1]</td></tr>
</table>

	with single PIFs and two combined PIFs and fitted the dataset to multiplicative versus additive models		indicators on error rates, while the additive rule can roughly interpret the results 2. The individual and combined effects of PIF indicators can behave differently if the indicators show a demand on cognitive resources that exceeds the cognitive limits	
Unsp	Review of studies about the effect of combined environmental factors on human errors	Environmental factors: Noise, temperature, sleep deprivation, and others	Combined effect is no more than the added single effects and can be predicted from single effects	[94]
Unsp	Combined environmental stress	Environmental stresses: Noise, temperature, ambient light, vibration, sleep deprivation.	For possible effects of joint stressors, with Outcome 1 and 2 are prevalent while number 3 is rare but is important to hazard: 1. No effect. Combinations produce no effects greater than those of any of the included stressors individually 2. Additive effect. Combinations produce effects greater than any single stressor, but not greater than addition of effects from single stressors 3. Great than additive effect 4. Subtractive effect	[95]
Unsp.	This paper examines the combined effects of heat and noise upon behavioral measures of human performance. Specifically, capabilities on a variety of neuromuscular and mental tasks are reviewed with respect to their vulnerability to joint thermal and acoustic action.		Most of the evidence indicates that heat and noise do not interact significantly within the ranges experienced commonly in the industrial setting. However, various experimental and methodological inadequacies in the data caution against a simple interpretation of this apparent insensitivity.	[92]

Step	*Variable*	R^2	R^2 *Change*	*B-Weight*	*95% CI for B*
1.	Cognitive Ability	.058	.058*	0.021*	$.009 \leq .021 \leq .033$
2.	Conscientiousness	.122	.064*	0.531*	$.281 \leq .531 \leq .781$
3.	Ability × Conscientiousness	.123	.001	0.003	$-.032 \leq .003 \leq .038$
	(Constant)			1.533	

$N = 121$
*95% confidence interval does not include zero.

Figure A22-1 Results of Hierarchical Regression Analyses of Cognitive Ability, Conscientiousness, and their Interaction for District Managers

Ref.	Task	PIF indicators	Error rates
Lee[15]	Driving simulation	A: Warning time B: Driving speed	R (1, 22, 3, 31) W (22, 3, 31)
Cummings[17]	Air traffic control simulation	A: HSI format B: Task load	R (4,8,12,20) W (2, 3, 5)
Colquloun[24]	Detection	A: Signal saliency B: Criterion	R (25,45, 45,70) W (1.8,1.8, 2.8)
Strayer[25]	Driving simulation	A: dual-task B: Distraction	R (2.5,4,5, 6) W(1.8,2,2.4)
Xing[26]	Color use tests	A: Complexity B: Color vision deficiency	R (2,4,32,34) W(2,16,17)

Figure A22-2 Error rates for individual and combined PIF indicators

Appendix A23 Probability Distribution of Time Available

Table A23-1 IDHEAS-DATA IDTABLE-23 – Distribution of Time Needed

1		2		3	4
Task Description		Mean (min)	SD (min)	Note	Ref.
Basic SGTR Events The time operator spent on from beginning of the SGTR to the ruptured SG isolated.	6 actual SGTR events of U.S. nuclear power plants with the rupture flow rate greater than 300 gpm.	18.5	5.5		[96]
	23 Korean crews performed simulator re-training of SGTR events in a Korea standard nuclear power plant (KSNP) simulator, a 1000MWe CE pressurized water reactor (PWR) with conventional control interfaces.	19.8	3.0		[96]
	6 Korean crews performed simulator re-training of SGTR events in a KSNP simulator (a 950MWe Westinghouse 3-loop PWR) with conventional control interfaces. Most crews identified SGTR symptoms before reactor trip and implemented procedures quickly.	13.8	3.6		[98]
	3 US crews performed simulator runs of a basic SGTR events in their home simulator, a 4-loop Westinghouse PWR with conventional control interfaces. The tube rupture flow rate is 500 gpm. Basic SGTR event in the US HRA Benchmark Study.	19.0	3.5		[309]
	14 Swedish crews performed simulator runs of basic SGTR events in the HAMMLAB simulation facility, a 3-loop Westinghouse French PWR (CP0 series) with digitalized control interfaces. Basic SGTR event in the International HRA Benchmark Study.	15.9	3.6		[310]
Complex SGTR Events The time operator spent on from beginning of the SGTR to the ruptured SG isolated.	3 US crews performed simulator runs of a complex SGTR events in their home simulator, a 4-loop Westinghouse PWR with conventional control interfaces. The time operators spent from the beginning of the SGTR to the isolation of the ruptured SG. The SGTR occurred when the restored Auxiliary Feed Water was injected into the SG during a feed-and-bleed operation. Complex SGTR event in the US HRA Benchmark Study.	22.9	11.0		[309]
	14 Swedish crews performed simulator runs of complex SGTR events in HAMMLAB simulation facility, a 3-loop Westinghouse French PWR (CP0 series) with digitalized control interfaces. The time operators spent from the beginning of the SGTR to the isolation of the ruptured SG. The complication is the SGTR occurred immediately following a major main steamline break event. Complex SGTR event in the International HRA Benchmark Study.	26.9	6.4		[310]
	5 US crews of different plants performed simulation experiment at HAMMLAB on an event with a SG tube leak and SG tube rupture event with additional scenario complications. The time-required is from the time of the tube rupture to the ruptured SG being isolated.	45.8	6.5		[106]
Point Beach 1 (Westinghouse, 2-loop, 1800MWt) SGTR (rupture flow rate 125 gpm), occurred in 1975		58.0	NA		[96]
Fort Calhoun (CE, 1136 MWt) SGTR (rupture flow 112 gpm), occurred in 1984		40.0	NA		[96]
Based on 36 training records of an APR-1400 full-scope simulator, it was found that the **log-normal distribution** has the best fit (in comparison with normal, Gamma and Weibull distributions) on the time-required from reactor trip to complete the diagnosis procedure and transition to the event/function recovery procedure (i.e., diagnosis time) with the use of computerized emergency operating procedures.					[311, 312]

Appendix A24 Probability Distribution of Time Needed

Table A24-1 IDHEAS-DATA IDTABLE-24 – Modification to Time Needed to Complete a Human Action

1	2	3		4	5	6	7
CFM	**Time-Factor**	**Task completion time (mean and standard deviation, s- second, m-minute)**		**Task**	**PIF or Time Factor measure**	**Note**	**REF**
		Factor-Lo	**Factor-Hi**				
DM	MT2	110.3 (27.59)s	90.8 (30.83)s	Simple decisionmaking	Lo – No interruption Hi – With interruption	None	[102]
DM	MT2	608.3 (284.39)s	760.8 (293.76)s	Complex decisionmaking	Lo – No interruption Hi – With interruption	None	[102]
DM	MT2	831.3 (238.70)s	1702.5 (526.80)s	Complex decisionmaking	Lo- low interruption freq. Hi- High interruption freq.	None	[102]
DM	MT2	1317.4 (613.85)s	1842.0 (741.59)s	Complex decisionmaking	Lo- Different content Hi- Similar content	None	[102]
D	TC	38.11(5)s	46.44(4)s	Acquire information from radar visualization	Lo – 3-dimension info Hi – 7-dimension info	None	[313]
D	TC	30 (3)s	41.06(4)s	Acquire visualization information from flow charts	Lo – 3-dimension info Hi – 7 dimension info	None	[313]
D & U & E	TC	7.75 (4.76)s	62.33 (19.46) s	Perform procedure steps in NPP operator emergency training	Lo – complexity index = 1.279 Hi – complexity index = 2.58	None	[314]
D & U & E	TC	10.06 (5.31)s	74.60 (26.83s)	Perform procedure steps in NPP operator qualifying examination	Lo – complexity index = 1.279 Hi – complexity index = 2.58	None	[314]
		Time=44.76 x (complexity index) - 44.6					
D	MT1	N/A	88(25)s	Security-critical detection task requiring reading, comparing, and confirming Bluetooth numbers	Lo – No distraction Hi – static red visual stimuli for distraction	169 college students	[17]
D	MT1	35(12)s	90(16)s	Security-critical detection task	Lo – No distraction Hi - flickering red visual stimuli for distraction	169 college students	[17]
D	TMP	Effect size = -0.91 on response time		Perception tasks	Effect size of heat on response time	(meta-analysis)	[54]
U/ DM	TMP	Effect size = 0.02 on response time		Cognition tasks	Effect size of heat on response time	(meta-analysis)	[54]
E	TMP	Effect size = 0.68 on response time		Psych-motor Tasks	Effect size of heat on response time	(meta-analysis)	[54]
D	TMP	Effect size = -0.85 on response time		Perception tasks	Effect size of cold on response time	(meta-analysis)	[54]
U/ DM	TMP	Effect size = 0.64 on response time		Cognition Tasks	Effect size of cold on response time	(meta-analysis)	[54]

E	TMP	Effect size = -1.1 on response time		Psych-motor Tasks	Effect size of cold on response time	(meta-analysis)	[54]
E	PR	392(59)s	438(92)s	Soldiers on simple reaction time tasks	Lo – No protective suit Hi - Wearing protective suit	None	[100]
E	PR	73.5min	125.9min	Crews performed “Remove and Replace M60A3 Transmission"	Lo – Battle dress uniform Hi – Wearing MOPP 4 suit	None	[101]
Unsp	TE	9(1.5)s /per instruction	16(2)s	4 NPP crews perform EOP scenarios	Lo – Experienced with AP1400 Hi – No experience with AP1400	(4 crews)	[99]
Unsp	TPS	13(2.5)m	12(4)m	EOP scenarios	Lo - Urgent Hi - Less urgent	(4 crews)	[99]
Unsp	SF/ INF	12(5m)	14(2)m	EOP scenarios	Lo - Design basis event Hi - Design basis event and masking	(4 crews)	[99]

Appendix A25 Dependency

Table A25-1 IDHEAS-DATA IDTABLE-25 – Instances and Data on Dependency of Human Actions

1	2	3
Dependency Type	**Narrative/Explanation**	**Ref**
Consequential	**Narrative:** On April 17, 2005, at 8:29 a.m., Millstone Power Station, Unit 3, a four-loop pressurized-water reactor, experienced a reactor trip from 100-percent power [315]. The trip was caused by an unexpected “A” train safety injection (SI) actuation signal and main steamline isolation caused by a spurious “Steam Line Pressure Low Isolation SI” signal. As a result of the main steam isolation signal, the main steam isolation valves and two of the four main steamline atmospheric dump valves automatically closed. With the closure of the main steam isolation valves, the main steamline safety valves opened to relieve secondary plant pressure. Control room operators entered Emergency Operating Procedure (EOP) E-0, “Reactor Trip or Safety Injection,” and manually actuated the “B” train of SI and actuated the “B” main steam isolation train in accordance with station procedures. Both motor-driven auxiliary feedwater (AFW) pumps started to maintain steam generator (SG) levels. The turbine-driven AFW pump attempted to start but immediately tripped on overspeed. Operators were dispatched to investigate the cause of the turbine-driven AFW pump trip. At approximately 8:42 a.m., the shift manager noted that a “B” main steam safety valve had remained open for an extended time. In consultation with the unit supervisor and shift technical advisor, the shift manager declared an alert based on a stuck open main steam safety valve. The crew determined that the stuck open main steam safety valve represented a non-isolable steamline break outside containment. The main steam safety valves were in fact functioning as designed to relieve post-reactor-trip decay heat with a main steamline isolation signal present. In this event, the main steam safety valves closed once the operators took positive control of decay heat removal by remotely opening the atmospheric dump bypass valves. At 8:45 a.m., because of the addition of the inventory from the SI, the pressurizer reached water solid conditions and the pressurizer power-operated relief valves cycled many times to relieve RCS pressure and divert the additional RCS inventory to the pressurizer relief tank. No pressurizer safety valve actuations occurred, and the pressurizer relief tank rupture diaphragm remained intact. At approximately 8:59 a.m., the operating crew transitioned from EOP E-0 to ES-1.1, “Safety Injection Termination.” The SI was reset, the crew terminated SI at 9:12 a.m., and normal RCS letdown was reestablished at 9:20 a.m. [315] **Explanation:** Failure to control RCS inventory resulted in a liquid-solid pressurizer that complicated the situation. Managing the complexity delayed the operators from entering ES-1.1 to terminate safety injection.	[315]
Consequential	**Narrative:** On October 4, 1990, at 1:24 a.m., Braidwood Unit 1 experienced a loss of approximately 600 gallons of water from the reactor coolant system (RCS) while in cold shutdown. Braidwood 1 technical staff was conducting two residual heat removal (RHR) system surveillances concurrently, an isolation valve leakage test and valve stroke test. After completing a leakage measurement per one surveillance procedure, a technical staff engineer (TSE) in the control room directed an equipment attendant to close an RHR system vent valve. However, before those instructions could be carried out, another TSE in the control room directed that an RHR isolation valve be opened per another surveillance procedure. While the equipment attendant was still closing the vent valve, RCS coolant at 360 psig and 180 °F exited the vent valve, ruptured a Tygon tube line and sprayed two engineers and the equipment attendant in the vicinity of the vent valve. This loss of coolant was reported to the control room and the control room personnel quickly identified the cause and isolated the leak.	[104]

	Explanation: The isolation valve leakage test (Test1) affected the boundary condition of the valve stroke test (Test 2). Failure to complete Task 1 (in this case, the RHR vent valve was not closed completely) made Task 2 impossible to be complete.	
Resource-Sharing	**Narrative:** On May 7, 2004, Palo Verde simultaneous tested the atmospheric dump valve and boron injection systems resulting in a loss of letdown event on high regenerative heat exchanger temperature. The procedures of the two surveillances were "atmospheric dump valve (ADV) 30% Partial Stroke Test" and "Boron Injection Flow Test." The simultaneous performance of these evolutions caused a loss of letdown due to the high regenerative heat exchanger outlet temperature. This condition occurred due to a single charging pump operation per "Boron Injection Flow Test" procedure and the combined excessive letdown flow to accommodate the RCS heat up following ADV partial stroke testing. **Explanation:** The two tests, one limited the charging flow and the other demanded excessive letdown, affecting the regenerative heat exchanger outlet temperature. Combination of the two tests resulted in exceeding the threshold of the exit temperature.	[316]
Cognitive Dependency	**Narrative:** On March 20, 1990, at about 09:30, Catawba Station Unit I experienced an over-pressurization of the Residual Heat Removal System (RHR) and Reactor Coolant System (RCS) during the procedure to initially pressurize the RCS to 100psig following a refueling outage. The operators had three indicators for monitoring RCS pressure (two wide range indicators, 0-3000psig, and one low range indicator, 0-800psig) which were being closely monitored for a detectable rise in RCS pressure. However, unknown to the control room operators on duty, all three RCS pressure instrument transmitters were still isolated after the welding of tube fittings during the refueling outage. **Explanation:** Deisolation of the three indicators (two wide range and one low range) requires a common cue. Failure to deisolate any indicator would result in failing to deisolate all three indicators.	[105]

