

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

Survey issues for ecological valuation: current best practices and recommendations for research

Survey methods support many of the approaches for eliciting and measuring information about values discussed in the C-VPES report. Although scientific and technical issues concerning survey design and administration can affect some aspects of ecological valuation, they are distinct from the science and value assessment issues that are the main focus of the C-VPES report.

The C-VPES recognizes, however, that issues related to survey methods are important to some methods of ecological valuation and learned they were of particular concern to EPA representatives participating in the SAB's workshop "Science for Valuation of EPA's Ecological Protection Decisions and Programs," held December 13-15, 2005. After that workshop, the committee requested that this appendix be commissioned to supplement the main body of the committee's report. This appendix provides an introduction for EPA staff to questions posed to the C-VPES pertaining to survey use for ecological valuation. It provides an overview of how recent research and evolving practice relating to those questions might assist the Agency.

Defining survey research

Survey research entails collecting data via a questionnaire from a sample of elements (e.g., individuals or households) systematically drawn from a defined population (see Babbie, 1990; Fowler, 1988; Frey, 1989; Lavrakas, 1993; Weisberg, et al., 1996). Conducting a survey involves: drawing a sample from a population; collecting data from the elements in that sample; and then analyzing the data generated. Survey research is a well-established and respected scientific approach to measuring the behavior, attitudes, and beliefs (and much more) of populations of individuals.¹ Surveys are usually done for one or more reasons:

- To document the prevalence of some characteristic in a population
- To compare the prevalence of some characteristic across subgroups in a population
- To document causal processes that produce behaviors, beliefs, or attitudes

Because scientific surveys involve probability sampling, their results can be used to estimate population parameters. This appendix addresses issues of survey methodology that cut across

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

many different applications including: monetary valuations (e.g., CVM); measures of preference, importance, or acceptability; and determinations of the assumptions, beliefs, and motives that might underlie these expression of value.

Designs of surveys

Surveys can take on a variety of designs to address various types of research questions. For example, cross-sectional surveys are useful for measuring a variable at a given point in time, whereas repeated cross-section surveys are more useful for observing change over time in a population. Panel surveys are more useful for examining change over time in a sample of respondents, and surveys that implement experiments may be more useful for establishing causality. Many types of information can be derived from the data from each of these types of surveys.

Cross-sectional surveys involve the collection of data at a single point in time from a sample drawn systematically from a population, and are often used to document the prevalence of particular characteristics in a population. Cross-sectional surveys allow researchers to assess relations between variables and differences between subgroups of respondents. Data from cross-sectional surveys can also be used to provide evidence about causal hypotheses using statistical techniques (e.g., two-stage least squares regression or path analysis; Baron & Kenny, 1986; James & Singh, 1978; Kenny, 1979) by identifying moderators of relations between variables (e.g., Krosnick, 1988) or by studying the impact of an event occurring in the middle of data collection (e.g., Krosnick & Kinder, 1990).

Repeated cross-sectional surveys involve collecting data from independent samples drawn from the same population at two or more points in time. Such data can be used to provide evidence about causality by gauging whether changes in an outcome variable parallel changes in a purported cause of it. Repeated cross-sectional surveys can also be used to study the impact of social events that occurred between the surveys (e.g., Weisberg, et al., 1995).

Panel surveys involve collecting data from the same sample of respondents at two or more points in time and can be used to gauge the stability of a construct over time and identify the determinants of stability (e.g., Krosnick, 1988; Krosnick & Alwin, 1989). Panel surveys can also be used to test causal hypotheses. This can be done by examining whether changes over

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

time in a purported case correspond to changes in an outcome variable, by assessing whether changes over time in the outcome variable can be predicted by prior levels of the purported cause, or by testing the effects of events that occur between waves (see, e.g., Blalock, 1985; Kessler & Greenberg, 1981, on the methods; see Rahn, et al., 1994, for an example).

There are a number of challenges associated with conducting panel surveys, including respondent attrition (or "panel mortality"). This occurs when some of the people who provide data during the first wave of interviewing cannot or choose not to participate in subsequent waves. Attrition reduces a panel's effective sample size and it is particularly undesirable if a non-random subset of respondents drop out. However, the literature suggests that panel attrition minimally affects sample composition (Beckett et al., 1988; Clinton, 2001; Falaris & Peters, 1998; Fitzgerald et al., 1998a; 1998b; Price & Zaller, 1993; Rahn et al., 1994; Traugott, 1990; Zabel, 1998; Zagorsky & Rhoton, 1999; and Ziliak & Kniesner, 1998 ; although see Groves et al., 2000; Lubin et al., 1962; and Sobel, 1959).

A second methodological issue in panel research is panel conditioning, or the possibility that interviewing people repeatedly may change them and thereby make the sample less representative of the larger population to which investigators wish to generalize. But again, the literature is, for the most part, reassuring. A number of studies have found either no evidence of panel conditioning effects or very small effects (Clinton, 2001; Cordell & Rahmel, 1962; Himmelfarb & Norris, 1987; Sobol, 1959; Willson & Putnam, 1982). Particularly if repeated interviews with panel members touch on a wide variety of topics, each wave may blend in with memories of prior waves via what psychologists call "retroactive interference," thus minimizing the likelihood of stimulated interest in any one topic. However, some evidence suggests that interviewing people on a particular topic may cause them to become more cognitively engaged in that topic (Bridge et al., 1977; Granberg & Holmberg, 1992; Kraut & McConahay, 1973; Willson & Putnam, 1982; Yalch, 1976; although see Mann, 2005). Other studies have documented that asking people just one question about their behavioral intentions can affect their subsequent behavior (see, e.g., Greenwald et al., 1987; Gregory, et al., 1982).

Interestingly, membership in a long-term panel survey may actually be beneficial to the quality of data collected because of "practice effects" (e.g., Chang & Krosnick, 2001). The more a person performs any task, the more facile and effective he or she becomes at doing so. In our

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

case, the tasks of interest include question interpretation, introspection, recollection, information integration, and verbal reporting (see Tourangeau, et al., 2000).

Mixed designs are used when researchers can capitalize on the strengths of more than one of these designs by incorporating elements of two or more into a single investigation. For example, a researcher interested in conducting a two-wave panel survey but concerned about conditioning effects, could concurrently administer the second-wave questionnaire to both the panel and to an independent cross-sectional sample drawn from the same population. Differences between the data collected from these two second-wave samples would suggest that carry-over effects were a problem in the panel survey.

