White Paper: Practical Insights and Lessons Learned on Implementing Expert Elicitation

Completed: October 13, 2016

Prepared by:
Jing Xing & Stephanie Morrow
Office of Nuclear Regulatory Research
U.S. Nuclear Regulatory Commission

ABSTRACT

This report provides insights and guidance on the process of eliciting and integrating expert judgment, referred to as expert elicitation. It is intended to be used by NRC staff when performing an expert elicitation to derive expert judgment that will be used as an input to decision-making activities. This report is based primarily on literature review and lessons learned from three pilot applications where existing guidelines for eliciting expert judgment were adapted to support the elicitation of parameter inputs for risk analysis methods and models. Expert judgment is often needed when the available data about a technical issue is sparse or not applicable, the issue has great uncertainty, or is too complex to model accurately. This report includes an introduction to expert elicitation and its use within the NRC, a brief technical basis for eliciting expert judgment, guidance and insights for conducting a formal expert elicitation, and supplemental information based on lessons learned from research literature, existing guidance documents published by the NRC and other organizations, and recent expert elicitations performed by NRC staff.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

EXECUTIVE SUMMARY

In response to Staff Requirements Memorandum (SRM) COMGEA-11-0001, "Utilization of Expert Judgment in Regulatory Decision Making," the current report provides insights and practical guidance for staff when using expert elicitation to support decision-making activities.

Expert judgment is the information provided by experts in response to a technical question (Meyer & Booker, 1990). It represents an informed opinion or belief about the state of knowledge of a technical issue based on experts' training and background. Expert judgment obtained through a formal, structured process is referred to as expert elicitation.

The U.S. Nuclear Regulatory Commission (NRC) has used expert elicitation to inform many important regulatory decisions. Notably, expert elicitation has been an essential part of the NRC's use of probabilistic risk assessment (PRA) methods. The NRC's recent Level 3 PRA study uses expert elicitation in many key modeling areas, such as interfacing system loss of coolant accident (ISLOCA) modeling, low-power shutdown modeling, and human reliability analysis.

The NRC staff have customarily relied on three major documents for conducting expert elicitation:

- NUREG-1563, "Branch Technical Position on the Use of Expert Elicitation in the High-Level Radioactive Waste Programs," referred as BTP, provides an acceptable procedure for conducting expert elicitation in compliance with the NRC's geologic disposal regulations.
- NUREG/CR-6372, "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts," developed by the Senior Seismic Hazard Analysis Committee (SSHAC), describe a process to guide the performance of probabilistic seismic hazard analysis (PSHA) through formal expert judgment.
- NUREG-2117, "Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies," a companion to NUREG/CR-6372, provides practical implementation guidelines consistent with the framework and higher-level guidance of SSHAC.

These documents were developed for specific applications, but include many good practices that have been used to guide expert elicitation in other technical areas. Nevertheless, use of these documents often required interpretation and tailoring to apply the information to different applications and different types of expert judgment. The current report provides practical insights and lessons for adapting and implementing the SSHAC and BTP guidance in other technical areas. In particular, the current report provides insights based on lessons learned from three pilot applications where expert elicitation was used to develop parameter inputs for risk analysis methods and models. The guidance in this report does not supplant existing guidance in NUREG-1563, NUREG/CR-6372, or supporting documents (e.g., NUREG-2117). For projects that lack specific guidance on conducting expert elicitation, this report provides a first step for the staff to plan and conduct an elicitation project.

This report includes three parts:

- **The technical basis for using expert judgment.** This part describes the cognitive processing of formulating expert judgment, common cognitive vulnerabilities and biases in judgment, and strategies tactics and strategies to limit/mitigate these vulnerabilities and biases, and the mechanisms of using an interactive panel and structured process to achieve consistent results.

- **The expert elicitation guidance.** This part consists of (1) the conditions for conducting expert elicitation, (2) basic principles to follow when conducting expert elicitation, and (3) insights and guidance on the expert elicitation process that are consistent with the basic principles. The guidance is primarily adapted from the SSHAC guidelines and BTP report, and provides a generic framework that staff can use to tailor their expert elicitation projects to other technical areas.
- **Supplemental guidance for implementing the expert elicitation process.** This part identifies good practices for implementation based on the literature, techniques employed by past expert elicitation projects in the NRC and other organizations, and lessons learned from the three pilot projects.

The insights and guidance described in this report provides a structured, systematic approach to conducting expert elicitation. Depending on the intended use of the expert judgment and available resources, the staff may choose to implement an expert elicitation process at varying levels of effort, or develop application-specific procedures for conducting the elicitation. The basic principles and flexible process presented in this report should assist the staff in tailoring their expert elicitation to different applications, types of expert judgment, and levels of effort, particularly for technical areas that lack specific regulatory guidance such as SSHAC.

ACKNOWLEDGMENTS

This study was performed by the staff in the Office of Nuclear Regulatory Research (RES) of the U.S. Nuclear Regulatory Commission (NRC). The authors of this report sincerely thank the RES management and staff who supported the development and piloting of this guidance. Special thanks go to Nathan Siu, who provided direction and advice throughout the study. Thanks also go to Mark Salley, Kevin Coyne, and Sean Peters for their support of piloting the guidance in the three expert elicitation projects. The authors also appreciate the support of Christiana Lui, Mark King, Joseph Kanney, Susan Cooper, Mary Drouin, Douglas Coe, and Richard Correia for their input on the NRC staff's needs for the guidance of expert judgment. We also want to thank the NRC staff who reviewed this report and provided constructive inputs to greatly improve the quality of the report: Brittain Hill, Jon Ake, Donnie Harrison, Timothy McCartin, Todd Smith, Steve Ruffin, David Desaulniers, Ismael Garcia, Dan Widrevitz, and April Smith.

The initial guidance was piloted in three NRC expert elicitation projects; the project managers, peer reviewers, technical integrators/facilitators, and many experts of these projects provided verbal and/or written comments and suggestions on conducting expert elicitation. The lessons learned from piloting were incorporated into the final guidance document. The piloting work was extremely valuable to the study. The authors want to give a warm thanks to all of the individuals who participated in or supported the piloting. Special thanks go to the following individuals (listed alphabetically below):

Peter Appignani, NRC

Dennis Bley, The WreathWood Group

Robert Cavedo, Jr., Constellation Energy

Susan Cooper, NRC

Raymond Gallucci, NRC

Alan Kuritzky, NRC

Gerardo Martinez-Guridi, Brookhaven National Laboratory

Nicholas Melly, NRC

Dave Miskiewicz, Engineering Planning & Management, Inc.

Mary Presley, Electrical Power Research Institute

Steve M. Short, Pacific North National Laboratory

Nathan Siu, NRC

Mano Subudhi, Brookhaven National Laboratory

Gabriel Taylor, NRC

Michael D. Zentner, Pacific North National Laboratory

ACRONYMS AND ABBREVIATIONS

ACRS	Advisory Committee on Reactor Safeguards
ASME	American Society of Mechanical Engineers
ANS	American Nuclear Society
ATHEANA	A Technique for Human Event Analysis
BTP	Branch Technical Position
CFM	Crew Failure Mode
CFR	Code of Federal Regulations
CV	Check Valve
EPA	Environmental Protection Agency
EPRI	Electric Power Research Institute
FR	Federal Register
FTC	Failure to Close
HRA	Human Reliability Analysis
IDHEAS	Integrated Human Event Analysis System
ISLOCA	Interfacing System Loss of Coolant Accident
JACQUE-FIRE	Joint Assessment of Cable Damage and Quantification of Effects from Fire
LOCA	Loss of Coolant Accident
MOV	Motor Operated Valve
NASA	National Aeronautics and Space Administration
NRC	United States Nuclear Regulatory Commission
NUREG	U.S. Nuclear Regulatory Commission Staff Report
NUREG/CR	U.S. Nuclear Regulatory Commission Contractor Report
PIRT	Phenomena Identification and Ranking Table
PRA	Probabilistic Risk Assessment
PSHA	Probabilistic Seismic Hazard Analysis
PPRP	Participatory Peer Review Panel
RCP	Reactor Coolant Pump
RCS	Reactor Coolant System
RES	Office of Nuclear Regulatory Research
RHR	Residual Heat Removal
SI	Safety Injection
SRM	Staff Requirements Memorandum
SSHAC	Senior Seismic Hazard Analysis Committee

THERP	Technique for Human Error-Rate Prediction
TI	Technical Integrator
TFI	Technical Facilitator Integrator

1. INTRODUCTION

1.1 Background

Expert judgment is the information provided by experts in response to a technical question (Meyer & Booker, 1990). It represents an informed opinion or belief about the state of knowledge of a technical issue based on experts' training and background. Expert judgment obtained through a formal, structured process is referred to as expert elicitation. Expert elicitation has been a part of the technical basis for many regulatory decisions made by the U.S. Nuclear Regulatory Commission (NRC). Further, the use of expert elicitation to obtain expert judgment may be needed for new and emerging issues. For example, expert elicitation could play an important role in the resolution of difficult regulatory challenges involving new technologies, material aging issues, and severe accident risk assessment where little data or operational experience is available. In this report, we use the term "expert elicitation" to refer to the formal process of eliciting expert judgment.

In an expert elicitation, a technical expert panel provides information about their beliefs through the integration of available evidence on a technical issue (Morgan et al., 2014). The term "evidence" is used broadly to refer to such things as theories, models, experimental results, and empirical information from operational experience. The product of expert elicitation, (i.e., expert judgment) can be either (1) quantitative estimates of the frequency significance of technical issues, or (2) qualitative insights into the nature, scope, or significance of technical issues. Expert elicitation does not necessarily create knowledge; rather, it is the evaluation and integration of existing knowledge, to the extent that the knowledge base exists. Although expert elicitation is not a substitute for collecting data, it is a viable alternative when additional data collection or model development is impractical. Expert elicitation involves the following essential attributes:

- Identifying available evidence relevant to the technical issue
- Disseminating and sharing of common databases by the experts
- Evaluating the available evidence and views of the larger technical community
- Challenging experts' judgment through interactive workshops, feedback, and peer review
- Integrating and documenting the interactions of the expert panel

The NRC staff, licensees, and applicants have used expert judgment to support many important regulatory activities. Expert judgments can be a direct input to decision-making, or an indirect input where the elicited judgment feeds into an analysis whose output is then used for decision-making. An example of the direct use of expert judgment to support decision-making is through the application of Phenomena Identification and Ranking Tables (PIRT). The NRC has used PIRT in activities involving systematic identification of technical issues and qualitative determination of their significance, for example:

- Identifying and prioritizing technical issues associated with emerging technologies in advanced NPP control room design (NUREG/CR-6947).
- Prioritizing technical issues in associated with probabilistic risk assessment of cable fire in NPPs (NUREG/CR-7150, Vol.1).
- Understanding emergency core cooling system performance under LOCA scenarios. Specifically, PIRTs were used to investigate containment coating performance, evaluate debris transport in wet and dry containments, and to identify outstanding chemical effect issues (NUREG/CR-5249).

The NRC has also used expert elicitation to derive estimates of various parameters in risk analysis models. The outputs of those risk analysis models were then used as inputs to support regulatory decisions. The following are some examples of the indirect use of expert judgment:

- The landmark severe accident risk assessment study used expert elicitation to assess failure frequencies, failure modes, recovery actions, accident progression, and source term behavior, among other phenomena to characterize the risk associated with severe accidents in operating reactors (NUREG-1150).
- Expert elicitation has been used by applicants and licensees for probabilistic seismic hazard assessment (PSHA) for siting and design of nuclear facilities due to large uncertainties in the geoscience data and in their modeling. The Senior Seismic Hazard Analysis Committee (SSHAC) proposed a formal method for using expert judgment in PSHA (NUREG/CR-6372).
- Expert elicitation has been used by the U.S. Department of Energy to support its technical basis for the probability of volcanic events in its license application for a high-level waste repository. The NRC developed the Branch Technical Position (BTP) guidance document to provide an acceptable procedure for conducting expert elicitation in compliance with the NRC's geologic disposal regulations (NUREG-1563).
- Expert elicitation has been used to estimate loss-of-coolant-accident (LOCA) frequencies (NUREG-1829). These frequencies provided the basis for selecting the transition break size proposed in the risk-informed revision of the emergency core cooling system acceptance criteria (10 CFR 50.46a).
- Expert elicitation has been used to estimate human error probabilities for methods of human reliability analysis (e.g., NUREG-2199, NUREG-1624, NUREG/CR-2255, NUREG/CR-2743, and NUREG/CR-3688).

In 2011, the Commission directed the U.S. Nuclear Regulatory Commission (NRC) staff, via Staff Requirements Memorandum (SRM) COMGEA-11-0001, “Utilization of Expert Judgment in Regulatory Decision Making,” to develop a guidance document that would promote the consistent use of expert judgment in regulatory decision-making. The Commission requested that the guidance incorporate experience from past NRC activities that used expert judgment, adapt state-of-practices and lessons learned from the activities of other agencies relying on expert judgment, and integrate research on new approaches or methods to elicit expert judgment. Further, SRM–SECY-11-0172 directed the staff to pilot the draft guidance in Level 3 probabilistic risk assessment (PRA), including such areas as human reliability analysis (HRA) and severe accident analysis. In responding to these SRMs, the staff in the NRC’s Office of Nuclear Regulatory Research (RES) initiated the effort to review research and experience related to expert judgment, develop general guidance for the agency to use, and pilot the guidance in the NRC’s regulatory decision-making activities. This report describes the expert elicitation guidance developed as a result of these activities.

1.2 Purpose and Scope of the Report

The primary purpose of this report is to provide practical guidance to the NRC staff for conducting expert elicitation in the absence of other applicable guidance. NRC staff have customarily relied on three major documents for conducting expert elicitation:

- NUREG-1563, “Branch Technical Position on the Use of Expert Elicitation in the High-Level Radioactive Waste Programs,” referred as BTP, provides an acceptable procedure

for conducting expert elicitation in compliance with the NRC's geologic disposal regulations.

- NUREG/CR-6372, “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts,” developed by the Senior Seismic Hazard Analysis Committee (SSHAC), describe a process to guide the performance of probabilistic seismic hazard analysis (PSHA) through formal expert judgment.
- NUREG-2117, “Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies,” a companion to NUREG/CR-6372, provides practical implementation guidelines consistent with the framework and higher-level guidance of SSHAC.

These documents were developed for specific technical applications and have been incorporated into the NRC’s regulatory guidance and review plans (e.g., Regulatory Guide 1.208 is organized specifically for use of the established SSHAC process (NUREG/CR-6372 and NUREG-2117), and the Standard Review Plan for Yucca Mountain (NUREG-1804) is organized specifically for use of the BTP report (NUREG-1563)). The documents include many good practices for conducting expert elicitation and have been used to guide expert elicitation in other technical areas. Nevertheless, use of these documents often requires interpretation and tailoring to apply the information to different applications and different types of expert judgment. The current report provides practical insights and lessons for adapting and implementing the SSHAC and BTP guidance in other technical areas. Use of the guidance in this report can also help support the use of expert judgment as outlined in the PRA Standard (ASME/ANS RA-Sa-2009). The guidance in this report does not supplant existing guidance in NUREG-1563, NUREG/CR-6372, or supporting documents (e.g., NUREG-2117). For projects that lack specific guidance on conducting expert elicitation, this report provides a first step for the staff to plan and conduct an elicitation project.

This report consist of the following elements:

- **Chapter 1:** Introduction to expert elicitation and purpose of the report.
- **Chapter 2:** A brief description of the fundamental concepts in expert elicitation.
- **Chapter 3:** The basic principles of expert elicitation and a process for staff to follow when conducting expert elicitation.
- **Chapters 4:** Questions and answers for various issues staff have encountered when conducting expert elicitation.
- **Appendix A:** The Commission SRM on utilization of expert judgment.
- **Appendix B:** Summary of unaddressed comments from the NRC user offices’ review of the draft of this report.

Note that this report focuses on the process for what should be included in expert elicitation, while still leaving flexibility in implementing the process. The report is intended to be generally applicable to different types of expert judgment (e.g., probabilistic, numeric, preference ranking, open-ended opinions), which can then be tailored to specific applications. For example, PIRT is a special type of expert elicitation where the expected outcomes are typically not pre-defined and may vary in the process. While the general guidance in this report is applicable to PIRT, it may still be desirable to have a specific procedure for implementing PIRT.

Additionally, although the general guidance in this report is compatible with the BTP and SSHAC guidelines, this guidance is not a substitute for the detailed processes developed in NUREG-1563, NUREG/CR-6372 and associated regulatory guidance. The insights and lessons

learned in this report do not specifically address complex expert elicitation where a model of complex physical phenomenon needs to be developed or multiple alternative models need to be evaluated; NUREG/CR-6372 and 2117 provide explicit guidance for such applications. The report also does not provide specific techniques for eliciting knowledge from individual experts, or mathematical methods for combining elicited probability distributions. The report does provide additional references to tailor the guidance to specific applications, as discussed in the next section.

1.3 Relationship between This Report and Other Expert Elicitation Guidance

After a thorough review of various guidance documents on expert elicitation, the NRC staff concluded that the recommendations detailed in the SSHAC guidelines provide a reasonable framework for implementing expert elicitation. The SHAAC guidelines are compatible with the guidance in the PRA Standard and BTP report. Furthermore, the SSHAC guidelines are consistent with guidelines and processes for expert elicitation used by other organizations, such as the National Aeronautics and Space Administration (NASA) and the U.S. Environmental Protection Agency (EPA; Monroe, 1997; EPA, 2011; Meyer et al., 2002; Simola et al., 2005; Hukki, 2008; Cooke & Goossens, 2000). Although the SSHAC guidelines were initially developed for Probabilistic Seismic Hazard Analysis (PSHA), the SSHAC recognized that its technical guidelines and procedures were generally applicable to expert judgment in other domains. This has been confirmed by many applications that successfully used the SSHAC process for expert elicitation in the past two decades. While most expert elicitation guidance documents describe a similar process but differ in procedural details, the SSHAC guidelines provide a framework with basic principles, a highly structured process, and recommendations for implementation. As a result, the SSHAC guidelines were used as the foundation for this report, with the addition of useful aspects from other guidance documents, scientific literature, and lessons learned from the NRC's recent implementations.

The guidance in this report is based primarily on lessons learned from three pilot applications where the SSHAC guidelines were adapted to support the elicitation of parameter inputs for risk analysis methods and models. At the time of writing this report NUREG-2117 was also in the process of being updated. That update is intended to augment the existing implementation guidance and capture lessons learned from the many applications of the SSHAC guidelines since the 1990's. Topics that will be covered in the update include criteria for SSHAC level 1 and 2 studies, criteria for updating existing PSHAs, and application of SSHAC to non-seismic topics. Table 1-1 summarizes many of the references that provide specific recommendations for implementing expert elicitation, or detailed descriptions of projects that used expert elicitation. In particular, the last three references describe the three pilot applications that informed the practical implementation techniques in this report.

Table 1-1. Expert Elicitation References

Document	Document Description	Notes for Reference
PRA Standard (2009)	*Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications (ASME/ANS RA-Sa-2009)*	Provides broad requirements governing the use of expert judgment in probabilistic risk assessment.

BTP Report NUREG-1563 (1996)	*Branch Technical Position on the Use of Expert Elicitation in the High-Level Radioactive Waste Program (NUREG-1563)* Developed by the NRC staff to set forth technical positions that: (1) provide general guidelines on those circumstances that may warrant the use of a formal process for obtaining the judgments of more than one expert (i.e., expert elicitation); and (2) describe acceptable procedures for conducting expert elicitation when formally elicited judgments are used to support a demonstration of compliance with NRC's geologic disposal regulation, currently set forth in 10 CFR Part 60.	Describes an expert elicitation process in compliance with the NRC's geologic waste disposal regulations. • The BTP process was incorporated into the Yucca Mountain review plan, NUREG-1804. • The BTP guidelines specific to the application of PSHA were superseded by NUREG/CR-6372 and NUREG-2117.
SSHAC Guidelines NUREG/CR-6372 (1997)	*Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts (NUREG/CR-6372)* Developed by the Senior Seismic Hazard Analysis Committee (SSHAC) to provide technical guidance on addressing uncertainties in Probabilistic Seismic Hazard Analysis (PSHA) using expert judgment. It describe a process, referred to as SSHAC guidelines, to guide the performance of PSHA for seismic regulation of NPPs and other critical facilities through formal expert judgment.	Detailed guidelines and recommendations for complex expert elicitations involving developing or interpreting models with high uncertainties. • Includes applications of the expert judgment to PSHA. • Technical guidance and procedures are generally applicable to other domains.
NUREG-2117 (2012)	*Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies. Kammerer, A. M., & Ake, J. P.* As a companion to NUREG/CR-6372, this NUREG provides additional practical implementation guidelines consistent with the framework and higher-level guidance of the SSHAC Guidelines. It includes lessons learned from past PSHA studies that have used the SSHAC Guidelines.	Implementation guidance based on lessons learned from applications of the SSHAC process. • A future update of NUREG-2117 plans to include insights and lessons from application of the SSHAC Guidelines in a number of recent studies.
NUREG/CR-5424 (1990)	*Eliciting and Analyzing Expert Judgment, A Practical Guide, Meyer, M. A., & Booker, J. M.* The report provides background on the uses of expert judgment and on the processes by which humans solve problems, including those that lead to bias. Detailed guidance is offered on how to elicit expert judgment ranging from selecting the questions to be posed of the experts to selecting and motivating the experts to setting up for and conducting the elicitation. Analysis procedures are introduced and guidance is given on how to understand the database structure, detect bias and correlation, form models, and aggregate the expert judgment.	Guidelines and technique for expert elicitation. • Detailed discussion of cognitive biases • Techniques for combining individual judgments • Mathematical basis and techniques of expressing probabilities and uncertainties • Q & A on implementing the expert elicitation process

NUREG/CR-4962 (1987)	*Methods for the Elicitation and Use of Expert Opinion in Risk Assessment. Mosleh, A., Bier, V. M., & Apostolakis, G.*	The content of this report served as a basis for the BTP and SSHAC reports.
NUREG/CR-7150 (2015)	*Joint Assessment of Cable Damage and Quantification of Effects from Fire (JACQUE-FIRE), Volume 2: Expert Elicitation Exercise for Nuclear Power Plant Fire-Induced Electrical Circuit Failure* The Joint Assessment of Cable Damage and Quantification of Effects from Fire (JACQUE-FIRE) is a program in using an expert elicitation to build upon information developed in an electrical engineering panel's PIRT exercise. Volume 1 of the report details the findings of the PIRT exercise. Volume 2 documents the results of the PRA panel's expert elicitation.	Expert elicitation application. • Demonstration of SSHAC process implementation • Elicitation of probabilities and quantities • A good template of an expert elicitation project plan • Training materials
NUREG/CR (draft) PNNL-24783 (2016)	*Expert Elicitation to Support Interfacing System Loss of Coolant Accident (ISLOCA) Modeling. Short, S. M. et al.* This report describes the process and results of the expert elicitation to resolve the ISLOCA modeling issues in the NRC's full-scope, site Level 3 PRA.	Expert elicitation application. • Simplification of SSHAC process • Used web-based meeting to reduce cost • Experience with very low probabilities (< 1E-5) • Lessons learned on expert elicitation training
NUREG-2199, Vol. 2 (2017)	*IDHEAS Internal At-power Application, Vol. 2. Expert Judgments of Human Error Probabilities.* The NRC and EPRI developed a new HRA method, the Integrated Decision-tree Human Event Analysis System (IDHEAS) for nuclear power plant internal at-power applications. The method uses 14 crew failure modes (CFM) to characterize human failure events. Each CFM has a decision tree to represent the combinations of the factors that influence the error probability of the CFM. The objectives of the study were to (1) evaluate and refine the decision trees of the CFMs, (2) estimate the human error probabilities of the CFMs for every combination of the factors in the decision tree, and (3) pilot the expert elicitation process documented in this report.	Expert elicitation application. • Elicitation of probabilities and conceptual models • Template for a detailed project plan • Discussion of pitfalls and caveats in following and deviating from SSHAC guidelines

1.4 Development Approach

The NRC Commission's Staff Requirements Memorandum (SRM) COMGEA-11-0001directed staff to develop a guidance document that would promote the consistent use of expert judgment in regulatory decision-making. SRM–SECY-11-0172 directed staff to pilot the draft guidance in Level 3 PRA project. Following the direction provided by the Commission in SRM

COMGEA-11-001 and SECY-11-0172, staff in the NRC's Office of Nuclear Regulatory Research (RES) took the following approach to develop the expert elicitation guidance:

1) Reviewed guidance on expert elicitation from the NRC and other organizations, and research literature in areas including decision-making, expert judgment and expert elicitation, knowledge elicitation, probability calibration, and cognitive biases. The main findings are:

   - The guidelines and practices describe essentially the same process, with variations in the details for implementing the process.
   - The SSHAC approach is a formal, structured process that essentially encompasses the components found in other expert elicitation guidance. It also incorporated significant research findings on tactics and strategies of mitigating cognitive vulnerabilities and biases inherent to expert judgment. There have been many successful experiences of applying SSHAC (as documented in NUREG-2117) and adapting the SSHAC approach to areas other than PSHA.
   - The research literature provides a technical basis for using expert judgment in decision-making, which helps to define the basic principles of the elicitation process. The literature also provides information that helps with implementing the various components of the elicitation process, such as estimating probabilities or aggregating individual assessments.

