

## Guest Editorial: Special Issue on Expert Judgment

R.M. Cooke  
Chauncey Starr Chair for Risk Analysis  
Resources for the Future, and  
Department of Mathematics  
Delft University of Technology  
PO Box 5031, 2600 GA, Delft  
The Netherlands  
[cooke@rff.org](mailto:cooke@rff.org)

We are indeed gratified to be able to present eight impressive papers on the subject of expert judgment. It is well-known that the calls for taking uncertainty more seriously in quantitative decision support, at all levels, become ever more persistent<sup>1</sup>. Quantifying uncertainty means, proximally and for the most part, using structured expert judgment. The qualifier "structured" means that expert judgment is treated as scientific data, albeit scientific data of a new type. Elicitation and representation of uncertainties, processing the expert judgment data, and utilization of results must be subjected to transparent methodological rules grounded in the scientific method itself. The first article announces the availability of the TU Delft expert judgment database to all researchers in this field. Three other articles in this volume illustrate the use of this data. The articles will be mentioned and briefly summarized in the order of their appearance.

### **1. *TU Delft Expert Judgment Data Base*, Roger M. Cooke, and Louis L.H.J. Goossens**

This article reviews the "classical model" for performance based combinations expert judgments and summarizes the results of 45 contracted and (for the most part) peer reviewed applications. Robustness and choice of calibration variables are discussed, and lessons learned are summarized.

### **2. *Expert Judgement Combination using Moment Methods*, Bram Wisse, Tim Bedford, John Quigley**

The authors use linear Bayes moment based techniques to derive a linear pooling of expert judgments. Performance measures adapted to this form of data are derived and compared with the classical model<sup>2</sup> on data from the first article.

---

<sup>1</sup> As examples we may mention *Science and Judgment*, National Academy of Sciences, 1994, Eckerman K.F., Leggett R.W., Nelson C.B., Puskin J.S., Richardson A.C.B. *Cancer Risk Coefficients for Environmental Exposure to Radionuclides*. Washington, DC: U.S. Environmental Protection Agency. Federal Guidance Report No.13; EPA 402-R-99-001, 1999. Circular A-4 Office of Management and Budget, Sept. 17, 2003; *Guidelines for Carcinogen Risk Assessment*, EPA/639/P-03/001F, March 2005; *Proposed Risk Assessment Bulletin*, Office of Management and Budget, released for review 2006.

<sup>2</sup> Cooke, R.M. *Experts in Uncertainty*, Oxford University Press, 1991.

***3. Informing Risk-Mitigation Priorities Using Uncertainty Measures Derived from Heterogeneous Expert Panels: A demonstration using foodborne pathogens, Sandra Hoffmann, Paul Fischbeck, Alan Krupnick, and Michael McWilliams***

This article illustrates a different approach to expert judgment. that combines elements of conventional expert elicitation protocols with formal survey methods and larger, heterogeneous expert panels. Knowing when to rely on alternative data sources is often critical for successful risk management. The authors show how a large, heterogeneous sample can support internal validation of the experts' assessments, and also of prior information that is based on limited historical data.

***4. Eliciting Conditional and Unconditional Rank Correlations from Conditional Probabilities, O.Morales, D.Kurowicka and A.Roelen***

The subject of this article is dependence elicitation. Whereas the scientific community has realized the importance of uncertainty and structured expert judgment, they have only begun to acknowledge the role of dependence in uncertainty analysis, and have yet to confront problem of dependence elicitation. The authors have broken new ground in this area, both theoretically and in applications.

***5. A Study of Expert Overconfidence, Shi-Woei Lin and Vicki M. Bier***

This article also analyzes expert judgment from the TU Delft data base, assessing the overall extent of overconfidence in the data set. Significant differences in the extent of overconfidence were found among studies, among experts, and among questions within a study. Analysis of replications (multiple realizations for the same question) allowed a preliminary conclusion that there are significant differences between experts and that an apparent 'question effect' may be due to noise rather than systematic differences in the difficulty of achieving good calibration for different questions.

