Journal of the Korean Nuclear Society
Volume 35, Number 4, pp.318~335, August, 2003

On the Balanced Blending of Formally Structured and Simplified Approaches for Utilizing Judgments of Experts in the Assessment of Uncertain Issues

Kwang-Il Ahn, Joon-Eon Yang, and Jae-Joo Ha
Korea Atomic Energy Research Institute
1150 Dukjin-dong Yuseung-gu, Daejeon 305-353, Korea
kiahn@kaeri.re.kr

(Received February 27, 2003)

Abstract

Expert judgment is frequently employed in the search for the solution to various engineering and decision-making problems where relevant data is not sufficient or where there is little consensus as to the correct models to apply. When expert judgments are required to solve the underlying problem, our main concern is how to formally derive their technical expertise and their personal degree of familiarity about the related questions. Formal methods for gathering judgments from experts and assessing the effects of the judgments on the results of the analysis have been developed in a variety of ways. The most important interest of such methods is to establish the robustness of an expert' s knowledge upon which the elicitation of judgments is made and an effective trace of the elicitation process as possible as one can. While the resultant expert judgments can remain to a large extent substantiated with formal elicitation methods, their applicability however is often limited due to restriction of available resources (e.g., time, budget, and number of qualified experts, etc) as well as a scope of the analysis. For this reason, many engineering and decision-making problems have not always performed with a formal/structured pattern, but rather relied on a pertinent transition of the formal process to the simplified approach. The purpose of this paper is (a) to address some insights into the balanced use of formally structured and simplified approaches for the explicit use of expert judgments under resource constraints and (b) to discuss related decision-theoretic issues.

Key Words : engineering and decision-making problems, uncertainty assessment, expert judgment approach, balanced use, decision-theoretic issue

1. Introduction

When relevant data is not sufficient and there is little consensus as to the correct models to apply, subjective judgment of experts (often called engineering judgment or expert opinion) have been considered in every step of the analysis of engineering and decision-making problems. For example, experts are often required to give their own judgments about what problems need to be analyzed, what models and methods to use, what data sources to analyze, how to interpret results, and what actions to recommend. An additional use of the expert judgment is in making quantitative statements about various uncertain issues that would be often addressed in the modeling process of the aforementioned problems, whose impact on the final analysis result is very high. In that case, combining judgmental information of experts with all available evidence is a necessary step in reducing relevant uncertainties to the maximum extent allowable from our state-of-knowledge about the underlying problem [1-6]. Though the subjective judgment of experts may be reflected in various ways in practical situations, its essential information can be characterized as the experts' technical insight and expertise regarding the underlying problems, their personal degree of familiarity with the relevant subjects, and their reasoned expectations.

In convention, the elicitation of the foregoing judgmental information from experts has been made in simple, direct, and informal ways, with an implicit reasoning process. Whereas, the recent trend is to make it formal through an explicit reasoning process, where one can most methods developed for the formal elicitation [7-12] basically follow guidelines established by decision theory. In particular, the latter approach (i.e., formal approach for expert judgment elicitation) has been extensively used as an essential means for quantifying various uncertainties addressed in probabilistic safety assessments (PSA) of nuclear power plants such as NUREG-1150 risk studies [13] and European benchmark exercise on expert judgment techniques [14]. The aforementioned two applications of the formal approach showed a fairly high degree of formality for the detailed analyses about the assessment of phenomenological parameters and documentation. The most important interest in the use of such methods was to establish the robustness of the expert's knowledge upon which the elicitation of judgments is made and an effective trace of the elicitation process as possible as one can. However, it is also noted that while the judgmental information from experts can remain to a large extent substantiated with the formal elicitation processes, there are also corresponding disadvantages. This is mainly due to the difficulty in designing a formal approach for eliciting expert judgments. The second reason is that practical implementation of the formal approach is a time-consuming process. The last reason is that because expert judgment is fundamentally subject to a subjective human reasoning process there is no generally accepted, scientifically correct, and technically defensible approach, applicable to variable situations consistently. For these reasons, many probabilistic safety studies have not always been performed with a formal pattern in every step of the analysis, but rather a pertinent mixture of formally structured and less formally structured approaches. For example, the previous application of expert judgment approach for probabilistic analysis of a complex system [15] demonstrates that the influence of informal, qualitative expert judgments, especially in structuring the problem, choosing an analytical approach, and choosing applicable models and data may play a great role. Based on the foregoing result, it was noted that experts are generally confident of their expertise and are not easily swayed in the early phases of analysis, even when their results differ from those of other experts. However, as they use the same analysis models and similar data experts' diverse opinions are easily conversed with a small degree of difference. This fact points out the importance of qualified information in the elicitation process of expert judgments, rather than the formality of the elicitation process.

The purpose of this paper is (a) to address some insights into the balanced use of highly structured and simplified approaches for the explicit use of expert judgments under resource constraints and (b) to discuss related decision-theoretic issues. The primary object for discussion in this paper is related to the assessment of uncertainty issues that would be employed in risk and reliability analysis of complex technological systems. For that purpose, the existing approaches for the formal elicitation of expert judgments are briefly reviewed in the next section. The present characterizations of formal and simplified elicitation processes for practical use are given in Sections 3 and 4, respectively. In the final section, we complete this paper with some concluding remarks on the balanced use of formal and simplified elicitation processes under resource constraints.