2) Interviewed NRC staff working on risk-informed applications to identify the agency's needs for specific guidance. Highlights from the interviews include:

   - The guidance should provide a basis for determining when to use expert elicitation.
   - The guidance should address all types of applications of expert judgment (e.g., quantification, probabilities, fuzzy logic, conceptual model, PIRT).
   - The guidance should outline the basic principles to follow while also giving staff flexibility to implement the process. The staff needs to understand the implications of simplifying or deviating from the recommended expert elicitation process.
   - The guidance should point out the key areas for properly conducting expert elicitation and potential pitfalls and caveats in an expert elicitation process.
   - The guidance should address a common problem in past expert elicitation applications: Ensuring the experts have a common and correct understanding of the questions being asked (i.e., having everyone on the same page).

3) Developed draft guidance for piloting. The draft consists of three parts: (1) a set of basic principles with which the expert elicitation process should comply, (2) insights and guidance on adapting the SSHAC approach, and (3) templates and tips for implementing the elicitation process.

4) Piloted draft guidance in three NRC projects that used expert elicitation.

5) Developed the final report by incorporating the experience and lessons learned from piloting.

The NRC staff from five offices (RES, NRR, NRO, NMSS, and NSIR) reviewed the draft version of the final report and provided numerous comments that significantly improved the report. Appendix B of this report documents some of the comments that were not addressed in the report because they requires substantial additional work.

2. TECHNICAL BASIS FOR EXPERT ELICITATION

2.1 Expert Judgment

Expert judgment is information provided by an expert in response to a technical question (Meyer & Booker, 1990). It represents an informed opinion or belief about the state of knowledge of a technical issue based on the expert's training and background. Expert judgment can be qualitative or quantitative. For example:

- Quantities of parameters or variables.
- Likelihood or frequency of events.
- Prior distributions for Bayesian statistical models or interpretations of observed data.
- Quantitative bounds on subjective judgments (e.g., interpretations of qualitative terms such as "likely" and "rare")
- Conceptual models (formal or informal descriptions of cause-and-effect in a system or problem)
- Consensus among experts regarding a technical view or a contentious decision (Cooke and Goossens, 2000).
- Identification of phenomena and inputs of relative importance or significance for the prioritization of potential research options or potential decision options, (e.g., PIRT).

Expert judgment is often used as one form of the evidence or in combination with other forms of evidence because the "true answer" to a technical question cannot be established by other means. This may be because the issue is complex, novel, rare, or poorly understood. The quality of a judgment is indicated by the extent to which it is statistically accurate, precise, and complete:

- Statistically accurate (Cooke, 1991) - a good judgment is one that mimics the underlying probability of predicting the "truth" if it were known. In other words, the distribution of values presented by the experts should capture the "true" value within the expressed confidence intervals (i.e., 90% confidence intervals should include 90% of the true values). Furthermore, the estimates should be balanced. For example, 50 percent of any "true" values should be above, and 50 percent should be below an expert's estimated median values.
- Precise - a good judgment is one in which the probability mass is concentrated in a region (preferably near the true value) relative to the overall distribution.
- Complete - a good judgment should capture the range or distribution of knowledge about the technical question.

In reality, since the truth of the answer to the technical problem is unknown, these three characteristics cannot be analytically demonstrated. Instead, they are expressed in terms of the center, body, and range of the state of knowledge about a technical issue, and they are demonstrated through the formal, structured process employed to obtain the judgment.

2.2 Definition of Experts

To elicit expert judgment, there must be experts who have the knowledge and capability to support informed judgment and prediction about the technical issue of interest. An expert must have (1) specific knowledge and expertise in the technical area of interest, and (2) the ability to predict the qualitative or quantitative aspects of the technical issue. The predictive ability of an expert is demonstrated by his or her mental frames of interpreting the evidence and relating the evidence to the questions asked.

Meyer and Booker (NUREG/CR-5424,1990) define an expert as someone who is knowledgeable about a technical issue at an appropriate level of detail given the information that is being elicited, and who is capable of communicating his/her knowledge and expertise. Subject matter expertise is technical training, experience, knowledge of data and theory. Normative expertise is the ability to communicate, involving knowledge of the conventions and jargon of a field.

In decision-making theories, experts are also referred to as subject matter experts and normative experts. The former are the ones who possess expertise in the technical area(s) under investigation; normative experts are individuals who possess expertise in decision-making theories, cognitive science, and statistical and probabilistic theories.

2.3 Cognitive Basis for Generating Expert Judgment

The basis for making judgments resides in the human cognitive process for understanding a problem and making decisions based on understanding. Figure 2-1 is a diagram of the cognitive decision-making process. A key element in the decision-making process is the individual's mental models. People organize their knowledge into structured, meaningful patterns, referred to as mental models, store them in the memory, and use them to interact with the environment. Mental models serve to help people describe, explain, and predict system behaviors. The inputs to the decision-making process are the question being asked, data relevant to the question, and models that relate the data to the question. One then uses his or her mental models to process the inputs. The output of the process is one's belief about the state of the inputs (i.e., the decision or judgment).

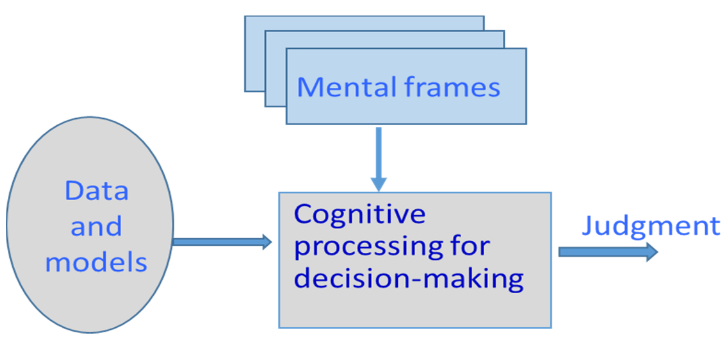

Figure 2-1. Illustration of human decision-making (making judgment).

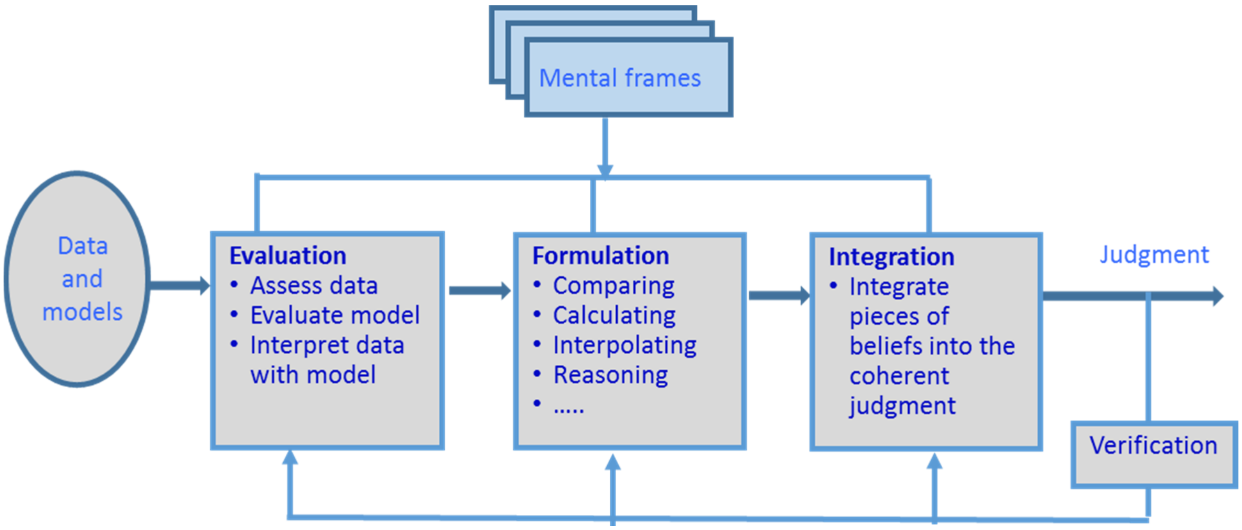

Figure 2-2. Detailed cognitive process of decision-making.

Figure 2- is an expanded diagram of Figure 2-1. We developed this diagram based on literature from the cognitive sciences to depict the detailed cognitive process of decision-making (Klein 1993; NUREG-2114). The diagram includes the following key elements:

1) Evaluation - understand and make sense of the inputs; this includes using one's mental frames to assess the data, evaluate the models, and interpret data with the model;
2) Formulation - formulate or generate beliefs or interpretations of the data and models through reasoning, inferencing, calculating, and comparing;
3) Integration - combine pieces of beliefs or interpretations into a coherent representation that answers the question; that is one's preliminary judgment;
4) Verification and Iteration - examine or mentally simulate whether the preliminary judgment answers the question and iterate the process if one is not satisfied with the judgment.

All of these elements involve using one's mental frames to process information. Cognitive research has revealed many vulnerabilities in human cognitive information processing, for example:

- Selective perception - This is the process of interpreting something the way you expect to see it (Marr 1982). People tend to filter the information they perceive based on their expectations. Moreover, people often are not aware of their expectations or not aware of using their expectations to filter information. In an expert elicitation, the challenge comes in recognizing the filters that the experts are using.
- Cognitive capacity limits - Since a person's cognitive capacity in processing information is limited (Cowan, 2012), people naturally tend to simplify the process by focusing only on the significant parts of the data and models, and fitting the data to a familiar mental frame instead of examining all of the frames (Klein, 1993). Such simplifications, referred to as heuristics, can introduce cognitive biases in judgment.
- Framing effect - A change in the presentation of a choice influences a person's choice behavior, even when the objective characteristics (outcomes or probabilities) are not changed (Kahneman and Tversky, 1979, 1984).

- Limited mental frames - Even the most experienced expert may not possess sufficient experience and expertise to have developed mental frames for thoroughly comprehending the evidence about the question at hand, especially when the question addresses rare events (Kahneman and Tversky 2002).
- Inconsistency - Mental frames, and the use of those frames, are subject to change given the context within which a question in presented and a person's mental state at the time a question is presented (e.g., stressed vs. relax, vigilant vs. fatigued).

Because of these kinds of vulnerabilities, the resultant judgment may have the following limitations:

- Biased - does not accurately represent the center of the distribution
- Incomplete - only captures part of the distribution
- Imprecise - is overly broad in representing the overall distribution

Team decision-making has been the natural choice for overcoming individual cognitive vulnerabilities and limitations as it engages the application of various classes of expertise by individual team members to expand individual's mental models and cognitive capacity. Morgan et al (1986) defined a team as: A group of individuals who must interact in pursuit of shared valued objectives. Team decision-making refers to a process that involves gathering, processing, integrating, and communication information in support of arriving to a task relevant decision. One basic principle for team decision-making is that team decision-making does not require that a consensus is reached among the team members. Instead, team decision-making requires that team members process and filter "raw" data, apply individual expertise, communicate relevant information, and make recommendations or challenges to other members. As such, the process of team decision-making has been employed in formal expert elicitation to generate consistent expert judgment. A formal expert elicitation means that an expert panel employs a formal, structured process to obtain the judgment that represents the state of knowledge and views of the overall technical community, As the panel members interact and reach a common understanding of the information shared by the members, the quality of the judgment is often far better than individual judgments .

In team decision-making, while the team members hold unique expertise that they bring to bear with the team, it is essential that the team members must hold shared mental models about the task in order to perform effectively (Klein et al 1993). Research found that establishing and maintaining shared mental models among the team members requires balancing the team member expertise, training the team members, and effectively facilitating the decision-making process (Cannon-Bowers & Salas, 1990). Such findings can be insightful for conducting expert elicitation. For example, when team members' knowledge and expertise overlap too much, their uniqueness of expertise gets lost and members tend to become group-thinking; as the result, the decision-making process leads to a single-minded view of the problem. This finding suggests the importance of having balanced expertise in an expert panel.

Mechanisms and strategies for effective team decision-making have been extensively studied. Incorporating those mechanisms and strategies into a structured process is essential to manage the various biases. Below we summarize the essential elements for effective team decision-making:

- Use balanced expertise to expand individuals' mental models and stimulate discussion
- Assign different roles and responsibilities to the members

- Train the members to establish shared mental models of the technical problem and the task;
- Maintain shared mental models throughout the process through interaction and facilitation
- Make available the range of data and models relevant to the technical problem and have the team thoroughly evaluate them to reveal diversities and uncertainties;
- Integrate information rather than reach a consensus
- Separate the different parts of the information processing (evaluation, formulation, and integration) to minimize use of cognitive heuristics (i.e., the experts have to thoroughly perform one part before moving to the next).

Over the years, the various guidance for conducting formal expert elicitation has adopted many principles and strategies of team decision-making. For example, the Delphi technique (Linstone and Turoff 1975) uses a group process in which individual experts make initial judgments, a facilitator aggregates the judgments, and the two steps iterate until a consensus is made. Moreover, the SSHAC guidelines advanced expert elicitation by imposing a structured process to facilitate interaction among the experts throughout the elicitation process to reduce inconsistency and minimize biases.

2.4 Cognitive Biases and Debiasing Strategies

Kahneman and Tversky (2002) stated, "We humans are not equipped with a competent mental statistical processor. Rather, in making judgments in the face of uncertainty, we unconsciously use a variety of cognitive heuristics. As a consequence, when asked to make probabilistic judgments, either in a formal elicitation or in a less formal setting, people's judgments are often biased." Cognitive biases in judgments are inherent from individual's cognitive vulnerabilities (as discussed in the previous section) and the process of team decision-making. Cognitive biases have been extensively studied in the literature. NUREG/CR-5424 and other guidance documents discuss many types of heuristics and biases. Below are some cognitive biases that are most relevant to expert elicitation.

Availability bias - Humans, including experts, often make judgments based on the ease with which instances or occurrences of an event can be brought to mind. For instance:

- Events that have occurred recently, or that have particular personal significance will come to mind more easily.
- When evaluating, one may be inclined toward using models that one is more familiar with or feels an affinity for because of knowing personally or by reputation of the authors of a given model.
- Rare events that attract disproportionate press attention will often be judged to be more probable than is actually the case. Therefore, the relative frequencies of well-publicized events (e.g., tornados) tend to be overestimated, whereas the relative frequencies of more mundane events (e.g., common influenza) are often underestimated. Slovic (1987) coined the term 'dread risks' to refer to high consequence events, where many people are killed at once, but which have relatively low probability of occurrence.

Anchoring and adjustment - When required to estimate a number, percentage or range of values, people often anchor on values that have either been previously suggested, or have emerged from previous judgments. Once anchored to a particular value, people adjust their estimates up or down from that point, in the direction that seems intuitive given the question

(Tversky and Kahneman, 1974). The strategy is particularly prevalent for repeated judgments of the same kind of situation, where the value from a previous instance can serve as an anchor and then adjustment can be made according to the differences between the previous and the current situation. Unfortunately, people tend to stick too closely to the initial value, not adjusting sufficiently through fully exploring the variabilities and uncertainties in the event.

Over-confidence - Overconfidence describes a situation in which the expert's own confidence in their judgment does not correspond to the *accuracy* of that judgment or estimate. This is often seen in experts' underestimation of uncertainty, which is the failure to account for the actual amount of uncertainty in the answers given.

For example, a researcher found that when people are asked to put a range around an answer such that they are 90% sure that the range encompasses the correct answer, their ranges only cover 30 to 60% of the total. A popular explanation for this effect is that people are uncomfortable with the amount of uncertainty in life, and thus minimize it. In particular, people may avoid confronting the large uncertainties in their judgment.

This is not an extensive list of heuristics and biases. Others include attribution and fallacies related to representativeness, correlation, and causation. There are also non-cognitive biases, such as those associated with individual behaviors, culture, and motivation. Moreover, eliciting expert judgment from a panel of experts may introduce different types of cognitive inconsistency, such as group polarization, amplification of errors, and cascade effect (NUREG/CR-5424).

While cognitive biases are inherent to the underlying cognitive process of making judgments, they can be managed through an effective team decision-making process and various strategies that combat with cognitive vulnerabilities. A formal, structured process incorporating team decision-making principles is the essential mechanism of reducing biases. In addition, the literature also explored varies strategies of debiasing. Gigerenzer and others (Gigerenzer et al, 1999) provides some insight on understanding and managing biases in perceptions of probability and risk. Kirkeboen (2009) reviewed de-biasing strategies in improving expert judgment. He classified them into two types:

1) Motivational and cognitive strategies, including:
   - Drawing people's attention to decision bias
   - Training in normative decision rules (e.g., Bayes' formula)
   - Learning to reformulating decision problems (e.g., taking an outsider's view, considering the opposite)
2) Technical strategies, including:
   - Incorporating a team decision-making process
   - Fostering deliberation by using different kinds of technological means (e.g., knowledge-management tools, problem-probing software)
   - Automating judgments

The most prevalent de-biasing strategy in this category is making experts aware of the biases in order to make them pay attention to potential biases and thereby avoid them. Unfortunately, this strategy seems to have little effect, as noted in Fischhoff's studies (Fischhoff et al, 1982; Fischhoff and MacGregor 1982) and the authors' observations in several expert elicitation projects sponsored by the NRC. Training on normative decision rules has some effect in certain circumstances, if conducted properly, such as accompanying with exercise and

feedback (Larrick, 2004). Reformatting decision problems requires quite a lot of effort on behalf of the experts and the effect seems to be case-specific (Gigernzer, 1996).

It should be noted that cognitive biasing cannot be completely removed due to cognitive vulnerabilities and uncertainties in the physical nature of the problems and human's knowledge about the problem. Thus, an important part of debiasing is to document potential sources of biases in resulted expert judgment. The disclosure of such information helps to manage the impact of the biases when the expert judgment is used as the input to decision-making activities.

2.5 Review of Guidance documents in Expert Elicitation

We reviewed available guidance documents on conducting expert elicitation (to the extent that we could find in public domains) and the literature documenting expert elicitation projects. Essentially, all of the guidance provided a process consistent to that of team decision-making:

- defining the technical problem,
- assembling the expert panel,
- training the panel members,
- gathering information,
- communicating and evaluating information,
- interacting and developing judgments,
- integrating the judgments, and
- documenting the process and outputs.

The various guidance differs in the details of implementing the process. Some guidance documents focus on specific elements of the elicitation process and incorporated many tactics and strategies on specific implementation issues from scientific literature. For example, NUREG/CR-5424 presented a collection of various methods of aggregating individual experts' judgment; the guidance by the Environmental Protection Agency (EPA) provides strategies for many practical issues in implementing elicitation, from selecting experts to documenting the outputs. On the other hand, some guidance documents present the elicitation process with little information on the basic principles and guidelines on how to construct a formal, structured process; implementing such guidance largely depends on the users' knowledge and experience to ensure the quality of the resulting outputs.

Among the guidance we reviewed, SSHAC provides guidelines on constructing a formal, structured process that incorporated the basic cognitive principles of individual and team decision-making. Some notable features in SSHAC include:

1) One overarching principle for SSHAC is that it does not require a consensus for a single answer to the technical problem; instead, it explicitly requires that the output represents the distribution of the views of the technical community;
2) SSHAC explicitly defines a formal, structured process regarding the roles and responsibilities of the panel members as well as the ways elicitation is conducted (e.g., separate workshops, facilitation);
3) SSHAC emphasizes the importance of having independent, participatory peer review throughout the elicitation process;

4) SSHAC process goes beyond the elicitation process in most guidance documents; SSHAC provides guidance on incorporating expert judgment into the decision-making process or models.

We evaluated SSHAC guidance against the cognitive basis of understanding and decision-making. SSHAC incorporated the essential elements for effective team decision-making. Although it was originally developed for seismic hazard analysis, its basic principles and framework are consistent with the scientific foundation of expert judgment, therefore, they are applicable to other technical areas. Following these principles and guidance engages the expert panel in a process that produces meaningful output of suitable quality to support decision-making activities.

The literature that reported expert elicitation activities generally demonstrated the success and applicability of the guidance used. Many of them also documented lessons learned from conducting the expert elicitation and sources of biases. Such information often provides valuable insights to practical issues in implementing the guidance. Thus, we recommend that reviewing past experience is an important step before starting an expert elicitation project.

