When is a method credible?

- Proven to be accurate
- Buy-in/involved a large number of experts
- Followed an accepted method
  - Is there an “acceptable” EE method?
  - Documentation
    - Of what?
    - To what detail?
  - Transparency
- To whom does the method have to be credible?
  - Consulting/forecasting
  - Government policy

Increasing Credibility

- EE is not only an EPA activity
- How to bring together a community EE?
- Web community Wiki?
- Database of EE
  - Annotated bibliography
  - Archiving assumptions, data, models, results
  - Retrospective analysis of accuracy
- Need for life-cycle analysis of EE

Wikipedia: Expert elicitation

- In science, engineering, and research, **expert elicitation** is the synthesis of opinions of experts of a subject where there is uncertainty due to insufficient data, when such data is unattainable because of physical constraints or lack of resources. Expert elicitation is essentially a scientific consensus methodology. It is often used in the study of rare events. Expert elicitation allows for parameterization, an "educated guess," for the respective topic under study. Expert elicitation generally quantifies uncertainty.

Wikipedia

- Expert elicitation tends to be multidisciplinary as well as interdisciplinary, with practically universal applicability, and is used in a broad range of fields. Prominent recent expert elicitation applications are to climate change, modeling seismic hazard and damage, association of tornado damage to wind speed in developing the Enhanced Fujita Scale, and risk analysis for nuclear waste storage.

References fro Wikipedia

- Apostolakis, G., 7 December 1990: The concept of probability in safety assessments of technological systems. *Science*, 250 (4986): 1359-1364. DOI: 10.1126/science.2255906
- Arkes, Hal R., Jeryl L. Mumpower, and Thomas R. Stewart, 24 January 1997: Combining Expert Opinions. *Science*, 275: 461-465. DOI: 10.1126/science.275.5299.461e
- Boissonnade, A., Hossain, Q., Kimbell, J., Mensing, R., and Savy, J., 2000: Development of a probabilistic tornado wind hazard model for the Continental United States, UCRL-ID-140922 Vol. I, Lawrence Livermore National Laboratory, Livermore, CA, 131pp.
- Kerr, Richard A., 8 November 1996: Risk Assessment: A New Way to Ask the Experts: Rating Radioactive Waste Risks. *Science*, 274 (5289): 913-914. DOI: 10.1126/science.274.5289.913
- SSHAC, 1997: Recommendations for probabilistic seismic hazard analysis: guidelines on uncertainty and use of experts, NUREG/CR-6372, UCRL-ID-122160, Vol. I, Lawrence Livermore National Laboratory, Livermore, CA, 131 pp.

Question 1

- Need to frame the role, objectives, boundaries of guidance
  - Make clear the broader context of other guidance under development.
  - Avoid tendency to be too proscriptive (cookbook)
  - Consider a role for EE that is integrated into discussions about future research directions, value of future research to decisions at hand, rather than just for questions about particular quantities
  - Discussion of when EE is appropriate/not appropriate will be challenging
- Remain open to new research, applications that demonstrate benefits
  - Evolving field,
  - methods would benefit from additional research, innovations, improvements to existing approaches
- Avoid tendency of guidance to be locked in time
- Life cycle of EE

Question 2

- Generally heard yesterday that the topics covered seemed appropriate, but not comprehensive.
- Needs to be built on, updated.
- Document needs sharpening of concepts, definitions, and careful resolution of conflicting redundancies, before it is put forward as a basis for guidance.

Charge A

- Consider literature on how to gather rationales (Bruine de Bruin). Could help with peer review and transparency
- Consider other ways to encode judgments that are not quantitative probability encoding
- Literature on performance measurement, scoring is missing from document
- Protocols for ex-post evaluation of judgments to evaluate coherence including longer term (i.e. if relevant data become available) follow up of experts' performance.

Charge C.1: Selecting Experts

- Consider what the goal of the elicitation is (consensus, range of views, etc)
  - See for example, Cooke's
    - Survey
    - Political consensus
    - Rational consensus
  - Factors into evaluation of what "balance" means (e.g. role of stakeholders)
  - Factors into decisions about how to combine
- Need to clarify impact of OMB's paperwork reduction act on numbers of experts

Charge C.2: Expert Aggregation

- Generally needs more comprehensive, accurate discussion of alternative methods, their strengths and limitations
- Literature needs broadening - (e.g., Tatlock (hedgehogs v. foxes); Cooke performance based combinations) and updating (e.g., Copulas?) Bayesian model averaging?
- Need more careful discussion of when it is even appropriate to combine experts' distributions.
- Need more careful discussion of dependence, independence of experts

Charge C.3: Problem Structure

- Clearer discussion of what is appropriate for elicitation: definitions of quantity parameter, relationship (i.e. don't generally elicit parameters, avoiding second order uncertainty, dealing with resistance to giving point estimates (preferences for ranges)
  - Add clearer discussion of epistemic v. aleatory uncertainty, uncertainty v. variability
- Need discussion of importance for expert of understanding the context and for the variable being elicited, what it is conditioned on
  - Presentation of the model itself (into which the elicited value is going)
  - Influence diagrams and mental models
  - Did we discuss Bayesian belief networks ?
- Role of stakeholders in development of question

Charge C.4: Findings & Conclusions

- Append or reference specific examples of EE, other concepts
- Consider alternative tools for characterizing uncertainties, strengths and limitations (e.g. p-boxes)
- Update literature on cognitive biases - lots of new literature
- Needs fuller discussion of consistency, coherence of judgments
  - Nature, role, and appropriate use of feed back information/tools