

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

**U.S. Environmental Protection Agency (EPA)
Science Advisory Board (SAB)
Expert Elicitation Advisory Panel**

The Honorable Lisa P. Jackson
Administrator
U.S. Environmental Protection Agency
1200 Pennsylvania Avenue, N.W.
Washington, D.C. 20460

Subject: Review of EPA's Draft Expert Elicitation Task Force White Paper.

Dear Administrator Jackson:

INSERT MAJOR POINTS

Sincerely yours,

Dr. Deborah L. Swackhamer,
Chair
Science Advisory Board

Dr. James K. Hammitt, Chair
Science Advisory Board Expert
Elicitation advisory Panel

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1
2
3 The panel commends EPA for preparing a comprehensive and thoughtful white paper on
4 the potential use of expert elicitation at the Agency. The white paper was written by a
5 task force charged by the EPA Science Policy Council “to initiate a dialogue within the
6 Agency about the conduct and use of EE and then to facilitate future development and
7 appropriate use of EE methods” (p. 2). The panel judges that the white paper succeeds in
8 providing much of the information needed for the proposed dialogue and to facilitate
9 future development and appropriate use of EE. The white paper provides a
10 comprehensive introduction to EE for readers who may be unfamiliar with it and careful
11 discussion of many of the issues that must be faced if the Agency is to use expert
12 elicitation (EE) in the future. This report offers some comments on the white paper and
13 suggestions for improvement.
14
15

16 **Charge question A - background and definition of expert elicitation**
17

18 *Does the white paper provide a comprehensive accounting of the potential*
19 *strengths, limitations, and uses of EE? Please provide comments that would help*
20 *to further elucidate these potential strengths, limitations, and uses. Please identify*
21 *others (especially EPA uses), that merit discussion.*
22

23 The white paper provides a good overview of EE and issues relevant to its use by EPA.
24 We offer some suggestions for improvement.
25

26 1. The white paper does not provide a critical analysis of the strengths and weaknesses of
27 other approaches that might be alternatives to EE in particular cases, including meta-
28 analysis, peer review, unstructured expert committees (e.g., SAB, National Research
29 Council committees), and additional original research (e.g., primary data collection). The
30 panel understands that EPA is preparing another white paper on the “Hierarchy of
31 Methods for Characterizing Uncertainty” that will discuss the choice among alternative
32 methods and recommends that the EE white paper reference this forthcoming document.
33

34 In characterizing the use of EE and other methods, attention should be given to the extent
35 to which EE is a complement rather than a potential substitute. EE does not create new
36 primary data, but is a structured and rigorous process for characterizing experts’
37 understanding of the implications of existing data and models. When predicting the
38 consequences of alternative policies, it is typically necessary to extrapolate from the
39 findings of empirical studies (e.g., animal to human, epidemiological cohort to general
40 population or to sensitive subgroup, past to future). EE (and other methods for
41 incorporating expert judgment) can be used to address this extrapolation whereas other
42 methods, such as meta-analysis, generally cannot. In short, EE should be presented as a
43 useful way to organize and understand whatever is already known about a matter and to
44 identify what remains to be studied.
45

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1 2. The white paper could be improved by adopting a more neutral, analytic tone. In parts,
2 it reads too much like an advocacy document for EE.

3
4 3. The white paper should include a discussion contrasting subjective (Bayesian) and
5 objective (frequentist) probabilities. Frequentist probabilities describe the chance of
6 various outcomes conditional on a hypothesis (e.g., that data follow a standard normal
7 distribution); subjective probabilities characterize an individual's degree of belief that a
8 certain event will occur. For regulatory purposes, EPA is generally interested in
9 predicting environmental and other outcomes conditional on alternative policies; hence
10 the subjectivist interpretation is often more relevant.

11
12 Recognition of the relevance of subjective probabilities has several implications. First,
13 EPA is generally interested in the probabilities of specific (environmental, health,
14 economic) outcomes, not in whether a particular scientific model (e.g., linear no-
15 threshold dose-response function) is "correct." Hence the objective when using EE
16 should be to elicit probabilities of events that are measurable in principle, if not
17 necessarily in practice (e.g., the number of deaths that will occur in a specific year from
18 airborne PM conditional on particular regulations). In other words, EE should be used to
19 elicit judgments about quantities about which people can know the truth.

20
21 Second, since subjective probabilities measure an individual's degree of belief, different
22 individuals (including experts) may legitimately attach different probabilities to the same
23 event. There may be no "correct" probability and, in general, no method for choosing
24 among probabilities held by equally well-qualified experts. EE is a method for eliciting
25 individual experts' judgments about a matter and characterizing their knowledge using
26 probabilities.