2.6 Consensus vs. Community Distribution

The cognitive basis and philosophical basis of expert judgment imply that proper use of multiple experts yields better judgment. Yet, individual judgments need to be combined to a single representation of the current state of knowledge that is agreed on by all of the participating experts. SSHAC identified four possible types of consensus:

- *Type 1.* Each expert believes in the same deterministic model or the same value for a variable or model parameter.
- *Type 2.* Each expert believes in the same probability distribution for an uncertain variable or model parameter.
- *Type 3.* All experts agree that a particular composite probability distribution represents them as a group.
- *Type 4.* All experts agree that a particular composite probability distribution represents the overall scientific community.

Although SSHAC describes these types of consensus with respect to quantity or probability judgment, they are applicable to any kind of judgment. SSHAC found that the first two types of consensus are difficult to achieve and often do not accurately represent the truth. Often groups may achieve an artificial consensus due to social pressures or fatigue, or groups may give up on reaching a judgment due to unintentional disagreements caused by semantics or confusion rather than substantive differences of opinion. SSHAC advocates that expert elicitation aims to achieve Type 4 consensus: expert judgment that represents the distribution (center, body, and, range) of the state of knowledge for the overall scientific (technical) community. With Type 4 consensus, the expert panel considers and evaluates other experts' inputs, yet they do not have to agree with each other because the final output of the elicitation is a community distribution (i.e., a distribution intended to represent the view of the technical community). Given the uncertainties in the technical problem, it is more reasonable to have a group of experts agree on how to represent the technical community's diversity of opinions about a technical issue (Type 4 consensus) than to have a group of experts agree on one specific answer to a technical issue (Type 1 consensus). Moreover, Type 4 is the appropriate input for decision-making. Decision-makers need to know what the overall technical community thinks, not what an individual expert thinks.

2.7 Summary

This chapter provides a highly condensed introduction to the technical basis for expert judgment. The concepts of human cognition, decision-making, and cognitive biases are reviewed and structured into a coherent framework to inform the formal expert elicitation process. The review demonstrates the following:

- There are potential vulnerabilities and biases (arising from social and cognitive mechanisms) associated with practical expert elicitations;
- There are tactics and strategies to limit/mitigate these vulnerabilities and biases;
- Embedding the cognitive basis of individual and team decision-making into a structured elicitation process is more important than other strategies in producing meaningful judgments and mitigating biases;
- Expert elicitation guidance such as SSHAC engages experts in a process that produces meaningful output of suitable quality to support risk-informed decisions).

The intent of this chapter is to help readers understand the basis for the expert elicitation guidance presented in Chapter 3 and implementation practices in Chapter 4. This technical basis may also assist readers with constructing a formal expert elicitation, deciding on the level of effort for implementing the process, and justify deviations or simplifications that may be necessary when implementing the expert elicitation process.

3. GUIDANCE AND INSIGHTS FOR CONDUCTING EXPERT ELICITATION

The process of using expert judgments for decision-making includes the following steps:

1) The need for an expert elicitation is identified (e.g., by a project team or decision-maker),
2) The specific technical issue(s) and elicitation process are defined,
3) The elicitation is performed,
4) The results of the elicitation are processed as necessary for use in a model or as direct input to a decision-making process,
5) The product of the elicitation (i.e., expert judgment) is used.

This chapter primarily addresses steps 1), 2), and 3). In contrast, the SSHAC guidelines and BTP report define specific applications of expert elicitation and therefore include guidance for steps 4) and 5) as well. This chapter outlines general conditions when an expert elicitation may be appropriate, basic principles to follow when defining the expert elicitation process, and practical insights and guidelines for performing the expert elicitation. Essentially, following the guidance described in this chapter helps to ensure confidence that the objectives of the expert elicitation have been met.

3.1 Conditions for Conducting Expert Elicitation

Determining whether an expert elicitation is needed typically involves a choice among a variety of alternatives. The general conditions for deciding to perform an elicitation are that:

a) There is a decision to be made or a decision support model to be developed.
    - Risk analysis studies may involve modeling actions where some inputs have well-defined bounds and acceptable values because there are requirements that must be met, or are supported by data, but other inputs may require use of expert judgment.

b) The currently available information base is sparse or uncertain.
    - Empirical data are sparse or not reasonably obtainable
    - Information may be available but there may be widely ranging views as to their applicability to the problem at hand (and these differences in views lead to a difference in results)
    - Uncertainties in the technical issues are potentially significant, and more than one conceptual model can explain, and be consistent with, the available information
    - Technical judgment is required to assess whether bounding assumptions are reasonable and appropriate, but existing conceptual models do not appropriately evaluate the problem

c) It is advantageous to perform an expert elicitation.
    - Considerations include value, cost, time limitations, availability of experts, etc.

Meeting all the three conditions supports the choice of conducting expert elicitation, yet it does not mean that expert elicitation is the only choice. In particular, there are situations where it would be ineffective or inefficient to perform an elicitation. For instance, expert elicitation may not be the best option when uncertainties in the available information base can be better addressed through analytic methods, or when there are not enough experts available to form a

panel that adequately represents the technical community. Note that expert elicitation can also be used in combination with other approaches to address the technical issue. For example, certain inputs to a decision model may developed using standard statistical methods, and expert elicitation may be reserved for inputs where data is sparse or non-existent. In all cases it is important to bear in mind that elicitation is not a substitute for data—if you can collect data to inform decision-making, then collecting data is the preferred option. Expert elicitation should be considered only in the absence of sufficient experiential/experimental data from which to develop acceptable models. In addition, expert elicitation should not be chosen solely due to cost considerations. Done properly, the elicitation will take significant time and resources.

3.2 Basic Principles When Eliciting Expert Judgment

The ultimate objective of conducting an expert elicitation is to appropriately represent the center, body, and range of the technical community's views about a technical problem. When expert judgment is used to support decision-making, the elicitation should be performed in a manner that ensures confidence in the results. As such, expert elicitation should conform to the following principles, regardless of the scale, level of effort, and the method or procedures employed for the elicitation process:

1) ***Representation of technical community*** – The purpose of expert elicitation is not to create new knowledge, but rather to obtain the center, body, and range of the views of the composite distribution of the technical community about the state of knowledge. While it is impractical to engage an entire technical community in the elicitation process, the expert panel should a) be an adequate sample of the overall technical community, b) have sufficient breadth of knowledge that it can evaluate the available data, and c) include leaders in the technical field who can capture the community's degree of consensus and diversity. The resultant expert judgment should represent the overall community's views and beliefs about the state of knowledge for the technical problem.

2) ***Independent intellectual ownership*** - While the project sponsors have legal ownership of the project deliverables, the expert panel collectively has intellectual ownership of the results. Intellectual ownership means that the expert panel takes responsibility for the robustness and defensibility of the results. To ensure intellectual ownership, all inputs to the elicitation should be shared with every expert. To maintain the independence of intellectual ownership, expert judgment must be based on the experts' knowledge and expertise, not the positions of the project sponsors or organizations the experts are associated with. The expert panel must clearly understand that they are not representing their employer or organization on the panel, but are serving in their own right as a recognized leader in their respective field. Each expert should also maintain independence from the other experts in the team in order to avoid (or mitigate) a groupthink bias risk.

3) ***Avoidance of conflicts of interest*** - To minimize bias in the elicitation, careful consideration should be given to potential conflicts of interest prior to selecting experts. Experts should be free from direct and potential conflicts of interest to the extent practical. In all cases, potential conflicts of interest or even the appearance of conflicts of interest should be disclosed up front.

4) ***Breadth of state of knowledge*** - The expert panel should evaluate a range of data and models that are representative of the overall technical community in order to obtain the range of knowledge and interpretations about the technical issue.

5) ***Interaction and integration*** - To represent the knowledge and interpretations of the technical community, experts should interact with each other as they accumulate and

evaluate existing knowledge and make interpretations. The expert panel cannot simply accumulate and evaluate inputs from the literature or elicit the judgment of one or more experts. Instead, individual experts should make their interpretations based on the integration of their own knowledge and inputs from other experts. The final results should be the integration of the individual judgments to represent the center, body, and range of the state of knowledge about the technical issue.

6) ***Structured process*** - An expert elicitation should employ a structured process to facilitate interaction and integration, and reduce biases in the outcomes.

7) ***Transparency*** - Often the results of an expert elicitation serve a range of users with different needs. To assure that the results are used appropriately, the information generated must be documented in a transparent way. Transparency includes the input data and models that were considered, the process employed, the results obtained, and the caveats and limitations of the inputs, process, and results. Transparency also helps to demonstrate the stability and integrity of the results as a whole.

3.3 Recommended Expert Elicitation Process

When formally eliciting expert judgment, a structured and systematic process should be used that encompasses all of the basic principles described above. This section describes a recommended systematic expert elicitation process that consists of ten steps across four phases. The level of effort in implementing the steps may vary with the nature of the technical issues and available resources. **Error! Reference source not found.** provides an overview of the expert elicitation process. Formal, structured expert elicitation should include all ten steps within the four phases.

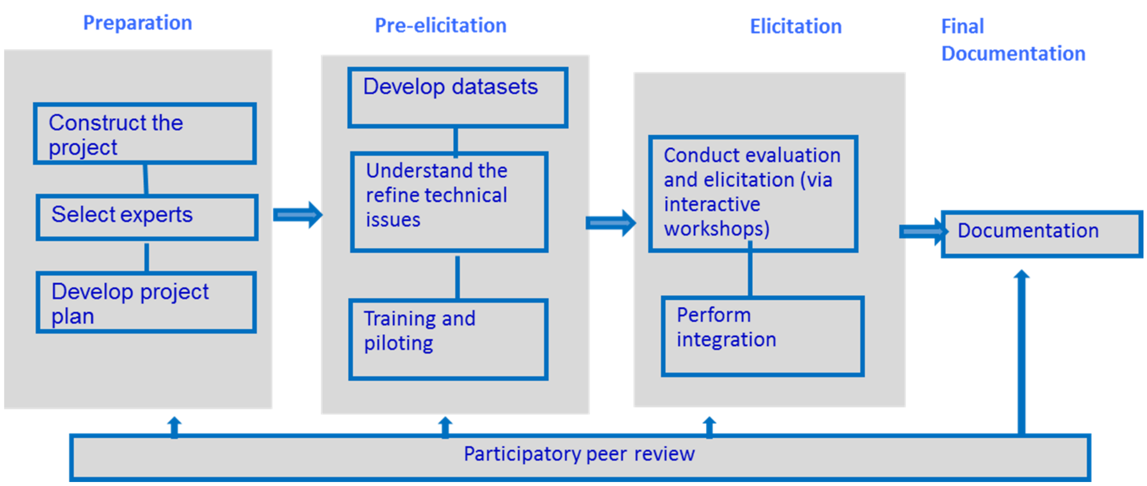

Figure 3-1. Expert elicitation process.

Phase 1: Planning and preparation. The purpose of this phase is to ensure the elicitation problem is sufficiently defined to address the regulatory application of interest; that the project team, expert panel and elicitation process are adequate to address the elicitation problem; and that the experts are provided with necessary information prior to the actual elicitation.

- Step 1. Define the expert elicitation
- Step 2. Form the expert panel
- Step 3. Develop the project plan

Phase 2: Pre-elicitation work. The purpose of this phase is to ensure that compiling the dataset is performed with the involvement of the expert pane, and that all of the team members understand the project, the technical problems, the individual's role/responsibilities, and the theories of probabilities and uncertainties.

- Step 4. Assemble and disseminate the dataset
- Step 5. Familiarize and refine the technical issues
- Step 6. Conduct training and piloting

Phase 3: Elicitation. The purpose of this phase is to elicit expert judgments through interactive workshops. The expert panel interacts to evaluate the data and models, make interpretations, form initial judgments, and integrate the judgments to represent the distribution of the views of the technical community.

- Step 7. Elicit expert judgments
- Step 8. Integrate expert judgments

Phase 4: Final documentation and sponsor review. The purpose of this phase is to develop final documentation of the process and results, and have the technical staff of the sponsor organization to review the documentation for regulatory assurance.

- Step 9. Document the process and results and conduct sponsors' technical review,

All-Phases: Participatory peer review. This is not a separate phase. Rather, the purpose of this activity is to ensure that the entire expert elicitation process is conducted with participatory peer review in all of the phases.

- Step 10. Conduct participatory peer review

The multi-step process described in the section should aid users in designing an expert elicitation that complies with the seven basic principles, regardless of the level of effort. For example:

- Selecting experts representing a range of expertise relevant to the problem from different organizations with different professional experiences while avoiding conflicts of interest to the extent practicable.
- Making all of the assembled data / models available to all of the experts.
- Providing training, as early as practical, to the experts on the potential sources of biases and how the elicitation process was being structured to minimize these biases.
- Balancing the influence of different experts by using a structured expert panel and facilitated, interactive workshops.
- Incorporating the judgments into a distribution that represents the uncertainties in the process and/or the results (The process does not require that the experts come to a consensus.)
- Using participatory peer review to monitor the elicitation to avoid significant systematic biases and enhance the breath of the knowledge on which the judgments are based. Peer review should include both the technical aspects and the process.

While Steps 1-9 are generally performed sequentially, they may be performed iteratively. For example, the experts can have different views on objectives and boundary conditions, and it's conceivable that the project might actually modify these following expert inputs in late phases.

The participatory peer review should be run through all of the phases. All of the ten process steps would be part of a procedure for an expert elicitation. If one or more of the process steps are omitted from the activity, additional information will be needed for the expert elicitation process to be considered complete.

3.3.1 Step 1. Define the Expert Elicitation

Defining the expert elicitation involves identifying the objectives and scope of the elicitation process. It is primarily performed by the sponsors and the project team, although inputs from the expert panel are useful after the panel has received training on avoiding bias and anchoring. In defining the expert elicitation, one must:

- Form the project team
- Define the technical problem
- Define boundary conditions
- Define the level of effort
- Identify and initially define the technical issues

Form the project team. The project team is responsible for constructing and managing the entire expert elicitation process. The project team should consist of individuals that fulfill the following roles (note that individuals may assume one or more roles on the project):

- Project manager - Person who manages the entire expert elicitation process.
- Representative of project sponsors - The sponsors are the organizations that have legal ownership of the results. Sponsor representatives are typically the domain experts who will apply the results to their decision-making or models. Together they are responsible for organizing, conducting, and documenting the elicitation process. They are individuals knowledgeable in one or more specific areas related to the problem and have a solid general understanding of the technical aspects of the problem. They should also *have* knowledge in areas such as probability theory, psychology, and decision analysis.

Define the technical problem(s). The technical problem of the elicitation comes from analysis or models that support decision-making activities. The technical problem should be defined explicitly and in a manner that reflects a clear understanding of how the judgments obtained will be used. The definition should then guide the entire elicitation process. In particular, the definition guides the determination of boundary conditions, the identification of technical issues of interest, the choice of experts, the information provided to them, and the form that the judgments should take. Defining the technical problem should include input from technical staff who are familiar with the specific information needs of the analysis or models for which the technical problem addresses and intended uses of the expert judgments.

Define boundary conditions. The boundary conditions of the elicitation are the basic assumptions underlying the judgments about the technical issues of interest. The boundary conditions guide the identification of technical issues, the scope of data to be evaluated, and the expert panel’s understanding of the technical issues. The boundary conditions should be clearly documented in the final report to inform users of the assumptions under which the judgments are made.

The boundary conditions may be iteratively refined with the identification of the technical issues and assembling of the data. They should be explicitly documented and made clear to the expert

panel at the end of Phase 1. They may be further modified in Phase 2 or even Phase 3. The project team should clearly document any modifications beyond Phase 1 and ensure that the expert panel is cognizant of the modifications throughout the interactive workshops in Phase 3.

Define the level of effort. The project team should estimate the level of effort to achieve the objectives. A formal expert elicitation can be expensive in cost, project management effort, and project duration. The determination of the level of effort is the balance between the technical demands and the available resources (funding, time, information overload, and project management burden). The technical demands are indicated by the following factors:

- The safety significance of the objectives and technical issues with respect to the intended use.
- The technical diversity.
- The amount of technical contention about the issues in the technical community.

Table 3-1 lists various options for conducting the elicitation process given different levels of effort. Note that the circles represent alternative choices for a component of the elicitation process and squares indicate that one or more choices are applicable. The project team should evaluate these factors against the available resources and technical demands. As a practical matter, the project team may design the process by identifying the “wants” (per the above bullets) and then trimming down the effort based on budget and actual availability of experts. There is also the consideration of the personalities of the experts (i.e., can they be constructive participants in the elicitation). The project team should recognize the pros and cons of different options; it is not always the more/bigger the better. Aside from the limitations of funding and time resources, the experts can become overloaded with too much information from the dataset and inputs from the panel, and the interactive workshops can become unmanageable with a large panel.

Table 3-1. Level of Effort Indicators for Designing the Expert Elicitation

Components of the elicitation process	**Level of effort**	**Justification**
Project team	○ A small team with individuals taking more than one role	
	○ One or more individuals fulfill each role	
Identification of technical issues	○ Via informal group meetings	
	○ Via a PIRT process	
Develop datasets	○ By the project team members	
	○ By the project team members and outside help enlisted	
	○ By additionally developing new data (e.g., through experiments) prior to conducting the elicitation	
	☐ A small panel with combined roles	

Composition of the expert panel	☐ 3-5 resource experts	
	☐ 4-6 proponent experts	
	☐ 1-2 combined technical integrators and lead integrator	
	☐ A team of technical integrators including the lead technical integrator(s)	
Expert panel's understanding of the issues, training, and piloting	○ Limited electronic/remote tutorial sessions	
	○ Face-to-face training and piloting sessions	
Workshops	○ Individual interviews and electronic/ remote meetings without face-to-face workshops	
	○ 1-2 face-to-face, structured, interactive workshops	
	○ 3 or more face-to-face structured, interactive workshops	
Peer review	○ Late-stage peer review of the full elicitation process	
	○ Participatory peer review of the full elicitation process	

Note that the original SSHAC guidelines defined four levels of elicitation with the number of participants, resources, and time requirements increasing progressively. It can be useful to assign a level of effort to an elicitation process for communicating with the technical community, sponsors, and the public. Table 3-2 is a brief description of the SSHAC levels of expert elicitation. The on-going updating of NUREG-2117 may re-define some elements in these levels.

Table 3-2. SSHAC Levels of Expert Elicitation

	Level 1	Level 2	Level 3	Level 4
Composition of expert panel	Small project team Small TI team Peer reviewers	Small project team Small TI team Peer reviewers Resource expert Proponent expert	Small project team Larger TI team with a TI lead Peer reviewers Resource expert Proponent expert Data team	Evaluator panel Technical facilitation integrator team Small project team Peer reviewers Resource expert Peer reviewers Proponent expert Data team

Interactive workshops	Limited interaction	Zero to two face-to-face workshops	At least 3 face-to-face TI team facilitated workshops	At least 3 face-to-face workshops for all participants
Peer review panel	Late stage*	Late stage*	Participatory	Participatory

* Note that in the current updating of NUREG-2117, all SSHAC Levels have participatory peer review.

Identify and define technical issues. Based on the boundary conditions, the project team may initially decompose the broad objectives of the elicitation by clearly and precisely specifying more focused technical issues. The description of a technical issue should include the problem to be answered, the expected format of the answers, and any boundary conditions or assumptions about the problem. The issues should be defined of sufficient clarity for the expert panel to be able to make the judgments in the expected form (e.g., a probability, quantity, or descriptive statements).

The initial identification of the technical issues guides the selection of experts. The selected experts in turn work with the project team to refine the issues to ensure that the issues are unambiguously defined, directly support the objectives, and can be objectively assessed by the expert panel. The technical issues are further refined in Phase 2. The project team and experts should agree on a common set of the technical issues. There may be times that some technical issues need to be revised during the interactive workshops in Phase 3. The expert panel and the project team should agree on the changes, and additional training may be needed to ensure all of the experts have a common understanding of the changes.

The project team and expert panel may identify and define the technical issues through group meetings. Since an expert elicitation process can only handle a limited amount and complexity of technical issues, a formal PIRT process may be employed to identify and prioritize the technical issues.

To make most efficient use of the experts, it is important for the project team to narrow and refine the technical problem and boundary conditions as they break the problem into technical issues. Often the panel can't (and shouldn't try) to address all potentially relevant issues. The project team should ensure the experts focus on those issues really needing an elicitation. Even in such cases, there may be too many sub-issues/conditions to address and the panel may only be able to practically address a subset. The project team has to recognize this and figure out what to do for those issues not addressed by the elicitation.

3.3.2 **Step 2. Form the expert panel**

A structured panel consists of experts assuming different roles. Experts may assume more than one role on the panel, but in all cases those roles should be clearly defined and delineated. All experts on the panel, regardless of role, should meet the following criteria:

- Possess the necessary knowledge and expertise to fulfill their roles and have demonstrated their ability to apply their knowledge and expertise
- Need to be an impartial evaluator of views other than your own to represent a broad diversity of independent opinions and approaches for addressing the technical issues
- Be willing to be identified publicly with their judgments

- Be free from direct and potential conflicts of interest to the extent practical and be willing to publicly disclose all potential conflicts of interest

- Be willing to follow the elicitation process and undergo training on avoiding cognitive and social biases

At the same time the experts are being selected for the expert panel, the project team should also choose to select the reviewers for the participatory peer review. The chosen reviewers should meet the same criteria as the experts. Once the expert panel is formed, the project team should make it clear that the expert panel collectively has intellectual ownership of the results and should make judgments independent from the organizations they are associated with. In addition, it should be made clear that the peer reviewers, although observing and reviewing the elicitation process, are independent from the process.

The project team should select the following three types of experts as the expert panel as well as participatory peer reviewers; the project team may also identify observers from the sponsor organization as needed:

- Proponent experts
- Resource experts
- Technical integrators (TIs) and Lead technical integrator (LTI)
- Participatory peer reviewers
- Observers

Proponent experts. Proponent experts are individuals who are at the forefront of a specialty relevant to the topic, and are recognized by their peers as authorities because of their sustained and significant expertise on the topic. A proponent expert often possesses extensive knowledge and experience in more than one key technical area involved in the topic. The proponent expert group, as a whole, should be balanced across all of the key technical areas needed to address the technical issue of interest.

Proponent experts serve as the primary subject matter experts who evaluate the data and make judgments. The responsibility of a proponent is to make judgments about the technical issues of interest. Each proponent is required to justify his/her assertions, to demonstrate the technical basis, and to defend the data or models used in the face of technical challenges by other experts. The proponent is also charged with making full disclosure about his/her judgments, including all underlying assumptions. A proponent may modify their initial judgments through consideration and incorporation of inputs from other experts. Thus, the proponent's final judgment should integrate the available data, their own evaluation, knowledge from resource experts, and the evaluation of other proponent experts. A part of a proponent's responsibility is to challenge but not deny other proponents' judgments. Proponents do not participate in the integration of the final results.

An individual who has another role in the project (such as resource expert or technical integrator) could adopt the role of a proponent expert at a specific moment during the project. This would require that everyone present is made very clearly aware of the switch of roles and that (in the case of an integrator) the individual is prepared to subsequently revert to the role of impartial integrator. However, as a psychological and social matter, complete switches between the roles is difficult. This is one of the downsides of having people adopt multiple roles. The project team should recognize the pros and cons of combining expert roles.