Experiments can also be implemented in surveys to test causal hypotheses. If respondents are randomly assigned to "treatment" and "control" groups that are asked different versions of a question or question sequence, differences between the two groups can then be attributed to the treatment.

Elements of a well-defined survey

Contexts

The title of the survey, the named source/investigators, introductions and other “context” information should be chosen consciously not bias who decides to participate in the survey, or how participants interpret and answer the questions posed. For example, the same set of questions might attract a different set of participants and a different set of answers when introduced as “the EPA wants to know” versus “the American Petroleum Institute wants to know.”

Sampling

Designing a survey sample requires decisions on: a definition of the sampling frame – that is, the complete list of elements in the population to which one wishes to generalize findings – must be defined; and selection of the subset of elements – the individual unit about which information is sought – in the population to be interviewed. These decisions have important implications for the results of the survey because they may affect both coverage and sampling

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

error (see, e.g., Laumann et al., 1994). Coverage error occurs when the sampling frame excludes some portion of the population. For example, telephone surveys usually exclude households without telephones. Sampling error is the discrepancy between the sample data and the true population values that is due to random differences between the sample and the sampling frame.

There are two broad classes of sampling methods: nonprobability and probability sampling. *Nonprobability sampling* refers to selection procedures such as haphazard sampling, purposive sampling, snowball sampling, and quota sampling in which elements are not randomly selected from the population or in which some elements have zero or unknown probabilities of selection. *Probability sampling* refers to selection procedures such as simple random sampling, systematic sampling, stratified sampling, or cluster sampling in which elements are randomly selected from the sampling frame and each element has an independent, known, nonzero chance of being selected. Unlike nonprobability sampling, probability sampling allows researchers to be confident that a selected sample is representative of the population from which it was drawn and to generalize beyond the specific elements included in the sample. Probability sampling also allows researchers to estimate sampling error, or the magnitude of uncertainty regarding obtained parameter estimates. Therefore, the best survey designs (and virtually all scientific surveys) use some form of probability sampling.

Sampling error can be minimized by surveying large samples. However, the relation between sampling error and sample size is not linear. A moderate sample size reduces sampling error substantially in comparison with a small sample size, but further increases in sample size produce smaller and smaller decrements in sampling error. Thus, researchers should recognize that beyond a moderate sample size, the funds necessary to produce a large sample might be better spent reducing other types of error.

Questionnaire design

High-quality, scientific surveys typically provide respondents with several key pieces of information when introducing the survey, whether through an introductory letter, an e-mail, or an introduction from a telephone or face-to-face interviewer. This information protects respondents' rights, helping to ensure that the survey is being conducted ethically. It may also help to increase the perceived validity of the survey and, as a result, respondent participation. The introduction

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

usually includes information about the sponsor of the survey, a brief description of the survey topic, and how the data from the survey will be used. It should also include a reassurance to respondents that their survey responses will be kept confidential and describe any other measures in place to protect respondents. Finally, the burden being placed on respondents and any risks to the respondent should also be described. This information allows respondents to give informed consent. That is, knowing this information, respondents can make an informed choice about whether or not to participate in the survey. However, it is important to also keep this introduction as short as possible, because longer introductions place a greater burden on respondents and may reduce survey participation.

Survey questions. All surveys include questions, and a series of decisions must be made to achieve optimal designs of those questions. First, a researcher must decide if each question will be open- or closed-ended. For closed-ended questions, a researcher interested in obtaining rank orders of objects must decide whether to ask respondents to report those rank orders directly or to rate each object separately. If respondents are asked to rate objects, the researcher must decide how many points to put on the rating scale, how to label the scale points, the order in which response options will be offered, and whether respondents should be explicitly offered the option to say they "don't know" or have no opinion. Once the questions are written, the researcher must determine the order in which they will be administered. Researchers must also decide how to optimize measurement on sensitive topics, where social desirability response bias may lead respondents to intentionally misreport answers in order to appear more respectable or admirable. There is a large body of relevant scientific studies about the questionnaire design that, when taken together, clearly suggest strategies for designing questionnaires to maximize the quality of measurement. Although a description of the entire literature is beyond the scope of this review, we provide a few examples here about survey questions using rating scales to provide a flavor of what this literature has to offer.

When designing a rating scale, one must begin by specifying the number of points on the scale (for a review of relevant literature, see Krosnick & Fabrigar, forthcoming). For bipolar scales that have a neutral point in the middle (e.g., running from positive to negative), reliability and validity are highest for about seven points (e.g., Matell & Jacoby, 1971). In contrast, the reliability and validity of unipolar scales, with a zero point at one end (e.g., from no importance

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

to very high importance), seem to be optimized for somewhat shorter scales, approximately five points long (e.g., Wikman & Warneryd, 1990).²

A number of studies show that data quality is better when all points on a rating scale are labeled with words, rather than just some of the points (e.g., Krosnick & Berent, 1993).

Researchers should try to select labels that have meanings that divide up the continuum into approximately equal units (e.g., Klockars & Yamagishi, 1988). For example, "very good, good, or poor" is a poor choice, because the meaning of "good" is much closer to the meaning of "very good" than it is to the meaning of "poor" (Myers & Warner, 1968).³

Researchers also must decide how to order the response alternatives, and people's answers to rating scale questions are sometimes influenced by this order. With most rating-scale questions, respondents are likely to begin to formulate a judgment after reading the question stem. For example, the question, "How effective do you think the cleanup plan will be?" would induce respondents to begin to generate an assessment of effectiveness. As respondents read or listen to the answer choices presented, some may settle for the first acceptable response option they encounter rather than considering all the response options and selecting the answer choice that best reflects their judgment. This results in primacy effects in ratings, which have been observed in many studies (e.g., Belson, 1966; Carp, 1974; Chan, 1991; Matthews, 1929). To minimize bias, it is usually best to rotate the order of response choices across respondents and to statistically control for that rotation when analyzing the data. [NOTE 33]

Pretesting. Even the most carefully designed questionnaires sometimes include items that respondents find ambiguous or difficult to comprehend, or items that respondents understand but interpret differently than the researcher intended. Researchers can conduct pretests of a draft questionnaire to identify these kinds of problems. Pretesting methods include conventional pretesting, in which interviewers conduct a series of interviews and report any problems with question interpretation or comprehension (see, e.g., Bischooping, 1989; Nelson, 1985); behavior coding, in which a researcher notes the occurrence of verbal events during the interview that might indicate problems with a question (e.g., Cannell, et al., 1981); and cognitive interviewing, in which a questionnaire is administered to individuals who either "think aloud" while answering or answer questions about the process by which they formulated their responses (e.g., Forsyth & Lessler, 1991). Each of these methods has advantages and disadvantages. When resources are

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

available, researchers can use multiple methods to pretest questionnaires because different methods identify different types of problems (see Presser et al., 2004).

