Comments from Members and Consultants of the SAB Committee on Valuing the Protection of Ecological Systems and Services (C-VPESS) on the 2/15/07 draft report for discussion at the 3/06/07 C-VPESS public teleconference call
Comments received as of 8 :a.m. ET, March 6, 2007

Comments Received

A. General Comments 1
Comments from Lou Pitelka 1
Comments from Paul Slovic 2
B. Comments on Appendix A: Survey Issues for Ecological Valuation 2
Comments from Bill Ascher 2
Comments from Terry Daniel 3
Comments from Rick Freeman 5
Comments from Lou Pitelka 6
Comments from Paul Slovic 6
C. Comments on Economic Methods (Part 3, Section 4; Part 2, Section 4.3) 8
Comments from Bill Ascher 8
Comments from Paul Risser 9
Comments from Lou Pitelka 10
D. Comments on Valuation by Decision Aiding/Structured Decision Making (Part 3, Section 5.1) 11
Comments from Rick Freeman 11
Comments from Paul Risser 11

A. General Comments

Comments from Lou Pitelka

General comment 1. This applies to the entire report, not just this section. There appears to be a tendency to using the terms “ecological” and “ecosystem” interchangeably. For instance, is it “Ecosystem Benefit Indicators” as in the report, or “Ecological Benefit Indicators” as in the Schedule for our teleconferences? I am not sure that in the case of EBIs it matters, but we should be consistent to avoid confusion.

General comment 2. The sections on Economic Methods and Survey Issues for Ecological Evaluation are well written and organized, and in general easy enough for a non-expert to follow and understand.

General comment 3. There continue (from last week) to be differences among sections in whether and how examples from the literature are cited. This sometimes may be related to the size of the literature, e.g., long history of studies on surveys or stated preference methods, vs. lack of examples of the Conservation Value Method. It might be helpful to EPA staff if every discussion of a major method had a sub-section or paragraph that recommended specific studies (not review papers) as appropriate examples of the method. Consistency among section/methods in explicitly citing a few key examples add value to the report. The problem in the case of methods that have been widely applied is

in objectively selecting a few examples without appearing to advocate specific aspects of one approach over another.

Comments from Paul Slovic

On January 9, 2006 OMB released its proposed Risk Assessment Bulletin, with a stated objective to "enhance the technical quality and objectivity of risk assessments prepared by federal agencies". This Bulletin was reviewed by a special committee of the National Research Council/National Academy of Sciences, which issued its report in January of 2007. Risk assessment, while not simple to do, is arguably much easier to do than ecosystem valuation. Nevertheless, according to the Academy report, the OMB guidelines were "fundamentally flawed". The criticisms were extensive and harsh. According to Science Magazine (Jan. 19, 2007) OMB has decided not to "finalize the Bulletin in its current form". One criticism of OMB risk assessment guidelines that might also apply to ecosystem valuation recommendations, is that the costs in terms of staff resources, timeliness of completing assessments, etc. are likely to be substantial. After detailed discussion of the cost of staff resources as well as the time requirements, the academy concluded that the potential for negative impacts would be very high if the OMB proposals were implemented. I believe that similar discussion is needed in our report to examine whether the costs of applying the various ecosystem valuation methods we present are likely to exceed their benefits.

B. Comments on Appendix A: Survey Issues for Ecological Valuation

Comments from Bill Ascher

There is a great deal of value in this appendix, but it is undermined by the exaggerations in defending the survey approach. Most of the points made below are in this vein.

Page/Line

298/1st para "Survey research is a well-established and respected scientific approach…" Why should our committee endorse this claim? To be sure, some survey research is terrific, but there are also survey research methods that are not. Such a statement seems to run counter to the balanced assessment of strengths and limitations that we apply to all methods. The general tone of the piece is over the top, and at points more propagandistic than balanced. On p. 300, a Pearson correlation of .85 for predicting election outcomes is deemed "nearly perfect"—it explains 72.25% of the variance. That's not nearly perfect in my book.