Appendix A26 Recovery

Table A26-1 IDHEAS-DATA IDTABLE-26 – Instances and Data on Recovery Actions

1	2	3
Narrative of recovery actions	**Notes**	**Ref**
In the course of the startup of the plant, it was discovered that the isolation valves in each of the three high pressure safety injection lines to the cold legs of the primary circuit were in the closed position. Their power supplies were disconnected. One day before startup, a leak-tight test of the check (isolation) valves in the high-pressure injection system was performed. The test requires that the isolation valves be closed but not disconnected from the electrical power supply. The test procedure did not provide specific instructions to restore or verify the proper line-up of the safety injection system after the test. The day following the completion of the test, the operators verified the line-up of the safety injection system as instructed in operating procedures.	The recovery action of the operator's verification of the safety injection system line-up is feasible because it was directed by procedures. No dependency between the failed action and its recovery action because the recovery action was performed a day later, and it is likely that the safety system line-up verification was performed by different operators than the one that performed the test using different procedures. Also, Section 3.1 of Reference [20] analyzed 17 human failure events. Eleven events occurred in the outage phase, and 5 of these during start up. Another might be during power operation. Scheduled periodical tests detected most (9) of the events. In 5 events, the deficiencies occurred on demand and 3 deficiencies were detected by chance. This reference provides a data point of error recovery in maintenance surveillance tests as 0.7 (= 12/17).	[20]
This study investigated human error recovery failure probabilities by conducting experiments in the operation mockup of advanced/digital main control rooms (MCRs) in NPPs. 48 subjects majoring in nuclear engineering participated in the experiments. In the experiments, using the developed accident scenario based on tasks from the standard post trip action (SPTA), the steam generator tube rupture (SGTR), and predominant soft control tasks derived from the LOCA and the excess steam demand event (ESDE). All subjects were trained theoretically and practically before the experiments regarding EOPs and interfaces. Once the experiments were performed, each subject executed the task written in the procedure without any supervisor's assistance and there was no time pressure when per forming the tasks. The results are summarized in Figures A26-1 and A26-2.	The experiment was designed such that human error recovery was feasible (tasks recoverable, adequate time, sufficient manpower, having procedures, sufficient cues). The results show that recovery failure probability regarding wrong screen selection was the lowest among human error modes, which means that most of the human error relating to wrong screen selection can be recovered. On the other hand, recovery failure probabilities of operation selection omission and delayed operation were 1.0. These results imply that once the subject omitted one task in the procedure, they had difficulties finding and recovering their errors without the supervisor's assistance. Although there were cues for detecting errors and initiating recovery, the student subjects might not use the cues as effective as licensed operators. Recognizing the cues requires understanding of event progression and context, while the students might not have good understanding of the scenario context.	[107]
The Halden Reactor Project conducted a simulation study for collecting HRA data. Five crews of licensed operators from three power plants in the U.S. participated in the study. The participants worked at Westinghouse PWR	Only 20% of errors were recovered. Scenario 3 had the highest number (30) of errors and lowest recovery rate (4/30). Detection and Execution errors had much	[106]

plants/units comparable to the one simulated by the Ringhals Plant Simulator (RIPS). The crews varied in the number of operators: three, four, and five. Three scenarios were used: **Scenario 1: Multiple Steam Generator Tube Rupture.** In the simulated scenario the loss of reactor coolant starts as a small leak in one steam generator (SG) to slowly increase up to a large tube rupture in another SG. In addition, the leaks are preceded by disturbances that interfere with the unique symptoms for steam generator tube ruptures events, i.e. abnormal radiation in the secondary system. The crew had to identify the leak in SG2 and the rupture in SG3 based on other indications. **Scenario 2: Loss of coolant outside containment** This scenario reproduces an Interfacing Systems Loss of Coolant Accident (ISLOCA). This event occurs when valves in series that isolate the reactor coolant system (RCS) from the Residual Heat Removal (RHR) system fail. **Scenario 3: Total loss of feedwater and induced steam generator tube rupture** This scenario is a loss of all feedwater event followed by an induced steam generator tube rupture that occurs when emergency feedwater flow is eventually restored. The five crews totally made 65 errors. The report described the details of every task with an error and its recovery. The overall recovery rate was 20%, and time between the error made to the initiation of recovery actions varied from 2mins to 35mins.	higher recovery rates (2/5 and 5/18) than those of Understanding and Decisionmaking (1/17 and 4/25). This might be due to less salient cues for operators recognizing Understanding and Decisionmaking errors. The report did not provide information on feasibility of error recovery; thus, it is unclear how many of the 80% unrecovered errors were feasible for recovery. The observed high rates of operator errors and low recovery frequencies must be understood in the context of the simulated scenarios as well as the data collection approach: • The emergency scenarios were characterized by multiple malfunctions. • The emergency operating procedures were inefficient at various occasions in the scenarios. • The majority of the unsafe acts reported are high-level cognitive identifications, decision and actions, rather than simple/basic tasks. • The crews were operating a different plant, albeit similar, to the one they work at, and in a new/unfamiliar control room.	

Human error mode	Operation selection omission E_0	Operation execution omission E_1	Wrong screen selection E_{2SS}	Wrong device selection E_{2DS}	Wrong operation E_3	Mode confusion E_4	Inadequate operation E_5	Delayed operation E_6
Number of tasks	27	109	46	53	31	14	13	63
Number of opportunities	1296	5232	2208	2544	1488	672	720	3024
Number of errors	5	15	44	19	11	22	11	15
Number of recoveries	0	10	40	9	6	17	5	0
Recovery failure probability, q_{50}	*0.948*	*0.291*	*0.073*	*0.500*	*0.403*	*0.196*	*0.500*	*0.984*
$[q_5, q_{95}]$	*[0.64, 1]*	*[0.13, 0.50]*	*[0.03,0.16]*	*[0.32, 0.68]*	*[0.19,0.65]*	*[0.08,0.36]*	*[0.26,0.74]*	*[0.87,1]*

Figure A26-1 Recovery failure probabilities according to human error modes in advanced MCRs using soft controls

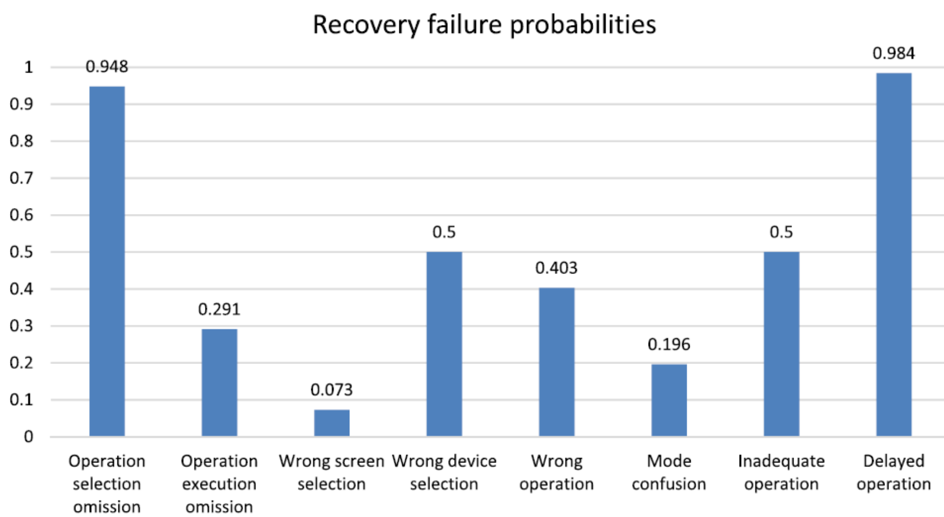

Figure A26-2 Recovery failure probabilities according to human error modes obtained from the experiments

Appendix A27 Main Drivers of Human Error

Table A27-1 IDHEAS-DATA IDTABLE-27 – Empirical Evidence on Main Drivers of Human Failure Events

1	2	3	4	5
CFMs	**PIFs**	**Error rate**	**Narrative of the event and Main drivers**	**Ref**
U	SF3, INF6	0.7 (7/10)	<u>Main Drivers: Inadequate knowledge, key information was cognitively masked.</u> This is HFE1B, i.e., initiating Bleed & Feed before steam generator (SG) dry-out in the complex Loss of Feed Water (LOFW) scenario, in the International HRA Benchmarking Study. The following are from section 2.3.2 of volume 3 of The International Benchmark Study report series: • The complex loss of feedwater (LOFW) scenario contained multiple issues. The first issue was that one condensate pump was successfully running, leading the crew to depressurize the SGs to establish condensate flow. However, the running condensate pump was degraded and gave a pressure so low that the SGs became empty before the pressure could be reduced enough to successfully inject water. • The procedure step to depressurize is complicated, and this action both kept the crew busy and gave them a concrete chance to re-establish feedwater to the SGs. The crews were directed by procedure FR-H.1 to depressurize the SGs to inject condensate flow. • Two of the three SGs had WR level indicators that would incorrectly show a steady (flat) value somewhat above 12% when the actual level would be 0% due to the degraded condensate pump. The two failing SG levels both indicated a level above the 12% criterion to start Bleed & Feed. To follow the criterion, the crews had to identify and diagnose the indicator failures, since the criterion, interpreted literally, would never be met.	[23]
D	INF, SF, HSI	N/A	<u>Main drivers: Scenario familiarity and Information reliability - the electric fault causes many indications to be momentarily unavailable.</u> In the event H.B. Robinson Steam Electric Plant electric fault with a near miss of reactor coolant pump (RCP) seal damage, an electrical fault occurred on a 4kV feeder cable and caused a fire that resulted in reactor trip. In the event, one key operator action was to reopen FCV626 to restore seal cooling or trip the RCPs to prevent RCP seal damage. The FCV-626 was located in the combined CCW return from the three RCP thermal barrier heat exchangers. In its normal open position, it allowed CCW flow to pass through the thermal barrier heat exchangers, providing backup cooling to the RCP seals in the event of a loss of the primary cooling flow (seal injection) from the charging pumps. The FCV-626 closed when power to the 480 V E-2 safety bus was transferred to the EDG. The valve remained closed for approximately 39 minutes before the operators recognized the condition, reopened FCV-626 at 19:31, and restored CCW cooling to the RCP thermal barrier heat exchangers. The crew failed to detect the RCP abnormal alarms. The key contributing factors are the following: • Information availability and reliability: The indications for this cue are genuine. However, the electric fault causes many indications to be momentarily unavailable. Some indications become available after the electric transition, and others remain unavailable throughout the event. The display reliability from the crew’s perspective is questionable. • Scenario familiarity: The MCR indications do not show a recognizable event pattern to the operating crew. Also, the operators’ expectation on information detection is biased, that is, when the crew was trained in the simulator for similar scenarios, the FCV-626 does not close. The crew would not expect the FCV-626 closure in this event; therefore, the operators do not have the motivation to check for the information. • Human-system-interface: The signal (cue) is weak or masked because there are simultaneously hundreds of alarms on the alarm panels. There	[317, 318]

			are also salience considerations about the information having a similar appearance with the surrounding information, that is, the alarm tiles relating to the cue are in the alarm panels with other similar alarm tiles.	
E	SF3	E-2 to E-1	<u>Main Drivers: Scenario familiarity - Tasks are rarely performed.</u> Significant events occurring in German nuclear installations are reported to the competent authorities if the notification criteria are fulfilled. After being reported, the events are analyzed and documented, and the event documentation is stored in the database BEVOR (6000+ events as of 2016, the year the analysis was performed). Preischl and Hellmich (2016) used a screening process to select a subset of events for analysis. Error rates were calculated for 67 types of tasks under different situations. The analysis shows that most of the high error rates are associated with rarely performed tasks. The snapshot table below, from the report, is a sample of error rates for carrying out a sequence of tasks. It shows that the error rates became larger as the number of times (the denominator m_i in the table) that the tasks were performed got smaller regardless of the presence or absence of other PIFs ("relevant PSFs" in Figure A27-1).	[5]
Unsp	Unsp	N/A	<u>Main drivers: Highly frequent error causes in NPP events: maintenance practices (54%), design deficiencies (49%), procedures (38%), communication and configuration management (27%).</u> Gertman et. al. (2002) studied the contributions of human performance to risk in operating events at commercial nuclear power plants. They reviewed 48 events described in licensee event reports (LERs) and Augmented Inspection Team reports. Human performance did not play a role in 11 of the events so they were excluded from the sample. In the remaining 37 events, 270 human errors were identified, and multiple human errors were involved in every event. The results show maintenance practices was highest (54%), followed by design deficiencies (49%), and procedures (38%). Errors in communication and errors in configuration management were each present in 27% of events. The numbers or percentages of error occurrences inform the prevalent types of human errors in the event sample analyzed.	[108]

Omission errors: task not remembered. HEP estimates resulting from sample 58, samples 30, 59 and 64, sample 35, samples 27, 34 and 65, samples 31 and 60, sample 28, and sample 66.

Task	*Error*	*Relevant PSFs*	m_i/n_i	q_{50}, $[q_5, q_{95}]$
Carrying out a sequence of tasks	Memorized task step not remembered	Highly trained, no error promoting factors	1/15,200	$7.78 \cdot 10^{-5}$, $[1.1, 26] \cdot 10^{-5}$
		Frequently performed, no error promoting factors	3/3067	$1.03 \cdot 10^{-3}$, $[0.3, 2.3] \cdot 10^{-3}$
		Rarely performed, no error promoting factors	1/48	$2.45 \cdot 10^{-2}$, $[0.3, 7.9] \cdot 10^{-2}$
		Rarely performed, moderately high level of stress	3/185	$1.71 \cdot 10^{-2}$, $[0.5, 3.8] \cdot 10^{-2}$
		Rarely performed, moderately high level of stress, ergonomically deficient work environment	2/41	$5.62 \cdot 10^{-2}$, $[1.4, 13] \cdot 10^{-2}$
		Rarely performed, moderately high level of stress, error prone PSFs and dynamic work environment	1/7	$1.61 \cdot 10^{-1}$, $[0.2, 4.4] \cdot 10^{-1}$
		Extremely rarely performed, no error promoting factors	1/3	$3.52 \cdot 10^{-1}$, $[0.6, 7.7] \cdot 10^{-1}$

Figure A27.1 Omission errors: task not remembered. HEP estimates resulting from sample 58, samples 30, 59 and 64, sample 35, samples 27, 34 and 65, samples 31 and 60, sample 28, and sample 66.