2. Brief Review of the Existing Approaches for Formal Use of Expert Judgments

When relevant data is not sufficient and there is little consensus as to the correct models to apply, approaches to obtain quantitative or qualitative judgments from experts have been developed in several ways [7-12, 16]. Through this paper, the expert refers to a person who has the background for the subject matter at the desired level of detail or those conducting the study being qualified to answer questions. Also, the expert judgment refers to both the expert' s answers to the underlying technical questions and his/her data on how this answer was reached (e.g., definitions, assumptions, and methods). Depending upon the degree of interaction among experts, they can be generally categorized into three classes: iterative, interactive, mechanical approaches. The classification of the forgoing approaches is based on the degree of formal meetings to revise the judgments in the elicitation process and to help alleviate some expert' s bias. The iterative approach is to gather individually the opinions of different experts that are averaged by the normative analyst. Delphi method [16] is a typical example using the iterative approach. The results are then sent back to the experts who are given the opportunity to revise their assessments on the basis of what others have to say. The process generally converges quickly because the time-consuming debating process among experts is no longer considered. The interactive approach basically follows guidelines established by a decision theory for both social and scientific processes and for reducing unnecessary bias in experts' opinions. In this approach, the experts are generally asked to debate through formal interactions (i.e., feedback via formal meetings, electric mail, telephone, etc.) and explain their judgmental background, to exchange information about the evidence base on which they rely, and to consider the probabilities of a full spectrum of alternative modeling approaches. Typical examples using the interactive approach are the SRI encoding process [8] and the NUREG-1150 expert opinion elicitation approach [9]. The mechanical approach utilizes either a mathematical model or a statistical combination for aggregating different opinions without taking into account any process for revising experts' assessments. While the first two approaches employ well-structured procedures for revising judgments in the elicitation process and helping alleviate some expert' s bias and dependency among the experts in making judgments, the mechanical approach does not have to be planned as early, as it is closely related to the choice of the elicitation methods. As a result, the mechanical approach can lead to the creation of a single answer that all of the experts would reject or a physical phenomena that could not occur physically. Thus, our primary concern has to be focused on the use of the explicit and well-structured formal approaches.

In light of practical implementation, on the other hand, most methods for characterizing reasonably uncertainties of complex technological systems with expert judgments often start with the selection of uncertain issues that are judged to be highly uncertain from the viewpoint of decision-making and can give a high impact on the final analysis result. In the experts' assessment of such uncertain issues, there are three major stages for obtaining their judgments regardless of the use of the aforementioned approaches. In the qualification of uncertain issues, experts have to cover all the aspects of the underlying issues in sufficient detail, including probabilistic modeling for uncertainty quantification, encoding process and scale, and aggregation methods. The first step for the qualification of these uncertain issues is to structure the problem into one or more logically related well-defined variables with which they are more familiar to quantify, that may be very helpful for identifying and quantifying uncertainties. This structuring process in making judgments [17] can be made from either a normative analyst (with/without assistance from relevant experts) or each expert whose judgment is required to be derived for the underlying issue. The decomposition can be used not only to combat motivational bias by producing a level of detail that disguises the connection between the subject' s judgments and personal interests, but also to reduce certain cognitive bias [1,10-11]. As for the next step, questions for the prepared variables are then suggested to experts whose judgments are required to be derived. In the process, the utmost care must be taken to be sure that the experts are knowledgeable in the area for which they are being questioned, and that they understand well the underlying questions they have to answer. In the stage of uncertainty assessment, probabilistic answers to each question are then obtained from experts, which are given in the form of his/her own probability distribution (e.g., uniform within intervals, piecewise uniform, and other type of probabilistic distribution function). Regarding the method for translation of the expert' s degree of belief in the probabilistic terms [10-11,18-20], however, relatively little guidance is available in literature [5,23]. Often, the piecewise uniform distribution is utilized to derive expert' s subjective probability distribution on unknown quantity since it fully complies with the experts' probability statements with a given finite and exhaustive set of intervals. It is known that in the case of independency among the percentiles, the use of fewer percentiles (or probability intervals) from each expert results in larger uncertainties and a broader range of the estimated percentiles (e.g., ratio of 95 to 5 percentile values). These probabilistic answers for each question (made in the decomposition level) can be integrated through the structured model to obtain an optimal solution for the underlying issue, from either a normative analysts or a relevant expert. In this stage, the potential dependency among the experts [10,19-22] is a crucial point in making their own judgments. In the stage of probabilistic integration, all alternative probabilistic answers obtained from experts for each issue are integrated into a compound distribution, to supply a composite probability distribution to be used in the course of a further evaluation to get the final outcome. In this process, the introduction of different 'weights' to experts' expertise (or experts' final answers for each issue) [7,10,21-23] is to enable the experts to quantitatively differentiate between answers to questions directly related to their particular fields of activity and those from adjacent technical fields. As for the dependency between experts in deriving their own estimates increases, the probability curve obtained by an integration model tends to move towards larger estimates.

Table 1. The Existing Approaches for the Formal Use of Expert Judgments

Delphi method [16]	Merkhofer approach [8] based on SRI Encoding Process (*)	NUREG-1150 Approach [9]	Keeney & von Winterfelt [11]	VTT-STUK Approach [12]
6 steps 1. Selection of issues and experts 2. Independent assessments from experts 3. Circulation of initial assessments to each anonymous experts with rationales and reasons for each opinion 4. Reassessment of each expert' s judgments 5. Interchange of assessments and written explanations several times 6. Convergence of opinion	7 steps 1. Motivating: reduce motivational and cognitive biases 2. Structuring: define and decompose the uncertain issue 3. Conditioning: relate all relevant knowledge to the uncertain variable by cause-specific and distributional information 4. Encoding: elicitation 5. Verifying: test the judgments obtained 6. Aggregating 7. Discretizing: discretized continuous variables for computational simplicity	10 steps 1. Selection of issues 2. Selection of experts 3. Preparation of issue statements 4. Elicitation training 5. Presentation of issues 6. Preparation of expert analyses by panel members 7. Discussion of analyses 8. Elicitation 9. Recomposition and aggregation 10. Review by the panel of experts	7 steps 1. Identification and selection of the issues 2. Identification and selection of the experts 3. Discussion and refinement of the issues 4. Training of elicitation 5. Elicitation 6. Analysis, aggregation, and resolution of disagreements 7. Documentation and communication	6 steps 1. Selection and training of experts 2. Elicitation of judgments 3. Modeling and combination of judgments 4. Sensitivity analyses 5. Discussion and feedback from experts 6. Documentation
A means for countering some of the biasing effects of interaction among experts	Application to decision analyses, public-policy analyses, decision analysis training seminar	Application to uncertainty issues addressed in PSAs of the five U.S. nuclear plants	Modification of NUREG-1150 draft report: application to performance assessment of high level waste repositories	Based on the NUREG-1150 final report: Application to a case study performed within the benchmark exercise on expert judgment techniques