Resource experts. Resource experts provide their technical knowledge to proponent experts but do not make interpretations or judgments. Resource experts should possess a deep and broad knowledge of one or more key areas relevant to the technical issue of interest, as evidenced by years of experience working on the topic and recognition in the technical community as a subject matter expert. The main responsibility of resource experts is to share technical knowledge in an impartial way to the expert panel. They should make full disclosure including all caveats, assumptions, limitations, and identify any potential conflicts of interest. . The resource expert is expected to present their understanding of a particular data set, including how the data were obtained, or to present a model or a method with their limitations and caveats. The resource expert is also expected to respond candidly and impartially to questions posed by the expert panel. Resource experts should not develop their own models to interpret data or make judgments of the technical issues. They do not have intellectual ownership of the final results.

Technical integrators (TIs). These are the individuals that lead the entire elicitation process. In particular, they are responsible for integrating proponents' judgments to form the final results (i.e., the center, body, and range of the judgments regarding the technical issues). They work with the LTI in resolving technical disagreements or considering alternative hypothesis involved.

TIs also have responsibility for challenging the technical basis of data presented by resource experts and judgments made by proponent experts and participatory peer reviewers. They must also subject their own assessments to challenges made by the expert panel. TIs serve in the role of evaluators when the available data can be interpreted differently using diverse models and individual resource experts do not possess the experience or expertise to compare the data against the diverse models. The responsibility of the TI is to identify existing data, models, and methods as well as alternative technical interpretations and to evaluate these in terms of their general quality/reliability and their specific applicability to the technical issue of interest. The process of evaluation includes identifying the issues and the applicable data, interacting among the experts (i.e., challenging other TIs and proponent experts, interrogating resource experts), and finally considering alternative models and proponent viewpoints.

The specific responsibilities of TIs include the following:

- Work with the project team to identify suitable resource and proponent experts.
- Identify and invite the TIs and provide them clear instructions on the scope of their participation
- Work with the project team to identify and refine the technical issues
- Review and provide critiques to the project plan
- Identify areas to be included in the dataset and review the dataset to ensure that it was properly assembled
- Lead the effort of ensuring a common understanding of the technical issues among the experts and ensure an adequate level of effort
- Supervise the training and piloting to ensure that the necessary training areas are covered and the piloting runs as expected, and also lead the recovery effort if the piloting does not run as expected
- Lead the interactive workshops that evaluate the data and develop individual judgments by performing the following tasks:

- Ensure that all participants clearly understand the workshop objectives, their individual roles, the required output from the workshops, and the intended use of the results
- Facilitate the meeting to ensure that uncertainties regarding the technical issues are considered, the resource experts' inputs are assessed, and proponent experts' individual judgments are justified.
- Provoke discussions and challenges to experts' inputs

TIs should possess extensive knowledge and experience in most, if not all, of the key technical areas involved. The main attributes of a TI are the ability to objectively evaluate the views of others and expressions of uncertainty and to deeply appreciate the influences of different hypotheses, models, and parameters on the technical issue being studied. The TIs must be willing and able to make a major commitment of time and effort to the project, because they need to deeply engage in the entire elicitation process.

Lead technical integrator (LTI). A lead technical integrator (LTI) is the most important role in a formal expert elicitation. The LTI leads the entire process and is the key decision-maker in resolving controversial problems or conflicts. The LTI has the highest authority in implementing every step of the elicitation. Further, the LTI ensures that the expert panel collectively has intellectual ownership for the results and is, therefore, responsible to advocate and defend the results and credibility of the elicitation process.

Typically, the LTI serves the role of a meeting facilitator for elicitation workshops. The LTI should be knowledgeable in the conduct of a formal elicitation and capable of doing this and is capable of taking the multiple expert views and constructing an aggregated representation. Alternatively, the LTI and the project team may choose a member from the project team or TIs to share the role of the meeting facilitator.

The LTI works with the team of TIs and has the same responsibilities as a TI, in addition to the following specific responsibilities:

- Lead the integration of individual judgments into the final results and be responsible for resolving controversial opinions and conflicts
- Ensure that all members of the TI team are made aware of the potential for cognitive bias and are alerted to when biases may be influencing their assessments
- Coordinate the activities of the TI team
- Ensure that all experts have full access to all of the available data and information
- Ensure that the project documentation is complete and comprehensive

The LTI should have a very strong technical background and experience in conducting studies in the topic areas; equally important, the LTI should have a facilitator skillset including negotiation skills (which could be developed via training). The LTI should also have a good standing in the technical community. The LTI must be willing and able to make a significant commitment of time and effort to the project. Other desirable attributes of the LTI include:

- Understanding of the expert elicitation goals, the regulatory framework for the study, and how the results will be used
- Technical expertise in the issues being addressed and in the kinds of analyses that use expert judgment
- Strong communication and interpersonal skills to work with the expert panel

- Experience or knowledge in conducting formal expert elicitations

Participatory peer reviewers. Although independent from the expert panel, peer reviewers should be identified at the same time as the expert panel is formed because they will likely share many of the attributes as the experts. Participatory peer reviewers review the elicitation process and results and evaluate whether the project has met the objectives of the expert elicitation. Peer reviewers are denoted as participatory because it is expected that they are involved throughout the entire project. They interact with the project team and the experts at numerous stages throughout the project. Their review includes determining whether the project is consistent with the basic principles of expert elicitation, whether it follows a formal elicitation process, and whether the technical assessment has been adequately defended and documented. The participatory peer reviewer helps identify problems early on so they can be corrected before the project end. However, it is important to emphasize that peer reviewers must have a well-defined role and preserve their independent status throughout the project, particularly because frequent interactions with the project can lead to a loss of objectivity.

Participatory peer reviewers fulfill two parallel roles, the first being technical review. This means that the reviewer is charged with ensuring that the range of data, models, and methods have been duly considered in the assessment, and that all technical decisions are adequately justified and documented. The second role of the reviewer is process review, which means ensuring that the project conforms to the basic principles and formal process for eliciting expert judgment. Collectively, these two roles imply oversight to ensure that the elicitation is performed in accordance with the guidance in this document.

The participatory peer reviewers are responsible for providing clear and timely feedback to the LTI and project manager to ensure that any technical or process deficiencies are identified at the earliest possible stage so they can be corrected. This is often in the form of the reviewers' perspectives and advice regarding how ongoing activities can be improved or carried out more effectively. In terms of technical review, a key responsibility of the peer reviewers is to highlight any data, models or proponents that have not been considered. Beyond completeness, it is not the responsibility of the reviewers to judge the weighting of the expert judgments in detail, but rather to judge the justifications provided for the data and models included or excluded, and the rationale for the weights applied (if any) to the expert judgments.

The participatory peer reviewers should perform all or most of the following:

- Participate in and review the construction of the project
- Review the project plan
- Attend all workshops and provide timely submission of written or verbal feedback
- Participate in debriefings with project leaders at workshops
- Highlight interface issues if these are not being adequately addressed
- Directly challenge the integration of the results
- Review the draft project report
- Prepare a review report as a part of the final project report

The attributes of the participatory peer reviewers can be defined both for individuals and in collective terms for all of the members of the review group. A key requirement is that each member of the group has an understanding of and commitment to the basic principles and process of formal expert elicitation. It is desirable that the members of the review group collectively cover all technical aspects of the topic.

For products that will have regulatory applications, we recommend that appropriate technical staff from the sponsor organizations are allowed to observe the elicitation workshops. This observation will help with staff's review of the elicitation products. The project sponsors should determine if other observers are allowed to attend workshops, and what (if any) role they are given. A ground consensus on the observers' role is that they only observe the workshops and do not participate in workshop discussion to avoid organizational biases. If observers have concerns on technical matters, such as misinterpreting some technical issues or background conditions of some data, they may make comments to the project team.

3.3.3 **Step 3. Develop the project plan**

A good project plan is needed because elicitations are often complex, resource-intensive, and subject to strong schedule pressure because of their decision-support role. The project team develops the project plan that describes the project objectives and all of the programmatic and technical activities in implementing the elicitation process, with clearly defined roles and tasks for all of the project participants. A project plan is the fundamental tool for documenting and communicating the specific elements and details of the study among the participants. It is also used for the proper management and monitoring of a study to ensure that all procedural steps are followed. A project plan should include, but is not limited to the following:

- Introduction and context of the study
- Objectives of the study (i.e., a clear definition of the problem statement)
- Project organization
- Key tasks
- Workshops
- Deliverables
- Risk identification and mitigation strategies
- The need for "checkpoints" throughout the process

Introduction and context of the study. This section should include a description of the context within which the study is being carried out including the sponsors of the study, past related studies and their applicability, and any new developments leading to the need to conduct the study.

Objectives of the study. This section should include a clear definition of the problem statement, a description of the expected results of the study, and the manner in which they will be used (e.g., design criteria, risk analyses). As applicable, the deliverables of the study should be described. The regulatory framework and the manner in which the study will be used should be discussed. An expert panel might be at risk of answering the wrong question without having a clear definition of the problem statement. Therefore, it would be prudent to explicitly define this item as part of the project plan.

Project organization. This section should define and describe the key components of the project. It should include the following aspects:

- Level of effort, as specified in Table 3-3. Specifically, the plan should specify the number of interactive workshops because those are the most resource demanding and management challenging.

- The overall schedule, with the project beginning dates, the required deliverable dates, and the dates for major activities, in particular, the interactive workshops.
- The team structure – This includes the types of the members in the project team and expert team, including their roles, participating steps or scope of work, and the communicating and reporting hierarchy for the project.

Key tasks. This section should describe the key tasks that make up the elicitation process. Task descriptions may be organized according to the elicitation steps, and each step may consists of one or more tasks. The task description should include the goal, who is responsible for the performing the task, the major inputs to the task, the expected outcomes, the time and duration of task performance, the specific requirements, and a "what-if" discussion as needed. Table 3-3 is a template that can be used for documenting each of the key tasks in the elicitation process.

In additional to the task table, a timeline of the tasks can help to manage the tasks and ensure their timely completion. The timeline should provide a description of the timing and duration of all of the key tasks with annotations or graphical depictions of the relationships among the tasks (e.g., Identifying when a given activity is dependent on the completion of preceding activities, or how future activities will use the results from a given activity).

Table 3-3. Task Description Template

Step	Tasks	Description
Step 1 Define the expert elicitation	1.1. Form project team: Project Manager, technical sponsors, data specialists 1.2. Define project 1.3. Determine the level of effort 1.4. Identify technical issues /Conduct a PIRT	
Step 2 Form the expert panel	2.1. Form expert team: Resource experts, proponent experts, TIs, and participatory peer reviewers 2.2. Determine lead TI 2.3. Select the experts 2.4. Set up contracts or paperwork with experts (e.g., consideration of conflict of interest)	
Step 3 Develop the project plan	3.1. Develop project plan 3.2. Develop the initial project plan 3.3. Modify the plan based on inputs from the lead TI, peer reviewer, and other project participants 3.4. Refine workshop procedures and elicitation worksheets	
Step 4 Assemble and disseminate the dataset	4.1. Compile and analyze available data 4.2. Collect new data if applicable 4.3. Disseminate the dataset to project participants 4.4. Compile and disseminate additional data in late phases of the project	
Step 5 Familiarize the technical issues	5.1. Conduct meetings to develop a common understanding of the technical issues 5.2. Refine the issues as needed	

	5.3. Modify workshop procedures and worksheets as needed	
Step 6 Conduct training and piloting	6.1. Prepare training materials 6.2. Perform initial general training 6.3. Perform specific training on issues revealed from the initial training 6.4. Perform piloting 6.5. Modify workshop procedures and worksheets as needed based on piloting	
Step 7 Elicit judgment	7.1. Before Workshop 1: Prepare workshop agent and logistics, compile and disseminate the information package for the experts 7.2. At Workshop 1: Evaluate the available data and models at the workshop 7.3. After Workshop 1: Compile and disseminate workshop documents to all of the experts 7.4. Before Workshop 2: Prepare workshop agent and information package for the experts 7.5. At Workshop 2: Elicit individual judgments 7.6. After Workshop 2: Project team compiles the results of Workshop 2 and communicates with the expert panel for verification and possible modification.	
Step 8 Integrate judgments	8.1. Conduct TI meetings 8.2. Perform calculation for combining individual judgments and sensitivity analysis 8.3. Conduct Workshop 3 (alternatively, communicate the integrated results with the experts for their verification and feedback)	
Step 9 Document the process and results	9.1. Develop the report, all of the experts review the report, and incorporate the feedback from the expert panel 9.2. The peer reviewers and sponsors review the report 9.3. Revise and finalize the report	
Step 10 Conduct participatory peer review	10.1. Describes all planned activities for the peer review including how the reviewers will observe and review key project activities during the course of the project 10.2. Perform participatory review and provide timely feedback 10.3. Prepare written comments or a peer reviewed report.	

Workshops. The workshop description should include the following:

- Workshop outlines, including the time, duration, focus, and expected outputs
- Workshop procedures, describing the activities before, during, and after each workshop, how the activities are to be implemented, key personnel's roles and responsibilities, lines of communications among the experts, and the ground rules of interaction.
- Experts' worksheets, including the worksheets for resource experts, proponent experts, and TIs to work with and documents their assessment, judgments, reasoning, and notes for other experts and TIs to pay attention.

Deliverables. This section describes the expectations or requirements for the final deliverables. Project deliverables should be described in sufficient detail to provide confidence that the project will meet the project objectives and realistic cost and schedule estimates can be developed. This description will also provide a basis for users of the results of the study to understand exactly what they can expect the project to deliver.

Risk identification and mitigation strategies. The plan should document the project's known risks in the expert elicitation process such as yielding biased results and suffering from mental fatigue (depending on the duration of the project) along with viable and executable mitigation strategies.

The need for "checkpoints" throughout the process. Teams may be at risk of working on an expert elicitation process from beginning to end with no "checkpoints" for briefing key stakeholders on the achieved process only to find out at the end that they fell short of the expectations. Therefore, having checkpoints (perhaps via deliverables and face-to-face meetings with key stakeholders) throughout the process (particularly for long projects) would be beneficial to avoid/minimize the risk of rework at the end.

3.3.4 **Step 4. Assemble and disseminate the dataset**

The goal of this step is to provide the expert panel the most complete and up-to-date information that adequately represents available data regarding the technical issue. Identifying and compiling data is critical for experts to develop the judgments representing the range of the views of the technical community. Data compilation is initially conducted by the project team members or assigned data specialists. As the elicitation process proceeds, the expert panel may recommend additional sources of information. The dataset is augmented based on data needs identified during the workshops. Where appropriate, data should be organized in formats that facilitate the experts' use of the data to make judgments. The sources of data, the conditions under which the data was originally collected, and the caveats in the data should be documented. The process of developing the dataset should be clearly documented and avoid biases in the selection of data. The compiled dataset should be checked against the following criteria:

- Representativeness - covering the most important (e.g., high risk significance) and recently available data
- Balanced - balancing the needs from the experts in different technical areas involved in the study
- Usability - readily accessible and searchable by the experts

While it is primarily the project team's task to assemble the dataset, the task should be performed with the input from the expert panel. The knowledge and experience of the experts are an important element in assembling the data and assessing the criteria above. In fact, training the expert panel should start prior to this step for the experts to understand the technical problems. That helps the experts to identify and balance the data needs. It is also important for the expert panel to become fully cognizant that they represent the overall technical community rather than the organizations they are associated with. Thus, experts' inputs to data assembling are not biased by their own experience and their organizations' benefit.

Although having complete and up-to-date data is desirable, data collection, analysis, and data generation comes at a cost. Aiming for completeness and currency may lead to inordinate efforts on the data collection side that take away time/resources from the elicitation. There has to be a balance between the completeness of data and the resource for assembling data as well

as for experts' capacity limits of processing the data. Also, risk significance can be used to prioritize data; some data may be excluded based on low risk significance.

3.3.5 Step 5. Familiarize and refine technical issues

It is essential for the experts to have a clear, precise, and thorough understanding of the technical issues and boundary conditions. The complexity and uncertainties in the questions being studied may lead to different understandings of the issues among the experts. The project team and LTI should ensure that all of the expert have the same understanding of the technical issues, including:

- The problems asked
- The intended use of the results
- The assumptions and boundary conditions of the issue

The project team and TIs should interact with the experts to achieve a common understanding of the technical issues. Strategies such as probing or feedback may be used to ensure the common understanding. The experts are encouraged to ask "what-if" questions on the technical issues. The experts may challenge the information provided, seek clarification, or more specification of the issues. These interactions among the project team and the expert panel members may lead to refining the definition of the technical issues.

3.3.6 Step 6. Conduct training and piloting

The expert panel should be provided training to:

1) Familiarize them with the subject matter (including the necessary background information on why the elicitation is being performed and how the results will be used) and the technical problems being asked
2) Familiarize them with the basic principles of elicitation and the elicitation process
3) Educate them on possible biases that could be present and influence the judgments (including introductions to de-biasing strategies)
4) Familiarize them with workshop purposes, procedures, worksheets, and good practices for workshops
5) Provide practice in formally articulating their judgments as well as explicitly identifying their associated assumptions, rationale, and factors contributing to uncertainties

In addition, if needed, experts may need tutorial training on both uncertainty and probability encoding and the expression of their judgments with probability. They should be provided with several exercises on calibrating probability distributions, especially on estimating very low probabilities.

Training on cognitive bias should be in sufficient detail to heighten awareness and overcome cognitive heuristics and vulnerabilities. It should include real examples of the various types of biases, exercises demonstrating biases, having the experts practice debiasing strategies, and using experts' feedback to reinforce the training.

Training and piloting are placed in Step 6 to emphasize that they are essential before the elicitation workshops. In practice, multiple training sessions should be conducted at different phases of the elicitation. For example, the expert panel should be given training on item 1-3) prior to assembling the dataset or as soon as the panel starts to work on the project. This helps to reduce biases from the very beginning of the elicitation. This is important because the biases

introduced in the early steps (e.g., assembling data) propagate to later steps. Without this training, the elicitation runs the risk of expert anchoring before even beginning the workshops.

A pilot study is recommended following the training. Piloting is a test before introducing the project in its full scope. It is a small-scale preliminary study conducted in order to evaluate feasibility, time, cost, and potentially adverse events and improve upon the study's design prior to performance of a full-scale project. The expert panel, at least a subset of it, should pilot the elicitation of one or more technical issues using the designed elicitation procedures. The piloting may lead the project team to refine the technical issues, elicitation procedure, or the elicitation worksheet. The piloting may also reveal areas that need further training.

3.3.7 **Step 7. Elicit judgment**

Workshops are the critical platform for key interactions to occur among the expert panel. This step should include two workshops:

- Workshop 1 for evaluating the available data and models
- Workshop 2 for eliciting individual judgments

A third workshop is discussed in Step 8 as part of the integration of the expert judgments. Each workshop serves a specific function in the elicitation process, and the interactions among the experts evolve as the process moves from evaluation to elicitation and, finally, to integration.

Workshop 1. The first workshop is focused on evaluating the data and models relevant to the technical issue. This workshop should seek to elicit experts' experience, knowledge, and interpretations of the technical issue, while emphasizing the uncertainties, limitations, and caveats in the data and models. The inputs to Workshop 1 include the complete dataset and resource experts' presentations of the data and models. The outputs of Workshop 1 include one or more of the following items:

- The technical community's understanding and interpretations of the available data/models
- Identification of the available data that will be needed to address the issues (identifying the data and information that should be made part of the project dataset, or additional information that is needed to understand the attributes of the available datasets such as the precision, drawbacks, assumptions, etc., to the extent possible.)
- Modification or refinement of the technical issues
- Identification of data gaps and significant issues where the dataset should be expanded
- Insights (if any) on how long the resulting expert judgment will be valid given the current and perceivable status of data and the conditions for updating results.

The workshop should begin with a clear definition of the goals of the workshop, an explanation for the process that will be followed, and a definition of the roles of all those who attend. The following are the key roles of workshop participants:

- The resource experts present the data and models.
- The proponent experts and TIs question and challenge the data/models for the purpose of understanding the attributes of data/models (note that the proponent experts should not begin to make and present their judgments of the technical issues at Workshop 1).
- The TIs in their evaluator role assess, debate, or defend the data/models.

- The LTI facilitates the discussion, fosters experts' challenging of the data/models, and resolves technical disagreements. The LTI also ensures that the interactions and discussion address the technical issues, the workshop procedures are followed, the experts use the worksheets as intended, and the ground rules of interaction are consistently enforced.

- The project sponsors and participatory peer reviewers observe the process and provide their comments to the LTI and project team. A good practice is that that a short period of time be set aside at the end of each day for participatory peer reviewers to make statements or to pose questions.

The project team and LTI determine the ground rules of interaction as a part of the workshop procedures. The general ground rules for Workshop 1 should include the following:

1) All of the experts should attend the entire workshop session.
2) All presentations should occur in an equitable manner.
3) If some experts serve more than one role in the project (e.g., an expert may serve as a resource expert and proponent expert), their particular roles during the workshop should be clearly specified to all the workshop participants.
4) The workshop should focus on presentation, understanding, and assessment of data and models, and should not move into discussions of interpreting data or making judgments.
5) All of the workshops should have consensus on the role of observers (especially for NRC staff observations for elicitations that likely will be used in regulatory products).

Workshop 2. The second workshop is for proponent experts to make individual judgments of the technical issues based on the inputs from Workshop 1. The outputs may include one or more of the following:

- Individual experts' judgments about the technical issues as documented on the elicitation worksheets, along with experts' justifications, reasoning, consideration of uncertainty factors, and/or the process that led to the judgments

- Justifications and selection, modification, or development of the model(s) on which the judgments are made

- Identification of significant issues that may lead to refinement of the technical issues or need for more research to make a judgment

The workshop should begin with a clear definition of the goals of the workshop, an explanation for the process that will be followed, and a definition of the roles of all those who attend. The following are the key roles of workshop participants:

- The LTI presents the workshop objectives and technical issues, facilitates consideration and discussion of uncertainties regarding the technical issues, fosters experts' challenging of others' judgments, and resolves technical disagreement. The LTI also ensures that the workshop procedures are followed, the experts use the worksheets as intended, and the ground rules of interaction are consistently enforced.

- The proponent experts present and defend their initial judgments, and may revise their judgments based on discussions. They also challenge other proponent experts' judgments.

- The resource experts may challenge the basis of proponent experts' judgments and, at the request of the proponent experts and TIs, provide information about the attributes of the data/models used by the proponent experts.
- The TIs assist the LTI in facilitating discussion on uncertainties regarding the technical issues and foster challenges and discussions of proponent experts' judgments and the basis for their judgments.
- The project sponsors and peer reviewers observe the process and provide their comments to the LTI and project team. A good practice is that a short period of time be set aside at the end of each day for sponsors and participatory peer reviewers to make statements or to pose questions.