Mode of data collection

Survey data can be collected in one of four primary modes: mail, telephone, face-to-face, and Internet. Interviewers administer telephone and face-to-face surveys, whereas mail and Internet surveys involve self-administered questionnaires. Mode choice can produce notable differences in survey findings. So mode choice must be made carefully in light of each project's goals, budget, and schedule. Each survey mode has advantages and disadvantages. When choosing a mode for a particular survey, researchers must consider cost, characteristics of the population, sampling strategy, desired response rate, question format, question content, questionnaire length, length of the data-collection period, availability of facilities, the purpose of the research, and the resources available to implement it.

Aspects of the population, including literacy, telephone coverage, and familiarity with and access to computers, are important in the decision about mode. Literacy is necessary for self-administered questionnaires. Broad telephone coverage of the population is necessary when conducting a telephone survey. Internet access and familiarity with computers is important for an Internet survey.

Coverage error is minimized in face-to-face household surveys, but is larger in random digit dial (RDD) telephone household surveys, because they exclude respondents without telephones and those with only cell phones. Coverage error for mail and Internet surveys depends upon the sampling strategy used and with list samples, the quality of the list that is used as the initial sample frame.

Although probability sampling is possible in all modes, mode affects the ease with which it can be implemented. Telephone and face-to-face surveys routinely use probability household sampling strategies, but mail and other self-administered surveys are more commonly used when a list of the entire population is available. In some Internet surveys, nonprobability sampling methods are used (e.g., inviting individuals to opt in through websites). This does not yield results that can be generalized to the population of interest (Malhotra & Krosnick, in press). Some researchers, however, have implemented probability sampling to recruit respondents to

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

complete questionnaires weekly via the Internet and provided Internet access to respondents who do not have it.

Mode also influences the response rates achieved in a survey, with face-to-face surveys typically achieving the highest response rates. Telephone surveys achieve somewhat lower response rates, and self-administered mail surveys achieve low response rates unless a sequence of multiple contacts are implemented at considerable cost and with considerable implementation time (see Dillman, 2006).

The types of information and questions researchers wish to present may also influence the choice of mode. If a survey includes open-ended questions, face-to-face or telephone interviewing is preferable because interviewers can probe incomplete or ambiguous respondent answers. If complex information will be presented as part of the survey, face-to-face interviewing or Internet questionnaires allow the presentation of both oral and visual information. If the researcher needs to ask questions about sensitive topics, self-administered questionnaires and computers provide respondents with a greater sense of privacy and therefore elicit more candid responses than interviewer-administered surveys (e.g., Bishop & Fisher, 1995; Cheng, 1988; Wiseman, 1972). Face-to-face interviewing is likely to elicit more honest answers than telephone interviewing because face-to-face interviewers can develop better rapport with respondents and more easily implement private response methods.

Face-to-face data collection permits interviews of an hour or more, whereas telephone interviews usually last no more than 30 minutes. With self-administered questionnaires, response rates typically decline as questionnaire length increases, so they are generally kept even shorter.

Telephone and Internet surveys can be completed in very short field periods, often within a matter of days (though at the cost of lower response rates). In contrast, mail surveys require significant amounts of time, and follow-up mailings to increase response rates further increase the overall turnaround time. Similarly, face-to-face interview surveys typically require a substantial length of time in the field.

Face-to-face interviews are usually considerably more expensive than telephone interviews, which are usually about as expensive as self-administered questionnaire surveys of comparable size using methods necessary to achieve high response rates. The cost of Internet

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

data collection from a probability sample is about equivalent to that of telephone RDD interviewing.

These differences between modes also contribute to differences in data quality. Face-to-face surveys have the highest response rates, are the most flexible in terms of interview length and presentation of complex information, and acquire more accurate reports than do telephone surveys (Holbrook et al., 2003). Internet surveys allow presentation of complex information, and reporting accuracy appears to be higher in Internet surveys than in telephone surveys (Chang & Krosnick, 2001). Although response rates from Internet surveys based on initial RDD telephone samples are quite low and have similar coverage error to telephone surveys, such difficulties may be reduced by recruiting probability samples of respondents face-to-face in their homes.

Assessing survey accuracy

In order to optimize survey design or to evaluate the quality of data from a particular survey, the accuracy (or conversely error) in survey data needs to be assessed. If optimal procedures are implemented a high level of accuracy can be achieved, but departures from such procedures can compromise the accuracy of a survey's findings. Usually, researchers have a fixed budget and must decide how to allocate those funds in order to maximize the quality of their data. The "total survey error" approach enables researchers to consider survey design issues within a cost-benefit framework that allows them to maximize data quality within budget constraints (cf. Dillman, 1978; Fowler, 1988; Groves, 1989; Hansen & Madow, 1953; Lavrakas, 1993).

The total survey error perspective recognizes that the goal of survey research is to accurately measure particular constructs in a sample of people who represent the population of interest. In any given survey, the overall deviation from the ideal is the cumulative result of several sources of survey error. The total survey error perspective disaggregates overall error into four components: coverage error, sampling error, nonresponse error, and measurement error. Coverage and sampling error are described above. Nonresponse error is the bias that can result when data are not collected from all members of a sample. Measurement error refers to all distortions in the assessment of the construct of interest, including systematic biases and random variance that can be brought about by respondents' own behavior (e.g., misreporting true

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

attitudes), interviewer behavior (e.g., misrecording responses), and the questionnaire (e.g., ambiguous or confusing question wording).

Nonresponse occurs when data are not collected from all of the eligible sample elements. Nonresponse occurs either because sampled elements are not contacted (e.g., no one is available at the time of contacted) or because members of sampled households decline to participate. The response rate for a survey is the proportion of eligible sample elements from whom data were collected and is almost always less than 100%. Lower response rate increase the risk that the sample is not representative of the population.