302, 2nd para on cross-sectional surveys: the claim that cross-sectional surveys "can be used to test causal hypotheses" is overblown. Cross-sectional surveys provide data, as do other measurement approaches. Surveys have no particular advantage in overcoming the "correlation is not causation" obstacle except in the special case of directly asking why people did what they did, in which case the problems of response bias and over-simplification arise. The implication in the discussion of cross-sectional surveys, and the

following discussion of panel surveys, is that there is something special about surveys in being able to capture causality. The case is not convincing.

304, 2nd para The good discussions of the various potential errors that may affect panel surveys are again subjected to exaggerated reassurance, such as stating that research on biases are "reassurance for the most part and helpful because they point to the most likely ways in which conditioning might occur, thereby allowing researchers to look out for, and correct for, conditioning if it occurs." Correcting for bias presumes that in a given application, the magnitude as well as direction of bias can be gauged; this is by no means easy.

306, bottom line: "researcher" rather than "research."

312, 1st full para: "Furthermore, although many surveys manifest substantial non-response error, there is little evidence that the observed amount of nonresponse error is related to the response rate for the survey"—This statement turns the very negative assessment—indeed many surveys have a lot of response error—into yet another reassurance: that response rates are not so responsible for response error. But that it also a negative: increasing response rate does not help very much. Again, the positive spin is not justified.

314, 1st para: the plug for the forthcoming Krosnick & Fabrigar book as the salvation from the need to rely on intuition is objectionable; the message of the earlier survey-methodology researchers is that the way questions are phrased requires subtle judgment; the teaser results that follow the plug refer to rather technical issues of how to structure the scale and how to avoid cognitive overload. They do not refer to the concerns of the earlier researchers, such as how to minimize the effect of the respondent's desire to look good (different from acquiescence) or how to avoid leading questions.

329 1st para, penultimate line: "OMB" rather than "OBM"

Of a more substantive vein:

323/2nd full para: Would the psychologists on our committee members agree with the statement that "[respondents'] experiences have cumulated into a set of beliefs and attitudes stored in long term memory"?

General comment: Why the complete disregard of small-n, intensive surveys, such as the very powerful Q-method (factor analysis of forced-distribution item rankings)?

<u>Comments from Terry Daniel</u>

The surveys most relevant to the C-VPESS report are those that provide quantitative measures of public "values" or "preferences" or "attitudes" toward ecosystems and ecosystem services, and policies to protect them. It would be very helpful to have some information about how these "value" surveys (value assessments) differ in technical

design, distribution, analysis or other requirements from surveys of behavior (past, present or future intentions), knowledge, beliefs or whatever other "targets" surveys typically address. There is certainly a role for many types of surveys in EPA policy making, but there is a detectable bias against all kinds (except perhaps the CVM kind), and a lack of appreciation for how non-dollar denominated quantitative measures of "values" (or preferences, etc) could be useful to EPA.

In the above context, the committee has often discussed and argued about the difference between value assessments (including surveys) that specify a "self constituency" (what do you want, like, prefer) versus a "civil constituency" (what should we as a society want or prefer). It would be helpful to have this issue further addressed in this appendix—or at least to have examples that illustrate both types of constituency. Also, within the civil constituency mode, what do we know about the relative efficacy of different scales of response/expression, such as preference versus importance versus social acceptability? Of course the largest research base in this regard focuses on the special issues that arise when a dollar-payment "vehicle" is specified for respondents to express their judgments—but that may open several cans of worms and exceed the scope of this appendix.

The present draft tends to imply (especially for those at EPA/OMB who are already predisposed) that dollar metrics, as in CVM, are the most appropriate (or even the only appropriate) surveys for quantitatively "assessing the value of ecosystems …" C-VPESS (or at least some of us) has argued that "psychological" (personal preference, judgments of importance, etc) and "social" (acceptability or importance as a social/public policy) metrics would also be very useful to EPA policy/decision making. Granted that these metrics provide only "relative" values (interval scale at best), they are still very useful for deciding among alternatives offered in a survey and, assuming the alternatives offered are relevant to the policy question at hand, also to policy making. If the appendix could address this issue, or at least present some examples of "non-dollar" quantification in value assessment surveys, that could be very helpful (assuming the authors agree that such surveys would be useful to EPA).