The project team and LTI determine the ground rules of interaction as a part of the workshop procedures. The general ground rules for Workshop 2 should include the following:

1) The intellectual ownership of the workshop outcomes should be clarified.
2) All of the proponent experts and TIs should attend the entire workshop sessions.
3) All proponent experts should present and be queried in a uniform manner and asked to provide specific answers to questions about the issues considered and the reasoning behind their responses.
4) If some experts serve more than one role in the project (e.g., an expert may serve as a resource expert and proponent expert), their particular roles during the workshop should be clearly specified to all of the workshop participants.
5) The workshop should focus on elicitation of experts' judgments and should not move into integration of judgments. Every proponent expert should have the opportunity to discuss his/her views without the pressure of reaching consensus with other experts' judgments.

Conducting Workshop 2 should conform to the following practices:

- At the beginning of the workshop, the LTI demonstrates how to derive the judgments in the expected format. This reinforces a common mental model for how the judgments should be made by the experts.
- All key definitions and assumptions about the technical issues should be reviewed.
- It should be emphasized throughout the workshop the importance of active listening and having the experts defend other points of view.
- During and after the workshop, each proponent expert should be asked whether they would like to revise or clarify their judgments based on the feedback received. The rationale for any revisions should be carefully documented.

While the workshop does not aim for "equal time" for individual proponent experts' presentation, enough time should be spent to achieve the workshop objectives. The LTI / meeting facilitator should address the social aspects of "uniform manner," such as experts' feeling slighted because of perceived inequities or dominating behaviors.

After the workshop, the project team compiles the elicitation results. The results may be represented in numerical, graphical, and/or other formats. As soon as practical the project team should distribute the results to the all of the experts for their confirmation and feedback. This allows the experts and project team to verify data codification and check for encoding errors.

While the project team should seek confirmation of the results from the experts, they should guard against attempting to force consensus or influence the outcome.

3.3.8 Step 8. Integrate individual judgments

Individual judgments, collectively, produce insights about the topic being studied. Yet, the intended use of expert elicitation results for decision-making often requires integrating multiple expert judgments into a single metric. Integration may involve simply combining individual judgments or processing individual judgments through mathematical means such as smoothing, interpolation, extrapolation, or aggregation.

Integration should be accomplished in two parts:

- TI integration to develop the distribution (center, body, and range) of the views that that represents the overall technical community
- Workshop 3 to evaluate the preliminary results and use expert feedback to finalize the results

TI integration. This stage is for the TIs, facilitated by the LTI, to perform a preliminary integration the judgments into a distribution representation. The TI integration includes the following functions:

1) Deliberate and determine ways of integrating the judgments of the expert panel based on the nature of the judgments and the intended use.
2) Resolve challenging issues in integration, such as treating alternative or conflicting viewpoints or incorporating uncertainties.
3) Perform the integration to generate the distribution.
4) Perform a sensitivity analysis or sanity-check by applying the preliminarily integrated results to their intended use for decision-making. This check may provide insights to the process used for integration and demonstrate the effect, if any, that disparate views would have on the final result.

Workshop 3. The third workshop is for the TIs, proponent experts, project team, and peer reviewers to evaluate the preliminarily integrated results and use the feedback to finalize the judgments. Workshop 3 performs the following functions:

1) The TI team presents their preliminary results with particular emphasis on the manner in which alternative viewpoints and uncertainties have been incorporated.
2) The expert panel questions and probes aspects of the preliminary results to understand the manner in which the views of the larger technical community have been considered and the range of technically defensible interpretations included.
3) The expert panel discusses significant contributors to uncertainty, potential significant biases in the results, and constraints to the results.
4) The expert panel revisits their insights from Workshop 1 on how long the results will be valid and the conditions upon which the results should be updated.

The feedback generated at Workshop 3 helps to ensure that no significant issues have been overlooked and allows for a comprehensive understanding of the implications of the results and uncertainties. The TI team should use the feedback from Workshop 3 to finalize the results, including the integrated judgments, constraints of use, and implications that disparate views would have on the final results.

There are a number of methodologies that aggregate expert judgments. Each methodology has advantages and shortcomings that vary with different types of judgment. The discussion in Chapter 4 presents methodologies and discusses the advantages and cautions of using them for different types of judgment.

The integration process should meet the following requirements:

1) Transparency and traceability - Regardless of what integration methodologies are used, the individual expert's inputs must be preserved and documented. In particular, treatment of disparate views or outliers should be clearly documented. This includes documentation of the rationales for the specific aggregation techniques employed, impact of disparate views on the final consolidated judgments, and the potential effect of disparate views on the intended use of the final judgments.

2) Completeness of representation - When widely disparate judgments arise among the experts and make it difficult to develop a single, representative metric, extra effort should be taken to document thoroughly the bases for the differing judgments. Each of the significantly varying views should be provided as output of the elicitation so that it may be incorporated directly into technical analyses or used to represent the extremes in a sensitivity analysis.

3.3.9 **Step 9. Document the process and results**

The final documentation of a formal expert elicitation should indicate what was done, why, and by whom. The final report should include, but is not limited to the following:

1) The elicitation process used - The project report should provide a detailed description of how the process was conducted and an explanation of selections when there are multiple alternatives (e.g., the justification for the level of effort, combination or omission of some steps, the approach for combining individual judgments). The discussion should include the activities, workshops, participants, schedules, and organizational structure used to achieve the project deliverables. If a process other than the one described here is used, explanation is needed to demonstrate that the method complies with the basic principles of expert elicitation, or a basis should be provided for why the method does not fully comply with the basic principles.

2) The dataset used - It is important for the reader of a project report to fully understand what data were considered at the time of the study and how those data were used by the expert panel in their deliberations. This includes the data that were provided to the experts in the initial and augmented datasets. A summary of the project database and references should be included or appended to the project report. It is important to document and inventory all data that were considered in the course of the project including those data that were not used. For those data that were relied upon, it is important to also document the manner in which those data were used.

3) Explicitly state key assumptions used in the elicitation.

4) The technical issues and the resulting judgments along with the reasoning supporting these judgments. Responses to the participatory peer reviewers' comments should also be preserved as part of the elicitation process documentation.

5) Any calculations that the experts considered important in determining judgments or models used.

6) A discussion of assumptions regarding how the results will be used for decision-making, how the process used provides confidence that the objectives of the elicitation were met.

The discussion may also include the factors that might have introduced some biases to the results or impose some limitations of the use of the results. The discussion should also include the expert panel's insights (if any) on how long the results will be valid and the conditions that results should be updated.

Note that final documentation does not mean that the activity only occurs at the end of the elicitation process. In fact, documentation occurs at every step. Documenting the process, intermediate results, and special considerations throughout helps capture all of the information in the final report.

Technical staff in the sponsor organizations should review the final documentation and understand the regulatory assurance in the resulting expert judgment. The staff review may challenge the process and request additional information to be documented. The review process helps to ensure that the results are technically defensible and are appropriate used as the input to decision-making activities.

3.3.10 **Step 10. Conduct participatory peer review**

Step 10, the participatory peer review, runs throughout the entire elicitation process. The purpose of the participatory peer review is to provide full, frequent, and independent input into the process, so that concerns can be addressed and corrections made before the project is complete. The selection of the peer reviewers should occur early in the project, thus a description of peer reviewer attributes and functions is included in Step 2. Peer reviewers should provide timely review and feedback at each phase of the process or periodically following key points in the project. This should include verbal and/or written comments following all workshops and other decision points as indicated by the project team.

The draft project report should be a review product for the reviewers, and comments should be provided in writing to the project team. At the conclusion of the project and after finalization of the final report, the peer reviewers should provide their final comments. These final comments should include the reviewers' final evaluation of whether the TI team has considered the overall technical community's viewpoints and made a concerted attempt to capture the center, body, and range of technically defensible views. The comments should also address their final assessment of the elicitation process used in the project and whether or not that process meets the objectives of the expert elicitation. The reviewers' assessment should be included in the final report of the project.

4. SUPPLEMENTAL GUIDANCE FOR IMPLEMENTING THE EXPERT ELICITATION PROCESS

This chapter presents supplemental guidance regarding good practices, lessons learned, and examples for implementing a formal expert elicitation process. This chapter is organized by the ten process steps outlined in Chapter 3 with common questions that may be asked within each step. The information comes from a literature review and lessons learned from some of the NRC's past experience conducting expert elicitation. In particular, three recent NRC projects (listed below) piloted the basic principles we identified in this report and used an expert elicitation process based on the SSHAC guidelines. Some examples are directly quoted or excerpted from the final reports of the three projects.

The JACQUE-FIRE project - The Joint Assessment of Cable Damage and Quantification of Effects from Fire (JACQUE-FIRE), Volume 2: Expert Elicitation Exercise for Nuclear Power Plant Fire-Induced Electrical Circuit Failure

The objective of the project was to estimate the conditional probabilities of the hot short-induced spurious operation failure mode of control circuits and its duration, given fire-induced cable damage. The technical issues for elicitation were pre-determined by a PIRT process. The results are intended for use in fire PRA applications.

- *Technical issues*: Parametric distributions of the probability of occurrence, and duration of fire-induced circuit-failure phenomena. Distributions are developed for control circuit configurations where applicable test data are available, as well as when there are none, but expert judgment can reasonably offer quantitative estimates.
- *Elicitation process*: SSHAC enhanced Level 2 process
- *Expert panel*: 3 resource experts, 4 proponent experts, 4 TIs (including the LTI)
- *Workshops*: Workshops 1, 2, 3, and two separate meetings for the TI team on integration
- *Peer review*: 3 participatory peer reviewers (referred to as the participatory peer review panel (PPRP))

The IDHEAS project - The NRC and EPRI developed a new human reliability analysis (HRA) method, the Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal Events At-Power Applications. The method uses 14 crew failure modes (CFM) to characterize human failure events. Each CFM has a decision tree to represent the combinations of factors that influence the error probability of the CFM (e.g., workload, human-system interface, compensatory work practices, recovery potential, etc.). The objectives of the study were to 1) refine the decision trees of the CFMs, and 2) estimate the human error probabilities for every combination of factors in each CFM decision tree.

- *Technical issues:* Refinement of the CFMs and human error probabilities of the CFMs
- *Elicitation process:* SSHAC Level 2 process
- *Expert panel:* 6 resource experts, 4 proponent experts, 1 LTI
- *Workshops:* Workshops 1 and 2
- *Peer review:* 2 participatory peer reviewers

The ISLOCA project - The objective of the project was to estimate inter-system loss of coolant accident (ISLOCA) model parameters for a Level 3 PRA in a U.S. nuclear power plant. The

technical issues included 1) Large internal leak failure rate for check valves (CVs) and motor operated valves (MOVs) that isolate the residual heat removal (RHR) and safety injection (SI) systems from the reactor coolant system (RCS) during plant at-power operation, 2) Common cause failure likelihood for pairs of in-series CVs and MOVs that isolate the RHR and SI systems from the RCS during plant at-power operation, 3) Probability of failure-to-close (FTC) against differential pressure for normally-open MOVs that could potentially mitigate certain ISLOCA scenarios in the RHR and SI systems during plant at-power operation, and 4) Likelihood of external break location due to a large internal valve leak relative to a reference location inside containment.

- *Technical issues:* Estimate ISLOCA model parameters for Level 3 PRA
- *Elicitation process*: A simplified SSHAC Level 2 process that complies with the basic principles
- *Expert panel*: 6 subject matter experts serving combined roles of resource and proponent experts, a facilitation team including 1 meeting facilitator, 1 technical integrator, and 1 project manager
- *Workshops*: Individual interviews through video conferences serving the functions of Workshop 1, and one group elicitation meeting serving the functions of Workshop 2
- *Peer review:* No formal reviewer was assigned, yet one NRC staff (author of this report) thoroughly participated in the project and performed the role of participatory peer reviewer

Over the last three decades, the NRC has sponsored many expert elicitation projects using SSHAC, NUREG-1563, and other guidance documents. Also, the NRC staff has performed review of license applications that utilized expert judgments. The on-going update of NUREG-2117 incorporates many insights gained from those experience.

4.1 Define the Expert Elicitation

4.1.1 How do the levels of effort in the current process compare to the SSHAC levels?

SSHAC defines four levels of effort. Each increase in level requires more resources (funding, time, management), and presumably produces an expert judgment with higher confidence. A lower level often means reduced resources or effort in every aspect of the process (e.g., number of participants, types of experts, and number of workshops) compared to a higher level. In practice, project teams may often need to make trade-offs between the available resources and the complexity of the question. For example, if the question being studied involves several technical areas it may require a large expert panel to represent experts in all of the key technical areas, yet the time and management resources may only allow for one interactive workshop. In fact, none of the three pilot projects employed a precise SSHAC Level (partly because none of these pilot projects represented the complex seismic hazard analysis process envisioned in SSHAC). Each project employed a combination of Level 2 and Level 3 elements, which were referred to as an “expanded Level-2” or “reduced Level-3” SSHAC process.

The expert elicitation process presented in this report does not employ a classification scheme for levels of effort. Instead, it recommends that the determination of level of effort be based on the complexity and safety significance of the question at hand, while also considering the resources available (i.e., consistent with guidance in NUREG-2117 for SSHAC). For any level of effort, all of the ten steps in the elicitation process should be performed and all of the key

roles should be filled (i.e., project team, expert panel, and participatory peer reviewers). On the other hand, given the limited resources, the roles may be combined, the number of members in each role may be reduced, and the workshops may be combined into fewer face-to-face meetings while still preserving the functions of the individual workshops. The checklist in Table 3-1 can be used for planning and communicating the level of effort in the elicitation process.

4.1.2 How should the technical issues be identified and prioritized?

After defining the project objective and boundary conditions, the project team and sponsors (intended users of the expert judgment) should break down the objective into technical issues that are relatively simple and more focused to facilitate expert judgment. Although the sponsors determine the preliminary technical issues (i.e., what is needed to meet the objective of the project), the expert panel should provide input on how the technical issues are defined, both to address the objective and be able to render a judgment. Typically, the project team, sponsors, and expert panel have several iterations or meetings to decide and agree on the technical issues.

Often the project objective involves many technical issues beyond what a project can handle given the limited resources, expert availability, and experts' mental capacity to focus on the issues (people typically become too mentally fatigued after three consecutive days of focused work). In such situations the project team and sponsors have to prioritize the issues. For very complex objectives, the project team may recruit additional subject matter experts to identify and prioritize the technical issues using a formal PIRT process. The results of the PIRT may help narrow boundary conditions to limit the number and the complexity of the technical issues. The JACQUE-FIRE project used a PIRT to determine the technical issues for the expert elicitation and also the relevancy of the available data for addressing the technical issues.

4.2 Form the Expert Panel

4.2.1 What if the technical issue to be addressed by the expert panel requires a variety of areas of expertise?

Some problems require knowledge and expertise in multiple technical areas. As one example, the ISLOCA project involved areas such as PRA modeling and common cause failure, valve internal leak failure modes, MOV performance under extreme conditions, and pipe failure due to over-pressurization. Each of the experts was asked to provide his/her judgement on all of the questions, but he/she was also given the option to not answer any question that he/she felt not qualified to answer. In a few instances the expert choose not to provide a judgement due to lack of expertise, but generally the experts answered all of the questions irrespective of the expertise for which they were chosen to be on the panel. At the end of the elicitation workshop, each proponent expert was then asked to rate their level of expertise on a scale of 1 to 5 in each of the technical areas. The project team chose this approach because 1) an expert in one area may have expertise relevant to another area and so including their elicitation results in the analysis improves the final distributions, and 2) each expert is the best judge of his/her level of expertise in an area. Nevertheless, the project team and participatory peer reviewers did not consider that this was a very effective solution to the problem. The report pointed out this problem as a potential source of biasing the results. An unpiloted alternative approach discussed is only having each expert answer the questions addressing the expertise for which the expert was chosen to be on the panel. A lesson learned is that the project team should try best to enlist the experts that are balanced in their expertise yet are comfortable with all of the technical areas involved. Alternatively, the project team should narrow the technical problem so that the adequate number of qualified experts can be enlisted.

4.2.2 Can one expert serve combined roles?

The number and diversity of experts are often limited by the expert availability and budget, therefore, it is often needed to have experts serve in combined roles. For example, a proponent expert may serve as a resource expert for their expertise in a particular technical area. To avoid biases in weighting the inputs from the panel, the individual's roles at different parts of the elicitation process should be clearly specified and the experts should only perform the responsibilities defined by that role. The LTI should ensure that interaction is structured with experts performing only their defined role at any time during the workshops.

4.2.3 How should experts' conflict of interest be addressed?

Experts' being free (to the extent practical) from conflict of interest is one of the basic principles of expert elicitation. Experts who are employees of licensees or applicants, their subsidiaries, or who receive substantive support from licensees or applicants, should not be elicited unless their knowledge is unique in the technical community. NUREG/CR-5411, "Elicitation and Use of Expert Judgment in Performance Assessment for High-Level Radioactive Waste Repositories", documented the staff's recommendations on addressing conflict of interest: "It is very important to avoid any potential conflict of interest between the specialists and the results of the performance assessment. A frequent concern is whether the prospective specialists derive their employment or any income from organizations charged with conducting the overall performance assessment or with constructing the repository. Other potential conflicts may involve close working relationships with individuals involved in the performance assessment or professional viewpoints viewed as unalterable by conflicting data or reason. Each possible specialist should be asked to provide a written statement of any potential or potentially perceived conflict of interest. Those available specialists with no conflicts should be chosen based on their expertise. Individuals with a perceived or real conflict of interest may not allow this conflict to influence their professional judgments."

4.3 Develop the Project Plan

4.3.1 How should time be estimated for expert participation?

The project plan should include an estimate how much time each expert should expect to devote to participating in the project. The time estimate should account for all time preparing for the elicitation, attending pre-elicitation meetings, participating in the workshops (including travel time), and performing all follow-up or documentation activities after the workshops. These estimates may vary for each type of expert participating in the project. In reality, it typically takes longer time than estimated.

4.3.2 What is a reasonable duration for a workshop session?

A workshop session should not last longer than three consecutive days. Beyond that, the experts may experience mental fatigue. Experiments have shown that the quantity and quality of team interaction and communications are reduced with increasing levels of mental fatigue. Thus, the project team may consider the need to split a workshop into two or more sessions if the goals of the workshop cannot be achieved in three days.

4.3.3 Can the workshops be combined to reduce the cost?

The formal expert elicitation process specifies three separate workshops for different stages of expert panel interaction: Workshop 1 for evaluating available data/models, Workshop 2 for eliciting individual judgments, and Workshop 3 for review and feedback on the integration of the

final results. These three workshops are distinct from each other by their functions and structures, yet they can be physically combined into one or two meetings as long as the structure of each workshop is maintained and the basic principles are adhered to. Below are some examples of combining workshops:

- In the IDHEAS expert elicitation project, the experts needed to estimate human error probabilities of all of the decision tree paths for 14 crew failure modes (CFMs), with each CFM comprised of 8-25 paths. It was difficult for the experts to maintain their awareness of the data for evaluating all 14 CFMs. Also, the data for each CFM was independent from the data for other CFMs. The project team combined Workshop 1 and 2 for each CFM: The expert panel first conducted the Workshop 1 tasks for a CFM, then after a recess break, the panel conducted the Workshop 2 tasks for the same CFM.
- The ISLOCA project employed web-meetings and individual interviews to accomplish the functions of Workshop 1, then used one face-to-face meeting to perform the functions of Workshop 2 and Workshop 3.

On the other hand, in many studies, much of the actual work (evaluation and integration) occurs outside of the workshops in working meetings (face-face, calls, or webinars). Such cost-efficient meetings fulfill the functions of the workshops provided they are conducted following the consensus on the rules and personal responsibilities set forth for the workshops.

4.4 Assemble and Disseminate the Dataset

4.4.1 How relevant and adequate are existing data?

Part of the decision to use an expert elicitation is based on assessment of the adequacy and relevancy of available data and models. A fundamental basis for convening an expert elicitation is that existing data are limited in quantity, quality, and/or relevancy to the technical issues. Quality problems may arise from imprecision or bias in the data. Relevancy problems may arise from the underlying assumptions of the data and models.

Rarely are there direct empirical data that are specific to the technical issues under the same context and boundary conditions of interest. For example, there is practically no direct empirical data for human error probabilities of nuclear power plant events because no two events have identical contexts. As a result, assessments of human error probabilities make inferences based on lines of evidence (e.g., ATHEANA, THERP, and IDHEAS). Crawford-Brown (2001) classified five types of evidence ordered by relevance as shown in Figure 4-1. This classification may be useful in annotating the data and models in the datasets for communication within the project and final project documentation.

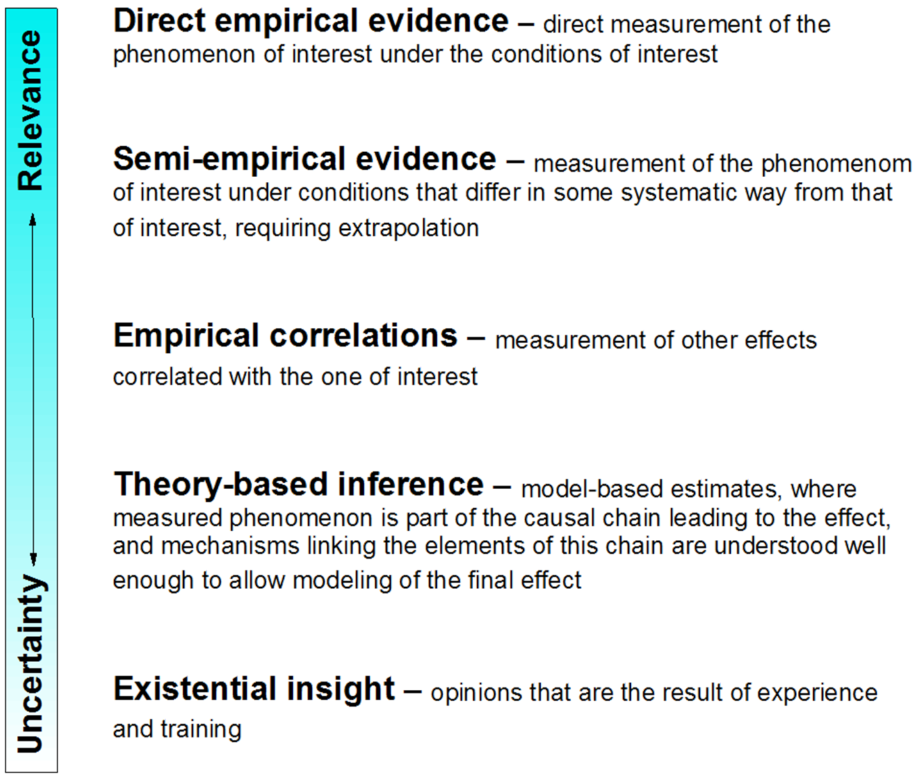

Figure 4-1. A classification of data relevancy for an expert elicitation.