To maximize response rates, researchers implement various procedures. For example, the field period during which potential respondents are contacted can be lengthened (e.g., Groves & Lyberg 1988; Keeter et al. 2000), the number of times an interviewer tries to contact a household member can be increased (Merkle, et al., 1993; O'Neil, 1979), financial incentives can be offered for participation (e.g., Singer et al., 1999; Singer, et al., 2000), advance letters can be mailed to households to inform residents about the survey (e.g., Camburn et al., 1995; Link & Mokdad 2005), and the questionnaire can be kept as short as possible (e.g., Collins et al. 1988). All of these strategies have been found to increase response rates in at least some studies in which these factors were considered one by one. However, some strategies, such as sending advance letters or leaving messages on answering machines of potential respondents, may not always be successful because they give advance notice that interviewers will try to contact respondents, and respondents may use this knowledge to avoid being interviewed.

Low response rates increase only the potential for nonresponse error, because nonresponse error is a function of two variables: the response rate and the size of the difference between respondents and nonrespondents. If respondents and nonrespondents do not differ substantially, response rates will be unrelated to nonresponse bias. That is, it is possible to conduct a survey with a response rate of 20 percent and end up with data that describe the population quite accurately.

A number of publications using a variety of methods have shown that as long as a representative sample is scientifically drawn from the population and professional efforts are made to collect data from all potential respondents, variation in response rates (between 20

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

percent and 65 percent) does not substantially increase the accuracy of the survey's results (Curtin et al., 2000; Holbrook et al., in press; Keeter et al., 2000). Furthermore, although many surveys manifest substantial nonresponse error, there is little evidence that the observed amount of nonresponse error is related to the response rate for the survey.

Measurement error includes any distortion or discrepancy between the theoretical construct of interest and the concrete measurement of that construct. One method for assessing measurement error is to compare responses to a survey to a known standard to assess their validity. For example, reports of whether or not a respondent voted in an election can be compared to public records of voting, or reports of drug use can be compared to the results of drug tests performed on hair, urine, or saliva samples. However, surveys often measure constructs for which there are no available standards. In these cases, the reliability or predictive validity of survey measures is often used to judge the quality of the measurement. One method for comparing different survey questions or question orders is to use split-ballot experiments. In these, half the respondents are randomly assigned to receive one form of a questionnaire (using one question wording or order) and the other half are randomly assigned to receive a different form of the questionnaire (using a second question wording or order). One or more of the approaches described above (e.g., comparison to a known standard, reliability, or predictive validity) can then be used to compare the reliability and/or validity of responses across questionnaire forms to determine if one question wording or order is better.

The total survey error perspective advocates explicitly taking into consideration each of these four sources of error and making decisions about the allocation of resources with the goal of reducing the total error. Many steps that do not cost real dollars can be taken to reduce error, but other steps to reduce error do cost money, and the more money spent on reducing one type of error, the less money is available to reduce other types of error. Researchers should make such tradeoffs explicitly, recognizing the opportunity costs they pay when making a particular move to maximize quality in a particular way, selecting approaches likely to yield the greatest overall impact

Challenges in using surveys for ecosystem protection valuation

Introduction. One application of the survey method is in assessing the value of

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

ecosystems and services. A variety of techniques have been developed to assess the monetary value of ecosystems, and these values can be used as input to required benefit-cost analyses by EPA in the policy-making process. When monetary values are not required, are too difficult to attain, or are deemed ethically or otherwise inappropriate to the problem at hand, surveys can be used effectively to determine quantitative measures of preference, importance, or acceptance of alternative policies, actions, and outcomes. When surveys are used for valuation, many respondents are asked to rank, rate, or place a monetary value on a change in the condition of ecosystems or services with which they may not be familiar prior to the survey. However, this does not mean that respondents lack a value for the ecosystem in question. Respondents' experiences have cumulated into beliefs and attitudes that influence their orientation toward objects or situations they encounter. Therefore, an important component of valuation survey design is to describe the ecosystem as fully as possible so that respondents can use these beliefs and attitudes to determine its value. Doing so helps to maximize the extent to which the values that respondents report validly reflect these underlying beliefs and opinions. This means that valuation surveys will be different from most other surveys because they must devote a considerable amount of time to educating the respondent about the ecosystem in question. This may require respondents to listen to or read relatively long passages of text and perhaps to observe visual presentations of nonverbal information as well, such as charts, maps, drawings, or photographs.

Conveying a large amount of information. The survey needs to provide all the information that respondents want in order to make the judgments being asked of them, and present that information in a way that is understandable to all respondents. To achieve these goals, researchers can conduct research with pretest respondents to assess what information they want to know and their understanding and interpretation of information presented to them. These procedures can be used iteratively to refine the presentation to enhance understanding and sufficiency of the information set.

In order to present a sizable set of information to respondents, a variety of techniques can be implemented to maximize comprehension. The principles of optimal design can be used to construct graphical displays of information (e.g., Kosslyn, 1994; Tufte, 2001). A single visual display can convey a great deal of information if an interviewer can explain or the respondent

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

has the opportunity to study the display. Information can also be presented in narrative form – for example, by telling respondents about: the state of an ecosystem as it used to exist 50 years ago; changes that have occurred to the ecosystem in the intervening years; the causes of those changes; what could be done to reverse those changes; and how this could be implemented. Instead of a long lecture, a questionnaire can maintain respondent engagement by presenting information in small chunks, separated by questions that allowing them to react briefly to the information they've been given (e.g., "Had you ever heard of the Golden River before today?"). As the story progresses, respondents can also be asked periodically to verbalize any information that they would like to have, to allow them to express their cognitive responses to the presentation.

The choice of survey mode also has an impact on the presentation of information about an ecosystem. Face-to-face interviewing is optimal because it allows any type of visual display and interviewers can create a strong sense of interpersonal connection with respondents. Telephone interviewing permits a similar connection, though probably less strongly, and visual displays are usually not possible. Computer administration of a questionnaire can include static and dynamic presentation of visual and verbal information, and questions can be interspersed with this information, but it may not be possible to create the strong sense of connection between the respondent and the researcher. Self-administered paper and pencil questionnaires allow only visual presentation of information and do not allow information to be presented in small chunks (because respondents can look ahead in the questionnaire). A large volume of information presented densely on a many pages may be intimidating or dispiriting, thus minimizing respondent motivation and provoking superficial processing of the information. For this reason, the self-administered mode may be the least desirable. For all modes, it is important to pretest the final instrument to be sure it is working as intended.