Note: (*) C.A. Stael von Holstein and J.E. Matheson, "A Manual for Encoding Probability Distributions," prepared for Defense Advanced Research Projects Agency under Subcontract for Decisions and Designs, Inc., SRI International, Menlo Park, CA, Final Rep., SRI Project 7078, 1979.

3. The Present Characterization of Formal Elicitation Process

As shown in Table 1, the well-known formalized approaches for eliciting expert judgments employ several steps which are different in nature to define clearly the nature of the expert judgments, to make explicit the assumptions and reasoning behind them, to expose their inherent uncertainties, and to counter possible judgmental bias. Then, open and documented judgment elicitation processes allow results to be evaluated and reproduced by other experts. In many cases, the essential parts of the formalized process can be summarized by closely related, but distinctive five steps: (a) selection and motivation of experts, (b) qualification of the underlying uncertain issue, (c) qualification of questions to be elicited, (d) elicitation of experts' answers for questions, and (e) recomposition and aggregation of experts' answers. A particular concern in this section is to investigate the respective steps in light of minimizing expert's bias and unnecessary dependency among experts, maximizing accuracy and traceability of the experts' estimates, and maintaining consistency of the experts' judgments. The present elicitation procedure should be regarded as a newly restructured version of the existing formal procedures, rather than a newly developed one in the present study.

3.1. Selection and Motivation of Experts

An elicitation process employed by the expert must conclude expert judgments based on purely technical and scientific criteria. Fundamentally, experts' own analysis experience and capability in modeling and assessment of the underlying uncertain issue plays a dominant role in the elicitation process. Thus, the quality of expert's judgment can vary quite a lot, depending upon their familiarities with the issue and dependency among the experts. Equally qualified experts that do not have a vested interest in the outcome of the underlying problem or in the particular methodological approaches used, conduct the best peer reviews. While the decision to obtain information from several experts for decision-making under uncertainty, is usually motivated by a desire to reduce the uncertainty and thus to increase precision, a shared knowledge of professional judgment may make it difficult to conduct a truly independent peer review in highly technical areas where there are a limited number of qualified specialists. The foregoing description means that it is important to pay appropriate attention to the selection of the experts to be included in the survey as well as to the elicitation strategy. In the NUREG-1150 study [13], the experts were selected on the basis of their recognized expertise of the technical issues, such as demonstrated by their publications in referred journals. Experts from various disciplines and organizations (3 to 10 experts for each uncertain issue) were assigned to each expert panel to ensure a balance of diverse perspectives.

In general, experts would often be subject to some degree of bias from what the experts think that they are confident in and some burdens from the use of probabilistic expressions of their knowledge. Thus, some kind of training [9,10] is often required in the practice of expressing better their beliefs and knowledge of the problem and to avoid psychological bias like *overconfidence* affecting judgmental accuracy. The review of previous uses of expert judgment in similar fields may assist considerably in reducing the problems of such judgmental bias and overconfidence. In addition, if more experts are required to obtain additional information, then it might be important to seek out experts that are believed not to be highly dependent on each other or with the prior experts. This is mainly true except for some cases of extremely dependent experts; independent experts provide more varied and valuable information in the context of decision-making under uncertainty.

3.2. Qualification of the Underlying Issue

As mentioned previously, it is well recognized that the identification of suitable logically related well-defined variables through the decomposition and structuring process of the underlying issue is critical for the success of its probabilistic formulation process. Since the errors from the parts should tend to compensate for one another, the decomposition generally improves accuracy. If the fundamental mechanisms for the issue have not been clearly identified, it may be difficult to attempt directly the structuring process at the level of fundamental mechanisms. As the depth of the decomposition of each uncertain issue increases, moreover, the number of variables to be estimated from experts increases tremendously. As a result, the experts' estimation process of the underlying issue can be a very expensive exercise and we may not have the opportunity to arrange an interactive meeting among them to facilitate a reduction of bias in their estimates. In addition, the experts might have already come to an agreement when they indirectly endorsed the common information source in decomposing the fine level of uncertain issues. In that case, all that remains to be done is to formulate a consensus state of knowledge distribution by assessing the credibility and bias of the common information source from the experts.

3.3. Qualification of Questions to be Elicited

Once the uncertain issue is structured and qualified with an appropriate level of decomposition, the next step is to prepare questions for eliciting uncertainties by characterizing each variable of the issue. In order to improve judgmental accuracy, questions should be presented to the experts without ambiguity and without the potential for preconditioned or biased responses. The problem of developing an expert's subjective probability distribution is not new, but there is no approach that could be called the only correct one. For example, questions for the respective intervals of each uncertain variable to be quantified by probabilities might be explicitly given to experts. For this, the subject must first specify values for each uncertain variable that serve as the boundaries for the intervals over the range of possible values [7,11,13]. Then, each interval is a component of a finite and exhaustive set of intervals for the possible values of each variable. In another approach, questions can require the underlying subject to respond by specifying points on the value scale while the probabilities (e.g., 5-, 50-, 95-percentiles, etc.) remain fixed. As mentioned in the previous section, the main advantage of the percentile technique lies in the fact that it fully complies with the experts' probability statements. The results of both approaches are subject to a piecewise uniform distribution, characterized as lower and upper bounds and various percentile values.