4.4.2 When should the datasets be made available to the expert panel?

The project team develops the datasets that contain all of the available data/models relevant to the technical issue. The experts should thoroughly review the datasets before the start of the workshops. In fact, the initial datasets often help experts understand and refine the technical issues. Yet, the datasets should be made available to the expert panel only after an initial training session in cognitive biases. Otherwise, the project runs the risk of experts anchoring their positions based on initial examination of the data.

Lessons learned from past projects is that all of the datasets, especially the ones provided by the experts, should preferably be made available to the experts at least one month prior to the start of Workshop 1. Developing datasets can take much longer than expected. Hence, the project team may well plan the effort, allocate time and labor resources, and start the development of the datasets as early as possible. The following examples demonstrate the effort needed and various factors that may hinge on the data assembly effort.

- In the JACQUE-FIRE project, because of the complexity of the problem, the project team first conducted a PIRT to identify phenomena that can affect the fire-induced failure modes of electrical circuits after cables are damaged by fire. The PIRT panel used the results from recent fire tests to identify and rank the parameters that can influence the hot short induced failure modes of electrical control circuits. Using these influencing parameters, the results of cable-fire tests, and operating experience, the PIRT panel identified thirteen (13) circuit configurations for which the expert judgments of cable damage probabilities were elicited. The PIRT panel also identified the data sources to be used and noted the lack of data in certain areas.

The PIRT results served as the basis for the project team to select and compile the datasets for subsequent expert elicitation.

- In the ISLOCA project, important design information about the plant was not available until about one week before the individual elicitation sessions began. Consequently, this limited the time available for the experts to evaluate this additional substantial and relevant information. This additional information provided design data that was critical to the valve experts' valve leak rate assessments. The critical path for this effort was to obtain plant-specific design and/or as-built information on the systems and components being evaluated. Because obtaining the plant-specific information requires that it be formally submitted to the NRC by the owner of the NPP, the information must first be approved via the plant's internal approval processes, which can take two weeks or longer. Backing up from this, the information must be located and compiled by the plant staff, which may take weeks, depending on plant staff availability and priority of the information gathering effort. Finally, it may take several iterations between the experts and NRC staff agree on the final information request. Hence, the development of this information request should be initiated as soon as possible after contracts with the experts have been signed.

- In the IDHEAS project, the dataset was made available to the experts about two week before Workshop 1. This led to a caveat in the elicitation that the experts did not have adequate time to thoroughly review the datasets before the Workshop. The IDHEAS project had four data specialists to identify, assemble, and disseminate available human error data several months before Workshop 1. Yet, it took much longer than planned to organize the large amount of data into a database that could be readily used by the experts.

4.5 Familiarize and Refine the Technical Issues

4.5.1 How can experts achieve common understanding of the technical issues or questions being asked?

Getting all of the experts to have the same understanding of the questions asked is probably the biggest challenge in expert elicitation. The literature and lessons learned from past experience suggest some tips for achieving a common and correct understanding of the technical issues:

1) Experts should be aware of others' thinking. Klein et al. (1989) found that a common situational awareness is important to the quality of expert judgment. They proposed using situational awareness probes to make the experts' understanding explicit. In a practical manner, workshops should be conducted in a way that facilitates open dialog between both the presenters and the experts, and among the experts themselves.

2) Use state-of-art methods to ask questions. Designing questions is a science, with scientific considerations on the number of questions, internal consistency, anchoring vs. not anchoring the questions, and different types of questions. Guidance on asking questions can be found in the literature (e.g., "Standards by Educational and Psychological testing").

3) Designing questions should consider how data are going to be aggregated and if the format of the answers is viable for the aggregation to be performed.

4) If the technical issues are about underlying relations among various contextual factors, the questions asked should focus on observable phenomena. Then the underlying, unobservable relations can be inferred given the observable judgments provided by the experts. For example, if the technical issue is to get $y=a*b^2$, experts may have difficultly providing estimates of a and b; instead, the questions may ask experts, "given such a condition, what is y," and then the project team and TIs can infer a and b.

The examples below may help to inform potential caveats in understanding the technical issues:

- *Appreciating only a part of the scope of a technical issue*. In the JACQUE-FIRE project, the panel did not, at the beginning of the project, have a common understanding of what constituted a "spurious operation." This common understanding was developed after considerable discussion during the workshops. Recognizing that analogous situations might arise dynamically over the course of any project, the project team recommended that the subjects of the elicitation are clearly defined prior to engaging the expert panel.

- *Misunderstanding the boundary conditions of a technical issue*. In the IDHEAS project, the technical issues are the probabilities of the set of the pre-defined crew failure modes. The experts did not realize that every crew failure mode only addresses human failures under the specified conditions. Recognizing this, the project team developed graphic charts to illustrate the conditions for every failure mode and the boundaries between the failure modes. The LTI emphasized the specific conditions at the beginning of the elicitation for each new crew failure mode. The LTI also reinforced the experts' understanding by demonstrating how the individual judgments would be integrated and eventually used in human reliability analysis.

- *Different understanding or use of terminology defining the technical issues.* In the ISLOCA project, proponent experts made their initial judgments through individual interviews. One technical issue was to assess the probabilities of common cause failures of two valves. As the experts presented their judgments at the interactive workshop, it was clear that the experts and project sponsor had several different interpretations of what common cause failure was. After substantial discussions, the LTI requested that the experts use the project sponsor's interpretation because that was how the judgments would be used. This problem was not revealed earlier in group meetings, training, or individual interviews, because common cause failure is a term widely used in probabilistic risk assessment (PRA).

- *Misunderstanding the question asked*. In the IDHEAS project, the question asked was "If the condition X happens, what is the likelihood that X leads to the consequence Y?" Yet, several experts interpreted the question as "What is the frequency of X happening?" This misunderstanding happened because some current practices in human reliability analysis use the frequency X as an indication for the frequency of Y. The project team clarified the confusion at the beginning of Workshop 1, and the LTI emphasized it throughout Workshop 1 and 2. Yet, one proponent expert still misunderstood the question and his/her judgment was for the wrong interpretation of the question. As a result, the LTI did not include this expert's judgment in the integration.

- *Misunderstanding the questions due to complex mathematical models involved*. In a number of cases of the ISLOCA project, experts first misunderstood the questions because the elicitation process required estimates for a parameter that was not

directly elicited. Instead, the parameter was represented as the output of a mathematical model whose parameters were the ones actually elicited.

- *Misinterpretation of worksheets*. The project team designs worksheets to guide experts' elicitation. In one case in the ISLOCA project, there was confusion amongst the experts as to what was specifically being elicited in the worksheet (i.e., whether valve failure rates corresponded to leak sizes (equivalent diameter in inches) or to leakage rates (gpm), because each of these values were shown on different columns in the datasheets). Exacerbating the issue was that the leakage rate was not correct for the leak size for the intended use of the judgments. The lesson learned from this case is that the worksheets should be better described and the parameters being elicited better defined prior to the elicitation sessions. While the datasheets were tested by the project team, the discrepancy in the two leakage rate columns was not identified. This problem can be addressed by doing the following: (1) providing a definition/description of each field in the worksheets, (2) having the sponsor's subject matter experts perform a quality review of the worksheets via a simulated elicitation session, and (3) reviewing the worksheets with the expert panel members in a simulated group elicitation session.

4.6 <u>Conduct Training and Piloting</u>

4.6.1 **Can experts be trained to avoid cognitive biases?**

Cognitive biases have been extensively studied, yet very few training strategies have been shown to be effective at reducing or avoiding cognitive biases. In fact, because of human limitations in retaining and processing the full span of available information when making judgments, there is simply no way to eliminate cognitive biases. The best one can hope to do is to work diligently to minimize its influence through a structured process which includes sensitivity to this issue by the participatory peer reviewers. One of the goals of the training session held before the workshops is to develop awareness of natural cognitive biases that can skew judgments. This appears achievable by simply providing presentations on common cognitive biases. However, the challenge is to have the expert maintain such awareness throughout the process of assessing data and making judgments. Good practices include (1) using vivid, graphic examples demonstrating the biases, and (2) having the experts perform exercises with known biases. Examples and exercises can be found in the literature (Burgman et al 2006, Kirkebøen 2009, NUREG/CR-5424).

4.6.2 **What should be included in training about anchoring bias?**

The human judgment process, just like any decision-making process, includes fast heuristics and detailed analytic iteration. While a structured elicitation process enforces the detailed analysis, experts still use heuristics to simplify the judgment process. One of the most common heuristics used is "anchoring and adjustment" through:

- Identifying a similar situation for which an estimate was already available, and then adjusting that estimate (the "anchor") to account for differences between the case at hand and the anchoring case, or
- Identifying a specific case or example of the technical issue, making the estimate for that case, then adjusting the estimates to the broad scope of the technical issue.

Anchoring and adjustment heuristic is a potentially important source of bias; the anchor may be incorrectly selected, the adjustment may be insufficient, or both. However, when experts appear to have little data to rely on, anchoring is unavoidable. In the IDHEAS project, for

every technical issue (i.e., the human error probability of a crew failure mode), the experts first worked on coming up with one or more examples from their experience, then they had deep discussions on the examples, and then they started to make estimates. Hence, their judgments were anchored to the examples. Since anchoring is unavoidable, training should emphasize the implications of different adjustment strategies (e.g., using absolute values versus relative values or percentages, or such approaches as “splitting the difference” between bounding estimates) and warn the experts to avoid significant systematic biases when they have to use anchoring and adjustments.

4.6.3 **What are some strategies for training experts to estimate probabilities?**

The biggest challenge in training is to have the experts represent their state of knowledge in terms of a probability distribution. Often the experts are engineers who may be used to thinking about and dealing with technical issues in a deterministic manner. Here are some tips from past experience:

Let experts know that uncertainty is to be expected. The expert elicitation process characterizes epistemic (state of knowledge) uncertainties using probability distributions. Training should convey the following concepts to experts:

- In a deterministic world, specific conditions are important influences to behavior. Yet, those conditions do not completely dictate behavior. That is where uncertainty resides. The training should remind the experts that even in situations where empirical data are available, their expert judgment is needed in assessing the relevance of that data to the practical situations of interests, because the data may not result in appropriate representations of parameter central tendencies, let alone uncertainties.
- Variations in unspecified conditions lead to variability in behavior, which we cannot predict deterministically. We choose to model this variability (aleatory uncertainty) using probability distributions.

In the JACQUE-FIRE projects, some technical issues had relatively ample data while others have very limited data, weakly relevant to the issue or poorly observed. The panel spent a considerable amount of time discussing situations for which there were relatively ample data yet less amount of time on situations with limited data. The training for such projects should address the sensitivity of parameter estimates to changes in datasets. For example, for a given sample size, what kinds of changes are needed to make a large difference in the estimate? Such training can also help avoid expert overcompensation for overconfidence biases identified in current training.

Help experts to become comfortable with probability. Probabilities measure “degrees of belief.” Experts may be uncomfortable with this concept because they have been used to working with objective data. Training should use examples to demonstrate that the degree of belief can represent the center, body, and range of the truth, and ensure that probabilities are consistent with the laws of probability (i.e., “coherence”).

Here is an example of experts’ reluctance to express judgment with probabilities from the JACQUE-FIRE project, “Initially, a number of PRA panel members were reluctant to exercise their judgment for cases for which test data were lacking. Some appeared to question the premise of the elicitation process, dismissing their own expertise (e.g., via statements such as “I have no basis to quantify”) even when they had opinions regarding phenomenology and expected trends. Over the course of the project, most of the experts appeared to become more comfortable with the process. Nevertheless, as shown in the main report, the panel

employed various strategies (e.g., grouping of cases, excluding cases) to avoid providing estimates for a number of data-poor situations. Recognizing that the experts' reluctance to provide estimates based on holistic judgments (rather than discrete, easily identified sets of test data or mathematical model predictions) may be difficult to change, the training portion of future elicitation projects should nevertheless spend more time addressing panelist qualms and empowerment."

Calibrate intuitive notions of likelihood (and unlikelihood) with observable events. Calibration is a measure of the performance of subjective probabilistic judgments. According to Lichtenstein et al. (1982), a judgment is well calibrated "if, over the long run, for all propositions assigned a given probability, the proportion that is true equals the probability that is assigned. Judges' calibration can be empirically evaluated by observing their probability assessments, verifying the associated propositions, and then observing the proportion that is true in each response category." Training should include one or several calibration exercises for experts to practice quantifying their beliefs regarding event probabilities. Past experience has shown that this was a very useful part of the training. Kahneman's (2011) book, "Thinking, Fast and Slow," provides examples from very poorly calibrated results (e.g., clinical diagnosis of pneumonia) to well-calibrated results (e.g., probabilistic precipitation judgments by US weather forecasters).

One project did not have such an exercise in the initial training. At the end of the first day of Workshop 2, it was apparent that some proponent experts were still uncomfortable developing probability distributions. The expert panel was given a calibration exercise at the beginning of the next day and consequently the panel got much better in discussing and expressing probability distributions. Also noted was the relevance of the exercise to the problem at hand. For example, one exercise in the JACQUE-FIRE project was to predict the results for a number of events in the 2012 Summer Olympic Games. This exercise was not particularly analogous to the problem faced by the project since it involved the outcome of a one-time event (as opposed to the outcome of a repeated series of trials). The exercise was not performed by a number of panelists because they did not feel it was relevant or useful.

Help experts understand the context and meaning of very low probabilities. Often, probability training tends to focus on normal (observable) probability levels, not on extremely unlikely events. People typically do not understand what "teeny-tiny numbers" mean. Nevertheless, PRAs typically consider very low probability events, and many models are sensitive to the tails of parameter distributions. In such situations, the training needs to broaden the statistics and look at observable examples of low probabilities (e.g. various health problems in the U.S. population).

The ISLOCA project report noted, "a significant topic of discussion during the workshop was the challenge associated with understanding the relative meaning of very low probabilities or numbers, such as the large internal leakage failure rate for valves that are generally 1E-08 failures per hour or even smaller. While the PRA experts were quite familiar with very low probabilities because of their extensive use in PRA models, most of the experts were not PRA specialists and did not generally work with very low probability events. To provide some context and training for the meaning of low probabilities, the workshop facilitator used a graphic chart on the lifetime odds of death for various causes in the United States to demonstrate the meaning of very low probabilities. While the experts agreed the graphic and ensuing discussion provided clarity about the meaning of very low probability events and was useful in the development of their elicitation input, a lesson learned is that this topic needs more emphasis during the expert training session conducted prior to the elicitation sessions. Furthermore, while the training did include a specific exercise to elicit probabilities on issues that have known answers based on scientific studies, the probabilities elicited were very large in comparison to the very small internal leakage failure rates for valves in nuclear power plants.

Conducting training exercises that are more representative of the magnitude of the probabilities being elicited is a lesson learned."

The IDHEAS project did not provide specific training on very low probabilities. Yet, the questions asked involve estimating very low probabilities (typically less than 10E-3). During the first session of Workshop 2, while the LTI explained and fostered discussion on the context and meaning of very low probabilities for the given questions, most experts were still uncomfortable with their estimates of very low probabilities (e.g., Some experts said, "I am not sure if it has any meaning," or "I tried my best to put some numbers there but I had no idea what they actually meant"). Some experts did not perform estimates for cases where they expected the probabilities to be very low, but not zero. Before the second session of Workshop 2, the LTI trained the experts on calibrating probabilities with a specific focus on very low probabilities. It turned out that the experts were able to articulate their thinking process for very low probabilities and made more consistent estimates on very low probability cases.

4.7 Elicit Judgments

4.7.1 Why are the workshops important?

The essential parts of an expert elicitation process are achieved in the activities of Workshop 1 to evaluate the data related to the technical issue and the activities of Workshop 2 to elicit the expert judgments. Interaction and facilitation are key to the effectiveness of both workshops. Both the SSHAC guidelines and the process outlined in this report recommend that the workshops be conducted as face-to-face meetings. If the level of effort for the elicitation must be adjusted, then the workshops may be combined into one face-to-face meeting, or the workshops may be conducted via video conferencing. However, in all cases, it is critical to facilitate interactions among the experts.

4.7.2 Why are facilitated interactive workshops preferred over dynamic workshops?

The workshops should be facilitated by the LTI acting as the facilitator or the assigned workshop technical facilitator-integrator(s) (as defined in the SSHAC guidelines). The workshop should not be conducted dynamically without structured facilitation. The problem with dynamic interaction is that the interaction may end up with endless debating without converging to achieve the goals of the elicitation. A lesson learned from the IDHEAS project is that the goal should be made explicit and emphasized throughout the interaction by the LTI/facilitator.

4.7.3 Can one person fulfill all of the lead technical integrator's (LTI's) responsibilities?

Past elicitation projects using the SSHAC process demonstrated that the LTI/facilitator is critical to the success of the elicitation workshops. The LTI's functions at workshops include running the meeting, facilitating discussion, and resolving technical and procedural (e.g., deviations from the guidelines) disagreements. Ideally, the LTI should possess expertise in all of the technical areas involved as well as expertise in utilizing expert judgment. Hence, using facilitators in additional to the LTI to share the responsibilities may help the LTI focus on the most important parts of his/her role.

4.7.4 Why is challenging and defending important during the workshops?

Workshop interaction is achieved through presenting one's evaluation or judgment, challenging other experts' results, and defending one's own results. Effective interaction depends on experts' attributes and the effectiveness of the workshop facilitators/LTI. Also, the depth of interaction decreases as the experts experience more mental fatigue in later parts of a day and later days of the workshop. To prompt challenging and defending, workshop preparation may include developing procedures and templates for experts to write down their thinking process. For example, if an expert lists his assumptions, factors, data sources, uncertainty/biases in the data, calibration of the data, and rules used to come up with the probability (and duration), then when the expert hears another expert presenting their assumptions it will be easier to identify differences and determine whether revision to the judgment is needed. This technique was exercised and considered useful in the ISLOCA and IDHEAS projects.

4.7.5 What cautions should be observed when screening out data?

The purpose of Workshop 1 is for experts to assess and evaluate the available data/models. As a part of the outcome, some data/models may be considered as irrelevant or not important to the technical issues and thereby ruled out. When eliminating data/models, experts need to carefully consider the broader implications of removing the data from further consideration. For example, the peer reviewers in the JACQUE-FIRE project noted, "Another simplification employed by the expert panel involved the pooling of hot short duration data. This simplification, which reduces the numbers of parameters to be estimated, was facilitated through the use of statistical tests determining whether it is appropriate to model a set of data as coming from the same underlying population. These same tests identified statistical outliers (i.e., data that don't appear to belong to the population), which were explicitly discussed and dispositioned by the panel. In general, the PPRP notes that empirical data can reveal realistic behaviors not included in models and that, therefore, data outliers should not be discarded based solely on statistical arguments. Similar to this project, future elicitation projects should ensure that any data screening be done with caution, that any screening have a sound phenomenological or practical applications basis, and that the basis be carefully documented."

4.7.6 What is subjective probability?

In daily life, people often use qualitative words such as "likely" and "unlikely" to describe the likelihood (i.e., the degree of uncertainty) of an event. However, without some quantification, the use of qualitative words to describe uncertainty can mask important, often critical, differences between the views of different experts. The problem arises because the same words can mean very different things to different people, as well as different things to the same person in different contexts. In the engineering and scientific world, probability is used to quantify likelihood.

Expert judgment often involves the estimates of the probabilities of certain events. The concept of probability is usually connected with long-run relative frequency. When the frequency is unobservable for rare events, engineering judgment of the frequency can be expressed using subjective Bayesian probability. Kirkeboen (2009) noted, "A subjectivist or Bayesian interpretation of probability is used when one makes subjective probabilistic assessments of the present or future value of uncertain quantities, the state of the world, or the nature of the processes that govern the world. In such situations, probability is viewed as a statement of an individual's belief, informed by formal and informal evidence that he or she has available." Hence, subjective probability is a measure of the degree of belief given the state of imperfect knowledge. Such a measure is often desired for analysis or models regarding rare

but highly significant events. The need to quantify beliefs of risk regarding nuclear power plants led to the foundation of probabilistic risk assessment (WASH-1400, 1975).

4.7.7 What strategies are useful to express likelihood in probabilities?

Experts who has been using deterministic methods in their work tend to have difficulties express their beliefs in subjective probabilities. One useful strategy is to begin with the extremes of the probability distribution and gradually approach the median or best estimate (e.g., ask questions such as, "How bad can it be? What are the reasons?" and then use similar questions for the "best case"). The scientific literature strongly supports this approach to probability distribution as better than starting from the center of the distribution. This strategy is particular useful in reducing anchoring biases.

Kahneman (1974) describes this strategy as the following, "When presented with an estimation task, if people start with a first value (i.e., an anchor) and then adjust up and down from that value, they typically do not adjust sufficiently." Kahneman and Tversky (1982) call this second heuristic "anchoring and adjustment." To minimize the influence of this heuristic when eliciting probability distributions, it is standard procedure, as decribed by Morgan (2014), not to begin with questions that ask about "best" or most probable values but rather to first ask about extremes: "What is the highest (lowest) value you can imagine for coefficient X?" or "Please give me a value for coefficient X for which you think there is only one chance in 100 that actual value could be larger (smaller)." Having obtained an estimate of an upper (lower) bound, it is then standard practice to ask the expert to imagine that the uncertainty about the coefficient's value has been resolved and the actual value has turned out to be 10% or 15% larger (smaller) than the bound they offered. We then ask the expert, "Can you offer an explanation of how that might be possible?" Sometimes experts can offer a perfectly plausible physical explanation, at which point we ask them to revise their bound. After obtaining estimates of upper and lower bounds on the value of a coefficient of interest, we then go on to elicit intermediate values across the probability distribution (e.g., "What is the probability that the value of X is greater (less) than Y?"). If the results seem to be unduly scattered, changing the question format may help (e.g., "Give me a value of X such that the odds that the true value is greater (less) than 1 in Z (or probability P)").

4.7.8 How should time be managed during the workshops?

Often it takes longer than planned to work on individual technical issues, therefore, there is not enough time to go through all data/models at Workshop 1, or all of the planned technical issues at Workshop 2. A remedy to running out of time is to have more sessions of the workshop. This is often limited by the funding resource and experts' availability. Another option is for the experts to take the uncompleted issues as homework, but this may affect the quality of the judgment due to lack of face-to-face interaction. Thus, the project team should carefully plan the time for workshops and develop remedy strategies ahead of time.

An example of time management is noted in the peer reviewers' report for the JACQUE-FIRE project, "Although the PRA (the proponent experts) panel was charged with estimating the likelihood of both spurious operation occurrences and durations, the panel spent most of its time on the former (the treatment of occurrences). Partly for this reason, while agreeing with the PIRT panel that the duration of a spurious operation is dependent on a number of scenario-specific factors, the PRA panel chose to develop only two distributions for spurious operation duration: one for AC circuits and one for DC circuits. Moreover, the parameters characterizing these distributions were treated as single values (estimated by averaging inputs from proponents), rather than as distributed quantities. The PPRP expects that the stated uncertainties in the early (pre-"floor") portion of the spurious operation duration distribution likely

under-represent the range of uncertainty of the technical community. We note that as a matter of general principles, further efforts to characterize uncertainty might shift the central tendency of the distribution as well as increase its breadth."