Communicating uncertainty. Because of the uncertainty inherent in estimating the effect a policy might have on an ecosystem or service, researchers using surveys for valuation may not only want to convey large amounts of information to respondents, but they may also want to convey their level of certainty or uncertainty about that information. Such uncertainty could be conveyed to respondents in a number of ways, including: providing ranges or confidence intervals for the information provided (e.g., the estimated cost of maintaining the ecosystem is

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

between 1 million and 3.3 million dollars per year); providing a verbal description of scientists' confidence in the information (e.g., scientists are very confident that a policy will protect an ecosystem); communicating the degree of consensus about the information among scientists (e.g., 75 percent of scientists agree that a particular policy will protect the ecosystem); or conveying the probability that an outcome or benefit will occur (e.g., scientists believe this policy has a 75 percent probability of protecting the ecosystem). There is substantial evidence that people have difficulty accurately interpreting this last type of evidence (e.g., Tversky & Kahneman, 1974), but the EPA may want to explore these various methods for conveying uncertainty to determine the extent to which people understand and use different types of information about uncertainty in valuation.

Scale and spatial issues. Because the spatial and temporal scale of ecological systems and services may affect valuation processes, these dimensions should be incorporated into the communication of information and the measurement of value. For example, the information that respondents receive during the survey interview should, if possible, explicitly describe the scale of a proposed policy or the ecosystem or service for valuation. This is particularly true if the scale is fixed and can be described consistently across presentation of information, evaluation of policies, and valuation of ecological systems and services. In other cases, the physical or temporal scale may be variables of interest, so researchers may want to measure whether these features affect respondents' evaluation of the policy. This could be accomplished by manipulating the physical or temporal scales of a proposed policy (either between- or within-subjects) to determine whether and how these features influence support for the policy.

Transfer issues. The most effective way to use surveys for valuation applicable to a particular ecosystem is to use a survey tailored specifically to that situation. However, this requires that time and material resources be devoted each time EPA must complete a value assessment. A more efficient approach might be to design studies to test whether the findings from a survey about one set of environmental conditions can be extrapolated to a different set of environmental conditions. For example, if a survey measures the ecosystem values affected by one oil spill, would it be possible to multiply these losses by three to anticipate the comparable losses caused by three comparable oil spills to three comparable ecosystems? Even if such extrapolations must be done using more complex transformations, it may be possible to conduct

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

parametric research to ascertain how such predictions can be made.

Implementing survey research at EPA. Whatever the value measure being sought, the design and conduct of surveys is best done when informed by the literature on survey methods. Therefore, it is important that EPA surveys be implemented at least partly by individuals who are well-versed and up-to-date in this literature. This is probably best accomplished by teams of researchers composed partly of EPA employees who specialize in surveys and outside consultants who are experts in survey methods. EPA may, therefore, want to assess its current capacity to conduct or oversee contractor design and implementation of high-quality surveys.

OMB clearance is required for all EPA surveys, and achieving this clearance requires that a survey meet high standards of quality. In order to maximize the likelihood of approval, it is important that a proposed survey meet a set of criteria:

- Representative sampling of the population of interest with minimal non-coverage error
- A very high response rate or a plan to assess the presence of nonresponse bias
- A measuring instrument that has been developed according to optimal design and pretesting practices
- A measurement approach for which a body of empirical evidence documents validity

Probability sampling is relatively easy to do for general population samples, but more challenging for smaller, more specific subpopulations which require specialized sampling procedures currently under development (e.g., Blair & Blair, 2006; Rocco, 2003). If EPA is interested in conducting surveys of such specialized subpopulations, it may be of value to commission a group of sampling statisticians to develop a series of guidelines that can be consulted and followed when conducting sampling for such studies.

The recent literature on response rates has focused on exploring the impact of response rates on data accuracy and exploring the effectiveness of various data-collection techniques for enhancing response rates. Although lower response rates are generally not associated with substantially decreased accuracy, it may be useful for EPA to reanalyze a set of its own past surveys simulating lower response rates and observing the impact on the survey results. If systematic bias is detected, it may be possible to build correction algorithms to adjust the results

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological Systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

of future surveys to correct for such bias.

It might seem obvious that when EPA conducts surveys, all possible steps should be taken to increase response rates. According to federal convention, that cannot include offering financial incentives to respondents, but EPA can implement other techniques to enhance response rates, including lengthening the field period during which data are collected, and more attempts to contact potential respondents. However, to justify resources to implement such techniques, it is important to have empirical evidence documenting the effectiveness of these techniques for EPA surveys. It is also important to be sure that efforts to increase the response rate of a survey do not inadvertently decrease the representativeness of the sample. For example, telling respondents that a survey is about the environment may increase response rates among people interested in the environment and may decrease response rates by a smaller margin among less-interested people, thus increasing nonresponse bias. EPA may want to conduct studies assessing whether efforts to increase response rates unintentionally decrease sample representativeness.

Another approach to facilitating OMB approval may be to gather evidence documenting the effectiveness of particular measurement techniques. For example, there is considerable controversy surrounding the use of contingent valuation (CV) methods in surveys. However, a National Oceanic and Atmospheric Administration blue ribbon panel concluded that CV is a viable method of valuation (Arrow et al, 1993). It may be of value for EPA to identify the optimal elements and implementation of a CV survey and to assess the validity of CV measurement in surveys by comparisons with other monetary measures (e.g., from revealed-preference studies) or with measures based on judgments of preference, importance, or acceptability. This same sort of developmental work can be conducted with other valuation techniques such as conjoint analysis, about which there is little consensus (e.g., Dennis, 1998; Stevens, et al., 2000; Wainright, 2003). This may help to reassure OMB evaluators of the merit of value measurements produced by the various methods when they are implemented well. EPA could also consider conducting research comparing the validity of value assessments by these and other techniques to identify the technique(s) that yield the most valid data.

Finally, new OMB guidelines on surveys suggest that when a survey is expected to obtain a relatively low response rate, investigators should plan to implement techniques to assess

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

sample representativeness. Rather than outlining what such procedures would look like, OMB has left it to investigators to propose and justify such techniques. EPA could therefore commission work to design procedures for this purpose and conduct studies to validate the effectiveness of the procedures.