Regardless of any process, a short schematic description of each question to be quantified explicitly by the experts must be provided to help experts in deriving their estimates, in addition to the meaning of the relevant probability and uncertainty (i.e., aleatory variation or epistemic uncertainty). However, the probability for the respective interval does not yet determine the type of distribution for the underlying variable. Whenever required, the analyst can specify his/her state of knowledge prior to eliciting the expert opinions in the form of the probability distribution (e.g., normal or lognormal distribution). It would be possible to find out the influence of a particular distribution type through a distributional sensitivity analysis. Prior to the elicitation of judgments for each question, discussion of the underlying and alternative perspectives among experts can take place in the structured and controlled meetings that encourage the exploration of alternative beliefs through the exchange of relevant information while inhibiting pressure to confirm a particular response.

3.4. Elicitation of Experts' Answers for Questions

The experts are asked to make a probabilistic assessment of their familiarity with the subject matter of each of the prepared questions based on the assessment scale contained and carefully explained in the questions. Whenever necessary, these assessments could be made separately for some of the questions. The revision of experts' answers [10,13-14] can be made with formalized face-to-face meetings or with indirect communication via telephone and electronic mail. In the process, the experts are asked to debate and explain their judgments, to exchange information about the evidence base on which they rely, and to consider the probabilities of a full spectrum of alternative models. If possible, however, such interactions among experts must be minimized to avoid an experts' bias in their judgments. For example, the experts' initial answers (including their reference distribution and confidence interval) are then sent back to the experts who are given the opportunity to revise their assessments on the basis of what others have to say. The process generally converges quickly, unless an expert is convinced that he/she knows something that the others don't. The reference distribution can be obtained by combining probabilistically alternative distributions obtained from every expert into a single composite distribution. Then, the confidence interval of the reference distribution can also be obtained by a continuous connection of the upper and lower estimates of the distributions obtained from the experts. The use of both distributions is very important in light of giving a reference to experts while revising their opinions in the intermediate stage of the judgmental process.

3.4.1. Consistency and Diversity of Experts' Answers

In the process of obtaining the expert's answers, two different evaluation procedures of expert's estimates, can be taken as a basis for a consistency check with respect to the judgment of each expert in expressing his/her degree of belief: expert's answers for each uncertain variable made at the level of fundamental mechanisms and expert's answers made at the direct evaluation of the underlying uncertain issue. When the probabilistic answers for each variable are surveyed through various experts, they may be subject to a large scatter. The key contributors to the diversity of these experts' answers within an individual question as well as for all the questions characterizing the underlying issue, primarily come from various sources: (a) uncertainties of experts' knowledge and predictive capabilities with respect to a given subject, (b) normative analyst's questions that do not cover all aspects of the physical phenomena under consideration in sufficient detail, and (c) a partially uneven pattern of questions that may influence the experts' answers to some extent. If one expert's answers are inconsistent throughout the process of elicitation, his/her estimates should be discarded in the aggregation stage of experts' answers.

3.4.2. Judgmental Bias and Dependency Among Answers

Experts' answers are often subject to bias affecting the judgmental accuracy, and they may influence the experts to express a response that does not reflect their true knowledge and beliefs. There are various sources of bias that can be addressed in the judgmental process [1,10-11]: (a) the availability or ease with which the underlying question can be brought to mind, (b) the anchoring to initial information, (c) institutional affiliation, (d) educational background, (e) motivational bias and economic interest in the final outcome, etc. Therefore the effect of these diverse sources of expert bias has often been characterized as overconfidence bias and location bias, without any loss of generality. The former refers to the tendency of the experts to overstate the confidence of their beliefs about the uncertain variable more than their actual state of knowledge would permit. The latter exhibits an extent that the judgment of each expert is systematically over or under the true value. The training of experts is sometimes carried out to study and familiarize experts with their biases.

On the other hand, experts' answers are often not independent when their judgments are obtained from a number of information sources. The dependency among experts' answers is caused by, in part, shared assumptions, experience, and problem solving approaches. The dependency may also occur among an individual expert' s estimates for one more variables, even if the variables are physically independent. As a result, the assessment of inter-expert dependency allows the normative analyst to describe the extent of similarities and differences among the experts' assumptions in deriving their degrees of belief. For these reasons, it is not always practical to have feedback since it may also bias the experts' answers towards the dominant expert within the group of experts. Approaches for using expert judgment should therefore guide the normative analyst responsible for the expert judgment process, in detecting and interpreting dependencies in experts' assessments. The effect of inter-expert dependency has often been ignored in many practical situations. Regarding the impact of dependency on the precision of information, however, past studies [19,20,23] show that positive dependency among information sources can have a detrimental effect on the precision and value of the information. In the context of decision-making under uncertainty, positive dependency among information sources reduces the overall information content and a gain in precision afforded by the information. This means that a set of dependent information sources produces information that, in the aggregated result, is equivalent to that produced by a smaller number of independent sources. In other words, answers from a set of dependent experts can be expected to be less valuable than answers from independent experts with the same precisions. In case that the inter-expert dependency is high, employing a large number of experts can yield the similar level confidence as employing fewer than two independent experts. Whereas, ignoring sources of dependency may cause an underestimation of uncertainty in the aggregated results and the mean for each variable while the spread of the uncertainty band is much less [23].