27
28 4. Perhaps because it is a relatively transparent process, EE highlights several issues that
29 are common to many methods that could be used to obtain judgments from domain
30 experts (or other individuals). The white paper would benefit from greater
31 acknowledgment of this fact, perhaps distinguishing between issues that are common to
32 any method of eliciting judgments from individuals and those that are specific to EE. For
33 example, selection of experts is likely to be critical to any process for eliciting expert
34 judgments, whether it is a survey, an expert committee (e.g., SAB, National Research
35 Council), Delphi method, or others. Similarly, structuring the analysis and defining the
36 parameters for which probabilities are elicited is critical even when parameter values will
37 be based on literature review and the analyst's own judgment.

38
39 5. The white paper should address methods for evaluating and ensuring the quality of
40 expert judgments, including tests for coherence of judgments over multiple factors and
41 performance (tested using judgments for seed quantities, the values of which will become
42 known).

43
44 6. The white paper should reference more recent literature. A list of suggested references
45 by topic is appended to the end of this report.

1
2
3 **Charge question B – transparency**
4

5 *Transparency is important for analyses that support Agency scientific*
6 *assessments and for characterization of uncertainties that inform Agency decision*
7 *making. Please comment on whether the white paper presents adequate*
8 *mechanisms for ensuring transparency when 1) considering the use of EE*
9 *(chapter 4), 2) selecting experts (chapter 5); and 3) and presenting and using EE*
10 *results (chapter 6). Please identify any additional strategies that could improve*
11 *transparency.*
12
13

14 In general, EE is at least as transparent as most alternative methods for obtaining expert
15 judgments. Unlike committee processes, each expert provides a set of judgments about
16 the quantities that are elicited and so the degree of overlap or disagreement among
17 experts can be made readily apparent. It is argued that transparency would be further
18 enhanced by associating each distribution with the expert who provided it, but the panel
19 concludes that the disadvantages of identification (e.g., implicit biases to provide a
20 distribution consistent with an employer's or other's position) more than offset the
21 advantages.
22

23 Transparency is important to 1) identify the range of expertise; 2) identify the experts'
24 rationales for the quantitative judgments (for credibility and to decide when the effort is
25 obsolete; 3) evaluate strengths and weaknesses of the study in the future; 4) evaluate
26 credibility and increase credibility by demonstrating that the approach was applied
27 rigorously; 5) withstand litigation.
28

29 In terms of what should be transparent, it should be methodology (how the problems
30 were framed, how judgments were elicited and combined), the distributions provided by
31 the experts, characteristics of the experts, the assumptions they hold, and the specific
32 interactions between elicitor and expert.
33

34 The white paper does a good job in urging transparency in methodology, judgments, and
35 characteristics of experts (this does not requiring identifying which expert provided
36 which specific probabilities). However, the white paper should say more on how to
37 capture the assumptions and "basis of judgment" relied on by each expert. It should also
38 discuss how deepening the interactions between elicitors and experts makes it more
39 difficult to be fully transparent. In other words, the report should assess each method in
40 terms of how much of the process is intrinsically a black box. E.g., the extended in-
41 person interviews often used for EE may be the best approach in terms of getting the
42 most accurate estimates, but the interaction between the elicitor and the expert is difficult
43 to chronicle and may influence the results in ways that are difficult to identify. For
44 example, if the elicitor has to point out to the expert that the expert's responses are
45 incoherent (i.e., violate certain properties of probability distributions), this interaction

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1 and how the expert modifies his or her distributions is very difficult to chronicle. A
2 remotely-conducted Delphi, in contrast, is much easier to describe with full transparency.
3 It would be useful to consider creating a table that lists the aspects that can be easily
4 conveyed transparently and those that cannot. That is, put transparency in the context of
5 tradeoffs. In short, the more help that the experts get, the more difficult it is to document
6 and convey all of the interactions.

7
8 Although the white paper makes the correct point that the “basis for judgment” (the
9 experts’ assumptions) should be made transparent, it does not explain how this can be
10 done under conditions of fairly limited resources. It should also be emphasized that
11 putting huge resources into chronicling the interactions between elicitor and expert may
12 not be worthwhile in light of the impossibility, in many cases, to be fully transparent. The
13 suggestions for documenting seem extreme in terms of time and costs involved.