4.7.9 How should mental fatigue be managed during the workshops?

Each technical issue likely involves many dimensions of information for experts to assess, compare, relate, and integrate. These activities are very cognitively demanding and, therefore, the experts may quickly experience mental fatigue. Moreover, the time-constrained thinking process aggregates mental fatigue. Experts receive overwhelming amounts of information from other experts at the workshop and their information processing is most likely time-constrained, thus they have to speed up their cognitive activities. The following tips were based on the cognitive science literature and were demonstrated as good practices in the pilot projects:

- Use pre-organized hard-memory materials as memory aids to relieve experts' mental fatigue. For example, formatting experts' write-ups in a consistent framework, organizing data to make them readily accessible, giving experts plenty of white paper for them to take notes.
- Cool-down rule: After one expert makes their presentation, there should be a 2-3 minute quiet time for people to think, without speaking/asking immediately. Similarly, experts are encouraged to write down differences between their judgment and the presenter's judgment (assumption, factors, rules, probabilities, etc.), without trying to assess or immediately modify their own judgment.

4.7.10 How should social pressure be managed during the elicitation process?

Peer pressure. Some experts are more experienced or more persuasive than others, thus they can influence other experts' thinking process in a persuasive way rather than through technical challenges to the thinking process. While it is the LTI's job to manage situations like this, a good practice is for the experts to take turns presenting their judgments. Consequently, every expert gets an opportunity to present their views, and gets chances to ask questions and add information. This strategy was observed as extremely useful in relieving the peer-pressure bias and prevent the expert panel from engaging in off-track talks. Another good practice is for the expert panel to review the basic principles of expert elicitation and become cognizant that the final result should represent the distribution of the views of the technical community.

Social Pressure from Observers. Project sponsors often choose to observe the workshop meetings. The general guidance to observers is that since they are not a member of the expert panel they should not interfere with the discussions between panel members or to try to influence the judgement of the experts. However, observers may become actively involved in certain technical discussion topics, sometime to the extent that observers' opinion may dominate the discussion. One example in the ISLOCA project was the NRC project sponsors' involvement in the discussions of the applicability of common cause failure assessments to the valve internal large leak failure mode. Because the experts had different understandings of the meaning of common cause failure, the sponsors wanted to make sure that the experts assessed common cause failure in the context of the intended use of the results in the NRC's Level 3 PRA model. While it was acceptable for observers to comment on the technical aspects of the topics being discussed when clarity may be enhanced, the experts experienced pressure from the observers. The SSHAC expert elicitation process (NUREG-2117) provides two mechanisms for accomplishing this:

1) The NRC SMEs attend the elicitation sessions strictly as observers and should be precluded from the technical discussions. However, the facilitator should "allot some time at a specified time at the end of each day or each workshop to open the floor to questions and comments from observers. In such a context, the regulator could provide feedback, raise concerns, or ask for points of clarification."
2) A sponsor subject matter expert may be included on the expert panel as a resource expert. The role of a resource expert is to present data, models and methods in an impartial manner. The resource expert does not participate in the elicitation of data/models or attempt to influence the other experts' judgements.

4.7.11 How should deviations from the workshop procedures be addressed?

The JACQUE-FIRE project had an instance in which some experts were not comfortable with the format of the worksheet and requested to choose alternative formats. The report noted: "although the technical integration team requested that the proponents provide their estimates in the form of a set of pre-designated quantiles, some of the proponents had alternate, preferred ways of expressing their uncertainties and didn't always comply with the TI requests. Over the course of the project, this difference led to some angst on the part of some TI team members. The PPRP observes that quantitatively expressing one's uncertainties on a particular matter involves a certain degree of introspection, and that many analysts do not think naturally in terms of the quantiles requested by the TI team. The PPRP recommends that future elicitation projects continue to provide guidance as to the desired format for reporting uncertainties, but remain flexible in accepting expert input in the form the expert feels most representative of his/her beliefs. Furthermore, as indicated earlier, the project should ensure that any mathematical representations of these beliefs do not unduly distort the beliefs for the sake of convenience."

4.7.12 How should expert "no shows" be managed during the workshops?

At times, not all of the selected experts may be physically present at the workshops for various uncontrollable reasons. Often the absent expert's expertise cannot be fully covered by other experts in the panel. Then, a "remote-show" may be considered as better than "no-show." The ISLOCA report noted, "Although the expert may participate in the workshop but via GoToMeeting and conference call, this means of participation, at times, severely limited the expert's involvement and contribution to the group discussions because of technology-related deficiencies, in particular, communication disruptions (lost GoToMeeting connection), difficulty in hearing discussions when the phone speaker system was not close to the speaker, and the facilitator presentation and the GoToMeeting presentation not always being in sync with one another. This may result in some degradation in the effectiveness of the panel."

Therefore, it is best if all experts attend the workshops in person to maximize the value that each brings to the issue(s) being addressed. The attendance by all experts is a necessity because of the enhanced technical interaction and dialogue that occurs amongst the experts. However, because of the limited number of experts on the panel, this would have necessitated replacing the expert. The consequential significant delay in the project schedule from replacing the expert would need to be balanced against the potential degradation in the effectiveness of the panel from retaining the expert.

In order to minimize the risk of "no shows" the number of expert team members should be minimized as much as possible (by using, for example, the guidance in Table 3-2 (SSHAC Levels of Expert Elicitation) while ensuring that the required expertise would be available.

4.7.13 **What are good practices for running the workshops?**

The peer reviewer's report on the JACQUE-FIRE project noted that, "the project demonstrated a number of good elicitation practices. These practices are worth emulating in future elicitation projects." Below are some of the recommended practices:

- At every elicitation session, remind the participants regarding:
  - The basic SSHAC philosophy of developing a community distribution (not the "right answer"); and
  - The participants' assigned SSHAC roles (be it technical integrator, resource expert, proponent, or participatory peer reviewer).
- At every elicitation session, remind the panel about the expert elicitation risks such as biased results and what behavioral attributes or "ground rules" should be followed during the workshop in order to avoid/minimize them.
- Encourage participants to openly share their point of view, even if it is not shared by others. Find ways to help an expert overcome his/her reluctance to quantify his/her beliefs even when test data or mathematical model predictions are not available.
- Ensure panel discussions include adequate consideration of application issues, including: a) the effect of potentially key differences between test and field conditions when interpreting test data (e.g., hundreds of conductors in a typical tray versus 10 in a typical test; a predominance of one particular polarity in the field versus an even polarity distribution in a typical test), and b) how the elicitation results are likely to be used in PRA studies.
- Ensure panel discussions include adequate consideration of previous, relevant efforts (yet be alert that this may introducing anchoring bias).
- Perform sanity checks to ensure the estimates are consistent with the experts' beliefs.
  - Check elicited and derived estimates for implications regarding the random variable (e.g., the number of failures in a set of trials), not just the parameter (e.g., the failure probability).
  - Check estimates for trends across cases and determine if there is a justification for those trends.
  - Check uncertainty representations for consistency with the amount of information. (In general, the uncertainties should be larger when less information, including data from tests and information regarding actual plant conditions, is available.)
- Acknowledge cognitive biases, brief experts on the issue, and design elicitation procedures that work to achieve this objective. Of course, the same cognitive biases arise in the deliberations of less formal consensus panels, but in those cases they are virtually never acknowledged or addressed. The performance of consensus expert panels might be improved if panel members first performed individual elicitations before they begin their group deliberations.

4.8 Integrate Judgments

4.8.1 Why is the integration approach important?

EPA guidance noted, "A typical expert elicitation obtains judgments from at least three proponent experts because diversity is more likely to reflect all relevant knowledge. In addition to their knowledge, however, each expert brings different experiences, perspectives, and biases to the question(s) of interest. Combining expert judgments may present several pitfalls, including the potential for misrepresenting expert judgments, drawing misleading conclusions about the scientific information, and adding biases to the conclusions (Hora, 2004; Keith, 1996; O'Hagan, 2005). Therefore, expert elicitation practitioners must be cautious about aggregating expert judgments and presenting combined conclusions about expert elicitation results." Therefore, the project team and LTI need to carefully choose an integration approach and make it an agreed-on procedure among the TIs, proponent experts, as well as the project sponsors.

4.8.2 How should the available resources be balanced with the thoroughness of the integration?

For complex technical issues, integration may involve many iterations among the TIs and with different techniques for aggregating individual judgments. The TI team may deliberate integration through multiple group meetings. The project team and LTI should plan ahead to determine the type (face-to-face vs. web-based) and number of integration meetings to ensure that all the functions in the TI integration are achieved. Underestimating the level of effort needed for integration in past elicitation projects has led to a depletion in project time or funding after several deliberation meetings without reaching the final integration.

Generally speaking, the functions of Workshop 3 have been under-appreciated as compared to the functions of Workshop 1 and 2. Yet, the integration function is an integral part of a formal expert elicitation process. It is desirable to have a face-to-face meeting for Workshop 3. If the resources are not available for a face-to-face meeting, the project team should at least use tele-conferencing or web-based communications to perform the functions of Workshop 3.

4.8.3 What are some mathematical and behavioral approaches for combining individual judgments?

The methodologies of combining individual judgments can be classified into mathematical and behavioral approaches. The EPA guidance provides the following discussion on the two approaches:

- "Mathematical aggregation methods involve processes or analytical models that operate on the individual probability distributions to obtain a single combined probability distribution. Mathematical approaches range from simple averaging using equal weights (Keeney and von Winterfeldt, 1991) to a variety of more complex Bayesian aggregation models. Although the Bayesian aggregation methods are theoretically appealing, difficult issues remain concerning how to characterize the degree of dependence among the experts and how to determine the quality of the expert judgments (e.g., how to adjust for such factors as overconfidence). Clemen and Winkler (1999) reviewed both mathematical and behavioral approaches for combining individual judgments along with empirical evidence on the performance of these methods. Using mathematical methods to combine expert opinions relies on an assumption that the individual expert opinions are independent (O'Hagan, 1998)."

- "Behavioral aggregation approaches "attempt to generate agreement among the experts by having them interact in some way" (Clemen and Winkler, 1999). Based on their review of the empirical evidence evaluating both mathematical and behavioral aggregation methods, Clemen and Winkler found both approaches tended to be similar in performance and that "simple combination rules (e.g., simple averaging) tend to perform quite well." They also indicated the need for further work in the development and evaluation of combination methods and suggest that the best approaches might involve aspects of both the mathematical and behavioral methods."

4.8.4 Which combination rule is better, linear (equal-weight) or geometric combination?

Probably the most commonly used algorithmic combination rules are the linear equal-weight combination (the average of the experts' distributions) and geometric combination. Hora (2004) and others (Gneiting and Ranjan 2011) found that any linear combination formula with strictly positive coefficients fails to be coherent. Hora (2010) evaluated the performance of various combination rules and compared equal-weight and geometric combinations of experts' distributions. He found that geometric combination performs best when experts are independent and well calibrated but its performance deteriorates compared with the equal-weight combination as dependence increases or expert calibration decreases. He also found that geometric combinations tend to have lower variance than the equal-weight combination. Liechtendahl et al. (2013) provided analytical properties of quantile aggregation and the linear opinion pool in terms of calibration, sharpness and shape; the results suggested that averaging experts' quantiles might give a better decision maker than an equal weight, or "averaging probabilities" combination of their distribution functions. Busetti (2015) compared quantile aggregation against the linear and the logarithmic opinion pool as methods for combining density forecasts. Overall the properties of quantile aggregation are in between those of the linear and the logarithmic pool. Hammitt and Zhang (2013) compared various algorithmic methods for combining expert judgments. They evaluated the properties of five combination methods (equal-weight, best-expert, performance, frequentist, and copula) using simulated expert-judgment data. They examined cases in which two well-calibrated experts were of equal or unequal quality and their judgments were independent, positively or negatively dependent. In this setting, the copula, frequentist, and best-expert approaches performed better and the equal-weight combination method performed worse than the alternative approaches. In general, the decision to use a fitted aggregate distribution should be informed by how the elicitation results will be used. For example, some computer codes have the ability to input complex (e.g., user-defined) parameter distributions, rather than common statistical forms (e.g., beta distribution). In instances where unfitted distributions can be used, the unfitted results of the elicitation should be used for the probability distribution.

4.8.5 What are some techniques or methods for mathematical aggregation?

NUREG/CR-5424 describes various methods and techniques for mathematical aggregation as well as the pros and cons of the techniques. The report has a good collection of techniques and should be the first reference to use when choosing how to perform a mathematical aggregation.

One popular technique is to first fit each individual judgment into a probabilistic function then aggregate the fitted functions. Literature has been controversial on fitting vs. no-fitting. The

ISLOCA project has an example on using the fitting strategy, “In order to develop results that could be easily used by current PRA software, the project placed a heavy emphasis on using a standard probability function (the beta distribution) to approximate the spurious operation probability estimates of individual experts and of the panel as a whole. In some cases, the fitting process was straightforward. In others, because of the limited flexibility of the beta distribution, choices had to be made as to which portions of the experts’ distributions should be emphasized in the fitting process. In a few cases, the use of the beta distribution may have masked situations where multi- modal distributions (representing the possibility of distinct, competing “models of the world”) accurately represent the current state of knowledge. Overall, it is not clear whether the use of the beta distribution has a positive or negative effect when representing the range of uncertainty of the technical community.”

4.8.6 Can software tools be used to combine probability distribution?

When using a mathematical model for the parameter of interest, the probability distribution for that parameter is a mathematical function of the distributions for the elicited parameters. When developing a project team, care should be taken that the team has sufficient expertise (or access to such expertise) to understand how the project’s chosen software tool develops the output distribution, what potential cautions should be observed when using such tools, and how to develop the distribution if a software solution is not readily available.

4.8.7 Can a consensus process be used to combine individual judgments?

A consensus process is often used to combine individual judgments for situations where the goal of the assessment is to obtain consensus views, rather than the distribution of views of the technical community. Obtaining the consensus views is particularly popular for problems that cannot be directly measured. Under the consensus approach, the aggregation of individual expert judgments requires the experts to adjust their judgments and move toward consensus. By defining the quantitative issues of interest and removing ambiguous judgments, this process can help experts to refine their understanding of the problem and potentially narrow their differences. Yet, without performing the functions of TI integration and Workshop 3, the consensus may only capture a narrow part of the truth in the stated problem and may not represent the center of the state of knowledge.

4.8.8 What are some challenges in integrating expert judgments?

The EPA report discussed some challenges in integration of individual judgment:

“According to Keith (1996), combining judgments could be problematic because an underlying methodological assumption is that the experts chosen for the elicitation represent the entire continuum of “truth” with respect to the technical question. He cautions that the “fraction of experts who hold a given view is not proportional to the probability of that view being correct” (Keith, 1996). In part, this is caused by how the experts are selected and the extent to which any of the selected experts possess knowledge that approximates the “truth.” As mentioned in Section 3.1, expert opinions are not necessarily evenly distributed across the entire spectrum of potential opinions. Furthermore, prior to interviewing experts, it may not be possible to determine the range of expert opinion on a particular question. Consequently, depending on which experts are selected (and agree) to participate in the EE, the fraction of experts used for the elicitation cannot be assumed to be proportional to the probability of that view or opinion being correct. In addition, if all else is equal and because a true value cannot be known, there is no objective basis to value the opinion of any one expert over any other.

Combining expert judgments requires the relative weighting of individual expert judgments to each other. They may be weighted equally or in some differential manner—for example, by social persuasion (as might occur in Delphi consensus building methods), by expert credentials, or by some form of calibration or performance assessment (Cooke, 1991). Keith (1996) argues that equal weighting of expert judgments is generally inappropriate because it is not possible to obtain a sufficiently large sample of in-depth EEs to ensure that all possible expert views are represented. Others argue that equal weighting is often as effective as more sophisticated differential weighting approaches (Clemen and Winkler, 1999).

Other EE practitioners also urge caution about combining the individual judgments (Wallsten et al., 1997). In a methodology similar to Jevolsek et al. (1990), Wallsten et al. (1997) proposed a model that considers both "the structure of the information base supporting the estimates and the cognitive processes of the judges who are providing them." Wallsten et al. determined areas in which experts agree and derived rules that satisfy those points of agreement. The resulting model avoids some of the criticisms of combining expert judgments by specifically considering subjective inputs for data and the processes used by the experts in the elicitation."

4.9 Document the Process and Results

4.9.1 Should the experts be identified in the documentation?

Users of expert judgments want to know where the information comes from. Throughout an expert elicitation process, experts should stand behind their views and judgment. It is important that the experts attempt to represent the technical community, not just themselves. Hence, documentation should comply with the transparency principle of expert elicitation and clearly document each expert's inputs to the final results. For this reason, the documentation should provide the name and background of the experts instead of presenting the results anomalously.

4.10 Conduct Participatory Peer Review

4.10.1 Why use a participatory peer review?

A standard peer review typically reviews the final report, while a participatory peer review is involved throughout all phases of the project to provide clear and timely feedback on both technical and process matters. The participatory peer review provides feedback on the project's conformance to the requirements of a formal expert elicitation process and the extent to which all technical assessments were adequately defended and documented. Hence, the formal expert elicitation process in Chapter 3 recommends participatory peer review as an integral part of the process. By reviewing the process and making timely inputs to the project, participatory peer review provides assurance that:

- A structured, systematic process has been followed that conforms to the basic principles of expert elicitation
- The study has considered the diversity of prevailing views within the technical community
- The uncertainties have been properly evaluated and incorporated into the analysis
- The documentation of the study is clear and complete

4.10.2 **How does a participatory peer review work?**

The participatory peer reviewers provide input throughout the course of the project. The peer reviewers may perform all or as much as possible the following functions:

- Participate in project workshops and other meetings
- Review the development of technical issues, project plan, and training materials
- Provide comments and suggestions during the periods set aside for reviewer inputs. Reviewer comments should address the basic principles of the process, implementation of the process, and whether the process is appropriate given the intended use of the results.
- Review the final project documentation
- Write a review report on the process and results of the project

Past experience has found it useful to have an informal daily debrief with the peer reviewers during the workshops to allow for mid-workshop corrections and discussion of comments. During the workshops, with the consent of the LTI or meeting facilitator, the reviewers may also provide input on the process or note deviation from the basic principles. The three pilot projects also revealed some areas that participatory peer reviewers should specifically pay attention to, including:

- Sufficiency of the datasets
- Thorough conformance to the boundary conditions or assumptions made at the beginning of the project
- Adequate and thorough evaluation of available data, models, and methods
- The treatment of variability and uncertainty
- Validation and verification of models and methods used
- Appropriateness of methods used for aggregating the judgments

5. SUMMARY

This report provides practical insights and lessons learned for eliciting expert judgments as the inputs to analysis and models in the NRC's decision-making activities. It is intended to be used to assist NRC staff with applying expert elicitation in applications that require expert judgment, yet do not have application-specific guidance. The report presents an introduction to expert elicitation and its use within the NRC, a technical basis for expert judgment, practical insights and guidelines for conducting expert elicitation, and supplemental information based on the literature review and lessons learned from recent expert elicitation pilot studies performed by NRC staff. The guidelines consists of (1) the conditions for conducting expert elicitation, (2) basic principles to follow when conducting expert elicitation, and (3) a recommended expert elicitation process that complies with the basic principles.

This white paper is the first release of the work performed to date. Future applications of the guidance will likely yield additional lessons learned and identify areas where improvement or additional information is needed. For example, the staff intends to use expert judgment to conduct a PIRT of the technical issues involved in modeling low-power shutdown for Level 3 PRA. This experience may lead to additional guidance specific to the PIRT process. Also, it is desirable to expand the scope of the report with guidance on the NRC staff's performing review of expert elicitation in the technical areas where no regulatory guidance exists. New developments and lessons learned can be incorporated into future revisions of this report.

6. REFERENCES

American Society of Mechanical Engineers/American Nuclear Society (ASME/ANS), ASME/ANS RA-Sa-2009, "Standard for Level 1/Large Early Release Frequency Probabilistic Risk Assessment for Nuclear Power Plant Applications," Addendum A to RA-S-2008, ASME, New York, NY, February 2009.

Apostolakis, G., "Probability and Risk Assessment: The Subjectivistic View Points and Some Suggestions," *Nuclear Safety*, 19(3): 305–315, 1978.

Apostolakis, G., "The Concept of Probability in Safety Assessments of Technological Systems," *Science*, 250(4986):1359–1364, 1990.

Bar-Hillel, M., "The Base-Rate Fallacy in Probability Judgments," *Acta Psychologica*, 44:211–33, 1980.

Bar-Hillel, M., "Representativeness and Fallacies of Probability Judgment," *Acta Psychologica*, 55(2):91–107, 1984.

Budnitz, R.J., G. Apostolakis, D.M. Boore, L.S. Cluff, K.J. Coppersmith, C.A. Cornell, and P.A. Morris, "The Use of Technical Expert Panels: Applications to Probabilistic Seismic Hazard Analysis," *Risk Analysis*, 18(4):463–470, August 1998.

Cannon-Bowers J.A. and Salas E., "Shared Mental Models in Expert Team Decision Making," in Individual and group decision making: current issues. N. John Castellan (Ed.), Psychology Press, 1993.

Cooke, R.M., *Experts in Uncertainty: Opinion and Subjective Probability in Science*, Oxford University Press, New York, NY, 1991.

Cooke, R.M., and L.H.J. Goossens, "A Procedures Guide for Structured Expert Judgment," EUR 18820, European Commission Directorate-General for Research, 2000.

Cowan, N., *Working Memory Capacity*, Psychology Press, 2012.

Crawford-Brown, D., "Scientific Models of Human Health Risk Analysis in Legal and Policy Decisions, *Law and Contemporary Problems*, 64:63–82, 2001.

Ericsson, K.A., R.T. Krampe, and C. Tesch-Römer, "The Role of Deliberate Practice in the Acquisition of Expert Performance," *Psychological Review*, 100:363–406, 1993.

Fischhoff, B., and D. MacGregor, "Subjective Confidence in Forecasts," *Journal of Forecasting*, 1(2):155–172, 1982.

Fischhoff, B., P. Slovic, and S. Lichtenstein, "Lay Foibles and Expert Fables in Judgments about Risk," *American Statistician*, 36:240–255, 1982.

Gigerenzer, G., "The Psychology of Good Judgment: Frequency Formats and Simple Algorithms," *Medical Decision Making*, 16(3):273–280, 1996.

Gigerenzer, G., and U. Hoffrage, "How to Improve Bayesian Reasoning without Instruction: Frequency Formats," *Psychological Review*, 102:684–704, 1995.

Gigerenzer, G., P.M. Todd, and the ABC Research Group, *Simple Heuristics That Make Us Smart*, Oxford University Press, New York, NY, 1999.