***6. An Expert Judgment Paired Comparison Analysis for Aircraft Wiring Risk Assessment, Thomas A Mazzuchi, William G. Linzey, and Armin Brunin***

Aircraft wiring is illustrative of problems requiring innovative approaches. Modern aircraft contain tens of thousands of kilometers of wiring differing in e.g. gauge, insulation, conducting threads, loads, and environments. Knowledgeable experts are available, but data is very sparse, owing in part to relatively few wiring failures. Nonetheless, wiring increases with greater automation and weight constraints become more critical. Thus, more attention must be paid to reliability in wiring design. The authors apply paired comparisons to obtain severity scale values, calibrate these on some known data, and regress the scores on explanatory variables to guide wiring design decisions.

***7. Uncertainty in Mortality Response to Airborne Fine Particulate Matter: Elicitation of European Air Pollution Experts Jouni T. Tuomisto, Andrew Wilson, John S. Evans, and Marko Tainio***

This is a showcase application of the classical model to assess the health consequences of the Kuwaiti oil fires ignited by Iraq during the first Iraq war, performed under contract with the Kuwait government, for the United Nations Claims Commission. Six well-known international experts assessed the risk of exposure to fine particulate mater

(PM<sub>2.5</sub>). Their assessments were calibrated using air quality and mortality measurements from Athens and London. In general, experts agreed with published point estimates of the effects of PM<sub>2.5</sub> on mortality, but used larger confidence bounds. The equal weight decision maker was less accurate statistically and considerably less informative than the performance based combination.

### ***8. On the Performance of Social Network and Likelihood Based Expert Weighting Schemes, Roger M. Cooke, Susie, ElSaadany, Xinzheng Huang***

Health Canada representatives posed the interesting question whether weights based on measurable 'importance' of experts could reap the advantages of performance based weighting at less cost. This question could be studied using a particular implementation of social network theory in conjunction with some of the TUD database. A similar exercise was simultaneously carried with regard to likelihood weights, that are proposed from time to time.

Various expert judgment methodologies have achieved acceptance, have passed peer review and have been used "in anger", that is in real applications to real problems with real oversight committees and real stakeholders. On the other hand, none of these methods have achieved universal acceptance. Despite an effort to include the widest possible diversity of operational methods, there are significant omissions in the selection for the present volume. In particular, there is no representative of Bayesian methods, and no example of the SSHAC methodology<sup>3</sup>.

Given that expert judgment methods are still operating in the 'research mode', it was judged appropriate to ask well-known critical analysts to publicly review the articles in this volume. Prof. O'Hagan, Prof. Clemen, and Prof. French agreed to do so. The charge to reviewers was to select articles from this volume according to their own interests, and write an authored review. The authors of the articles reviewed were then given the opportunity for a short rejoinder. These reviews and rejoinders are placed at the end of this volume. They constitute a scientific contribution of the first order, and illustrate the research agenda of expert judgment methodology.

Is there a concluding verdict regarding expert judgment? The question is not, "should we or should we not use expert judgment?" Expert judgment has always been used in one form or another. The question is whether to use expert judgment well or badly, and we are gaining a better idea of what 'doing well' and 'doing badly' might mean. The state of the art is that expert judgment methods have crossed the threshold of scientific acceptance, but have not yet reached standardization. Standardization requires scientific

---

<sup>3</sup> R. J. Budnitz, D. M. Boore, G. Apostolakis, L. S. Cluff, K.J. Coppersmith, C. A. Cornell, and P. A. Morris, Recommendations for Probabilistic Seismic Hazard Analysis: Guidance on Uncertainty and Use of Experts, NUREGKR-6372, U.S. Nuclear Regulatory Commission, 1995; Budnitz, R.J., Apostolakis, G., Boore, D.M., Cluff, L.S., Coppersmith, K.J., Cornell, C.A., and Morris, P.A., "Use of Technical Expert Panels: Applications to Probabilistic Seismic Hazard Analysis," *Risk Analysis*, 18:463-469, 1998.

consensus, and scientific consensus is reached by, and only by, the free and proper operation of the scientific method. Speaking for the authors and reviewers, we hope that this volume contributes to that end.