14
15
16 **Charge question C.1 – selecting experts**

17
18 *Section 5.2 considers the process of selecting of experts.*

19 *a) Although it is agreed that this process should seek a balanced group of experts*
20 *who possess all appropriate expertise, there are multiple criteria that can be used*
21 *to achieve these objectives. Does this white paper adequately address the*
22 *different criteria and strategies that may be used for nominating and selecting*
23 *experts?*

24 *b) Are there additional technical aspects about this topic that should be included?*
25

26 The panel notes that for an EE study to succeed, the experts selected must be credible, the
27 set of experts must be acceptable to stakeholders, and the process for selection should be
28 clearly documented and replicable. To enhance the quality of the results, experts should
29 have the ability to characterize their beliefs in terms of probability distributions that are
30 well-calibrated. To enhance the transparency and credibility of the study, it may be
31 helpful if the experts are capable of articulating the basis of their judgments.

32
33 Although it seems intuitive that the set of experts should span the set of reasonable
34 perspectives in the domain, the panel cautions that it is difficult to evaluate this criterion
35 in advance (e.g., to determine whether an outlying perspective is “reasonable”). Some
36 panelists are concerned that in some domains the set of reasonable perspectives may not
37 be adequately represented without including more than ten experts (hence requiring
38 approval from OMB).

39
40 The panel suggests that the white paper should distinguish criteria for establishing the
41 pool of experts from criteria for selecting from the pool, given that it is larger than the
42 number of experts required. Another approach to establishing the pool is to invite
43 stakeholders to nominate experts who meet the criteria specified above.

44
45 With regard to the question of whether EPA should conduct the expert selection process

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1 or contract it out (p. 71 of the white paper), the panel notes that contracting out may
2 increase the perception of greater objectivity in cases where EPA is viewed by some to
3 have a prior bias. This may be important for certain high-profile, highly-contested issues.
4 Whether EE is conducted by EPA or by a contractor, the process should be judged on its
5 perceived quality. This may be enhanced by prior review of the EE study design
6 (including public comment) before experts are selected.

7
8
9
10 **Charge question C.2 – multi-expert aggregation**

11 *Sections 5.4 and 6.7 present multi-expert aggregation.*

12 *a) Among prominent EE practitioners there are varied opinions on the*
13 *validity and approaches to aggregating the judgments obtained from multiple*
14 *experts. Does this white paper capture sufficiently the range of important*
15 *views on this topic?*

16 *b) Are there additional technical aspects about this topic that should be*
17 *included?*

18
19
20 The panel recognizes that there is disagreement among EE scholars about the extent to
21 which multi-expert aggregation is desirable, and the most appropriate methods for
22 aggregation when it is conducted. It offers the following remarks.

23
24 1. Whether experts' judgments are combined or not, each judgment should be reported
25 individually. This allows readers to see the individual judgments, to evaluate their
26 similarities and differences, and potentially to aggregate them using alternative
27 approaches.

28
29 2. Some form of aggregation is usually required, whether explicit or implicit. For
30 example, a policy maker cannot choose different policies conditional on which expert is
31 most accurate; she must choose a single policy. When expert judgments are obtained
32 about multiple parameters in a model, it is neither feasible nor useful to report model
33 outputs for each combination of judgments (e.g., Expert A's judgment on parameter 1,
34 Expert B's judgment on parameter 2, etc.) because the number of combinations grows
35 geometrically with the number of parameters and experts.

36
37 3. The white paper devotes inadequate attention to methods of aggregating experts'
38 judgments. It should discuss performance-based methods (e.g., Cooke's "classical
39 method" 1991) and other recent work (e.g., Jouini and Clemen 1996).

40
41
42 **Charge question C.3 – problem structure**

43 *Section 5.2.2 discusses how the problem of an EE assessment is structured and*
44 *decomposed using an "aggregated" or "disaggregated" approach.*
45

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

- 1 *a) The preferred approach may be influenced by the experts available and the*
2 *analyst's judgment. Does this discussion address the appropriate factors to*
3 *consider when developing the structure for questions to be used in an EE*
4 *assessment?*
5 *b) Are there additional technical aspects about this topic that should be included?*
6

7 The panel agrees that the problem structure must be acceptable to the experts, specifically
8 that it accords with their knowledge. It urges that the quantities for which judgments are
9 elicited be quantities that are measurable (at least in principle, if not necessarily in
10 practice). To the extent that experts use a common model that permits unambiguous
11 translation between a model parameter and a quantity that is measurable (in principle),
12 elicitation of judgments about the parameter is acceptable.

13
14 The white paper should give more attention to conditionality among the model
15 parameters. Conditionality is important for at least two reasons. First, for experts to
16 provide judgments about the value of some parameter, they must understand which of the
17 factors on which its value is conditional are specified (and their specified values) and
18 which are unspecified. Second, when experts are asked to provide judgments about
19 multiple parameters, dependencies among these parameters may be relevant; i.e., using
20 marginal distributions for multiple uncertain parameters in a model can produce
21 misleading outputs.