5 REFERENCES

1. Xing, J., Y.J. Chang, and J. DeJesus, *The General Methodology of an Integrated Human Event Analysis System (IDHEAS-G) — Draft Report*, 2019, NUREG-2198 (ADAMS Accession No. ML19235A161), U.S. Nuclear Regulatory Commission, https://adamswebsearch2.nrc.gov/webSearch2/main.jsp?AccessionNumber=ML19235A161.
2. Xing, J., J. Chang, and J. DeJesus, *Integrated Human Event Analysis System for Event and Condition Assessment (IDHEAS-ECA)*, 2020, RIL-2020-02 (ADAMS Accession No. ML20016A481), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML2001/ML20016A481.pdf.
3. Xing, J., M. Kichline, J. Hughey, and M. Humberstone, *Utilization of Expert Judgment to Support Human Reliability Analysis of Flexible Coping Strategies (FLEX)*, 2020, NRC Research Information Letter RIL-2020-xx U.S. Nuclear Regulatory Commission,
4. Preischl, W. and M. Hellmich, *Human error probabilities from operational experience of German nuclear power plants.* Reliability Engineering & System Safety, 2013. **109**: p. 150-159.https://doi.org/10.1016/j.ress.2012.08.004.
5. Preischl, W. and M. Hellmich, *Human error probabilities from operational experience of German nuclear power plants, Part II.* Reliability Engineering & System Safety, 2016. **148**: p. 44-56.https://doi.org/10.1016/j.ress.2015.11.011.
6. Xing, J., G. Parry, M. Presley, J. Forester, S. Hendrickson, and V. Dang, *An Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal Events At-Power Application, Volume 1*, 2017, NUREG-2199, U.S. Nuclear Regulatory Commission; Electric Power Research Institute, NRC ADAMS Assession Number: ML17073A041
7. O'Hara, J., C. Ruger, J. Higgins, W. Luckas, and D. Crouch, *An Evaluation of the Effects of Local Control Station Design Configurations on Human Performance and Nuclear Power Plant Risk*, 1990, NUREG/CR-5572, U.S. Nuclear Regulatory Commission, https://adamsxt.nrc.gov/AdamsXT/content/downloadContent.faces?objectStoreName=MainLibrary&ForceBrowserDownloadMgrPrompt=false&vsId={8C24E27A-92CA-428B-8FFC-B1FBFB208F37}.
8. Prinzo, O.V., A.M. Hendrix, and R. Hendrix, *The Outcome of [Air Traffic Control] Message Complexity on Pilot Readback Performance*, 2006, DOT/FAA/AM-06-25, Federal Aviation Administration, https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/200625.pdf.
9. Prinzo, O.V., A.M. Hendrix, and R. Hendrix, *The Outcome of ATC Message Complexity on Pilot Readback Performance*, 2006, DOT/FAA/AM-06/25, Civil Aerospace Medical Institute, Federal Aviation Administration, https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2000s/media/200625.pdf.
10. Patten, C.J.D., A. Kircher, J. Östlund, L. Nilsson, and O. Svenson, *Driver experience and cognitive workload in different traffic environments.* Accident Analysis & Prevention, 2006. **38**(5): p. 887-894.https://doi.org/10.1016/j.aap.2006.02.014.
11. Xing, J., Y.J. Chang, and N. Siu. *Insights on human error probability from cognitive experiment literature*. in *2015 International Topical Meeting on Probabilistic Safety Assessment and Analysis (PSA 2015)*. 2015.2015/04/26/30.
12. Virovac, D., A. Domitrović, and E. Bazijanac, *The Influence of Human Factor in Aircraft Maintenance.* PROMET - Traffic Transp., 2017. **29**(3): p. 257-266.https://www.researchgate.net/publication/319673282_The_Influence_of_Human_Factor_in_Aircraft_Maintenance.

13. Kyriakidis, M., K.T. Pak, and A. Majumdar, *Railway Accidents Caused by Human Error: Historic Analysis of UK Railways, 1945 to 2012.* Transportation Research Record: Journal of the Transportation Research Board, 2015. **2476**(1): p. 126-136.https://doi.org/10.3141/2476-17.
14. Altmann, E.M., J.G. Trafton, and D.Z. Hambrick, *Effects of interruption length on procedural errors.* Journal of Experimental Psychology: Applied, 2017. **23**(2): p. 216-229.http://doi.apa.org/getdoi.cfm?doi=10.1037/xap0000117.
15. Braunstein, M.L. and W.J. White, *Effects of Acceleration on Brightness Discrimination*.* Journal of the Optical Society of America, 1962. **52**(8): p. 931-933.http://www.osapublishing.org/abstract.cfm?URI=josa-52-8-931.
16. Forester, J., H. Liao, V.N. Dang, A. Bye, E. Lois, M. Presley, J. Marble, R. Nowell, H. Broberg, M. Hildebrandt, B. Hallbert, and T. Morgan, *The U.S. HRA Empirical Study – Assessment of HRA Method Predictions against Operating Crew Performance on a U.S. Nuclear Power Plant Simulator*, 2016, NUREG-2156 (ADAMS Accession No. ML16179A124), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1617/ML16179A124.pdf.
17. Berg, B., T. Kaczmarek, A. Kobsa, and G. Tsudik, *Lights, Camera, Action! Exploring Effects of Visual Distractions on Completion of Security Tasks.* arXiv:1706.00056 [cs], 2017.https://arxiv.org/pdf/1706.00056v1.
18. Hunter, C.S., D.I. Gertman, C.D. Gentillon, M.R. Holbrook, and D.G. Marksberry, *Review of Human Error Contribution to Operational Events - Summary Report*, 2006, Nuclear Regulatory Commission, NRC ADAMS Accession Number: ML062700429
19. Commission, U.S.N.R., *Palo Verde Nuclear Generating Station - NRC Integrated Inspection Report 05000528/2004003, 05000529/2004003, 05000530/2004003*, 2004, IR 05000528/2004003, 05000529/2004003, and 05000530/2004003 (ADAMS Accession No. ML042220267), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0422/ML042220267.pdf.
20. OECD/NEA, *Human Factor Related Common Cause Failure - Part 1, Report from the Expanded Task Force on Human Factors*, 1996, NEA/CSNI/R(95)10/PART1, Organisation for Economic Co-operation and Development, Nuclear Energy Agency, https://www.oecd-nea.org/nsd/docs/1995/csni-r95-10-part1.pdf.
21. Lois, E., V.N. Dang, J. Forester, H. Broberg, S. Massaiu, M. Hildebrandt, P.Ø. Braarud, G. Parry, J. Julius, R. Boring, I. Männistö, and A. Bye, *International HRA Empirical Study - Phase 1 Report: Description of Overall Approach and Pilot Phase Results from Comparing HRA Methods to Simulator Performance Data*, 2009, NUREG/IA-0216, Vol. 1 (ADAMS Accession No. ML093380283), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0933/ML093380283.pdf.
22. Bye, A., E. Lois, V.N. Dang, G. Parry, J. Forester, S. Massaiu, R. Boring, P.Ø. Braarud, H. Broberg, J. Julius, I. Männistö, and P. Nelson, *International HRA Empirical Study - Phase 2 Report: Results from Comparing HRA Method Predictions to Simulator Data from SGTR Scenarios*, 2011, NUREG/IA-0216, Vol. 2 (ADAMS Accession No. ML11250A010), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1125/ML11250A010.pdf.
23. Dang, V.N., J. Forester, R. Boring, H. Broberg, S. Massaiu, J. Julius, I. Männistö, H. Liao, P. Nelson, E. Lois, and A. Bye, *International HRA Empirical Study - Phase 3 Report: Results from Comparing HRA Methods Predictions to HAMMLAB Simulator Data on LOFW Scenarios*, 2014, NUREG/IA-0216, Vol. 3 (ADAMS Accession No. ML14358A254), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1435/ML14358A254.pdf.
24. Chang, Y.J. *Preliminary Analysis Results of SACADA Data for HEPs*. in *HRA Data Workshop*. 2020. Rockville Maryland, March 12-13, 2020.

25. Chang, Y.J., D. Bley, L. Criscione, B. Kirwan, A. Mosleh, T. Madary, R. Nowell, R. Richards, E.M. Roth, S. Sieben, and A. Zoulis, *The SACADA database for human reliability and human performance.* Reliability Engineering & System Safety, 2014. **125**: p. 117-133.https://doi.org/10.1016/j.ress.2013.07.014.
26. U. S. Nuclear Regulatory Commission. *Human Reliability Analysis Data - SACADA data*. 2019; Available from: https://www.nrc.gov/data/ (Human Reliability Analysis).
27. Robert, G., J. Hockey, A. Healey, M. Crawshaw, D.G. Wastell, and J. Sauer, *Cognitive Demands of Collision Avoidance in Simulated Ship Control.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2003. **45**(2): p. 252-265.http://journals.sagepub.com/doi/10.1518/hfes.45.2.252.27240.
28. Dodds, C.M., S. Morein-Zamir, and T.W. Robbins, *Dissociating inhibition, attention, and response control in the frontoparietal network using functional magnetic resonance imaging.* Cerebral Cortex, 2011. **21**(5): p. 1155-1165.
29. Xing, J. and S. Morrow, *White Paper: Practice Insights and Lessons Learned on Implementing Expert Elicitation*, 2016, NRC ADAMS Access No. ML16287A734, U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1628/ML16287A734.pdf.
30. Sarter, N.B. and B. Schroeder, *Supporting Decision Making and Action Selection under Time Pressure and Uncertainty: The Case of In-Flight Icing.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 2001. **43**(4): p. 573-583.https://doi.org/10.1518/001872001775870403.
31. Albantakis, L. and G. Deco, *The encoding of alternatives in multiple-choice decision making.* Proceedings of the National Academy of Sciences, 2009: p. pnas.0901621106.http://www.pnas.org/content/early/2009/06/03/0901621106.abstract.
32. Xing, J. *Information Complexity in Air Traffic Control Displays*. in *Human-Computer Interaction. HCI Applications and Services. HCI 2007*. 2007. Beijing, China, July 22-27, 2007.https://doi.org/10.1007/978-3-540-73111-5_89.
33. Xing, J., *Information Complexity in Air Traffic Control Displays*, 2007, DOT/FAA/AM-07/26, Civil Aerospace Medical Institute, Federal Aviation Administration, Department of Transportation, http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.94.8718&rep=rep1&type=pdf.
34. Jung, W., J. Park, Y. Kim, S.Y. Choi, and S. Kim, *HuREX – A framework of HRA data collection from simulators in nuclear power plants.* Reliability Engineering & System Safety, 2020. **194**: p. 106235.http://www.sciencedirect.com/science/article/pii/S0951832017309857.
35. Munger, S.J., R.W. Smith, and D. Payne, *An index of electronic equipment operability: Data Store*, 1962, (AIR-C43-I/62-RP[I]); DTIC No. AD607161, American Institute for Research, Pittsburgh, PA, https://apps.dtic.mil/dtic/tr/fulltext/u2/607165.pdf.
36. Nicholson, R.M., *Maximum Information-Handling Rates for Sequentially Presented Visual Stimuli.* Human Factors, 1962. **4**(6): p. 367-373.https://doi.org/10.1177/001872086200400605.
37. Department of Energy, *Savannah River Site Human Error Data Base Development for Nonreactor Nuclear Facilities (U)*, 1994, WSRC-TR-93-581, Westinghouse Savannah River Co.,
38. Jung, W. *HuREX–Human Reliability Data Extraction - A Framework for Simulator Data Collection and Analysis to Generate Human Error Probability*. in *HRA data Workshop*. 2018. Rockville, Maryland, U.S. Nuclear Regulatory Commission, March 15-16, 2018.
39. Just, M.A. and P.A. Carpenter, *A capacity theory of comprehension: Individual differences in working memory.* Psychological Review, 1992. **99**(1): p. 122-149.
40. Turner, M.L. and R.W. Engle, *Is working memory capacity task dependent?* Journal of Memory and Language, 1989. **28**(2): p. 127-154.http://www.sciencedirect.com/science/article/pii/0749596X89900405.

41. Cowan, N., *The Magical Number 4 in Short-Term Memory: A Reconsideration of Mental Storage Capacity.* Behavioral and Brain Sciences, 2000. **24**: p. 87-185.
42. Strom, D.J., *Health Impacts from Acute Radiation Exposure*, 2003, https://www.osti.gov/biblio/15020969.
43. Cucinotta, F.A., M. Alp, F.M. Sulzman, and M. Wang, *Space radiation risks to the central nervous system.* Life Sciences in Space Research, 2014. **2**: p. 54-69.http://www.sciencedirect.com/science/article/pii/S2214552414000339.
44. Strangman, G.E., W. Sipes, and G. Beven, *Human cognitive performance in spaceflight and analogue environments.* Aviation, space, and environmental medicine, 2014. **85**(10): p. 1033-1048.https://doi.org/10.3357/ASEM.3961.2014.
45. Barkaszi, I., *Effects of extreme environments on cognitive processes using the method of evenr-related potentials* in *Psychology*. 2017, University of Debreen.https://dea.lib.unideb.hu/dea/bitstream/handle/2437/242179/Disszertacio_titkositott.pdf?sequence=1&isAllowed=y.
46. Haskin, F.E., F.T. Harper, L.H.J. Goossens, and B.C.P. Kraan, *Probabilistic accident consequence uncertainty analysis -- Early health effects uncertainty assessment. Volume 2: Appendices*, 1997, NUREG/CR-6545, Vol. 2, https://www.osti.gov/biblio/291010; https://www.osti.gov/servlets/purl/291010.
47. Haskin, F.E., F.T. Harper, L.H.J. Goossens, B.C.P. Kraan, and J.B. Grupa, *Probabilistic accident consequence uncertainty analysis -- Early health effects uncertainty assessment. Volume 1: Main report*, 1997, NUREG/CR-6545, https://www.osti.gov/biblio/291010.
48. Beer, J.d. and R.A. Gallaway. *Effects of Continuous and Strobing Laser Glare on Performance in a Visually Simulated Flight Task*. 1999.
49. Schlittmeier, S., A. Feil, A. Liebl, and J. Ck, *The impact of road traffic noise on cognitive performance in attention-based tasks depends on noise level even within moderate-level ranges.* Noise & health, 2015. **17**: p. 148-157.
50. Szalma, J.L. and P.A. Hancock, *Noise effects on human performance: A meta-analytic synthesis.* Psychological Bulletin, 2011. **137**(4): p. 682-707.
51. Chase, B., W. Karwowski, M. Benedict, P. Queseda, B. Karwowski, and W. Benedict, *Effects of Thermal Stress on Dual Task Performance and Attention Allocation.* Journal of Human Performance in Extreme Environments, 2005. **8**.
52. Basra, G. and B. Kirwan, *Collection of offshore human error probability data.* Reliability Engineering & System Safety, 1998. **61**(1): p. 77-93.http://www.sciencedirect.com/science/article/pii/S0951832097000641.
53. Pilcher, J.J., E. Nadler, and C. Busch, *Effects of hot and cold temperature exposure on performance: a meta-analytic review.* Ergonomics, 2002. **45**(10): p. 682-698.https://doi.org/10.1080/00140130210158419.
54. Hancock, P., J. Ross, and J. Szalma, *A Meta-Analysis of Performance Response Under Thermal Stressors.* Human factors, 2007. **49**: p. 851-77.
55. Hancock, P.A. and E.K. Milner, *Mental and Psychomotor Task Performance in an Open Ocean Underwater Environment.* Research Quarterly for Exercise and Sport, 1982. **53**(3): p. 247-251.https://doi.org/10.1080/02701367.1982.10609349.
56. Conway, G.E., J. Szalma, and P. Hancock, *A quantitative meta-analytic examination of whole-body vibration effects on human performance.* Ergonomics, 2007. **50**: p. 228 - 245.https://www.semanticscholar.org/paper/A-quantitative-meta-analytic-examination-of-effects-Conway-Szalma/ce88a58a4c43b95c4fd3861dfba02418053e4650.
57. Pregnolato, M., A. Ford, S.M. Wilkinson, and R.J. Dawson, *The impact of flooding on road transport: A depth-disruption function.* Transportation Research Part D: Transport and Environment, 2017. **55**: p. 67-81.http://www.sciencedirect.com/science/article/pii/S1361920916308367.