3.5. Recomposition and Aggregation of Experts' Answers

The recomposition process is not necessary for an uncertainty issue that is not decomposed for direct evaluation. For the issue that was decomposed previously, however, the experts' answers for the underlying variables (i.e., probability distributions) must be recomposed to obtain an overall distribution for the issue. Recomposing the probability distributions provided by the experts for each variable of the issue can be made using direct calculations or Monte Carlo sampling. The latter approach is normally required to converge several types of distributions. After the recomposition for each issue, the results of all the experts considering the issue are aggregated to obtain a single composite distribution for the issue. One should consider that the aggregation of experts' probabilistic answers should preserve the uncertainty that exists among alternative points of view for each expert (e.g., inter-expert dependency and relative importance of each answer).

3.5.1. Rating of Expertise and Inter-expert Dependency

While an equal weight (based on an assumption equally qualified experts) is most often used in the assessment of the expertise for each expert [4], the introduction of different ratings by the normative analyst is to enable the experts to quantitatively differentiate between experts' answers to questions, on the basis of their particular fields of activity and those from adjacent fields. The spread of expert-generated measures, evidence and reasoning of estimated measures may also give a basis for assigning the relative weights for each expert: higher weights are assigned for a smaller spread or when evidence or reasoning is presented to the normative analyst. Then, the analyst' s opinion of the relative expertise of the experts is used in connection with the weighted aggregation for experts' answers for the underlying issue. The most effective approach for quantifying the extent of dependency is to use the concept of correlation coefficients. Regarding the impact of different weights for experts' judgments on the final results, a study [13] showed that the analyst' s confidence in the aggregated result (i.e., expert' s answers) is more sensitive to the analyst-assessed confidence in the experts than their inter-dependency. In that case, the increased confidence of the analyst in such experts (i.e., greater weight) is more than offset by any increase in posterior uncertainty that dependency will bring about. These results indicate that the analyst does not have to assess exactly the extent of dependency among the experts, and approximations are acceptable. Methods to address differences among experts and probabilistic dependency among experts can be found in other literature [19-20,22,23].

3.5.2. Probabilistic Aggregation of Experts' Answers

In order to aggregate probabilistically different probability distributions for the underlying issue obtained from experts into a single composite distribution, the analyst may use some mathematical rules (by either weighted average technique or Bayesian approach) or statistical approach (by Monte Carlo technique). Table 2 summarizes the different capabilities of these methods in aggregating the judgments of experts. The weighted-average approach [11,13,23] usually computes an average of the individual probabilities (i.e., average of the probabilities for each value, not the values for each probability). An example study [23] showed that averaging the probability results in an aggregated distribution with a greater spread compared favorably with that obtained by averaging the value increments in the probability distributions. The Bayesian aggregation approach explicitly attempts to incorporate expert' s bias, inter-expert dependency, and the analyst' s relative confidence in the experts [21]. With a Bayesian approach,

Table 2. Different Capabilities of Methods for the Aggregation of Expert Judgments

Approaches / Capability		Credibility (weight) of the expert' s estimate	Inter-expert dependency	Expert' s judgmental bias	Final form of uncertainty distribution
Statistical approach		O	O	X	Single composite
Simple weighted average	Fixed value, Weighted probabilities	O	X	X	Single composite
	Fixed probability, Weighted values	O	X	X	Single composite
Geometric averaging technique	Fixed value, Weighted probabilities	O	X	X	Single composite
	Fixed probability, Weighted values	O	X	X	Single composite
Bayesian approach		O	O	O	Single composite and uncertainty

probabilities assessed from experts are regarded as a new sample evidence for the analyst that is then used to update his/her prior assessments. The unique versatility of the Bayesian model and the fact that it encompasses weighted averaging techniques make it a very attractive choice for aggregating opinions. The drawback of the Bayesian model has been that there is little guidance to the practitioners on the selection of parameter values for the aggregation model. A few studies [21,23] attempted to provide practical guidance on the selection of parameters for the Bayesian aggregation model and on the use of the versatile expert judgments, by utilizing two representative approaches. The first approach uses the experts' opinions as evidence to update the decision-maker' s state of knowledge. The second approach, in recognition of the fact that the experts are highly dependent on a common information source, assumes that the common information source is the actual expert and the participants are assessing its bias and credibility. Another example study [23] showed that the posterior median obtained by the Bayesian aggregation model becomes the weighted geometric mean of the normative analyst' s prior median and the experts' estimates. In addition, the resulting posterior median obtained under the assumption of equally credible, unbiased and independent experts is within the experts' stated medians as opposed to the one obtained by the foregoing weighted-average approach (averaging probabilities) that provides a median value smaller than the stated medians of all the experts. If the normative analyst is not familiar with the use of the Bayesian aggregation model, a statistical propagation approach can be directly utilized for the aggregation. The expert' s weights and inter-expert dependencies can be limitedly taken into account in the statistical process.

3.5.3. Documentation of the Elicitation Process

In general, the expert judgment is not as technically defensible as analysis using detailed and validated codes. Thus, the reproducibility of expert' s judgmental results is regarded as an area of concern. In order to assure appropriate scrutability of the experts' estimates and further utilize the estimated results, it is necessary to document all the processes for the elicitation of experts' judgments and the relevant technical basis for experts' estimates (such as assumptions and methodologies used to obtain probabilistic answers, key factors affecting experts' bias and inter-expert dependency, strategy to improve the consistency of the elicitation process, relevant uncertainty sources).