The discussion of "total survey error" and much of the content of the appendix emphasizes sampling error issues—which seems to fit with the predispositions of EPA/OMB and many "consumers" of survey research. The message that some concerns about sampling error may be over-blown is enlightening and encourages greater and more sophisticated use of surveys in EPA value assessments. However, other issues, especially survey design (identifying and crafting questions, organizing them in a survey, etc) might need a little greater emphasis, as there is some real concern that too many people think that a survey is just writing down some questions and getting people's answers, and so not such a big technical deal. The current text clearly identifies survey design, including choice of rating scale formats, etc, as important issues, but a stronger statement of the importance of care and expertise in crafting surveys would help to counter some presumptions and predispositions toward surveys at EPA.

There is a substantial "qualitative analysis" movement in social science and in political science that is in some sense in opposition to what has previously been defined as rigorous quantitative survey research (and sometimes science more generally!). The proponents of qualitative analysis (focus groups, individual interviews, etc) are increasing in natural resources and environmental management domains (especially in the context of "community based planning"), and they can be quite evangelical in pushing their methods (which often entails derogation of quantitative survey methods). It would be helpful if this appendix anticipated the eventual arrival of these methods at the doorstep of EPA, and assisted analysts and decision makers in determining when and how these methods, along with surveys and economic and ecological analyses, would be most helpful—and/or what pitfalls are associated with them. Such a discussion might also help to clarify the sometimes fuzzied boundary between qualitative methods and systematic quantitative survey research, which could turn out to be a deterrent to greater use of systematic surveys at EPA.

Comments from Rick Freeman

On Appendix A:

1. I think that pp. 298-338 are relatively uncontroversial This section is probably useful for informing non-experts. But do we need it to fulfill our charge, especially since not all of our methods are based on surveys?

2. pp. 339-350 on "do preferences exist" is controversial. I don't think that this belongs in an Appendix. It probably should be integrated into an expanded version of the paragraph at the top of p. 19 (Section 2.4). And it would be more useful if it dealt more directly with Paul Slovic's arguments and evidence.

3. Similarly for pp. 351-353, it fits better in section 2.4 where this issue is first raised. But I have reservations about getting into this issue in any meaningful way. It strays way over the boundary pf our charge. And it is really about policy and social goals not science, and especially not the science of ecosystem services. Since the issue has come up, I guess we have to say something. But I recommend simply identifying the issue and saying it is beyond our charge and area of expertise.

4. On educating respondents prior to surveying them, Paul Slovic wonders if it has ever been done. I am aware of one careful attempt to do this. There is an RFF report and a subsequent journal article. See:

A. Valuation of Natural Resource Improvements in the Adirondacks ... NATURAL RESOURCE IMPROVEMENTS IN THE ADIRONDACKS Spencer Banzhaf, Dallas Burtraw, David Evans, Alan Krupnick. At http://www.rff.org/Documents/RFF-RPT-Adirondacks.pdf

B. "Valuation of Natural Resource Improvements in the Adirondacks" Spencer Banzhaf, Dallas Burtraw, David Evans, and Alan Krupnick. Land Economics | August 2006 | Vol. 82, No. 3 | pp. 445-464

Comments from Lou Pitelka

There is a tone of self-congratulation on the great accomplishments of survey research and use. It is mild and perhaps deserved but still is a different tone from most other sections.

This is an appendix, so maybe the high level of detail is okay. On the other hand, one might ask why we include an appendix with so much detail on surveys when we don't include appendices on other methods? Also, in some cases there appears to be uneven treatment of sub-topics within the document. In particular, on pages 302-306 the design of surveys is discussed. Three types of design are mentioned, but less than a page is devoted to explanations of the first two, while almost four pages are devoted to the third, panel surveys. I have no idea whether this is arbitrary or justified. It seems as though the discussion of panel surveys goes into extreme detail even for an appendix.