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

References

- Adorno, T. W., Frenkel-Brunswick, E., Levinson, D. J., & Sanford, R. N. 1950. *The authoritarian personality*. New York: Harper.
- Ahlgren, A. 1983. Sex differences in the correlates of cooperative and competitive school attitudes. *Developmental Psychology* 19: 881-888.
- Babbie, E. 1990. *Survey research methods*. Belmont, CA: Wadsworth.
- Baron, R. M. & Kenny, D. A. 1986. The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology* 51:1173-1182.
- Beckett, S., Gould, W., Lillard, L., & Welch, F. 1988. The PSID after fourteen years: An evaluation. *Journal of Labor Economics* 6: 472-92.
- Belson, W. A. 1966. The effects of reversing the presentation order of verbal rating scales. *Journal of Advertising Research* 6: 30-37.
- Bischooping, K. 1989. An evaluation of interviewer debriefing in survey pretests. In C.F. Cannell, L. Oskenberg, F. J., Fowler, G. Kalton, & K. Bischooping (Eds.), *New techniques for pretesting survey questionnaires* (pp. 15-29) Ann Arbor, MI: Survey Research Center.
- Bishop, G. F. & Fisher, B. S. 1995. "Secret ballots" and self-reports in an exit-poll experiment. *Public Opinion Quarterly* 59: 568-588.
- Blair, E. & Blair, J. 2006. Dual frame Web/telephone sampling for rare groups. *Journal of Official Statistics* 22: 211-220.
- Blalock, H. M. 1985. *Causal models in panel and experimental designs*. New York: Aldine.
- Bracht, G. H. & Glass, G. V. 1968. The external validity of experiments. *American Education Research Journal* 45: 437-474.
- Bridge, R. G., Reeder, L. G., Kanouse, D., Kinder, D. R., Nagy, V. T., & Judd, C. M. 1977. Interviewing changes attitudes – sometimes. *Public Opinion Quarterly* 41: 56-64.
- Camburn, D., Lavrakas, P. J., Battaglia, M. J., Massey, J. T. & Wright, R. A. 1995. Using advance response letters in random-digit-dialing telephone surveys. *Proceedings of the American Statistical Association, Section on Survey Research Methods* pp. 969-974. Washington, D.C.: American Statistical Association.
- Cannell, C. F., Miller, P. & Oksenberg, L. 1981. Research on interviewing techniques. In S. Leinhardt (Ed.), *Sociological methodology* (pp.389-437). San Francisco: Jossey-Bass.
- Cantril, H. 1951. *Public opinion 1935-1946*. Princeton, NJ: Princeton University

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

Press.

- Carp, F. M. 1974. Position effects on interview responses. *Journal of Gerontology* 29: 581-587.
- Chan, J. C. 1991. Response order effects in Likert-type scales. *Educational and Psychological Measurement* 51: 531-541.
- Chang, L., & Krosnick, J. A. 2001. The accuracy of self-reports: Comparisons of an RDD telephone survey with Internet Surveys by Harris Interactive and Knowledge Networks. Paper presented at the American Association for Public Opinion Research Annual Meeting, Montreal, Canada.
- Cheng, S. 1988. Subjective quality of life in the planning and evaluation of programs. *Evaluation and Program Planning* 11: 123-134.
- Clausen, A. R. 1968. Response validity: Vote report. *Public Opinion Quarterly* 32: 588-606.
- Clinton, J. D. 2001. Panel bias from attrition and conditioning: A case study of the Knowledge Networks Panel. Stanford.
- Collins, M., Sykes, W., Wilson, P., & Blackshaw, N. 1988.. Nonresponse: The UK Experience. In R. M. Groves, P. B. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, & J. Waksberg (Eds.), *Telephone Survey Methodology* (pp. 213-232).
- Cooper, W. E. & Ross, J. R. 1975. World order. In R.E. Grossman, L. J. San, and T. J. Vance (Eds.), *Papers from the parasession on functionalism* (pp. 63-111). Chicago: Chicago Linguistic Society.
- Cordell, W. N. & Rahmel, H. A. 1962. Are Nielsen ratings affected by non-cooperation conditioning or response error? *Journal of Advertising Research* 2: 45-9.
- Couch, A. & Keniston, K. 1960.. Yeasayers and naysayers: Agreeing response set as a personality variable. *Journal of Abnormal Social Psychology* 60: 151-174.
- Curtin, R., Presser, S., & Singer, E. 2000. The effects of response rate changes on the index of consumer sentiment. *Public Opinion Quarterly* 64, 413-428.
- Dennis, D. F. 1998. Analyzing public inputs to multiple objective decisions on national forests using conjoint analysis. *Forest Science* 44, 421-429.
- Dickinson, T. L. & Zellinger, P. M. 1980. A comparison of the behaviorally anchored rating and mixed standard scale formats. *Journal of Applied Psychology* 65, 147-154.
- Dillman, D. A. 1978. *Mail and telephone surveys: The total design method*. New York: Wiley.
- Dillman, D. A. 2006.. *Mail and internet surveys: The tailored design method*. New York: Wiley.