4. The Present Characterization of Simplified Elicitation Process for Actual Practices

The first part of this section summarizes some insights obtained from careful investigations of the present formal process in light of the advantages and disadvantages of the formal elicitation process over the informal process. The following section discusses the necessity for simplified elicitation approaches under resource constraints and their potential characterization for actual use.

4.1. Two Facets of Formality

As mentioned in the previous section, key factors making a difference between formal and informal elicitation processes are (a) whether the problem-solving process well-structured and explicit, and (b) whether expert' s estimates can be reproduced by the other experts. While the former factor controls degree of available evidence, expert' s judgmental bias, and inter-expert dependency, the latter factor is controlled by the degree of documentation. If it is fully implemented, formality of the expert judgment elicitation process is valuable in the assessment of uncertain issues. Past experiences [9,13] showed that the uncertainty bands might be broader when they are estimated through the formal elicitation process than when they are estimated informally. This is mainly due to the fact that experts are often less confident of their expertise in the underlying issue when asked to reflect formally on it. In addition, avoiding the formal elicitation process may lead to expensive data collection or to elaborate problem-solving process. Additional advantages and disadvantages of the formal elicitation process over the informal process are as follows,

<u>Main Advantages</u>

- Formal elicitation process increases the understanding of issues and the drawing out of diverse judgments from experts;
- Formal elicitation process helps alleviate some experts' biases in real implementation and elicit their probabilistic expressions for uncertainty. The resultant expert judgments can remain to a large extent substantiated. In other words, the formal elicitation process gives a deserved degree of objectivity, justification, and stability to the analysis results;
- Formal elicitation process enhances the clarity of judgments by carefully probing the underlying assumptions and thinking process of the experts. As a result, it is possible to identify areas of agreement and disagreement, making it possible to identify sources of disagreement and to seek potential resolutions;
- Formal elicitation process provides the analysis' s transparency in exposing all aspects of the elicitation process.

<u>Disadvantages</u>

- The full implementation of the formal elicitation process is a cost- and time-consuming job because it requires several interactions among experts, explicit consideration of expert's bias in judgments, ratings for expertise and inter-expert dependency, and qualified documentation of results. Moreover, all steps of the present formal elicitation process cannot be satisfactorily implemented under resource constraints (such as time, budget, and number of experts);

- Formal elicitation process often fails to recognize several biases (e.g., structural bias occurring when experts are unduly influenced by the way the underlying issue has been structured before it is presented to experts, motivational bias occurring when the experts are affected by the outcome of the assessment, cognitive biases caused by overconfidence and the availability of relevant information);
- Even when formal elicitation process is used, there is a possibility that when generating judgments commonly used information can introduce bias from what the experts think that they are confident in the quantities produced.

4.2. The Present Characterization of Simplified Elicitation Processes

The above shortcomings of the formal elicitation process can be avoided to some extent in practical problems in various ways. For example, we can elicit as much from the information of the expert's problem-solving process (such as definitions and assumptions for the experts' answers). Also we can control appropriately key factors that can enter into the elicitation process and influence the expert's problem-solving process (e.g., level of the issue structuring, formal interaction among experts). Where possible, we can avoid the blind assumption about particular properties of expert opinion (e.g., an assumption that the expert answers are independent, unbiased, and equally weighted for each another). When the scope for the use of expert judgment is limited for a particular purpose such as the raw data generation from experts, the use of a less structured and simplified approach is often preferred. For a probabilistic safety study (PSA) of a nuclear power plant, for example, we want to set up a complete data set with probabilistic information of a particular event from experts. In that case, we would often employ a two-step approach [24]: (a) the first step is for discussion of an event by the experts and (b) the second step is for assessment of the event by each expert. When experts depend upon highly qualified reference information (e.g., computer code results or relevant reports), we can regard experts' credibility of the reference information as a type of uncertainty [25]. If possible to obtain the results of a highly qualified computer code, the code can be taken into account as a qualified expert because it reflects knowledge from various experts. As a result, they may greatly influence the judgment of experts and there are diminishing returns in employing a large number of experts. In that case, the human experts can take part in estimating the biases and the credibility of the computer code results. In other words, if experts rely on a common source of information, the dependency among experts would be very high and thus employing a large number of experts will yield the similar confidence in the aggregated results as employing fewer than two independent experts.

On the basis of the foregoing insights, it is possible to characterize a potentially applicable simplified approach for utilizing expert judgments in the assessment of uncertain issues when little information is available and under limited resources. The latter part of Table 3 gives the present characterization of explicit expert judgments structured at some level of the formal elicitation process, which are compared with the

Table 3. The Present Characterization of Formal and Simplified Elicitation Processes for the Assessment of Uncertainties in PSA

Formality	Major Implementation Steps	Key Features and Contents
Well-structured Formal Procedure	***Key considerations:*** greater availability of resources, consideration of expert' s bias and inter-expert dependency, disciplined use of experts, judgmental consistency, qualification of the problem solving process, high-level documentation	
5-Step Procedure (High level of formality)	(1) Selection and motivation of experts	- Formal training of experts and disciplined use
	(2) Qualification of the underlying issue	- Decomposition of uncertain issue - Presentation of uncertain issue to experts - Several meetings among experts
	(3) Qualification of questions to be elicited	- Decomposed elements, Probability distributions
	(4) Elicitation of experts' answers(1) for questions: Multiple stages of judgment revision and aggregation	- Check of consistency and diversity of experts' answers - Qualification of bias and dependency among experts - Reflection of the other experts' judgments
	(5) Recomposition and aggregation of experts' answers	- Rating of expertise and inter-experts dependency - Probabilistic/weighted aggregation - Detailed documentation of the elicitation processes
Simplified Procedures	***Key considerations:*** limited scope of the analysis, limited resources, ignorance of expert' s bias and inter-expert independence, indirect interaction among experts, qualification of information sources and estimated results, and medium or low-level documentation	
4-Step Procedure (Medium level of formality)	(1) Presentation of the underlying issue to experts	- Limited level of expert training
	(1) Elicitation of experts' answers(1) and their aggregation: The first stage of judgment elicitation	- First aggregation of uncertainty distributions - Limited number of meetings among experts
	(3) Modification and revision of experts' answers	- Reflection of the aggregated result (ignorance of expertise)
	(4) Elicitation of experts' answers(1) and their aggregation: The final stage of judgment elicitation	- Second aggregation (ignorance of expertise) - Documentation of uncertainty sources and results
2-Step Procedure (Low level of formality)	(1) Presentation of the underlying issue to experts	- Minimization of expert training (or no expert training) - Limited number of meetings among experts - Utilization of a reference value (RV) on the uncertain issue
	(2) Elicitation of experts' answers(1) and their aggregation	- Probabilistic credibility or sensitivity about the RV - Final aggregation (ignorance of expertise) - Documentation of uncertainty sources and results