On page 308 I would think that most readers of our report would understand the requirements for probability (random) sampling and the problems with non-probability sampling. Thus, almost a page explaining this might be too much. Also, while I understood how systematic sampling could be a form of probability sampling once it was explain, at first I was quite confused. The term is confusing since to most ecologists it means non-random sampling. It might be clearer if these sub-categories of probability sampling were called "systematic random sampling" and "stratified random sampling" to distinguish them from non-random systematic and stratified sampling.

The section on Challenges in Using Surveys for Ecosystem Protection Valuation is perhaps the most important for our report and should not be cut.

A typo that may be important: Page 323, second paragraph under Challenges…, last line. I think the word should be "balanced", not "unbalanced".

Comments from Paul Slovic

Before offering some specific comments on the section of the report pertaining to surveys, I would like to thank Kathy, Buzz, and others who put together this extensive first draft. The careful thought and hard work that went into this draft is evident and appreciated.

This same appreciation also goes for the section on survey methodology, which contains much worthwhile information, very clearly presented. I will keep this handy as a reference on survey methods. That said, I believe much of the section on survey methods should be deleted from the report and published as a separate addendum. I say this for two reasons. First, the level of detail is far greater than that offered for other methods. It adds bulk to a report that is already so lengthy as to possibly deter many

readers. Perhaps more important, given the caveats expressed in other sections (and in my comments below) about the problems of surveying poorly informed persons about complex valuation issues pertaining to complex scientific issues, the role of surveys in ecosystem valuation may be quite limited.

There is a section on challenges to using surveys for ecosystem valuation beginning on p. 323 that recognizes some of the problems, such as the need to first inform or educate respondents. But I question whether the elaborate protocol described at the top of p. 324 has ever been done in conjunction with an ecosystem valuation exercise.

The discussion of survey issues is framed around the common use of surveys to assess attitudes and preferences regarding social and political (and perhaps consumer) issues. This carries over to the attempt to answer, on p. 339, the question pertaining to the stability of the tradeoffs for the component attributes of ecosystem valuation. The text of this answer, covering pages 339-350, seems to me to miss the point of the concern (perhaps because the question was too vague – my fault).

The issue is not about general attitudes being bounced about in small ways by question wording, question order, etc. It is about people not holding stable, well-defined tradeoffs among unfamiliar and technically complex attributes of ecosystems, such as would be necessary to impute defensible quantitative values to those attributes. Underlying this challenge is the evolution of the concept of preference construction as described in some 38 chapters and 1200 references by Lichtenstein & Slovic (2006; The Construction of Preference). We're not talking about small effects, but often full reversals of preference triggered by logically equivalent ways of framing the questions, describing the attributes, and assessing preferences.

The material on p. 351 and following pages addresses the question – why should EPA survey the public etc? The answer centers around the importance of assessing political knowledge and attitudes pertaining to political judgments. Again, I think the response misses the point. The concern regards asking for public opinions about quantitative tradeoffs pertaining to valuation issues that involve scientific expertise pertaining to ecosystems. Although the response closes on p. 353 with the assertion that EPA could educate respondents prior to surveying them, this would probably be difficult and time consuming. To my knowledge this has rarely if ever been done for ecosystem valuation.

More specifically, the discussion of question wording, question-order effects, issues of sampling error, etc. is not "tuned in" to the increased understanding of preference and valuation that has emerged in recent years. To use an example from a different context, consider my values for important aspects of prescription drugs. I want my medicines to be 1) effective, 2) safe, and 3) inexpensive. At this general, non-quantitative level, my values are strong, clear, and stable. But when it comes to tradeoffs, things begin to unravel. I don't really have a firm idea about how much additional risk of various side effects I am willing to accept for an increased probability of an improvement in various symptoms. I can make such tradeoffs if forced to do so (hopefully while being guided by a skilled decision analyst), but this tradeoff will certainly be strongly determined by how the various probabilities and consequences are described to me. And logically equivalent descriptions may change my tradeoffs greatly. And it's not a matter of right vs. wrong valuations. My "real" preferences may truly change depending upon

the way they are assessed – something akin to the Heisenberg Principal for psychological values.