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

- Ebel, R. L. 1982. Proposed solutions to two problems of test construction. *Journal of Educational Measurement.*, 19: 267-278.
- Eiferman, R. R. 1961. Negation: A linguistic variable. *Acta Psychologica* 18:258-273.
- Falaris, E. M. & Peters, H. E. 1998. Survey attrition and schooling choices. *The Journal of Human Resources* 33: 531-54.
- Fitzgerald, J., Gottschalk, P., & Moffitt, R. 1998a. An analysis of sample attrition in panel data: The Michigan Panel Study of Income Dynamics. NBER Technical Working Papers National Bureau of Economic Research, Inc.
- Fitzgerald, J., Gottschalk, P., & Moffitt, R. 1998b. An analysis of sample attrition on the second generation of respondents in the Michigan Panel Study of Income Dynamics. *The Journal of Human Resources* 33: 300-344.
- Fitzsimons, G. J. & Morwitz, V. G. 1996. The effect of measuring intent on brand-level purchase behavior. *The Journal of Consumer Research* 23: 1-11.
- Forsyth, B. H. & Lessler, J. T. 1991. Cognitive laboratory methods: A taxonomy. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, & S. Sudman (Eds.), *Measurement error in surveys* (pp. 393-418). New York: Wiley.
- Fowler, F. J. 1988. *Survey research methods* (2nd ed.). Beverly Hills, CA: Sage.
- Frey, J. H. 1989. *Survey research by telephone* (2nd ed.). Newbury Park, CA: Sage.
- Granberg, D. & Holmberg, S. 1992. The Hawthorne Effect in election studies: The impact of survey participation on voting. *British Journal of Political Science* 22: 240-247.
- Gregory, W. L., Cialdini, R. B., & Carpenter, K. M. 1982. Self-relevant scenarios as mediators of likelihood estimates and compliance: Does imagining make it so? *Journal of Personality and Social Psychology* 43: 89-99.
- Greenwald, A.G., Carot, C. G., Beach, R., & Young, B. 1987. Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology* 72:315-318.
- Groves, R. M. 1989. *Survey errors and survey costs*. New York: Wiley.
- Groves, R. M. & Lyberg, L. E. 1988. An overview of nonresponse issues in telephone surveys. In R. M. Groves, P. B. Biemer, L. E. Lyberg, J. T. Massey, W. L. Nicholls II, & J. Waksberg (Eds.), *Telephone Survey Methodology* (pp. 191-212).
- Groves, R. M., Singer, E., & Corning, A. 2000. Leverage-salience theory of survey participation: Description and an illustration. *Public Opinion Quarterly* 64: 288-308.
- Guilford, J. P. *Psychometric Methods*. McGraw-Hill, New York, 1954
- Hansen, M. H. & Madow, W. G. 1953. *Survey methods and theory*. New York:

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

Wiley.

- Hastie, R. & Park, B. 1986. The relationship between memory and judgement depends on whether the judgment task is memory-based or on-line. *Psychological Review* 93: 258-268.
- Himmelfarb, S. & Norris, F. H. 1987. An examination of testing effects in a panel study of older persons. *Personality and Social Psychology Bulletin* 13: 188-209.
- Holbrook, A. L., Green, M. C., & Krosnick, J. A. 2003. Telephone vs. face-to-face interviewing of national probability samples with long questionnaires: Comparisons of respondent satisficing and social desirability response bias. *Public Opinion Quarterly* 67: 79-125.
- Holbrook, A. L., Krosnick, J. A., Carson, R. T., & Mitchell, R. C. 2000. Violating conversational conventions disrupts cognitive processing of attitude questions. *Journal of Experimental Social Psychology* 36, 465-494.
- Holbrook, A. L., Krosnick, J. A., & Pfent, A. M. 2007. Response rates in surveys by the news media and government contractor survey research firms. In J. Lepkowski, B. Harris-Kojetin, P. J. Lavrakas, C. Tucker, E. de Leeuw, M. Link, M. Brick, L. Japac, & R. Sangster (Eds.), *Advances in Telephone survey methodology*. New York: Wiley.
- Jackman, M. R. 1973. Education and prejudice or education and response-set? *American Sociological Review* 38: 327-339.
- Jackson, D. N. 1967. Acquiescence response styles: Problems of identification and control. In I. A. Berg (Ed.), *Response set in personality measurement*. Chicago: Aldine.
- Jagodzinski, W., Kuhnel, S. M., & Schmidt, P. 1987. Is there a "Socratic effect" in nonexperimental panel studies? *Sociological Methods and Research* 15: 259-302.
- James, L. R. & Singh, B. H. 1978. An introduction to the logic, assumptions, and the basic analytic procedures of two-stage least squares. *Psychological Bulletin* 85: 1104-1122.
- Keeter, S., Miller, C., Kohut, A., Groves, R. M., & Presser, S. 2000. Consequences of reducing nonresponse in a national telephone survey. *Public Opinion Quarterly* 64: 125-148.
- Kenny, D. A. 1979. *Correlation and causality*. New York: Wiley.
- Kessler, R. C. & Greenberg, D. F. 1981. *Linear panel analysis: Models of quantitative change*. New York: Academic Press.
- Klockars, A. J. & Yamagishi, M. 1988. The influence of labels and positions in rating scales. *Journal of Educational Measurement* 25: 85-96.

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

- Kosslyn, S. M. 1994. *Elements of graphic design*. New York: Freeman.
- Kraut, R. E. & McConahay, J. B. 1973. How being interviewed affects voting: An experiment. *Public Opinion Quarterly* 37: 398-406.
- Krosnick, J. A. 1988. The role of attitude importance in social evaluation: A study of policy preferences, presidential candidate evaluations, and voting behavior. *Journal of Personality and Social Psychology* 55:, 196-210.
- Krosnick, J. A. 1991. Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology* 5: 213-236.
- Krosnick, J. A. & Alwin, D. F. 1989. Aging and susceptibility to attitude change. *Journal of Personality and Social Psychology* 57: 416-425.
- Krosnick, J. A. & Berent, M. K. 1993. Comparisons of party identification and policy preferences: The impact of survey question format. *American Journal of Political Science*, 37: 941-964.
- Krosnick, J. A., & Bullock, J. 2005. Thought piece on survey participation. Paper presented at the conference entitled "New Approaches to Understanding Participation in Surveys," Belmont Conference Center, Elkridge, Maryland.
- Krosnick, J. A. & Fabrigar, L. R. (forthcoming). *Designing great questionnaires: Insights from psychology*. New York: Oxford University Press.
- Krosnick, J. A. & Kinder, D. R. 1990. Altering popular support for the president through priming: The Iran-Contra affair. *American Political Science Review* 84: 497-512.
- Laumann, E. O., Michael, R. T., Gagnon, J. H., & Michaels, S. 1994. *The social organization of sexuality: Sexual practices in the United States*. Chicago: University of Chicago Press.
- Lavrakas, P. J. 1993. *Telephone survey methods: Sampling, selection, and supervision* (2nd ed.). Newbury Park, CA: Sage.
- Likert, R. 1932. *A technique for the measurement of attitudes*. New York: Archives of Psychology.
- Link, M. W. & Mokdad, A. 2005. Advance letters as a means of improving respondent cooperation in RDD studies: A multi-state experiment. *Public Opinion Quarterly* 69: 572-587.
- Lodge, M., McGraw, K. M., & Stroh, P. 1989. An impression-driven model of candidate evaluation. *American Political Science Review* 83: 399-420.
- Lubin, B. E., Levitt, E. & Zuckerman, M. 1962. Some personality differences between responders and nonresponders to a survey questionnaire. *Journal of Consulting Psychology* 26: 192.
- Malhotra, N. & Krosnick, J. A. (in press). The effect of survey mode on inferences about political attitudes and behavior: Comparing the 2000 and 2004 ANES to

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

Internet surveys with non-probability samples. Political Analysis.