58. Sherwood, N. and M.J. Griffin, *Effects of whole-body vibration on short-term memory.* Aviation, Space, and Environmental Medicine, 1990. **61**(12): p. 1092-1097.
59. Skraaning Jr., G. and G.A. Jamieson, *Twenty Years of HRP Research on Human-Automation Interaction: Insights on Automation Transparency and Levels of Automation DRAFT*, 2017, HPR-387, OECD Halden Reactor Project,
60. Trapsilawati, F., C. Wickens, C. Chen, and X. Qu. *Transparency and Conflict Resolution Automation Reliability in Air Trffic Control*. 2017.https://www.semanticscholar.org/paper/Transparency-and-Conflict-Resolution-Automation-in-Trapsilawati-Wickens/50d72ac80051975f573852b900f0a0264b4f899a.
61. Kwak, Y.-P., Y.-C. Choi, and J. Choi, *Analysis between Aircraft Cockpit Automation and Human Error Related Accident Cases.* International Journal of Control and Automation, 2018. **11**: p. 179-192.http://article.nadiapub.com/IJCA/vol11_no3/16.pdf.
62. Eitrheim, M.H.R., G.S. Jr., N. Lau, T. Karlsson, C. Nilwing, M. Hoffmann, and J.E. Farbrot, *STAFFING STRATEGIES IN HIGHLY AUTOMATED FUTURE PLANTS - RESULTS FROM THE 2009 HAMMLAB EXPERIMENT*, 2010, HWER-938, OECD Halden Reactor Project,
63. Jacob, V., M. Bhasi, and R. Gopikakumari, *Effect of work related variables on human errors in measurement.* Journal of Electrical and Electronics Engineering Research, Accepted for Publication, 2011.https://www.researchgate.net/profile/M_Bhasi/publication/261179540_Effect_of_work_related_variables_on_human_errors_in_measurement/links/570f726908aec95f0613baa7/Effect-of-work-related-variables-on-human-errors-in-measurement.pdf.
64. U.S. Nuclear Regulatory Commission. *10 CFR Part 55 - Operators' Licenses*. 2020; Available from: https://www.nrc.gov/reading-rm/doc-collections/cfr/part055/.
65. Fire Reporting System (www.firereporting.org). *PHYSIOLOGICAL EFFECTS OF FIREFIGHTER CREW SIZE*. 2020; Available from: http://www.firereporting.org/documents/resources/50.pdf.
66. Fire Reporting System (www.firereporting.org). *TIME TO TASK COMPLETION*. 2020; Available from: http://www.firereporting.org/documents/resources/19.pdf.
67. Fire Reporting System (www.firereporting.org). *Experiment Results - EMS Field Experiments*. 2020; Available from: http://www.firereporting.org/documents/resources/42.pdf.
68. Fire Reporting System (www.firereporting.org). *THE EFFECT OF FIREFIGHTER CREW SIZE ON TIME TO COMPLETE ESSENTIAL FIRE FIGHTING TASKS*. 2020; Available from: http://www.firereporting.org/documents/resources/9.pdf.
69. Cummings, M., C. Mastracchio, K.M. Thornburg, and A. Mkrtchyan, *Boredom and Distraction in Multiple Unmanned Vehicle Supervisory Control.* Interacting with Computers, 2014. **25**.https://www.semanticscholar.org/paper/Boredom-and-Distraction-in-Multiple-Unmanned-Cummings-Mastracchio/10f9d98f2103ef8e08a783c51d5daa338406bb1a.
70. Hobbs, A., *An Overview of Human Factors in Aviation Maintenance*, 2008, AR-2008-055, Australian Transport Safety Bureau (ATSB), https://www.atsb.gov.au/media/27818/hf_ar-2008-055.pdf.
71. Kim, A.R., J. Park, J.T. Kim, J. Kim, and P.H. Seong, *Study on the identification of main drivers affecting the performance of human operators during low power and shutdown operation.* Annals of Nuclear Energy, 2016. **92**: p. 447-455.http://www.sciencedirect.com/science/article/pii/S0306454916300810.
72. Converse, S.A., *Operating Procedures: Do they Reduce Operator Errors?* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 1994. **38**: p. 205 - 209.https://journals.sagepub.com/doi/pdf/10.1177/154193129403800404.

73. Goodstein, L.P., *Training diagnostic skills for nuclear power plants*, 1986, Risø-M; No. 2587, Technical University of Denmark, https://orbit.dtu.dk/ws/files/104234448/ris_m_2587.pdf.
74. Ha, J.S. and P.H. Seong, *Experimental investigation between attentional-resource effectiveness and perception and diagnosis in nuclear power plants.* Nuclear Engineering and Design, 2014. **278**: p. 758-772.http://www.sciencedirect.com/science/article/pii/S0029549314004956.
75. De Dreu, C.K.W. and L.R. Weingart, *Task versus relationship conflict, team performance, and team member satisfaction: A meta-analysis.* Journal of Applied Psychology, 2003. **88**(4): p. 741-749.http://web.mit.edu/curhan/www/docs/Articles/15341_Readings/Negotiation_and_Conflict_Management/De_Dreu_Weingart_Task-conflict_Meta-analysis.pdf.
76. International Atomic Energy Agency, *Conduct of Operations at Nuclear Power Plants*, 2008, Safety Guide No. NS-G-2.14, IAEA, http://www-pub.iaea.org/MTCD/publications/PDF/Pub1339_web.pdf.
77. Skraaning, G., *A reanalysis of the Work Practice Experiments in HAMMLAB (2009-2014)*, 2017, HWR-1194, OECD Halden Reactor Project,
78. Nicholas, C.A. and A.L. Cohen, *The effect of interruption on the decision-making process.* Judgment and Decision Making, 2016. **11**(6): p. 611-626.http://journal.sjdm.org/16/16822/jdm16822.pdf.
79. Fysh, M., *Time Pressure and Human-Computer Interaction in Face Matching*, in *Psychology*. 2017, University of Kent.http://kar.kent.ac.uk/65773/.
80. Leon, M. and W. Revelle, *Effects of Anxiety on Analogical Reasoning. A Test of Three Theoretical Models.* Journal of personality and social psychology, 1985. **49**: p. 1302-15.https://www.researchgate.net/profile/William_Revelle/publication/19082836_Effects_of_Anxiety_on_Analogical_Reasoning_A_Test_of_Three_Theoretical_Models/links/09e41506f7d87d8b59000000/Effects-of-Anxiety-on-Analogical-Reasoning-A-Test-of-Three-Theoretical-Models.pdf.
81. Szalma, J., P. Hancock, and S. Quinn, *A Meta-Analysis of the Effect of Time Pressure on Human Performance*. Vol. 52. 2008.
82. Office of Oversight, *Independent Oversight Special Study of Hoisting and Rigging Incidents Within the Department of Energy*, 1996, Office of Environment, Safety and Health, U.S. Department of Energy, http://158.132.155.107/posh97/private/research/references-crane-safety/9610hoisting.pdf.
83. Brewer, J.D., P.J. Amico, S.E. Cooper, and S.M.L. Hendrickson, *Preliminary, Qualitative Human Reliability Analysis for Spent Fuel Handling*, 2012, NUREG/CR-7017; SAND2010-8464P, U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1105/ML110590883.pdf.
84. Working Group on Operating Experience (WGOE), *Heavy Load Accidents in Nuclear Installations*, 2017, NEA/CNRA/R(2017)4, NUCLEAR ENERGY AGENCY, COMMITTEE ON NUCLEAR REGULATORY ACTIVITIES, http://www.oecd.org/officialdocuments/publicdisplaydocumentpdf/?cote=NEA/CNRA/R(2017)4&docLanguage=En.
85. Prasad, R., G. Coles, A. Dalton, N. Kohn, K. Branch, A. Bittner, and R. Taylor, *Effects of Environmental Conditions on Manual Actions for Flood Protection and Mitigation: Identifying Conditions, Characterizing Actions and Assessing Impacts*, 2020, NUREG/CR-7256, Volume 1, U.S. Nuclear Regulatory Commission, https://www.nrc.gov/reading-rm/doc-collections/nuregs/contract/cr7256/v1/.
86. Park, J. and W. Jung, *OPERA—a human performance database under simulated emergencies of nuclear power plants.* Reliability Engineering & System Safety, 2007.

92(4): p. 503-519.https://www.sciencedirect.com/science/article/pii/S0951832006000329.

87. Kubicek, J. and J. Holy. *Simulator Data Collection in Czech Republic*. in *HRA Data Workshop*. 2018. Rockville, Maryland, U.S. Nuclear Regulatory Commission, March 15-16, 2018.
88. Knecht, W.R., *Predicting Accident Rates From General Aviation Pilot Total Flight Hours*, 2015, DOT/FAA/AM-15/3, Civil Aerospace Medical Institute, Federal Aviation Administration, https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2010s/media/201503.pdf.
89. Van Iddekinge, C.H., H. Aguinis, J.D. Mackey, and P.S. DeOrtentiis, *A Meta-Analysis of the Interactive, Additive, and Relative Effects of Cognitive Ability and Motivation on Performance.* Journal of Management, 2018. **44**(1): p. 249-279.http://journals.sagepub.com/doi/10.1177/0149206317702220.
90. Liu, P. and J. Liu, *Combined Effect of Multiple Performance Shaping Factors on Human Reliability: Multiplicative or Additive?* International Journal of Human–Computer Interaction, 2020. **36**(9): p. 828-838.https://www.tandfonline.com/doi/full/10.1080/10447318.2019.1695461.
91. Mount, M.K., M.R. Barrick, and J.P. Strauss, *The Joint Relationship of Conscientiousness and Ability with Performance: Test of the Interaction Hypothesis.* Journal of Management, 1999. **25**(5): p. 707-721.http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.455.967&rep=rep1&type=pdf#:~:text=Overall%2C%20the%20preponderance%20of%20evidence,than%20those%20who%20are%20not.
92. Hancock, P.A. and J.O. Pierce, *Combined Effects of Heat and Noise on Human Performance: A Review.* American Industrial Hygiene Association Journal, 1985. **46**(10): p. 555-566.https://www.researchgate.net/profile/Peter_Hancock2/publication/19101231_Combined_Effects_of_Heat_and_Noise_on_Human_Performance_A_Review/links/556db0e408aefcb861d9ad94/Combined-Effects-of-Heat-and-Noise-on-Human-Performance-A-Review.pdf.
93. Murray, R. and M. McCally, *Combined environmental stresses.* 1973.https://www.researchgate.net/publication/23622718_Combined_environmental_stresses.
94. Grether, W.F., *Effects on Human Performance of Combined Environmental Stress*, 1970, AMRL-TR-70-68, Aerospace Medical Research Laboratory, https://apps.dtic.mil/dtic/tr/fulltext/u2/727041.pdf.
95. Broadbent, D.E., *Differences and interactions between stresses.* Quarterly Journal of Experimental Psychology, 1963. **15**(3): p. 205-211.http://www.tandfonline.com/doi/abs/10.1080/17470216308416325.
96. Park, J., W. Jung, J. Kim, and J. Ha, *Analysis of human performance observed under simulated emergencies of nuclear power plants*, 2005, KAERI/TR-2895/2005, Korea Atomic Energy Research Institute,
97. Park, J., Y. Kim, J.H. Kim, W. Jung, and S.C. Jang, *Estimating the response times of human operators working in the main control room of nuclear power plants based on the context of a seismic event – A case study.* Annals of Nuclear Energy, 2015. **85**: p. 36-46.http://www.sciencedirect.com/science/article/pii/S0306454915002200.
98. Park, J. and W. Jung, *A study on the validity of a task complexity measure for emergency operating procedures of nuclear power plants - comparing task complexity scores with two sets of operator response time data obtained under a simulated SGTR.* Reliability Engineering and System Safety, 2008. **93**(4): p. 557-566.

99. Park, J., D. Lee, W. Jung, and J. Kim, *An experimental investigation on relationship between PSFs and operator performances in the digital main control room.* Annals of Nuclear Energy, 2017. **101**: p. 58-68.https://linkinghub.elsevier.com/retrieve/pii/S0306454916307381.
100. Kelly, T.L., C.E. Englund, D.H. Rayman, J.E. Yeager, and A.A. Sucec, *The Effects of 12 Hours of MOPP (Mission Oriented Protective Posture) IV Gear on Cognitive Performance under Non-Exercise Conditions*, 1987, 88-6, Naval Health Research Center, San Diego, CA, https://apps.dtic.mil/dtic/tr/fulltext/u2/a192527.pdf.
101. Taylor, H.L. and J. Orlansky, *The effects of wearing protective chemical warfare combat clothing on human performance.* Aviation, space, and environmental medicine, 1993. **64**(3 Pt 2): p. A1-41.http://europepmc.org/abstract/MED/8447813

https://pdfs.semanticscholar.org/c484/04db98b27ef5a32c3674b4baced9b00cb23f.pdf.
102. Speier, C., J.S. Valacich, and I. Vessey, *The Influence of Task Interruption on Individual Decision Making: An Information Overload Perspective.* Decision Sciences, 1999. **30**(2): p. 337-360.https://www.interruptions.net/literature/Speier-DS99.pdf.
103. Commission, U.S.N.R., *Millstone Power Station Unit 3 - NRC Special Inspection Report 05000423/2005012*, 2005, IR 05000423/2005012 (ADAMS Accession No. ML051860338), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0518/ML051860338.pdf.
104. U.S. Nuclear Regulatory Commission, *Onsite Analysis of the Human Factors of an Event at Braidwood Unit 1 on October 4, 1990 (Reactor Coolant System Loss)*, 1991, Idaho Engineering National Laboratory,
105. U.S. Nuclear Regulatory Commission, *Onsite Analysis of the Human Factors of an Event at Catawba Unit 1 on March 20, 1990 (Overpressurization of RHR System)*, 1990, Idaho Engineering National Laboratory,
106. Massaiu, S. and L. Holmgren, *A HAMMLAB HRA Data Collection with U.S. Operators*, 2016, HWR-1123, OECD Halden Reactor Project,
107. Jang, I., A.R. Kim, W. Jung, and P.H. Seong, *An empirical study on the human error recovery failure probability when using soft controls in NPP advanced MCRs.* Annals of Nuclear Energy, 2014. **73**: p. 373-381.https://linkinghub.elsevier.com/retrieve/pii/S0306454914003351.
108. Gertman, D.I., B.P. Halbert, M.W. Parrish, M.B. Sattison, D. Brownson, and J.P. Tortorelli, *Review of Findings for Human Performance Contribution to Risk in Operating Events*, 2002, NUREG/CR-6753 (ADAMS Accession No. ML020930077), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0209/ML020930077.pdf.
109. Swain, A.D. and H.E. Guttmann, *Handbook of Human Reliability Analysis with Emphasis on Nuclear Power Plant Applications*, 1983, NUREG/CR-1278 (ADAMS Accession No. ML071210299), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML0712/ML071210299.pdf.
110. Kirwan, B. *Human Reliability Assessment Practitioner Guidance on the Treatment of Human Error Limits*. in *UK Nuclear Industry workshop on Human Performance Limiting Values (HPLVs)*. 2007.
111. Smith, D.J., *Appendix 6 - Human Error Probabilities*, in *Reliability, Maintainability and Risk (Eighth Edition)*, D.J. Smith, Editor. 2011, Butterworth-Heinemann: Oxford. p. 395-397.http://www.sciencedirect.com/science/article/pii/B9780080969022000313.
112. Schoonard, J.W. and S.J. Boies, *Short-Type: A Behavioral Analysis of Typing and Text Entry.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 1975. **17**(2): p. 203-214.http://journals.sagepub.com/doi/10.1177/001872087501700211.