A few more examples. A study by Irwin et. al. (Preference reversals for environmental values; J. Risk and Uncertainity, 1993, 6, 5-18) found that people were more willing to pay more for a specified upgrade in a consumer product (washing machine, stereo) than for a specified reduction in Denver's air pollution. However, when asked directly to compare the two improvements, they indicated that the air quality improvement was more valuable to them.

Another example of the subtlety and complexity of valuation psychology: On pages 324-325, following a paragraph on the "principles of optimal design" it is suggested that respondents could be given information in narrative form; for example, by telling respondents about the state of an ecosystem as it existed 50 years ago, changes that have occurred, and what could reverse these changes. But the value function of Prospect Theory, for which Daniel Kahneman won the Nobel Prize in 2002, predicts that describing a specific ecosystem improvement as "restoring a loss" will be valued quite differently than describing the same change as an improvement from today's status quo. Indeed this prediction was verified empirically in a study by R. Gregory, S. Lichtenstien, and D.G. MacGregor (The role of past states in determining reference points for policy decisions; Organizational Behavior and Human Decision Processes, 1993, 55, 195-206). In sum, it would take a remarkably savvy survey constructor to navigate the complex waters of preference construction for ecosystem valuations in a fair and defensible way.

C. Comments on Economic Methods (Part 3, Section 4; Part 2, Section 4.3)

<u>Comments from Bill Ascher</u>
Page/line

227/19 After the word revealed, a word I missing: probably "preference."

227/24 Compensating surplus and equivalent surplus are labeled but not explained. Why is the distinction worth making, especially in light of the fact that the labels are counter-intuitive? (I.e., as the text reads, the compensating surplus is not how much people would demand to be compensated for a loss, but rather thow much they would be willing to pay for a gain.

230/23 Underline subtitle "Status as a Method".

233/6 It seems somewhat narrow to say that travel cost is relevant only to "outdoor recreation." People may go to various places to get general improvements in weather, relief from pollution/allergens, etc.

237/1 First letter "I" is missing.

238/1 I very much appreciate that the section on the strengths and limitations of hedonics points out the problem of model specification and data dependence, because it is important to convey that revealed preference approaches are not necessarily the gold

standard—such an assessment tends to devalue other approaches, such as stated preference approaches, that can capture the effects of public-regardedness.

239/18 The discussion of averting behavior models could use an elaboration of an environmental example (maybe the water filter). There is nothing on strengths and limitations, or uncertainty. Perhaps this is the way to handle methods that are not regarded as ready for prime time, but this ought to be discussed.

241/6 Is the primary advantage of stated preference approaches their flexibility, or would it be more useful to say that their primary advantage is that they have the potential to capture non-use values? This is mentioned at 243/27, but I think it deserves more emphasis.

241/26 "Choice experiment" is mentioned twice.

242/26 This and the following paragraph say that stated preference approaches are controversial, but not how and why. To say that the concerns are over the validity of the estimates doesn't say very much, and might leave the reader (and the EPA analyst) to worry that there are unknown dragons about. The most commonly expected problem is the exaggeration bias (alluded to in the para beginning at 243/7), but not stated explicitly. The meta-analysis by Murphy et al. (James J. Murphy, P. Geoffrey Allen, Thomas H. Stevens, and Darryl Weatherhead, "A Meta-Analysis of Hypothetical Bias In Stated Preference Valuation," Environmental and Resource Economics (2005) 30: 313–325) addresses this, but is not cited. Maybe Rick has good reason to believe that their analysis is flawed, but maybe not.

247/8 The section on ecosystem benefit indicators does a good job of giving examples of the services that EBIs might measure, but does not give actual examples of EBIs. A simple list of examples would go a long way to making this section more understandable.

249/21 "Principal" rather than "principle".