Note superscript (1): given in the form of probability distributions (such as probability density function or cumulative distribution function) present characterization of the formal procedure given in the former part of Table 3. While the aforementioned formal approach is based on a full implementation of the five-step procedure for the elicitation of expert judgments described in Section 4, the two simplified approaches mainly pay great attention to the optimal utilization of limited resources in the assessment of uncertain issues. While the aforementioned formal procedure is based on higher qualification of the sources of information and of the problem solving process, the simplified approaches pay more attention to the qualification of the information sources and expert' s estimates on the uncertainty of given issues rather than the elicitation process. In other words, the simplified elicitation approaches focus on the use of explicit expert judgment structured at some level of the formal elicitation process. Even for the simplified elicitation processes, however, an appropriate level of formality is needed to assure appropriate reproducibility of the experts' estimates. This is because the high degree of subjectivity in the elicitation process makes it difficult to accept the use of expert judgments.

In general, it is very difficult to perform the relative comparison between the formal and simplified processes in a quantitative way. This is mainly because there are no generally accepted evaluation criteria for breaking the formal procedures into technically defensible simpler processes as well as there is no generally accepted framework for evaluating expert' s subjectivity involved inherently in the problem-solving and judgmental process. Perhaps the most essential difficulties in developing evaluation criteria for the relative comparison between the two different approaches stem from the current lack of an adequate understanding and formal treatment of expert biases and dependency. There was an attempt to define the suitable criteria for qualitative comparison of the existing elicitation approaches in JRC-ISIS [26], by taking into account test case-independent and dependent situations. While the former part for comparison took into account factors such as compliance, applicability, robustness, traceability, justifiability, and knowledge-integration support, the latter part characterized factors such as understandability and operability, normative and meeting efficiency, and accuracy and precision. But, little guidance seems to be currently available for practical implementation of the criteria because of the aforementioned reasons.

5. Concluding Remarks

In various engineering and decision-making problems, past practices related to the use of expert judgments would often create an essential gap between the need of the practitioners in variable situations and the applicability of the formal elicitation processes available to them. From this point of view, the primary concern of this paper was to obtain some insights into the technical basis for the balanced use of formally structured and simplified approaches for the explicit use of expert judgments under resource constraints, and to identify related decision-theoretic basis. For this, first, we characterized the five-step approach for the elicitation of expert judgments under variable situations, which is based on careful investigation of the existing formal approaches. Second, we qualitatively assessed the merits and limitations of the formal approach for the use of experts' judgments. Finally, we characterized two simplified approaches for the optimal use of limited resources in the assessment of uncertain issues, by paying more attention to the use of explicit expert judgment structured at some level of the formal elicitation process. Whereas, we did not perform the relative comparison between the present formal and simplified elicitation processes in a quantitative way, because of the current lack of generally accepted and technically defensible evaluation criteria. Thus, the present characterization of simplified approaches must be regarded as an example approach on how to develop such an elicitation process in variable situations.

If it is fully implemented, formality of the expert judgment elicitation process is valuable in the assessment of uncertain issues. However, the design of a well-structured and formalized elicitation procedure is a highly complex and time-consuming job and depends on a normative analyst' s (or expert' s) subjectivity, even with different extents of assistance from relevant experts. Moreover, its practical design may differ greater or less in terms of (a) the resource constraints available for the study, (b) goals for obtaining particular data, and (c) various factors controlling the quality of expert judgments (such as interaction and dependency among experts, treatment of bias). Different approaches posses differing degrees of advantages as well as abilities to control particular factors (e.g., bias and dependency in the expert' s judgment). While the use of the formal approach developed in a particular way is generally recommended to obtain qualified information from experts, the strict requirements for the approach cannot be fulfilled with resource constraints that are often encountered in practical situations. When a limited resource is available at the time of expert judgment elicitation, the use of a less structured and simplified approach for eliciting expert judgments is inevitable. For these reasons, many probabilistic safety studies have not always been performed with a formal pattern in every step of the analysis, but a pertinent mixture of formally structured and less structured processes. Even in these cases, however, the use of any guidance that ensures a minimally acceptable standard of expert judgment or methods that promote their convergence must not be overlooked. In addition, the principles and goals of the formal procedure must be fulfilled even for these cases. Without satisfying the above conditions, the resultant expert judgments may remain largely unsubstantiated from the viewpoint of justification. This leads to the fact that taking an appropriate balance between formally structured and less structured approaches is the best way for reasonable decision-making under resource constraints. However, little guidance or technically defensible evaluation criteria for breaking the formal procedures into simpler parts seems to be currently available, and in turn this is the reason why we tried to obtain some insights for the use of the balanced approach through this work.