Gilovich, T., D. Griffin, and D. Kahneman, *Heuristics and Biases: The Psychology of Intuitive Judgment*, Cambridge University Press, New York, NY, 2002.

Gneiting, T., and R. Ranjan, “Comparing Density Forecasts Using Threshold- and Quantile-Weighted Scoring Rules,” *Journal of Business and Economic Statistics*, 29(3):411–422, 2011.

Hammitt, J.K., and Y. Zhang, “Combining Experts’ Judgments: Comparison of Algorithmic Methods Using Synthetic Data,” *Risk Analysis*, 33(1):109–120, 2013.

Hora, S.C., “Probability Judgments for Continuous Quantities: Linear Combinations and Calibration,” *Management Science*, 50(5):567–604, 2004.

Hora, S.C., “Eliciting Probabilities from Experts,” in W. Edwards, R.F. Miles, Jr., and D. von Winterfeldt, eds*., Advances in Decision Analysis: From Foundations to Applications*, pp.129–153, Cambridge University Press, New York, NY, 2007.

Hora, S.C., “An Analytic Method for Evaluating the Performance of Aggregation Rules for Probability Densities,” *Operations Research*, 58(5):1440–1449, 2010.

Hukki, K., “A Formal Process for Elicitation and Validation of Expert Judgments for Safety Case in the Context of Spent Nuclear Fuel Management,” Working Report 2008-60, Posiva Oy, Finland, September 2008.

Kahneman, D., *Thinking, Fast and Slow*, Farrar, Straus and Giroux, New York, 2011.

Kahneman, D., P. Slovic, and A. Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York, NY, 1982.

Kahneman, D., and A. Tversky, “Prospect Theory: An Analysis of Decisions under Risk,” *Econometrica*, 47(2):263–292, March 1979.

Kahneman, D., and A. Tvesky, “Evidential Impact of Base Rates,” in D. Kahneman, P. Slovic, and A. Tvesky, eds., *Judgement Under Uncertainty: Heuristics and Biases*, Cambridge University Press, New York, NY, 1982.

Kahneman, D., and A. Tversky, *Choices, Values and Frames*, Cambridge University Press, New York, NY, 2000.

Keeney, R.L., and D. von Winterfeldt, “Eliciting Probabilities from Experts in Complex Technical Problems,” *IEEE Transactions on Engineering Management*, 38(3):191–201, 1991.

Kirkebøen, G., “Decision Behaviour—Improving Expert Judgement,” in T.M. Williams, K. Samset, and K.J. Sunnevåg, eds., *Making Essential Choices with Scant Information: Front-End Decision Making in Major Projects*, pp. 169–192, Palgrave MacMillan, Chippenham and Eastbourne, UK, 2009.

Klein, G.A., “A Recognition-Primed Decision (RPD) Model of Rapid Decision Making, in G.A. Klein, J. Orasanu, R. Calderwood, and C.E. Zsambok, eds., *Decision Making in Action: Models and Methods*, pp. 138–147, Ablex Publishing, Westport, CT, 1993.

Klein, G., R. Calderwood, and D. MacGregor, “Critical Decision Method for Eliciting Knowledge,” *IEEE Transactions on Systems, Man and Cybernetics*, 19(3): 462–447, 1989.

Larrick, R.P., “Debiasing,” in D.J. Koehler & N.J. Harvey, eds., *Blackwell Handbook of Judgment and Decision Making*, pp. 316–338, Blackwell Publishing, Ltd., Oxford, 2004.

Lichtendahl, Jr., K.C., Y. Grushka-Cockayne, and R.L. Winkler, “Is It Better to Average Probabilities or Quantiles?” *Management Science*, 59(7):1594–1611, 2013.

Lichtenstein, S., B. Fischhoff, and L. Phillips, “Calibration of Probabilities: The State of the Art to 1980,” in D. Kahneman, P. Slovic, and A. Tversky, eds., *Judgment Under Uncertainty: Heuristics and Biases*, pp. 306–334, Cambridge University Press, New York, NY, 1982.

Linstone, H.A., and M. Turoff, *The Delphi Method: Techniques and Applications*, Addison-Wesley, Reading, MA, 1975.

Marr, D., *Vision: A Computational Investigation into the Human Representation and Processing of Visual Information*, Freeman, New York, NY, 1982.

Meyer, M. A., K.B. Butterfield, W.S. Murray, R.E. Smith, and J.M. Booker, "Guidelines for Eliciting Expert Judgment as Probabilities or Fuzzy Logic," in T.J. Ross, J.M. Booker, and W.J. Parkinson, eds*., Fuzzy Logic and Probability Applications: Bridging the Gap*, American Statistical Society Series on Statistics and Applied Mathematics, Los Alamos National Laboratory, LA-UR-00-218, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2002.

Monroe, R.W., "Development of an Expert Judgement Elicitation and Calibration Methodology for Risk Analysis in Conceptual Vehicle Design," Old Dominion University Project No. 130012, NASA NCC 1-02044, 1997.

Morgan, M.G., "Use (and Abuse) of Expert Elicitation in Support of Decision Making for Public Policy," *Proceedings of the National Academy of Sciences*, 111(20):7176–7184, 2014.

Schuman, H., and S. Presser, *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context*, Academic Press, New York, 1981.

Simola, K., A. Mengolini, and R. Bolado-Lavin, "Formal Expert Judgment: An Overview," EUR 21772 EN, European Commission Directorate-General Joint Research Centre, Institute for Energy, The Netherlands, July 2005.

Slovic, P., "Perception of Risk," *Science*, 236:280–285, 1987.

Surowiecki, J., *The Wisdom of Crowds: Why the Many Are Smarter than the Few and How Collective Wisdom Shapes Business, Economies, Societies and Nations*, Doubleday, New York, NY, 2004.

Tversky A., and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, New Series, 185(4157):1124–1131, 1974.

Tversky, A., and D. Kahneman, "The Framing of Decisions and the Psychology of Choice," *Science*, 211(4481):453–458, 1981.

Tversky, A., and D.J. Koehler, "Support Theory: A Nonextensional Representation of Subjective Probability," *Psychological Review*, 101:547–567, 1994.

U.S. Environmental Protection Agency (EPA), "Expert Elicitation Task Force White Paper," 2011. Available at www.epa.gov/stpc/pdfs/eewhite-paper-final.pdf.

U.S. Nuclear Regulatory Commission, "Reactor Safety Study: An Assessment of Accident Risks in U.S. Commercial Nuclear Power Plants," NUREG-75/014 (WASH-1400), October 1975.

U.S. Nuclear Regulatory Commission, "Expert Estimation of Human Error Probabilities in Nuclear Power Plant Operations: A Review of Probability Assessment and Scaling," NUREG/CR-2255, 1982.

U.S. Nuclear Regulatory Commission, "Procedures for Using Expert Judgment to Estimate Human Error Probabilities in Nuclear Power Plant Operations," NUREG/CR-2743, January 1983.

U.S. Nuclear Regulatory Commission, "Generating Human Reliability Estimates Using Expert Judgment," NUREG/CR-3688, November 1984.

U.S. Nuclear Regulatory Commission, “Analysis of Core Damage Frequency from Internal Events: Expert Judgment Elicitation,” Part 1, “Expert Panel Results,” NUREG/CR-4550, Vol. 2, April 1989.

U.S. Nuclear Regulatory Commission, “Severe Accidents in Spent Fuel Pools in Support of Generic Safety Issue 82,” NUREG/CR-4982, 1987.

U.S. Nuclear Regulatory Commission, “Methods for the Elicitation and Use of Expert Opinion in Risk Assessment,” NUREG/CR-4962, August 1987.

U.S. Nuclear Regulatory Commission, “Eliciting and Analyzing Expert Judgment: A Practical Guide,” NUREG/CR-5424, January 1990.

U.S. Nuclear Regulatory Commission, “Elicitation and Use of Expert Judgment in Performance Assessment for High-level Radioactive Waste Repositories,” NUREG/CR-5411, May 1990.

U.S. Nuclear Regulatory Commission, “Severe Accident Risks: An Assessment for Five U.S. Nuclear Power Plants,” NUREG-1150, December 1990.

U.S. Nuclear Regulatory Commission, “Branch Technical Position on the Use of Expert Elicitation in the High-Level Radioactive Waste Program,” NUREG-1563, November 1996.

U.S. Nuclear Regulatory Commission, “Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts,” NUREG/CR- 6372, Vols. 1 and 2, April 1997.

U.S. Nuclear Regulatory Commission, “Technical Basis and Implementation Guidelines for A Technique for Human Event Analysis (ATHEANA),” NUREG-1624, May 2000.

U.S. Nuclear Regulatory Commission, “Quantifying Reactor Safety Margins: Application of Code Scaling, Applicability, and Uncertainty Evaluation Methodology to a Large-Break, Loss-of-Coolant Accident,” NUREG/CR-5249, Rev. 4, 1989.

U.S. Nuclear Regulatory Commission, “Estimating Loss-of-Coolant Accident (LOCA) Frequencies through the Elicitation Process,” NUREG-1829, April 2008.

U.S. Nuclear Regulatory Commission, “Human Factors Considerations with Respect to Emerging Technology in Nuclear Power Plants,” NUREG/CR-6947, October 2008.

U.S. Nuclear Regulatory Commission, “Staff Requirements—COMGEA-11-0001, Utilization of Expert Judgment in Regulatory Decision Making,” SRM-COMGEA-11-0001, March 15, 2011.

U.S. Nuclear Regulatory Commission, “Response to Staff Requirements Memorandum COMGEA-11-0001, ‘Utilization of Expert Judgment in Regulatory Decision Making,’” Commission Paper SECY-11-0172, December 13, 2011.

U.S. Nuclear Regulatory Commission, “Staff Requirements—SECY-11-0172—Response to Staff Requirements Memorandum COMGEA-11-0001, ‘Utilization of Expert Judgment in Regulatory Decision Making,’” SRM-SECY-11-0172, February 7, 2012.

U.S. Nuclear Regulatory Commission, “Practical Implementation Guidelines for SSHAC Level 3 and 4 Hazard Studies,” NUREG-2117, Rev. 1, April 2012.

U.S. Nuclear Regulatory Commission, “Joint Assessment of Cable Damage and Quantification of Effects from Fire (JACQUE-FIRE),” NUREG/CR-7150, Vol. 1, “Phenomena Identification and Ranking Table (PIRT) Exercise for Nuclear Power Plant Fire-Induced Electrical Circuit Failure,” October 2012.

U.S. Nuclear Regulatory Commission, “Joint Assessment of Cable Damage and Quantification of Effects from Fire (JACQUE-FIRE),” NUREG/CR-7150, Vol. 2, “Expert Elicitation Exercise for Nuclear Power Plant Fire-Induced Electrical Circuit Failure,” May 2014.

U.S. Nuclear Regulatory Commission, “Cognitive Basis for Human Reliability Analysis,” NUREG-2114, January 2016.

U.S. Nuclear Regulatory Commission, “An Integrated Human Event Analysis System (IDHEAS) for Nuclear Power Plant Internal Events At-Power Application,” NUREG-2199, Vol. 1, Rev. 0, in press.

Appendix A. Staff Requirements Memorandum (SRM) COMGEA-11-0001, "Utilization of Expert Judgment in Regulatory Decision Making"

COMGEA-11-0001

January 19, 2011

SUBJECT: UTILIZATION OF EXPERT JUDGMENT IN REGULATORY DECISION MAKING

1. Objective

To ensure that the formal utilization of expert judgment incorporates lessons learned from past major studies and is applied consistently in regulatory decision making throughout the Agency.

2. Background

The formal utilization of expert judgment is a process that provides either (1) quantitative estimates for the frequency and/or significance of physical phenomena, or (2) qualitative insights into the nature, scope, and/or significance of physical phenomena. Expert judgment is used when the following conditions are present: the available data or operating experience is sparse or not applicable, the subject is too complex to model accurately, and the phenomena or issues have significant safety or regulatory implications.

Expert judgment has been a principal component of the technical basis for many important regulatory decisions and its use is expected to be more prevalent in the future as issues become more complex and as technology evolves. For example,

i. The landmark NUREG-1150[1] study utilized expert judgment to assess failure frequencies, failure modes, recovery actions, accident progression, and source term behavior, among other phenomena to characterize the risk associated with severe accidents in operating reactors.

ii. Expert judgment is used for probabilistic seismic hazard assessment (PSHA) for siting and design of nuclear facilities. Due to large uncertainties in the geoscience data and in their modeling, multiple model interpretations are often possible. This has led to expert disagreement on the selection of ground motion for design at a given site. The Senior Seismic Hazard Analysis Committee (SSHAC)[2] concluded that the differences in PSHA results were due to procedural rather than technical differences and proposed a formal method for utilizing expert judgment in PSHA.

---

[1] "Severe Accident Risks: An Assessment of Five Nuclear Power Plants, Final Summary Report," NUREG-1150, Vol. 1, 1990.

[2] "Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts," NUREG/CR-6372, 1997.

iii. Expert judgment has been required to assess the performance of a high-level waste repository. Specifically, it is used to predict future climates, characterize waste degradation and transport if the waste package is breeched, and to perform the volcanic hazard analysis.

iv. All analyses of human reliability performance rely on expert judgment.

v. More recently, expert judgment has been used to estimate loss-of-coolant-accident (LOCA) frequencies. These frequencies provided the basis for selecting the transition break size (TBS) proposed in the risk-informed revision of the emergency core cooling system (ECCS) acceptance criteria (10 CFR 50.46a).

vi. Informal expert judgment is used to determine safety and security requirements for commercial and medical uses of radioactive materials.

vii. The Phenomena Identification and Ranking Table (PIRT) process is another specific application of the formal utilization of expert judgment. Examples include:

   a) identifying the technical issues associated with advanced reactor licensing;

   b) understanding ECCS performance under LOCA scenarios (i.e., in GSI-191). Specifically, PIRTs were used to investigate containment coating performance, evaluate debris transport in wet and dry containments, and to identify outstanding chemical effect issues.

3. Motivation

There are many similarities but also significant differences in the approaches used in the above studies that can impact regulatory decision making. For example, the SSHAC approach introduces the concept of a technical facilitator/integrator (TFI) to develop the final aggregated results, but this approach may not be appropriate in all cases. For example, the LOCA frequency study did not utilize a TFI.

A unique feature of the LOCA frequency study was the adjustment of results to account for the well-known overconfidence that is typically present in individual expert judgments. The study also recommended a less-common scheme for aggregating the individual expert results into group estimates. Sensitivity studies indicated that the selection of the aggregation scheme affected the results significantly. When the recommended, but less- common, aggregation scheme is used, the TBS for a pressurized water reactor is approximately 6" while aggregating using more-common methods leads to a TBS of approximately 11"[3]. Selecting and documenting the appropriateness of the methods of analysis ahead of the regulatory decision should increase transparency, public confidence, and the objectivity of the results.

It is anticipated that formal use of the expert judgment process could play an important role in the resolution of difficult regulatory challenges including cyber security, digital instrumentation and control, small modular reactors, and material aging issues. The NRC would benefit from formal guidance to assist the staff in choosing the method for obtaining and utilizing expert judgment to avoid the pitfalls of the past and ensure the appropriate level of effort. The extensive elicitation used to develop the LOCA frequency estimates will not be appropriate for all cases. In some cases, it may only require consultation with several subject matter experts.

---

[3] A subtle yet important point is that the choice of aggregation scheme does not, in and of itself, lead to a TBS of either 6" or 11". Rather, these sizes correspond to the frequency of 10^{-5} /yr that was cited in the Commission's SRM. There are more considerations that went into the TBS, e.g., the inclusion of margin to account for uncertainties.

4. Recommendation

I recommend that the Commission direct the staff to provide, within 6 months, a plan for the development of guidance that will ensure that the formal utilization of expert judgment is applied consistently in regulatory decision making throughout the Agency. This plan should describe the staff's approach, schedule, and estimated resources. This plan should recognize the development of the guidance should include the following:

viii. a summary of past and ongoing significant NRC activities that utilized expert judgment to identify the lessons-learned, document the approaches[4], and identify significant differences among the approaches,

ix. a survey of recent research to identify promising new approaches (or techniques that can be applied within the broader approach) to expert judgment that may be appropriate for use in nuclear applications,

x. an evaluation of recent activities within other agencies that relied on expert judgment to identify the lessons-learned, document the approaches, and identify differences among the approaches and those used in NRC activities,

xi. options that match the approach with the nature and significance of the issue and the extent to which expert judgment is relied upon in regulatory decision making,

xii. estimates of resources associated with each option for planning purposes,

xiii. guidance that is prescriptive enough to ensure consistent application of expert judgment within the Agency, yet is sufficiently flexible to account for the wide diversity of issues that the Agency faces. The user should be able to tailor the approach to be applicable to the unique issue of concern, and

xiv. the possibility of developing national standards.

5. Benefits

This effort will promote a more consistent and transparent basis for regulatory decision making when expert judgment is required. It will also provide clear and consistent guidance to licensees and staff for both formally utilizing expert judgment and for reviewing licensing actions that are based, at least in part, on expert judgment. Finally, it is anticipated that this effort will improve the efficiency of Agency planning by identifying and prioritizing resources that are commensurate with the significance of the safety or security issue(s) and degree of reliance on expert judgment in the associated regulatory decision making.

---

[4] The expert judgment approach refers to the process used to elicit information from experts, analyze this information to develop results, and determine the implications of the results to support regulatory decision making.

Appendix B. Unaddressed Comments from the NRC Staff Review

A draft of this report was reviewed by the staff from five NRC offices: RES, NRR, NRO, NMSS, and NSIR. The reviewers provided numerous constructive comments to improve the report. The authors addressed most of the comments, except for the comments that needed substantial additional development work. Nevertheless, those unaddressed comments are important for readers to be aware of when reviewing the insights and guidance in this report. Below are the major unaddressed comments.

1. Scope of the report

1.1 Lack of guidance on reviewing expert elicitation

The report lacks guidance on when to exercise judgment during a regulatory review. For example, Section 1.1 (Background) discusses the use of expert judgment as input to important regulatory decisions such as those noted in Section 1.2 (Use of Expert Judgment in the NRC). However, it does not explain what other steps were followed prior to reaching the point of relying on an expert elicitation process. Dr. S. Birla (NRC-RES) provides a general idea of the guidance to be followed prior to exercising judgment during a regulatory review. Either such guidance or similar should be evaluated for inclusion in the report.

Sample guidance to be followed prior to exercising judgment in a regulatory review (Author: Dr. Sushil Birla 2016-09-21):

1) *Make it clear that the reviewer's judgment is not a substitute for inadequate quality of information.*
2) *It is also not a substitute for lack of sound reasoning in the applicant's safety analysis.*
3) *It is also not a substitute for an individual reviewer's inadequate knowledge.*
4) *First, the individual reviewer should identify these gaps, including gaps in one's own knowledge.*
5) *Then the individual reviewer should use the RAI process to fill gaps for which the applicant is responsible.*
6) *Then the individual reviewer should reach out within one's own work group to fill knowledge gaps.*
7) *Guidance should be included to form an internal experts' team to fill the knowledge gaps.*
8) *This team should determine gaps in its own collective expertise and identify external experts who could fill those gaps.*
9) *The team should formulate appropriate questions to be answered by the external experts (a) individually, and (b) collectively. Thus, narrow the scope of external elicitation to the gaps to be filled and the individual external expertise matching those gaps.*

1.2 Lack of guidance on utilization of expert judgment as a piece of evidence in regulatory decision making.

The use of the community knowledge representation as a piece of evidence in regulatory decision making is a different topic beyond the scope of this report. It requires consideration of how the expert judgment should be weighed against other forms of evidence, which is an issue for the broader topic of decision making under uncertainty.

1.3 Comparison and evaluation of other expert elicitation guidance

It would be useful to document the details of the work performed by the authors to compare the various expert elicitation guidance documents to SSHAC guidance. It is interesting for readers to see what SSHAC elements or guidelines are lack in other documents and understand the pros (if any) and cons of lack of those elements.

2. Technical basis – social biases

The Section 2.3 technical basis discussion is focused on the individual and team cognitive processes of the experts. The elicitation process involves social behaviors and also involves more than just the experts. These other aspects can give rise to additional vulnerabilities and

are the driving force for some elements of the elicitation process design. A separate section addressing social aspects seems warranted.

3. Guidance on data collection process

The reports lacks guidance in regards to the data collection process considering that it is an important part of the overall expert elicitation process. For example, there may be scenarios where an expert elicitation team may need to develop PDFs for modeling an operators' response in the PRA. Despite the fact that there may be seasoned operators in the expert panel, the team may still need to collect data from the operating fleet (e.g., via surveys, interviews) on how they judge different crews with different levels of proficiency may respond to a particular PRA scenario(s). However, the data collection process for this and similar scenarios is not captured in this report. Therefore, the report should provide guidance on the data collection process.

4. Explicit guidance on using Table 3.1 to plan an expert elicitation

The table provides options for the various elements (e.g., number of experts, training sessions, etc.) in the elicitation. Yet, there appears to be no guidance helping users decide among options. Pros and cons are needed. Note that this isn't just a matter of resources. In fact, the guidance should recognize that information overload on experts can have adverse impacts and that large panels can become unmanageable. In general, the report can be very useful in providing guidance that protects against hidden, "common-sense" assumptions.

5. Limited review of past experience in Chapter 4

The supplementary information in Chapter 4 heavily relies on the results of three relatively limited scope pilot projects and fails to capture the extensive use of this (SSHAC) process over the past 10 years. It should be made very clear that this section does not capture the insights from extensive experience gained in application and continuing development of the SSHAC guidelines.

[Note from the authors: The revision of Chapter 4 states that this section does not capture the insights from extensive experience gained in applications of SSHAC and the Branch Technical Position (NUREG-1563) guidelines. Readers are informed that many of the insights are being captured in the current updating of NUREG-2117.]

6. Guidance on weighting inputs from experts

Ideally, when an appropriately unbiased group of experts are enlisted and fulfill their roles and responsibilities, their inputs to the resulted judgment (i.e., the distribution of the state of knowledge) are treated in an equity manner. Practically, it happens that the group of experts may be unbalanced in the technical areas involved or some experts may heavily anchor to a subset of data / evidence without evaluating the range of available data. In situations like these, weighting the experts' inputs in the integration process becomes a consideration. What is the rationale for weighting experts' judgments differently? Chapter 4 discusses the literature on weighting inputs from expert but arrive at no effective recommendation. What value (if any) does NRC see in weighting individual judgments, and what criteria would we use to judge the validity of a weighting scheme? If the guidance implies that TIs may weight expert judgments, then there needs to be much more discussion on the need for clear rationale and technical basis for performing the weights (for example, a clear method for calibrating expert knowledge a weighting scheme). Otherwise, weighting experts' inputs without proper rationale may introduce particular biases into the aggregation process. The NRC staff should reach some determination herein on whether weighting is acceptable, or not, or only in some well-defined circumstances.