113. Dhillon, B.S., *Human reliability with human factors*. Reprinted ed. Pergamon international library of science, technology, engineering and social studies. 1987, New York u.a: Pergamon Press. 239.
114. Jung, W., J. Park, S. Kim, S.Y. Choi, Y. Kim, and J. Yang, *Handbook for Analysis of Human Error Probability based on Simulator Data of Nuclear Power Plants*, 2016, KAERI/TR-6649, Korea Atomic Research Institute,
115. Kim, Y., J. Park, S. Choi, S. Kim, and W. Jung, *Statistical Analysis of Simulation Data for Generating Quantitative Bases of Human Reliability Analyses*, 2016, KAERI/TR-6652, Korea Atomic Research Institute,
116. Park, J., Y. Kim, and W. Jung, *Calculating nominal human error probabilities from the operation experience of domestic nuclear power plants.* Reliability Engineering & System Safety, 2018. **170**: p. 215-225.http://www.sciencedirect.com/science/article/pii/S0951832016306500.
117. Nazeri, Z., *Data Mining of Air Traffic Control Operational Errors*. 2006. 321-324.
118. Bailey, L., *Analysis of En Route Operational Errors: Probability of Resolution and Time-On-Position*, 2012, DOT/FAA/AM-12/2, Federal Aviation Administration, https://www.faa.gov/data_research/research/med_humanfacs/oamtechreports/2010s/media/201202.pdf.
119. Chang, Y.J. and A. Cheng. *A Case Study of Human Error Probability in Commercial Aviation*. in *International Topical meeting on Probabilistic SAfety Assessment and Analysis (PSA 2015)*. 2015. Sun Valley, Idaho April 26-30, 2015.
120. Jung, W., J. Park, Y. Kim, S.Y. Choi, and S. Kim, *HuREX – A Framework of HRA Data Collection from Simulators in Nuclear Power Plants.* Reliability Engineering & System Safety, 2018.https://doi.org/10.1016/j.ress.2018.07.036.
121. Aalipour, M., Y.Z. Ayele, and A. Barabadi, *Human reliability assessment (HRA) in maintenance of production process: a case study.* International Journal of System Assurance Engineering and Management, 2016. **7**(2): p. 229-238.http://link.springer.com/10.1007/s13198-016-0453-z.
122. Institute for Safe Medication Practices (ISMP). *High-Alert Medication Modeling and Error-Reduction Scorecards (HAMMERS) for Community Pharmacies*. 2012; Available from: https://www.ismp.org/resources.
123. McGuirl, J.M. and N.B. Sarter, *Supporting Trust Calibration and the Effective Use of Decision Aids by Presenting Dynamic System Confidence Information.* Human Factors, 2006. **48**(4): p. 656-665.https://doi.org/10.1518/001872006779166334.
124. Cummings, M.L., C.G. Tsonis, and A.A. Rader, *The Impact of Multi-Layered Data-Blocks on Controller Performance.* Air Traffic Control Quarterly, 2008. **16**(2): p. 147-169.https://doi.org/10.2514/atcq.16.2.147.
125. Lee, J.D., D.V. McGehee, T.L. Brown, and M.L. Reyes, *Collision Warning Timing, Driver Distraction, and Driver Response to Imminent Rear-End Collisions in a High-Fidelity Driving Simulator.* Human Factors, 2002. **44**(2): p. 314-334.https://doi.org/10.1518/0018720024497844.
126. McRobert, A.P., J. Causer, J. Vassiliadis, L. Watterson, J. Kwan, and M.A. Williams, *Contextual information influences diagnosis accuracy and decision making in simulated emergency medicine emergencies.* BMJ Quality & Safety, 2013. **22**(6): p. 478.http://qualitysafety.bmj.com/content/22/6/478.abstract.
127. Mackworth, N., *Visual noise causes tunnel vision.* Psychon Sci., 1965. **3**.http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.963.1652&rep=rep1&type=pdf.
128. Vlassova, A., C. Donkin, and J. Pearson, *Unconscious information changes decision accuracy but not confidence.* PNAS Proceedings of the National Academy of Sciences of the United States of America, 2014. **111**(45): p. 16214-16218.

129. Ziegler, M.G., P. Lew, and B.C. Singer, *The Accuracy of Drug Information From Pharmaceutical Sales Representatives.* JAMA, 1995. **273**(16): p. 1296-1298.https://doi.org/10.1001/jama.1995.03520400066047.
130. Dashevsky, S., *Combining check-reading accuracy and quantitative information in a space-saving display.* Journal of Applied Psychology, 1964. **48**(6): p. 348-350.
131. Blanchard, R.E., M.B. Mitchell, and R.B. Smith. *Likelihood-of-accomplishment scale for a sample of man-machine activities*. 1966.
132. Cummings, M.L. and C. Tsonis, *Deconstructing Complexity in Air Traffic Control.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2005. **49**(1): p. 25-29.https://doi.org/10.1177/154193120504900107.
133. Harms, L., *Experimental studies of task demands and processing performance in driving and driving simulation*, 1990, 1102626X (ISSN), Statens Väg- och Trafikinstitut., VTI särtryck 155, http://vti.diva-portal.org/smash/get/diva2:671757/FULLTEXT01.pdf.
134. Howell, W.C. and J.D. Tate, *Research on display variables*, 1964, RADC-TR-64-266, Griffiss AFB, NY: Rome Air Development Center,
135. Dyer, W.R. and R.J. Christman, *Relative influence of time, complexity, and density on utilization of coded large-scale displays*, 1965, RADC-TR-65-325; DTIC No. AD 622786, Griffiss AFB, NY: Rome Air Development Center,
136. Hammer, C.H. and S. Ringel, *Information assimilation from updated alphanumeric displays.* Journal of Applied Psychology, 1966. **50**(5): p. 383-387.
137. Kosmider, G., M. Young, and G. Kinney, *Studies in display symbol legibility. 8. Legibility of common five-letter words*, 1966, ESD-TR-65-385; TM-4239, Bed-ford, MA: Mitre Corporation,
138. Ryum, K.A., J. Park, Y. Kim, J. Kim, and P.H. Seong, *Quantification of performance shaping factors (PSFs)' weightings for human reliability analysis (HRA) of low power and shutdown (LPSD) operations.* Annals of Nuclear Energy, 2017. **101**: p. 375-382.www.elsevier.com/locate/anucene.
139. Matthews, G., *Signal probability effects on high-workload vigilance tasks.* Psychonomic Bulletin & Review, 1996. **3**(3): p. 339-343.https://doi.org/10.3758/BF03210757.
140. Gorgoraptis, N., R.F.G. Catalao, P.M. Bays, and M. Husain, *Dynamic Updating of Working Memory Resources for Visual Objects.* The Journal of Neuroscience, 2011. **31**(23): p. 8502.http://www.jneurosci.org/content/31/23/8502.abstract.
141. Joos, D.W., Z.A. Sabri, and A.A. Husseiny, *Analysis of gross error rates in operation of commercial nuclear power stations.* Nuclear Engineering and Design, 1979. **52**(2): p. 265-300.http://www.sciencedirect.com/science/article/pii/0029549379900566.
142. Hameed, S., T. Ferris, S. Jayaraman, and N. Sarter, *Using Informative Peripheral Visual and Tactile Cues to Support Task and Interruption Management.* Human Factors, 2009. **51**(2): p. 126-135.https://doi.org/10.1177/0018720809336434.
143. Podofillini, L., J. Park, and V.N. Dang, *Measuring the influence of task complexity on human error probability: an empirical evaluation.* Nuclear Engineering and Technology, 2013. **45**(2): p. 151-164.http://www.sciencedirect.com/science/article/pii/S1738573315300589.
144. Sauer, J., D.G. Wastell, G. Robert, J. Hockey, and F. Earle, *Performance in a Complex Multiple-Task Environment during a Laboratory-Based Simulation of Occasional Night Work.* Human Factors, 2003. **45**(4): p. 657-670.https://doi.org/10.1518/hfes.45.4.657.27090.
145. Speier, C., I. Vessey, and J.S. Valacich, *The Effects of Interruptions, Task Complexity, and Information Presentation on Computer-Supported Decision-Making Performance.* Decision Sciences, 2003. **34**(4): p. 771-797.https://doi.org/10.1111/j.1540-5414.2003.02292.x.

146. Elias, M.F., A.M. Snadowsky, and E.F. Rizy, *The relation of number of scan lines per symbol height to recognition of televised alphanumerics*, 1964, RADC-TDR64-433; DTIC No. AD608789, Griffiss AFB, NY: Rome Air Development Center, AD0608789

147. Meister, D., *Tables for predicting the opera- tional perfonnance of personnel*, 1982, NUREG/CR-2744/1 of 2, SAND82-7057/I of 2, AN, RX, Sandia National Laboratories: N.R. Commission

148. Kramer, M., K. Terheiden, and S. Wieprecht, *Safety criteria for the trafficability of inundated roads in urban floodings.* International Journal of Disaster Risk Reduction, 2016. **17**: p. 77-84.http://www.sciencedirect.com/science/article/pii/S2212420915301783.

149. Tsapakis, I., T. Cheng, and A. Bolbol, *Impact of weather conditions on macroscopic urban travel times.* Journal of Transport Geography, 2013. **28**: p. 204-211.http://www.sciencedirect.com/science/article/pii/S0966692312002694.

150. Andrey, J., B. Mills, M. Leahy, and J. Suggett, *Weather as a Chronic Hazard for Road Transportation in Canadian Cities.* Natural Hazards, 2003. **28**(2): p. 319-343.https://doi.org/10.1023/A:1022934225431.

151. Carnell, L., S. Blattnig, S. Hu, J. Huff, M.-H. Kim, R. Norman, Z. Patel, L. Simonsen, and H. Wu, *Evidence Report: Risk of Acute Radiation Syndromes Due to Solar Particle Events*, 2016, JSC-CN-35747, NASA, https://ntrs.nasa.gov/citations/20160003870.

152. Terelak, J.F., *The emergence and development of space psychology in Poland. The significance of historical space flight of Mirosław Hermaszewski.* Pol J Aviat Med Psychol, 2013. **19**(3): p. 23-26.https://pdfs.semanticscholar.org/efba/c7cb1d480e261a247502b71ec37a5941bba3.pdf.

153. Becker, J., H.L. Taylor, B.J. Doody, K. Wright, E. Gruntfest, and D. Webber, *A Review of Peopl… s Behavior in and around Floodwater.* Weather, Climate, and Society, 2015. **7**: p. 321-332.

154. White, W.J. and M.B. Riley, *The Effects of Positive Acceleration on the Relation Between Illumination and Instrument Reading*, 1958, Air Force Aerospace Medical Research Lab Wright-Patterson AFB OH, AD0206663

155. Rausch, M. and M. Zehetleitner, *Visibility Is Not Equivalent to Confidence in a Low Contrast Orientation Discrimination Task.* Frontiers in Psychology, 2016. **7**: p. 591.https://www.frontiersin.org/article/10.3389/fpsyg.2016.00591.

156. Smith, S.L., *Display color coding for visual separability*, 1963, Bedford, MA: MITRE Corp.,

157. Ko, P., A. Mohapatra, I. Bailey, J. Sheedy, and D. Rempel, *Effects of Font Size and Reflective Glare on Text-Based Task Performance and Postural Change Behavior of Presbyopic and Nonpresbyopic Computer Users.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2012. **56**: p. 2378-2382.

158. Gandara, H., *Effect of glare over the performance in an inspection task*, in *Industrial Engineering*. 1971, Kansas State University.http://krex.k-state.edu/dspace/bitstream/handle/2097/26801/LD2668T41971G35.pdf?sequence=1&isAllowed=y.

159. Green, P., S. Goldstein, K. Zeltner, and S. Adams, *Legibility of text on instrument panels: a literature review. Final report*, 1988, UMTRI-88-34: T.R.I. The University of Machiganhttps://deepblue.lib.umich.edu/handle/2027.42/790.

160. Kang, J.J., R. Ni, and G.J. Andersen, *Effects of Reduced Visibility from Fog on Car-Following Performance.* Transportation Research Record, 2008. **2069**(1): p. 9-15.https://doi.org/10.3141/2069-02.