22
23 The “clairvoyance test,” which “demands that all of the significant assumptions and
24 conditions that could impact the expert’s response are well-specified” attempts to capture
25 the first issue, but it is inadequately articulated. A better approach is to describe the
26 measurement that one would make to determine the value of the parameter, including
27 which of the other factors would be controlled. To illustrate, consider the elicitation of
28 experts’ judgments about dry deposition velocity (ddv) in the EU-USNRC study. It is
29 known that ddv depends on at least 80 physical parameters ranging from the mean free
30 path of Brownian motion to the mixing layer of the atmosphere, but it is not know how it
31 depends on all these. The study in question distinguishes ddvs according to chemical
32 species, surface (e.g., grass, urban, skin), aerodynamic diameter, and wind speed. An
33 expert is not asked to build a model for ddv, but is asked about a potentially measurable
34 quantity, e.g., “Suppose we measure the ddv of aerosols with aerodynamic diameter of 1
35 µm on grass with wind speed 2 m/s. Please provide a probability distribution for the
36 result we will obtain.”

37
38 The expert is asked to conditionalize his uncertainty in a way that is conformable to the
39 model to which his judgment will be input. It is known that ddv can vary by an order of
40 magnitude according to the species of grass. The expert is not told the species of grass,
41 rather he is told that uncertainty arising from this factor should be “folded into his
42 distribution.” Similarly, many other variables (e.g., electrostatic potential, humidity,
43 temperature, surface roughness, insolation, wind profile) may be important. The
44 omniscient being implied by the clairvoyance test presumably would know the values of
45 these, but neither the analyst nor the expert do. Maintaining a consistent

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1 conditionalization across a large study is critical (and difficult to accomplish). For
2 example, when eliciting judgments about atmospheric dispersion and wet deposition, the
3 conditionalization must be consistent with that for dry deposition. The problem of
4 maintaining a consistent conditionalization across multiple factors is easily neglected
5 when thing about “clairvoyance.”
6

7 Problem structure and consistent conditionalization may be facilitated by use of an
8 influence diagram such as the one illustrated in Figure 6.1 of the white paper. (The panel
9 recommends that all figures be adequately labeled so they can be understood without
10 reference to additional text; Fig. 6.1 does not meet this standard.)
11

12 The white paper identifies four categories of uncertainty (parameter, model, scenario and
13 decision-rule and suggests that EE may be used to address each of them; pp. 50-51). The
14 panel suggests that scenario and decision-rule uncertainty are not suitable objects for EE.
15 Scenario uncertainty involves questions of designing the analysis; while scenario design
16 may affect experts’ judgments about quantities (because the quantity may be conditional
17 on factors that are specified by the scenario) EE is not an appropriate tool for obtaining
18 expert judgment about analytic design. Decision-rule uncertainty concerns the principles
19 that will be used to make a policy decision; this choice is one to be made by policy
20 makers subject to statute, guidance, and other applicable criteria, not by expert judgment
21 about what principles will (or should) be applied.
22

23 **Charge question C.4 & 5 – findings and recommendations**

24
25 4) *Sections 7.1 and 7.2, presents the Task Force’s findings and*
26 *recommendations regarding: 1) selecting EE as a method of analysis, 2) planning*
27 *and conducting EE, and 3) presenting and using results of an EE assessment. Are*
28 *these findings and recommendations supported by the document?*
29

30 5) *Please identify any additional findings and recommendations that should*
31 *be considered.*
32
33

34 Overall, the findings and recommendations are supported by the white paper. The panel
35 suggests that these sections should include a more balanced discussion of the strengths
36 and weaknesses of EE and compare its use with other tools. As noted above, the tone of
37 the white paper and the conclusions is too much one of advocacy for EE rather than a
38 balanced review of its advantages and disadvantages, and comparison with other
39 approaches.
40

41 An important topic that receives little attention in the white paper is that of the
42 consistency of judgments from a single expert. When an expert provides probability
43 distributions to characterize his knowledge about each of several quantities, he is
44 providing information about a multivariate probability distribution. When there are
45 dependencies among his knowledge about multiple variables, it can be very easy to report

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1 distributions that do not satisfy basic properties of multivariate distributions (e.g., that the
2 covariance matrix is positive semidefinite). This raises the question of the extent to which
3 experts' judgments should be subjected to consistency tests and how violations of these
4 tests should be rectified. Presumably experts should be made aware of these violations
5 and asked to adjust their distributions, though success in satisfying multiple consistency
6 tests may require some guidance from the elicitor or others. A danger when there is
7 extensive interaction of this type is that transparency of the process is compromised and
8 the experts' reports may be significantly influenced by the elicitor.