Acknowledgments

This study has been carried out under the Nuclear R&D Program planned by the Korean Ministry of Science and Technology (MOST).

References

1. A. Tversky and D. Kahneman, "Judgment under Uncertainty: Heuristics and Biases," *Science*, 185, pp.1124-1131 (1974).
2. G.W. Parry, "A Discussion on the Use of Judgment in Representing Uncertainty in PRAs", *Nuclear Engineering and Design*, 93, pp.135-144 (1986).
3. E. Hofer, "On Surveys of Expert Opinion," *Nuclear Engineering and Design*, 93, pp.153-160 (1986).
4. A. Mosleh, "Hidden Sources of Uncertainty: Judgment in the Collection and Analysis of Data", *Nuclear Engineering and Design*, 93,

pp.187-198 (1986).

5. G. Volta and H. Otway, "The Logic of Probabilistic Risk Assessment versus Decision Levels", *Nuclear Engineering and Design*, 93, pp.329-334 (1986).
6. S. Kaplan, "Expert Information vs. Expert Opinions - Another Approach to the Problem of Eliciting/Combining/Using Expert Opinion in PRA," *Reliability Engineering and System Safety*, 35, pp.61-72 (1992).
7. E.J. Bonano and G. Apostolakis, "Theoretical Foundations and Practical Issues for Using Expert Judgments in Uncertainty Analysis of High-Level Radioactive Waste Disposal," *Radioactive Waste Management on Nuclear Fuel Cycle*, 16 (2), pp.137-159 (1991).
8. M.W. Merkhofer, "Quantifying Judgmental Uncertainty: Methodology, Experiences, and Insights", *IEEE Trans. on System, Man, and Cybernetics*, SMC-17 (5), (September/October 1987).
9. S.C. Hora and R.L. Iman, "Expert Opinion in Risk Analysis: The NUREG-1150 Methodology," *Nuclear Science and Engineering*, 102, pp.323-331 (1989).
10. M.A. Meyer and J.M. Booker, Eliciting and Analyzing Expert Judgment, NUREG/CR-5424, U.S. Nuclear Regulatory Commission, Washington, DC (January 1990).
11. R.L. Keeny and D. von Winterfelt, "Eliciting Probabilities from Experts in Complex Technical Problems," *IEEE Trans. on Engineering Management*, 38 (3) (August 1991).
12. U. Pulkkinen and R. Virolainen, "A Formal Expert Judgment Methodology and Its Application to the Prediction of the Results of a Fuel Coolant Interaction Experiment", PSAM-4, September 13-18, New York, USA, pp.1717-1722 (1998).
13. US Nuclear Regulatory Commission, "Severe Accident Risks: An Assessment for Five US Nuclear Power Plants", NUREG-1150, Washington, DC (1989).
14. G. Cojazzi, L. Pinola, U. Pulkkinen, "The Context of the Phase I of the Benchmark Exercise on Expert Judgment Techniques in Level 2 PSA: The JRC FARO L-24 Experiment," PSAM-4, September 13-18, New York, USA, pp.1609-1704 (1998).
15. A. Amendola, "Uncertainties in Systems Reliability Modeling: Insights Gained through European Benchmark Exercises," *Nuclear Engineering and Design*, 93, pp.215-225 (1986).
16. H.A. Linstone and M. Turoff, Eds, The Delphi Method: Techniques and Applications, Reading, MA: Addison-Wesley (1975).
17. J.S. Armstrong, W.B. Denniston, and M.M. Gordon, "The Use of the Decomposition Principle in Making Judgments," *Operational Behavior Human Perform*, 14, pp.257-263 (1975).
18. D.A. Seaver, D. von Winterfelt, and W. Edwards, "Eliciting Subjective Probability Distributions on Continuous Variables," *Organizat. Behavior, Human Perform*, 21, pp.379-391 (1978).
19. R.L. Winkler, "Combining Probability Distributions from Dependent Information Sources", *Management Sciences*, 27, pp.479-488 (1981).
20. R.T. Clemen and R.L. Winkler, "Limits fort the Precision and Value of Information from Dependent Sources," *Operational Research*, 33, p.427-442 (1985).
21. *A. Mosleh and G. Apostolakis*, "The Assessment of Probability Distributions from Expert Opinions with an Application to Seismic Fragility Curves," *Risk Analysis*, 6(4), pp.447-461 (1986).
22. R.M. Cooke, "A Theory of Weights for

Combining Expert Opinion", Report 87-25, Faculty of Mathematics and Informatics/ Computer Science, Delft University of Techology, Delft, Netherlands (1987).

23. S. Chhibber, G. Apostolakis, and D. Okrent, "On the Use of Expert, Judgments to Estimate the Pressure Increment in the Sequoyah Containment at Vessel Breach," *Nuclear Technology*, 105, pp.87-103 (1994).
24. A. Kreuser, J. Peschke, and C. Verstegen, "Practical Application of Expert Judgment in Assessing Operation Experience for Generating Reliability Data," CD-Rom Proceeding of PSAM 6, San Juan, Puero Rico, USA, June 24-28 (2002).
25. H.Glaeser, E.Hofer, M.Kloss, T.Skorek, "Uncertainty and Sensitivity Analysis of a Post-Experiment Calculation in Thermal Hydraulics," *Reliability Engineering and System Safety*, 45, pp.19-33 (1994).
26. G. Cojazzi, L. Pinola, and R. Sardella, "A Benchmark Exercise on Expert Judgment Techniques in PSA Level 2: Design Criteria and General Framework," International Topical Meeting on Probabilistic Safety Assessment (PSA '96), September 29-October 3, Park City, Utah, USA, pp.1775-1778 (1996).