161. Horrey, W., A. Alexander, and C. Wickens, *Does workload modulate the effects of in-vehicle display location on concurrent driving and side task performance.* DSC North America Proceedings, 2003: p. 1546-5071.
162. Pepper, R.L. and R.E. Cole, *Display System Variables Affecting Operator Performance in Undersea Vehicles and Work Systems*, 1978, Naval Ocean Systems Center San Diego CA, https://apps.dtic.mil/dtic/tr/fulltext/u2/a056791.pdf.
163. Bell, P.A., S. Hess, E. Hill, S.L. Kukas, R.W. Richards, and D. Sargent, *Noise and context-dependent memory.* Bulletin of the Psychonomic Society, 1984. **22**(2): p. 99-100.https://doi.org/10.3758/BF03333774.
164. Poulton, E.C. and R.S. Edwards, *Interactions and range effects in experiments on pairs of stresses: mild heat and low-frequency noise.* Journal of Experimental Psychology, 1974. **102**(4): p. 621-628.
165. Echeverria, D., V. Barnes, A. Bittner, N. Durbin, J. Fawcett-Long, C. Moore, A. Slavich, B. Terrill, C. Westra, D. Wieringa, R. Wilson, D. Draper, D. Morisseau, and J. Persensky, *The impact of environmental conditions on human performance Vol. 1 - A handbook of environmental exposures*, 1994, NUREG/CR-5680; BHARC-700/91/005; PNL-7675, Vol. 1, U.S. Nuclear Regulatory Commission,
166. Echeverria, D., V. Barnes, A. Bittner, N. Durbin, J. Fawcett-Long, C. Moore, A. Slavich, B. Terrill, C. Westra, D. Wieringa, R. Wilson, D. Draper, D. Morisseau, and J. Persensky, *The impact of environmental conditions on human performance Vol. 2 - A critical review of teh literature*, 1994, NUREG/CR-5680; BHARC-700/91/005; PNL-7675 Vol. 2, U.S. Nuclear Regulatory Commission,
167. Watkins, W.H., *Effect of certain noises upon detection of visual signals.* Journal of Experimental Psychology, 1964. **67**(1): p. 72-75.
168. Woodhead, M.M., *The Effect of Bursts of Noise on an Arithmetic Task.* The American Journal of Psychology, 1964. **77**(4): p. 627-633.http://www.jstor.org/stable/1420774.
169. May, D.N. and C.G. Rice, *Effects of startle due to pistol shots on control precision performance.* Journal of Sound and Vibration, 1971. **15**(2): p. 197-202.http://www.sciencedirect.com/science/article/pii/0022460X71905347.
170. Parasuraman, R., *Vigilance, monitoring, and search*, in *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance.* 1986, John Wiley & Sons: Oxford, England. p. 1-39.
171. Parasuraman, R., *Vigilance, monitoring, and search*, in *Handbook of perception and human performance, Vol. 2: Cognitive processes and performance.*, K.R. Boff, L. Kaufman, and J.P. Thomas, Editors. 1986, John Wiley & Sons: Oxford, England. p. xxi, 1143-xxi, 1143.
172. Ross, J.M., J.L. Szalma, and P.A. Hancock, *A Meta-Analysis of Performance under Thermal Stress.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2006. **50**(17): p. 1736-1740.https://doi.org/10.1177/154193120605001703.
173. Hancock, P.A., J.M. Ross, and J.L. Szalma, *A Meta-Analysis of Performance Response Under Thermal Stressors.* Human Factors, 2007. **49**(5): p. 851-877.https://doi.org/10.1518/001872007X230226.
174. Morris, D., *The cold driver: Driving performance under thermal stress*, in *Psychology*. 2015, Clemson University.https://tigerprints.clemson.edu/cgi/viewcontent.cgi?article=3241&context=all_theses.
175. Noroozi, A., R. Abbassi, S. MacKinnon, F. Khan, and N. Khakzad, *Effects of Cold Environments on Human Reliability Assessment in Offshore Oil and Gas Facilities.* Human Factors, 2013. **56**(5): p. 825-839.https://doi.org/10.1177/0018720813512328.

176. Gaoua, N., S. Racinais, J. Grantham, and F. El Massioui, *Alterations in cognitive performance during passive hyperthermia are task dependent.* International Journal of Hyperthermia, 2011. **27**(1): p. 1-9.
177. Hancock, P.A., *Sustained attention under thermal stress.* Psychological bulletin, 1986. **99**(2): p. 263.
178. Syndicus, M., B. Wiese, and C. Treeck, *In the Heat and Noise of the Moment: Effects on Risky Decision Making.* Environment and Behavior, 2018. **50**: p. 3-27.
179. Mortan, T., F.P. Rudek, A. Miller, and D.G. Norman, *Handbook of human engineering design data for reduced gravity conditions*, 1971, NASA-CR-1726, https://www.lpi.usra.edu/lunar/strategies/NASA%20CR-1726.pdf.
180. Havener, P.W. and L.B. Johnson, *Experimental research studies on tools for extravehicular maintenance in space, phase 2 Final report*, 1967, NASA-CR-91052, NASA, https://ntrs.nasa.gov/citations/19680002480.
181. Armstrong, N.C.D., A. Ward, M. Lomax, M.J. Tipton, and J.R. House, *Wearing body armour and backpack loads increase the likelihood of expiratory flow limitation and respiratory muscle fatigue during marching.* Ergonomics, 2019. **62**(9): p. 1181-1192.https://doi.org/10.1080/00140139.2019.1629638.
182. McLeod, R.W. and M.J. Griffin, *Effects of whole-body vibration waveform and display collimation on the performance of a complex manual control task.* Aviation, Space, and Environmental Medicine, 1990. **61**(3): p. 211-219.
183. Griffin, M.J. and R.A. Hayward, *Effects of horizontal whole-body vibration on reading.* Applied Ergonomics, 1994. **25**(3): p. 165-169.http://www.sciencedirect.com/science/article/pii/0003687094900140.
184. Fine, B.J. and J.L. Kobrick, *Effect of heat and chemical protective clothing on cognitive performance.* Aviation, Space, and Environmental Medicine, 1987. **58**(2): p. 149-154.https://apps.dtic.mil/dtic/tr/fulltext/u2/a162001.pdf.
185. Simon, Y.L., *The effects of personal protective equipment Level A suits on human task performance*, in *Engineering Management*. 2010, Missouri University of Science and Technology.https://scholarsmine.mst.edu/masters_theses/4994.
186. SKITKA, L.J., K.L. MOSIER, and M. BURDICK, *Does automation bias decision-making?* Int. J. Hum.-Comput. Stud., 1999. **51**(5): p. 991–1006.https://doi.org/10.1006/ijhc.1999.0252.
187. Sarter, N.B. and D.D. Woods, *Pilot interaction with cockpit automation II: An experimental study of pilots' model and awareness of the flight management system.* The International Journal of Aviation Psychology, 1994. **4**(1): p. 1-28.
188. Parasuraman, R., R. Molloy, and I.L. Singh, *Performance Consequences of Automation-Induced 'Complacency'.* The International Journal of Aviation Psychology, 1993. **3**(1): p. 1-23.https://doi.org/10.1207/s15327108ijap0301_1

https://www.researchgate.net/publication/23595844_Performance_Consequences_of_Automation_Induced_Complacency.
189. Mosier, K.L., L.J. Skitka, S. Heers, and M. Burdick, *Automation Bias: Decision Making and Performance in High-Tech Cockpits.* The International Journal of Aviation Psychology, 1998. **8**(1): p. 47-63.https://doi.org/10.1207/s15327108ijap0801_3.
190. Gawron, V., *Nothing Can Go Wrong - A Review ofAutomation Induced Complacency Rese arch*, 2019, The MITRE Corporation, McLean, VA, https://www.mitre.org/sites/default/files/pdf/pr-3426-lessons-lost-nothing-can-go-wrong-automation-induced-complacency.pdf.
191. Scallen, S.F., P.A. Hancock, and J.A. Duley, *Pilot performance and preference for short cycles of automation in adaptive function allocation.* Applied Ergonomics, 1995. **26**(6): p. 397-403.http://www.sciencedirect.com/science/article/pii/0003687095000542.

192. Helldin, T., *Transparency for Future Semi-Automated Systems : Effects of transparency on operator performance, workload and trust*, in *Örebro Studies in Technology*. 2014, Örebro Universitet: Örebro. p. 266.http://urn.kb.se/resolve?urn=urn:nbn:se:his:diva-10164.
193. Brill, R.W. *Instrumentation and control system failures in nuclear power plants*. 2000.
194. Laakso, K., P. Pyy, and L. Reiman, *Human errors related to maintenance and modifications*, 1998, STUK-YTO-TR 139, Radiation and Nuclear Safety Authority, https://pdfs.semanticscholar.org/fdf1/2d83deca7ed0a6b247a223f82272e417eba3.pdf.
195. Laakso, K., *Systematic analysis of dependent human errors from the maintenance history at Finnish NPPs - A status report*, 2002, NKS-81, NKS,
196. Xing, J. and D.J. Heeger, *Center-surround interactions in foveal and peripheral vision.* Vision Research, 2000. **40**(22): p. 3065-3072.http://www.sciencedirect.com/science/article/pii/S0042698900001528.
197. Joe, J.C., C.R. Kovesdi, A. Mack, and T. Miyake, *The Effect of Information Organization on Cognitive Workload and Visual Search Performance.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2019. **63**(1): p. 302-306.https://doi.org/10.1177/1071181319631085.
198. Wu, X., M. She, Z. Li, F. Song, and W. Sang, *Effects of integrated designs of alarm and process information on diagnosis performance in digital nuclear power plants.* Ergonomics, 2017. **60**(12): p. 1653-1666.https://doi.org/10.1080/00140139.2017.1335884.
199. Hooey, B., C. Wickens, E. Salud, A. Sebok, S. Hutchins, and B. Gore, *PREDICTING THE UNPREDICTABLE: ESTIMATING HUMAN PERFORMANCE PARAMETERS FOR OFF-NOMINAL EVENTS*. 2009.
200. Rø Eitrheim, M.H., A. Fernandes, and H. Svengren, *Evaluation of design features in the HAMBO operator displays*, 2017, HWR-1212, OECD Halden Reactor Project,
201. Culley, K.E., *Effects of Signal Ambiguity and Signal Location on Target Detection Under Varying Degrees of Time Constraint*, in *Psychology*. 2013, Old Dominion University.https://digitalcommons.odu.edu/cgi/viewcontent.cgi?article=1143&context=psychology_etds.
202. Krebs, W.K., J. Xing, and A.J. Ahumada, *A Simple Tool for Predicting the Readability of a Monitor.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2002. **46**(17): p. 1659-1663.https://doi.org/10.1177/154193120204601728.
203. Human Factors Research and Engineering Group, *Standard Practice - Baseline Requirements for Color Use in Air Traffic Control Displays*, 2007, HF-STD-002, Federal Aviation Administration, U.S. Department of Transportation, http://www.tc.faa.gov/its/worldpac/standards/faa-hf-std-002.pdf.
204. Xing, J., *Developing the Federal Aviation Administration's Requirements for Color Use in Air Traffic Control Displays.* 2007: p. 19.
205. Byrne, M.D. and E.M. Davis, *Task Structure and Postcompletion Error in the Execution of a Routine Procedure.* Human Factors, 2006. **48**(4): p. 627-638.https://doi.org/10.1518/001872006779166398.
206. Rhee, C., J. Moon, and Y. Choe, *Web interface consistency in e‐learning.* Online Information Review, 2006. **30**(1): p. 53-69.https://doi.org/10.1108/14684520610650309.
207. Reyes, R.M., J. de la Riva, A. Maldonado, A. Woocay, and R. de la O, *Association between Human Error and Occupational Accidents' Contributing Factors for Hand Injuries in the Automotive Manufacturing Industry.* Procedia Manufacturing, 2015. **3**: p. 6498-6504.http://www.sciencedirect.com/science/article/pii/S2351978915009373.
208. Ngadiman, Y., B. Hussin, N.A.A. Hamid, R. Rohaizan, and L.K. Boon. *Relationship between Human Errors in Maintenance and Overall Equipment Effectiveness in Food Industries*. in *2016 International Conference on Industrial Engineering and Operations*

Management. 2016. Kuala Lumpur, Malaysia, March 8-10, 2016.http://ieomsociety.org/ieom_2016/pdfs/668.pdf.

209. Gibb, A., S. Hide, R. Haslam, D. Gyi, T. Pavitt, S. Atkinson, and R. Duff, *Construction tools and equipment – their influence on accident causality.* Journal of Engineering, Design and Technology, 2005. **3**(1): p. 12-23.https://doi.org/10.1108/17260530510815303.

210. Stewart, M.G., *Human error in structural reliability IV : efficiency in design checking / M.G. Stewart and R.E. Melchers*. Civil engineering research reports ; 1985/no. 3., ed. R.E. Melchers and E. Monash University. Department of Civil. 1985, [Clayton, Vic.]: Monash University.

211. Stewart, M.G. *Risk Management and the Human Dimension in Structural Design*. in *First Int. Symposium on Uncertainty Modeling and Analysis*. 1990. University of Maryland, College Park, Maryland.

212. Bea, R.G., *The role of human error in design, construction, and reliability of marine structures*, 1994, Ship Structure Committee, SSC-378, Marine and Transport Technology, TUDelf, http://resolver.tudelft.nl/uuid:ad762bd0-0ebc-4a89-9d62-48fcfefb49ee.

213. Kidam, K. and M. Hurme, *Analysis of equipment failures as contributors to chemical process accidents.* Process Safety and Environmental Protection, 2013. **91**(1): p. 61-78.http://www.sciencedirect.com/science/article/pii/S0957582012000079.

214. Hallbert, B.P., A. Sebok, and D. Morisseau, *A Study of Control Room Staffing Levels for Advanced Reactors*, 2000, NUREG/IA-0137, U.S. Nuclear Regulatory Commission,

215. Barnes, M.J. and J. Warner. *Crew Studies for Army UAV applications 1994-2004*. in *Human Factors of Unmanned Aerial Vehicles Workshop "Manning the Unmanned"*. 2007. Cognitive Engineering Research Institute, Mesa, AZ.https://cerici.org/documents/2004_workshop/UAVCrewSystemsResearchatFtHuachuca.pdf.

216. Cummings, M.L., F. Gao, and K.M. Thornburg, *Boredom in the Workplace: A New Look at an Old Problem.* Human Factors, 2015. **58**(2): p. 279-300.https://doi.org/10.1177/0018720815609503.

217. Garrett, P.W., *"Full-Crew" Lows Fail as Safety Measures - In the 22 states where they operate, investigation proves such legislation has no material relations to 97.6% of accidents.*, acceessed on.

218. Hershberger, M.L., L.A. Scanian, and D.W. Craig, *Crew size evaluation for tactical all-weather strike aircraft*, 1975, D4163, P76-48R, HAir Force Avionics Laboratory, Wright-Patterson Air Force Base, OH,

219. Fadden, D.M., P.M. Morton, R.W. Taylor, and T. Lindberg. *First-Hand Evolution of the 2-Person Crew Jet Transport Flight Deck*. 2008 [cited 2020.

220. Schumacher, S., M. Kleinmann, and K.G. Melchers, *Job requirements for control room jobs in nuclear power plants.* Safety Science, 2011. **49**(3): p. 394-405.http://www.sciencedirect.com/science/article/pii/S0925753510002535.

221. Smith, J., *Small modular reactor issue identification and ranking program - control room staffing, fanal report* 2011, U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1114/ML111470784.pdf.

222. Nneji, V.C., *A Workload Model for Designing & Staffing Future Transportation Network Operations*, in *Department of Mechanical Engineering & Materials Science*. 2019, Duke University.https://dukespace.lib.duke.edu/dspace/handle/10161/18694.