249/27 It seems a bit strange that the section on uncertainties involved in using EBIs focuses on the vagaries of visual depictions of the indicators rather than the data problems and how to address them

Comments from Paul Risser

Part 3, Section 4. Economic Methods

Page	Line	Comment
We should reduce the jargon by just using ES and CS, or preferably, WTA and WTP.		
229	14	Would it be helpful to provide an estimate, even a qualitative one, on the consequences of just using the

		mean rather than incorporating a distribution of the social welfare function?
231	1	Another way of stating limitations is that selecting one variable, e.g., valuing wetlands by the single variable of commercial fish production, fails to address the more fundamental multivariate value of ecosystem studies. Moreover, this simplistic approach may inadvertently dissuade policy makers from pursuing further more complete analyses.
233	29	Do these methods address second-order contingencies, for example, lake A might have a higher preference because lake B is in the vicinity and thus adds a regional option (value)?
238	11	…is unrealistic.
244	2	Somehow brushing off this HUGE limitation in one brief paragraph in the last section, and offering only an equally challenging solution, leaves the impression of a technique far more useful than it really is.

Comments from Lou Pitelka

There are a few places where brief examples could be helpful in explaining a concept to non-economists. On page 231, what is a good example of a relevant (to ecological values) market imperfection that could distort things as discussed under Limitations? Similarly, on page 232 in the sub-section on Revealed Preference, it would be helpful to have an example of a behavioral relationship between observable choice variables and the ecosystem service. I am not sure what this means. On page 239, an example of averting behavior would be helpful. Finally, on page 247, section 4.5, water is used as an example, but I still don't have a good sense of what the method involves. Perhaps going into a little more detail would be useful.

There is an issue I remember from years ago when I was involved in discussions of some stated preference studies of pollution. I don't know if this has ever been discussed by C-

VPESS or if the experts would all agree it is a non-issue. The issue is that in such studies the measured dollar costs of willingness to pay for an environmental improvement often are fairly close (at least same order of magnitude) to costs of being willing to accept a decline in some aspect of environmental quality. It seemed totally unreal and not credible to me and, as I recall, to some of the others involved in those discussions. Thus, a person might be willing to have their taxes go up $50 per year to pay for some highly desirable environmental benefit, but at least for me, someone would have to pay me thousands of dollars to accept (live with) a decline in environmental quality of similar magnitude. Fifty dollars per year would be meaningless. Anyway, I am curious as to whether this is an issue, and if it is, whether it should be mentioned.

On page 250 in the list of bullets, it seems as though the last bulleted item actually should be a separate concluding sentence, not a bulleted item; it is not a research need but refers to the already listed needs.

D. Comments on Valuation by Decision Aiding/Structured Decision Making (Part 3, Section 5.1)

Comments from Rick Freeman

One of the main issues with this method is the role of the facilitator or "skilled analyst" in the process and the effect of the facilitator on the outcome. This point has come up in several discussions in meetings over the past several years. And I think that it was Joe Arvai who included a discussion of it in a working paper for the committee, including references to studies documenting a "facilitator effect." This needs to be discussed.

My other main comment has to do not so much with the Method description but the way the overall problem is framed. On p. 252, lines 30-31, it says, "if an objective in a given decision is to improve environmental health (emphasis added) ..." But often for EPA the objective is to improve human well-being. And this might require a decrease in environmental health in order to achieve a larger increase in some other thing that increases human well-being. The assumption of an environmental health objective is arguably inappropriate and certainly not necessary to the rest of this section.

Finally, regarding the observation that how questions are asked can influence the answers one gets, Hanemann and Kaninen have said, "Some critics of CV have asserted that one should expect the same results regardless of the elicitation procedure .... the evidence in market research as well as environmental valuation overwhelmingly rejects procedural or task invariance in Bateman and Willis, Valuing Environmental Preferences, 1999. And Carson, et al. (unpublished paper on "Incentive and Informational Properties of Preference Questions") add "... differences between estimates obtained using different elicitation formats, if predicted by economic theory, should be taken as evidence ... that respondents are taking the scenario posed seriously."

Comments from Paul Risser

Part 3. 5.1 Valuation by decision aid/structured decision making

Page	Line	Comment
251	6	The committee also uses the term "value" for absolute numbers, both parameters and estimates.
253	3	One wonders how many readers would be willing to plow through this sentence.
254	24	Is the "lake problem" the same as the estuary in the previous paragraph?
256	3	Some readers will want a stronger argument for accepting the legitimacy extrapolating from one monetized value, especially following the earlier discussion of the constructive nature of environmental preferences.