9
10 The literature on cognitive biases is much richer than is indicated in the white paper. In
11 addition to well-known estimation biases such as anchoring and availability heuristics,
12 there are biases relating to uncertainty perception such as probability misperception, the
13 conjunction fallacy, pseudocertainty, overconfidence, base rate fallacy, and neglect of
14 probability, all of which may distort perceptions of experts (REFERENCES). Strategies
15 to elicit expert judgments beyond familiarizing them with some of these biases merit
16 some discussion.

17
18
19 **Charge question D – development of future guidance**

20
21 *As EPA considers the future development of guidance beyond this white paper,*
22 *what additional specific technical areas should be addressed? What potential*
23 *implications of having such guidance should be considered? Do the topics and*
24 *suggestions covered in the white paper regarding selection, conduct, and use of*
25 *this technique provide a constructive foundation for developing “best practices”*
26 *for EE methods?*

27
28
29 The topics and suggestions covered in the white paper regarding selection, conduct, and
30 use of EE provide a constructive foundation for developing a description of “best
31 practices” for EE, but some parts of the white paper should be revised to incorporate
32 newer literature than is currently included (e.g., cognitive biases and elicitation of
33 quantities, methods for assessing performance of experts and aggregation of judgments
34 across experts).

35
36 In considering moving to guidance, the panel counsels that EPA be careful not to stifle
37 innovation in EE methods and that it should encourage research in the performance of
38 alternative methods. It encourages the development of guidance characterized as a set of
39 goals and criteria for evaluating success that can be met by multiple approaches rather
40 than something that will be used as a checklist or “cookbook.”

41
42
43 **Suggested additional references for inclusion in a revised White Paper**

44 [should be added to white paper and discussed, organized by topic]

45 EE studies (EU-USNRC)

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

1 EE guidance documents
2 Reliability engineering journal special issue
3 Cognitive biases in estimating quantities and probabilities
4 Other
5
6 Clemen, R. T., Fischer, G. W., and Winkler, R. L. (2000). Assessing dependence: Some
7 experimental results. *Management Science*, **46**, 1100-1115.
8
9 Clemen, R.T., and R.L. Winkler (1999), Combining Probability Distributions from
10 Experts in Risk Analysis, *Risk Analysis* **19**, 187-203.
11
12 Cooke, R.M., and L.J.H. Goossens, *Procedures Guide for Structured Expert Judgment*,
13 European Commission Directorate-General for Research, EUR 18820, 2000.
14
15 Glimcher, P.W. 2003. Decisions, Uncertainty, and the Brain: The Science of
16 Neuroeconomics. MIT Press/Bradford Press.
17
18 Hsu, M., M. Bhatt, R. Adolphs, D. Tranel, and C.F. Camerer. 2005. Neural systems
19 responding to degrees of uncertainty in human decision-making. *Science*
20 310:1680-1683.
21
22 Karlin, S. and W. J. Studden. 1966. Tchebyshev Systems: With Applications in Analysis
23 and Statistics. Interscience, New York.
24
25 O'Hagan, Anthony Caitlin E. Buck, Alireza Daneshkhah, J. Richard Eiser, Paul H.
26 Garthwaite, David J. Jenkinson, Jeremy E. Oakley, and Tim Rakow. 2006.
27 *Uncertain Judgements: Eliciting Experts' Probabilities* .Wiley.
28
29 Smith, J.E. 1990. Moment Methods for Decision Analysis. Ph.D. Dissertation, Stanford
30 University, Stanford, California.
31
32 Tucker, W.T., S. Ferson, A. Finkel, and D. Slavin (eds.) 2008. Strategies for Risk
33 Communication: Evolution, Evidence, Experience. Annals of the New York
34 Academy of Sciences, Volume 1128, Blackwell Publishing, Boston.
35

**SAB Expert Elicitation Advisory Panel 04/06/09 Draft Report to Assist Meeting
Deliberations -- Do not Cite or Quote -- This draft is a work in progress, does not reflect
consensus advice or recommendations, has not been reviewed or approved by the
chartered SAB, and does not represent EPA policy.**

References

1
2
3
4
5
6

Cooke, R.M., *Experts in Uncertainty*, Oxford University Press, 1991.

Jouini, M.N., and R.T. Clemen, "Copula Models for Aggregating Expert Opinions,"
Operations Research 44: 444-457, 1996.