223. Liao, P.-C., M. Liu, Y.-S. Su, H. Shi, and X. Luo, *Estimating the Influence of Improper Workplace Environment on Human Error: Posterior Predictive Analysis.* Advances in Civil Engineering, 2018. **2018**: p. 5078906.https://doi.org/10.1155/2018/5078906.

224. Taylor, C., M. Hildebrandt, R. McDonald, and N. Hughes, *Operator Response to Failures of a Computerised Procedure System: Results from a Training Simulator Study*, 2017, HWR-1198, OECD Halden Reactor Project,
225. Huang, F.-H. and S.-L. Hwang, *Experimental studies of computerized procedures and team size in nuclear power plant operations.* Nuclear Engineering and Design, 2009. **239**(2): p. 373-380.http://www.sciencedirect.com/science/article/pii/S0029549308005323.
226. Park, J., W. Jung, and J.-E. Yang, *Investigating the appropriateness of a decision chart to characterize the level of task descriptions in nuclear power plants.* Progress in Nuclear Energy, 2013. **66**: p. 41-51.http://www.sciencedirect.com/science/article/pii/S0149197013000498.
227. Kim, Y., J. Park, and W. Jung, *Measuring variability of procedure progression in proceduralized scenarios.* Annals of Nuclear Energy, 2012. **49**: p. 41-47.http://www.sciencedirect.com/science/article/pii/S0306454912002228.
228. Langstrand, J.-P., *Nuclear control operator simulator training: How eye tracking can be used to support training in a simulated environment*, in *Applied Computer Science*. 2016, Høgskolen i Østfold.https://hiof.brage.unit.no/hiof-xmlui/handle/11250/2448856?show=full&locale-attribute=en.
229. Norros, L. and P. Sammatti, *Nuclear power plant operator errors during simulator training*. 1986.
230. Szalma, J., P. Hancock, W. Dember, and J. Warm, *Training for vigilance: The effect of knowledge of results format and dispositional optimism and pessimism on performance and stress.* British journal of psychology (London, England : 1953), 2006. **97**: p. 115-35.https://pdfs.semanticscholar.org/884b/63973bb4708082ddd921a2d0160d553371a9.pdf?_ga=2.166180222.1192834828.1598066331-572251539.1549904442.
231. Morey, A. and P. Sanderson, *Are Melodic Medical Equipment Alarms Easily Learned?* Anesthesia and analgesia, 2008. **106**: p. 501-8, table of contents.https://www.semanticscholar.org/paper/Are-melodic-medical-equipment-alarms-easily-learned-Wee-Sanderson/49a703267cb1fbb3c181a80c5aa840c9c4392cb3.
232. Burkolter, D., *Training in process control: supporting performance in consideration of operator characteristics*, in *Psychologie*. 2010, Université de Fribourg.http://doc.rero.ch/record/19837/files/BurkolterD.pdf.
233. Liu, K.-H. and S.-L. Hwang, *Human performance evaluation: The procedures of ultimate response guideline for nuclear power plants.* Nuclear Engineering and Design, 2014. **273**: p. 234-240.http://www.sciencedirect.com/science/article/pii/S0029549314001824.
234. Skitka, L., K. Mosier, M. Burdick, and B. Rosenblatt, *Automation Bias and Errors: Are Crews Better Than Individuals?* The International journal of aviation psychology, 2000. **10**: p. 85-97.http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.896.2670&rep=rep1&type=pdf.
235. Bilke, T., *CAUSE AND PREVENTION OF HUMAN ERROR IN ELECTRIC UTILITY OPERATIONS*, https://spp.org/documents/18058/human%20error%20in%20electric%20utility%20operation.pdf.
236. Juhász, M.r. and J.n.K. Soós, *Human Aspects of NPP Operator Teamwork*, in *Nuclear Power - Control, Reliability and Human Factors*, P. Tsvetkov, Editor. 2010, Texas A&M University.http://cdn.intechopen.com/pdfs/21066/InTech-Human_aspects_of_npp_operator_teamwork.pdf.
237. Yim, H.B. and P.H. Seong, *A Quantitative Team Situation Awareness Measurement Method Considering Technical and Nontechnical Skills of Teams.* Nuclear Engineering

and Technology, 2016. **48**(1): p. 144-152.http://www.sciencedirect.com/science/article/pii/S1738573315002193.
238. Park, J., W. Jung, and J.-E. Yang, *Investigating the effect of communication characteristics on crew performance under the simulated emergency condition of nuclear power plants.* Reliability Engineering & System Safety, 2012. **101**: p. 1-13.http://www.sciencedirect.com/science/article/pii/S0951832012000051.
239. Kecklund, L.J. and O. Svenson, *Human errors and work performance in a nuclear power plant control room: associations with work-related factors and behavioral coping.* Reliability Engineering & System Safety, 1997. **56**(1): p. 5-15.http://www.sciencedirect.com/science/article/pii/S0951832096001378.
240. Morrow, S.L., G. Kenneth Koves, and V.E. Barnes, *Exploring the relationship between safety culture and safety performance in U.S. nuclear power operations.* Safety Science, 2014. **69**: p. 37-47.http://www.sciencedirect.com/science/article/pii/S0925753514000484.
241. Morrow, S.L., *Independent Evaluation of INPO's Nuclear Safety Culture Survey and Construct Validation Study*, 2012, U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1217/ML12172A093.pdf.
242. Pow, R.E., C. Mello-Thoms, and P. Brennan, *Evaluation of the effect of double reporting on test accuracy in screening and diagnostic imaging studies: A review of the evidence.* Journal of Medical Imaging and Radiation Oncology, 2016. **60**(3): p. 306-314.https://onlinelibrary.wiley.com/doi/abs/10.1111/1754-9485.12450.
243. Massaiu, S. *Micro-tasks results on identification reliability*. in *NRC HRA Data Workshop*. 2020. Rockville, Maryland, March 12-13, 2020: U.S. Nuclear Regulatory Commission
244. Massaiu, S. and A. Fernandes, *The reliability of identification tasks in digital and analogue interfaces: A re-analysis of two micro-tasks studies*, 2019, HWR-1218, OECD Halden Reactor Project,
245. De Carvalho, P.V.R., *Ergonomic field studies in a nuclear power plant control room.* Progress in Nuclear Energy, 2006. **48**(1): p. 51-69.http://www.sciencedirect.com/science/article/pii/S0149197005000430.
246. Massaiu, S. and L. Holmgren, *Preliminary Results from the 2013 Resilient Procedure Use Study with Swedish Operators*, 2016, HWR-1122, OECD Halden Reactor Project,
247. Massaiu, S. and L. Holmgren, *The 2013 Resilient Procedure Use Study with Swedish Operators: Final Results*, 2017, HWR-1216, OECD Halden Reactor Project,
248. Förster, J., E.T. Higgins, and A.T. Bianco, *Speed/accuracy decisions in task performance: Built-in trade-off or separate strategic concerns?* Organizational Behavior and Human Decision Processes, 2003. **90**(1): p. 148-164.http://www.sciencedirect.com/science/article/pii/S0749597802005095.
249. Strayer, D.L. and W.A. Johnston, *Driven to Distraction: Dual-Task Studies of Simulated Driving and Conversing on a Cellular Telephone.* Psychological Science, 2001. **12**(6): p. 462-466.https://doi.org/10.1111/1467-9280.00386.
250. Richard, C.M., R.D. Wright, C. Ee, S.L. Prime, Y. Shimizu, and J. Vavrik, *Effect of a Concurrent Auditory Task on Visual Search Performance in a Driving-Related Image-Flicker Task.* Human Factors, 2002. **44**(1): p. 108-119.https://doi.org/10.1518/0018720024494874.
251. Lu, S.A., C.D. Wickens, J.C. Prinet, S.D. Hutchins, N. Sarter, and A. Sebok, *Supporting Interruption Management and Multimodal Interface Design: Three Meta-Analyses of Task Performance as a Function of Interrupting Task Modality.* Human Factors, 2013. **55**(4): p. 697-724.https://doi.org/10.1177/0018720813476298.
252. Speier, C., *The influence of task interruption on individual decision making : an information overload perspective 337-360.* Decision Sciences, 1999. **30**(2).

253. Schuh, A.J., *Effects of an early interruption and note taking on listening accuracy and decision making in the interview.* Bulletin of the Psychonomic Society, 1978. **12**(3): p. 242-244.https://link.springer.com/content/pdf/10.3758/BF03329683.pdf.
254. Bailey, B.P. and J.A. Konstan, *On the need for attention-aware systems: Measuring effects of interruption on task performance, error rate, and affective state.* Computers in Human Behavior, 2006. **22**(4): p. 685-708.http://www.sciencedirect.com/science/article/pii/S074756320500107X.
255. Altmann, E.M., J.G. Trafton, and D.Z. Hambrick, *Momentary interruptions can derail the train of thought.* Journal of Experimental Psychology: General, 2014. **143**(1): p. 215-226.
256. Eyrolle, H. and J.-M. Cellier, *The effects of interruptions in work activity: Field and laboratory Results.* Applied Ergonomics, 2000. **31**: p. 537-543.
257. Li, S., A. Blandford, P. Cairns, and R. Young, *The Effect of Interruptions on Postcompletion and Other Procedural Errors: An Account Based on the Activation-Based Goal Memory Model.* Journal of experimental psychology. Applied, 2009. **14**: p. 314-28.https://pdfs.semanticscholar.org/0410/8bee4a0d6bf687817296c736fcf836e9b94c.pdf.
258. Nicholas, C.A., *The Effect of Interruptions on Primary Task Performance in Safety-Critical Environments*, in *Industrial Engineering & Operations Research*. 2016, University of Massachusetts - Amherst.https://scholarworks.umass.edu/cgi/viewcontent.cgi?article=1798&context=dissertations_2.
259. Lee, B. and V. Duffy, *The Effects of Task Interruption on Human Performance: A Study of the Systematic Classification of Human Behavior and Interruption Frequency.* Human Factors and Ergonomics in Manufacturing & Service Industries, 2014. **25**.https://interruptions.net/literature/Lee-HumFactorsErgonManufServIndust15.pdf.
260. McGuirl, J.M. and N.B. Sarter. *Presenting in-flight icing information: a comparison of visual and tactile cues*. in *20th DASC. 20th Digital Avionics Systems Conference (Cat. No.01CH37219)*. 2001.14-18 Oct. 2001.
261. Curry, R.E. and A.R. Ephrath, *Monitoring and Control of Unreliable Systems*, in *Monitoring Behavior and Supervisory Control*, T.B. Sheridan and G. Johannsen, Editors. 1976, Springer US: Boston, MA. p. 193-203.https://doi.org/10.1007/978-1-4684-2523-9_17.
262. Wickens, C.D. and C. Kessel, *The Effects of Participatory Mode and Task Workload on the Detection of Dynamic System Failures.* IEEE Transactions on Systems, Man, and Cybernetics, 1979. **9**(1): p. 24-34.
263. Ephrath, A.R. and L.R. Young, *Monitoring vs. Man-in-the-Loop Detection of Aircraft Control Failures*, in *Human Detection and Diagnosis of System Failures*, J. Rasmussen and W.B. Rouse, Editors. 1981, Springer US: Boston, MA. p. 143-154.https://doi.org/10.1007/978-1-4615-9230-3_10.
264. Fasanya, B., *Effects of Multitasking on Operator Performance using Computational and Auditory Tasks.* International journal of occupational safety and ergonomics : JOSE, 2016: p. 1-16.https://www.researchgate.net/profile/Bankole_Fasanya/publication/294918055_Effects_of_Multitasking_on_Operator_Performance_using_Computational_and_Auditory_Tasks/links/5b6b2522a6fdcc87df6da366/Effects-of-Multitasking-on-Operator-Performance-using-Computational-and-Auditory-Tasks.pdf.
265. Olive, T., R. Kellogg, and A. Piolat, *Verbal, visual, and spatial working memory demands during text composition.* Applied Psycholinguistics, 2008. **29**: p. 669-687.https://www.researchgate.net/profile/Annie_Piolat/publication/232004462_Verbal_visual_and_spatial_working_memory_demands_during_text_composition/links/02bfe50f2

df31d269b000000/Verbal-visual-and-spatial-working-memory-demands-during-text-composition.pdf.
266. Wickens, C.D., J. Goh, J. Helleberg, W.J. Horrey, and D.A. Talleur, *Attentional Models of Multitask Pilot Performance Using Advanced Display Technology.* Human Factors, 2003. **45**(3): p. 360-380.https://doi.org/10.1518/hfes.45.3.360.27250.
267. See, J., S. Howe, J. Warm, and W. Dember, *Meta-analysis of the sensitivity decrement in vigilance.* Psychological Bulletin, 1995. **117**: p. 230-249.https://www.researchgate.net/profile/Judi_See/publication/232476815_Meta-analysis_of_the_sensitivity_decrement_in_vigilance/links/56d6f6e908aee73df6c2f442/Meta-analysis-of-the-sensitivity-decrement-in-vigilance.pdf.
268. Grier, R.A., J.S. Warm, W.N. Dember, G. Matthews, T.L. Galinsky, J.L. Szalma, and R. Parasuraman, *The Vigilance Decrement Reflects Limitations in Effortful Attention, Not Mindlessness.* Human Factors, 2003. **45**(3): p. 349-359.https://doi.org/10.1518/hfes.45.3.349.27253.
269. Parasuraman, R. and M. Mouloua, *Interaction of signal discriminability and task type in vigilance decrement.* Perception & Psychophysics, 1987. **41**(1): p. 17-22.https://doi.org/10.3758/BF03208208.
270. Gawron, V., *The Watching - A Review of the Vigilance Research*, 2019, The MITRE Corporation, McLean, VA, https://www.mitre.org/sites/default/files/pdf/pr-16-3426-lessons-lost-the-watching-a-review-of-the-vigilance-research.pdf.
271. Craig, A., D.R. Davies, and G. Matthews, *Diurnal variation, task characteristics, and vigilance performance.* Human factors, 1987. **29**(6): p. 675-684.http://europepmc.org/abstract/MED/3436623

https://doi.org/10.1177/001872088702900607.
272. Webster, D.M., L. Richter, and A.W. Kruglanski, *On Leaping to Conclusions When Feeling Tired: Mental Fatigue Effects on Impressional Primacy.* Journal of Experimental Social Psychology, 1996. **32**(2): p. 181-195.http://www.sciencedirect.com/science/article/pii/S0022103196900098.
273. Wickens, C., S. Hutchins, L. Laux, and A. Sebok, *The Impact of Sleep Disruption on Complex Cognitive Tasks: A Meta-Analysis.* Human Factors The Journal of the Human Factors and Ergonomics Society, 2015. **57**.https://www.researchgate.net/profile/Christopher_Wickens/publication/272622494_The_Impact_of_Sleep_Disruption_on_Complex_Cognitive_Tasks_A_Meta-Analysis/links/54ec8fbb0cf28f3e6534401f/The-Impact-of-Sleep-Disruption-on-Complex-Cognitive-Tasks-A-Meta-Analysis.pdf.
274. Boring, R., D. Gertman, T. Ulrich, and T. Mortenson, *Consideration of Fatigue as a Factor in Human Reliability Analysis for Long-Duration Spaceflight*, 2019, INL/LTD-19-55988, Rev. 1, Idaho National Laboratoty: NASA
275. Griffith, C.D. and S. Mahadevan, *Inclusion of fatigue effects in human reliability analysis.* Reliability Engineering & System Safety, 2011. **96**(11): p. 1437-1447.http://www.sciencedirect.com/science/article/pii/S0951832011001244.
276. Griffith, C.D. and S. Mahadevan, *Human reliability under sleep deprivation: Derivation of performance shaping factor multipliers from empirical data.* Reliability Engineering & System Safety, 2015. **144**: p. 23-34.http://www.sciencedirect.com/science/article/pii/S0951832015001453.
277. Harrison, Y. and J.A. Horne, *One Night of Sleep Loss Impairs Innovative Thinking and Flexible Decision Making.* Organizational Behavior and Human Decision Processes, 1999. **78**(2): p. 128-145.http://www.sciencedirect.com/science/article/pii/S0749597899928273.

278. Harrison, Y. and J.A. Horne, *The impact of sleep deprivation on decision making: A review.* Journal of Experimental Psychology: Applied, 2000. **6**(3): p. 236-249.https://www.semanticscholar.org/paper/The-impact-of-sleep-deprivation-on-decision-making%3A-Harrison-Horne/387c4ba8b0a5bd5533a52d63a2324f02d0183797.
279. Thornburg, K., H. Peterse, and A. Liu, *Operator Performance in Long Duration Control Operations: Switching from Low to High Task Load.* Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2012. **56**: p. 2002-2005.https://www.researchgate.net/profile/Andrew_Liu12/publication/273346567_Operator_Performance_in_Long_Duration_Control_Operations_Switching_from_Low_to_High_Task_Load/links/5633dccd08aebc003ffde482/Operator-Performance-in-Long-Duration-Control-Operations-Switching-from-Low-to-High-Task-Load.pdf.
280. Baumann, M.R. *Time Pressure, Performance, and Awareness of Group Members in Co-acting Groups*. 1998.http://www.comp.nus.edu.sg/~cs6380/AMCIS_Baumann.pdf.
281. Akisato, K., T.L.v. Thaden, and W.D. Geibel. *Influence of time pressure on aircraft maintenance errors*. 2008.https://pdfs.semanticscholar.org/d67f/1d1641ec80c4f93b6d179d838d5f51e5e311.pdf.
282. Kerstholt, J.H., *The effect of time pressure on decision-making behaviour in a dynamic task environment.* Acta Psychologica, 1994. **86**(1): p. 89-104.https://apps.dtic.mil/dtic/tr/fulltext/u2/a245417.pdf.
283. van der Lubbe, R.H.J., P. Jaśkowski, B. Wauschkuhn, and R. Verleger, *Influence of time pressure in a simple response task, a choice-by-location task, and the Simon task.* Journal of Psychophysiology, 2001. **15**(4): p. 241-255.https://www.researchgate.net/publication/46613570_Influence_of_Time_Pressure_in_a_Simple_Response_Task_a_Choice-by-Location_Task_and_the_Simon_Task.
284. Alqahtani, D.A., J.I. Rotgans, S. Mamede, I. Alalwan, M.E.M. Magzoub, F.M. Altayeb, M.A. Mohamedani, and H.G. Schmidt, *Does Time Pressure Have a Negative Effect on Diagnostic Accuracy?* Academic Medicine, 2016. **91**(5).https://journals.lww.com/academicmedicine/Fulltext/2016/05000/Does_Time_Pressure_Have_a_Negative_Effect_on.33.aspx.
285. Trippas, D., S.J. Handley, and M.F. Verde, *The SDT model of belief bias: Complexity, time, and cognitive ability mediate the effects of believability.* Journal of Experimental Psychology: Learning, Memory, and Cognition, 2013. **39**(5): p. 1393-1402.https://www.researchgate.net/publication/236137379_The_SDT_Model_of_Belief_Bias_Complexity_Time_and_Cognitive_Ability_Mediate_the_Effects_of_Believability.
286. Anan, K., K. Fujiwara, C. Yaguchi, and N. Kiyota, *Effect of time pressure on attentional shift and anticipatory postural control during unilateral shoulder abduction reactions in an oddball-like paradigm.* Journal of Physiological Anthropology, 2014. **33**(1): p. 17.https://doi.org/10.1186/1880-6805-33-17.
287. Warren, W.R., *The Effect of Shift Turnover Strategy and Time Pressure on Aviation Maintenance Technician Performance*, in *Department of Human Factors & Systems*. 2011, Embry-Riddle Aeronautical University - Daytona Beach.https://commons.erau.edu/edt/146.
288. Wright, P., *The harassed decision maker: Time pressures, distractions, and the use of evidence.* Journal of Applied Psychology, 1974. **59**(5): p. 555-561.https://core.ac.uk/download/pdf/10199431.pdf.
289. Cain, M.S., J.E. Dunsmoor, K.S. LaBar, and S.R. Mitroff, *Anticipatory Anxiety Hinders Detection of a Second Target in Dual-Target Search.* Psychological Science, 2011. **22**(7): p. 866-871.https://doi.org/10.1177/0956797611412393.

290. Helton, W.S. and P.N. Russell, *The Effects of Arousing Negative and Neutral Picture Stimuli on Target Detection in a Vigilance Task.* Human Factors, 2011. **53**(2): p. 132-141.https://doi.org/10.1177/0018720811401385.
291. Nieuwenhuys, A. and R.R.D. Oudejans, *Anxiety and perceptual-motor performance: toward an integrated model of concepts, mechanisms, and processes.* Psychological Research, 2012. **76**(6): p. 747-759.https://doi.org/10.1007/s00426-011-0384-x.
292. Nibbeling, N., *Effects of Anxiety and Exercise-induced Fatigue on Operational Performance*, in *Human Movement Sciences*. 2014, Vrije Universiteit Amsterdam.https://www.semanticscholar.org/paper/Effects-of-Anxiety-and-Exercise-induced-Fatigue-on-Nibbeling/8e014ebd9c4da0dfa031d2caf2a9b897b8608e1d.
293. Keinan, G., *Decision making under stress: scanning of alternatives under controllable and uncontrollable threats.* Journal of Personality and Social Psychology, 1987. **52**(3): p. 639-644.https://www.ncbi.nlm.nih.gov/pubmed/3572731

http://www.ncbi.nlm.nih.gov/pubmed/3572731.
294. Robinson, O., K. Vytal, B. Cornwell, and C. Grillon, *The impact of anxiety upon cognition: perspectives from human threat of shock studies.* Frontiers in Human Neuroscience, 2013. **7**(203).https://www.frontiersin.org/article/10.3389/fnhum.2013.00203.
295. Nibbeling, N., R.R.D. Oudejans, E.M. Ubink, and H.A.M. Daanen, *The effects of anxiety and exercise-induced fatigue on shooting accuracy and cognitive performance in infantry soldiers.* Ergonomics, 2014. **57**(9): p. 1366-1379.https://doi.org/10.1080/00140139.2014.924572.
296. Sharp, M., M. Rosenberger, and J. Knapik, *Chapter 5 - Common Military Task: Materials Handling*, 2006, ARMY RESEARCH INST OF ENVIRONMENTAL MEDICINE NATICK MA MILITARY NUTRITION DIV, ADA472464 https://apps.dtic.mil/dtic/tr/fulltext/u2/a472464.pdf.
297. Steinhauser, M., *How to correct a task error: Task-switch effects following different types of error correction.* Journal of Experimental Psychology: Learning, Memory, and Cognition, 2010. **36**(4): p. 1028-1035.https://www.researchgate.net/profile/Marco_Steinhauser/publication/44689777_How_to_Correct_a_Task_Error_Task-Switch_Effects_Following_Different_Types_of_Error_Correction/links/552fd10c0cf20ea0a06e9456/How-to-Correct-a-Task-Error-Task-Switch-Effects-Following-Different-Types-of-Error-Correction.pdf.
298. Vesper, C., E. Abramova, J. Bütepage, F. Ciardo, B. Crossey, A. Effenberg, D. Hristova, A. Karlinsky, L. McEllin, S.R.R. Nijssen, L. Schmitz, and B. Wahn, *Joint Action: Mental Representations, Shared Information and General Mechanisms for Coordinating with Others.* Frontiers in Psychology, 2017. **7**(2039).https://www.frontiersin.org/article/10.3389/fpsyg.2016.02039.
299. Greenfield, M.A. and T.J. Sargent, *Probability of Failure of the Trudock Crane System at the Waste Isolation Pilot Plant (WIPP)*, 2000, EEG-74; DOE/AL/58309, Department of Energy, http://www.tomsargent.com/research/EEG-74.pdf.
300. Chandler, F., I.A. Heard, M. Presley, A. Burg, E. Midden, and P. Mongan, *NASA Human Error Analysis*, 2010, National Aeronautics and Space Administration, https://ftaassociates.files.wordpress.com/2018/12/F.-Chandler-NASA-Human-Error-Analysis-NASA-September-2010.pdf.
301. Jang, I., A.R. Kim, M.A.S.A. Harbi, S.J. Lee, H.G. Kang, and P.H. Seong, *An empirical study on the basic human error probabilities for NPP advanced main control room operation using soft control.* Nuclear Engineering and Design, 2013. **257**: p. 79-87.https://linkinghub.elsevier.com/retrieve/pii/S0029549313000186.

302. Williams, J.C. and J. Willey. *Quantification of Human Error in Maintenance for Process Plant Probabilistic Risk Assessment*. 1985: Institution of Chemical Engineers.1985/04/22/24.https://www.icheme.org/media/12179/viii-paper-22.pdf.
303. Baddeley, A.D. and D.J.A. Longman, *The Influence of Length and Frequency of Training Session on the Rate of Learning to Type.* Ergonomics, 1978. **21**(8): p. 627-635.https://www.tandfonline.com/doi/full/10.1080/00140137808931764.
304. Klemmer, E.T., *Communication and Human Performance.* Human Factors: The Journal of the Human Factors and Ergonomics Society, 1962. **4**(2): p. 75-80.http://journals.sagepub.com/doi/10.1177/001872086200400202.
305. Kirwan, B., G. Basra, and S.E. Taylor-Adams. *CORE-DATA: a computerised human error database for human reliability support*. in *Global Perpectives of Human Factors in Power Generation. Proceeding of the 1997 IEEE Sixth Conference on Human Factors and Power Plants*. 1997: IEEE.1997.http://ieeexplore.ieee.org/document/624858/.
306. Ohtsuka, T., S. Yoshimura, R. Kawano, M. Fujiie, H. Ujita, and R. Kubota, *Nuclear Power Plant Operator Performance Analysis Using Training Simulators: Operator Performance under Abnormal Plant Conditions.* Journal of Nuclear Science and Technology, 1994. **31**(11): p. 1184-1193.http://www.tandfonline.com/doi/abs/10.1080/18811248.1994.9735275.
307. Hotopf, N., *Slips of the pen.* Cognitive processes in spelling, 1980. **287**: p. 307.
308. Liu, P. and Z. Li, *Quantitative relationship between time margin and human reliability.* International Journal of Industrial Ergonomics, 2020. **78**: p. 102977.http://www.sciencedirect.com/science/article/pii/S0169814119300770.
309. Forester, J., H. Liao, V. Dang, A. Bye, E. Lois, M. Presley, J. Marble, R. Nowell, H. Broberg, M. Hildebrandt, B. Hallbert, and T. Morgan, *The U.S. HRA Empirical Study*, 2016, NUREG-2156, U.S. Nuclear Regulatory Commission, ML16067A241
310. **Bye, A.**, **E. Lois**, **V.N. Dang**, **G. Parry**, **J. Forester**, **S. Massaiu**, **R. Boring**, **P.Ø. Braarud**, **H. Broberg**, **J. Julius**, **I. Männistö**, and **P. Nelson**, *The international HRA empirical study - phase 2 report Results from comparing HRA methods predictions to HAMMLAB simulator on SGTR scenarios*, 2010, HWR-915, OECD Halden Reactor Project,
311. Kim, Y., S.Y. Choi, J. Park, W. Jung, and S.J. Lee. *Estimating Diagnosis Time of Emergency Situations in Digitalized Control Rooms*. in *AHFE 2018 International Conference on Human Error, Reliability, and Performance*. 2018. July 21-25, Florida.
312. Kim, Y., J. Kim, J. Park, S.Y. Choi, S. Kim, W. Jung, H.E. Kim, and S.K. Shin, *An HRA Method for Digital Main Control Rooms — Part I: Estimating the Failure Probability of Timely Performance*, 2019, KAERI/TR-7607/2019, Korea Atomic Energy Research Institute,
313. Albo, Y., J. Lanir, P. Bak, and S. Rafaeli, *Off the Radar: Comparative Evaluation of Radial Visualization Solutions for Composite Indicators.* IEEE Transactions on Visualization and Computer Graphics, 2016. **22**(1): p. 569-578.https://www.researchgate.net/profile/Sheizaf_Rafaeli/publication/282832017_Off_the_Radar_Comparative_Evaluation_of_Radial_Visualization_Solutions_for_Composite_Indicators/links/5628f1c108ae518e347c6ba2/Off-the-Radar-Comparative-Evaluation-of-Radial-Visualization-Solutions-for-Composite-Indicators.pdf.
314. Park, J., K. Jeong, and W. Jung, *Identifying cognitive complexity factors affecting the complexity of procedural steps in emergency operating procedures of a nuclear power plant.* Reliability Engineering & System Safety, 2005. **89**(2): p. 121-136.https://linkinghub.elsevier.com/retrieve/pii/S0951832004001899.
315. U.S. Nuclear Regulatory Commission, *Millstone Power Station Unit 3 - NRC Special Inspection Report 05000423/2005012*, 2005, NRC ADAMS Accession number: ML051860338,

316. U.S. Nuclear Regulatory Commission, *Palo Verde Nuclear Generating Station - NRC Integrated Inspection Report 05000528/2004003, 05000529/2004003, 05000530/2004003*, 2004, https://www.nrc.gov/reactors/operating/oversight/listofrpts-body.html#palo,
317. Commission, U.S.N.R., *H. B. Robinson Steam Electric Plant - Augmented Inspection Team Report 05000261/2010009*, 2010, IR 05000261/2010009 (ADAMS Accession No. ML101830101), U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1018/ML101830101.pdf.
318. U. S. Nuclear Regulatory Commission, *Final Precursor Analysis—H. B. Robinson Electrical Fault Causes Fire and Subsequent Reactor Trip with a Loss of Reactor Coolant Pump Seal Injection and Cooling*, 2011, ADAMS Accession No. ML112411359, U.S. Nuclear Regulatory Commission, https://www.nrc.gov/docs/ML1124/ML112411359.pdf.