- Mann, Christopher B. 2005. "Unintentional Voter Mobilization: Does Participation in Pre-election Surveys Increase Voter Turnout?" *Annals of the American Academy of Political and Social Science* 601:155-168.
- Matell, M. S. & Jacoby, J. 1971. Is there an optimal number of alternatives for Likert scale items? Study I: Reliability and validity. *Educational and psychological measurement* 31: 657-74.
- Matthews, C.O. 1929. Erroneous first impressions on objective tests. *Journal of Educational Psychology* 20: 280-286.
- Merkle, D.M., S.L. Bauman, and P.J. Lavrakas. 1993. The impact of callbacks on survey estimates in an annual RDD survey. *Proceedings of the American Statistical Association, Section on Survey Research Methods*. 1070-1075. Washington, DC: American Statistical Association
- Mirowsky, J., and C.E. Ross. 1991. Eliminating defense and agreement bias from measures of the sense of control: A 2 x 2 index. *Social Psychology Quarterly* 54: 127-145.
- Morwitz, V.G., E. Johnson, and D. Schmittlein. 1993. Does measuring intent change behavior? *The Journal of Consumer Research* 20: 46-62.
- Myers, J.H., and W.G. Warner. 1968. Semantic properties of selected evaluation adjectives. *Journal of Marketing Research* 5: 409-412.
- Nelson, D. 1985. Informal testing as a means of questionnaire development. *Journal of Official Statistics* 1: 179-188.
- Nosanchuk, T. A., & Marchak, M. P. 1969. Pretest sensitization and attitude change. *Public Opinion Quarterly* 33: 107-111.
- Ofir, C. & Simonson, I. 2001. In search of negative customer feedback: The effect of expecting to evaluate on satisfaction evaluations. *Journal of Marketing Research* 38: 170-182.
- O'Neil, M.J. 1979. Estimating the nonresponse bias due to refusals in telephone surveys. *Public Opinion Quarterly* 43: 218-232.
- O'Neill, J. and C.L. Spash. 2000. Appendix: Policy research brief: onceptions of value in environmental decision-making. *Environmental Values* 9: 521-536.
- Payne, S. 1951. *The art of asking questions*. Princeton, NJ: Princeton University Press.
- Presser, S. 1984. The use of survey data in basic research in the social sciences. In *Surveying subjective phenomenon*, ed. C.F. Turner and E. Martin. New York: Russell Sage Foundation.
- Presser, S., J.M. Rothgeb, M.P. Couper, J.T. Lessler, E. Martin, J. Martin, and E. Singer, eds. 2004. *Methods for testing and evaluating survey questionnaires*.

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

New York: Wiley.

- Price, V. and J. Zaller. 1993. Who gets the news? Alternative measures of news reception and their implications for research. *Public Opinion Quarterly* 57: 133-64.
- Rahn, W.M., J.A. Krosnick, and M. Breuning. 1994. Rationalization and derivation processes in survey studies of political candidate evaluation. *American Journal of Political Science* 38: 582-600.
- Roberto, K. A. & Scott, J. P. 1986. Confronting widowhood: The influence of informal supports. *American Behavioral Scientist*, 29, 497-511.; Tziner, 1987; Stapp & Fulcher, 1983
- Rocco, E. 2003. Constrained inverse adaptive cluster sampling. *Journal of Official Statistics* 19: 45-57.
- Ross, M. W. 1988. Prevalence of classes of risk behaviors for HIV infection in a randomly selected Australian population. *Journal of Sex Research* 25:441-450.
- Ruch, G.M. and M.H. DeGraff. 1926. Correction for chance and "guess" versus "do not guess" instructions in multiple-response tests. *Journal of Educational Psychology* 17: 368-375.
- Saris, W.E., I. Gallhofer, W. van der Veld, and I. Corten. 2003. *A scientific method for questionnaire design: SQP*. Amsterdam: University of Amsterdam.
- Schuman, H., and M.P. Johnson. 1976. Attitudes and behavior. *Annual Review of Sociology* 2: 161-207.
- Sherif, M. & Sherif, C. W. 1967. Attitudes as the individual's own categories: The social judgment-involvement approach. In C. W. Sherif and M. Sherif (Eds.), *Attitude, ego-involvement and change* (pp. 105-139). New York: Wiley.
- Sherif, C. W., Sherif, M., & Nebergall, R. E. 1965. *Attitude and attitude change: The social judgment-involvement approach*. Philadelphia: W.B. Saunders.
- Sherman, S. J. 1980. On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology* 39: 211-221.
- Singer, E., J. van Hoewyk, N. Gebler, T. Raghunathan, and K. McGonagle. 1999. The effect of incentives on response rates in interviewer-mediated surveys. *Journal of Official Statistics* 15: 217-230.
- Spangenberg, E. R. & Greenwald, A. G. 1999. Social influence by requesting self-prophecy. *Journal of Consumer Psychology* 8: 61-89.
- Stapp, J. & Fulcher, R. 1983. The employment of APA members: 1981. *American Psychologist* 38: 1298-1320.
- Stevens, T.H., R. Belkner, D. Dennis, D. Kittredge, and C. Willis. 2000. Comparison of contingent valuation and conjoint analysis in ecosystem management.

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

Ecological Economics 32: 63-74.

- Sudman, S. & Bradburn, N. M. 1982. *Asking questions: A practical guide to questionnaire design*. New York: Jossey-Bass.
- Tourangeau, R., L.J. Rips, and K. Rasinski, K. 2000. *The psychology of survey response*. Cambridge: Cambridge University Press.
- Traugott, M.W. 1990. Memo to pilot study committee: Understanding campaign effects on candidate recall and recognition. American National Election Studies (ANES) Pilot Study Report, No. nes002284.
- Tufte, E. R. 2001. *The visual display of quantitative information*. Cheshire, CT: Graphics Press.
- Tziner, A. 1987. Congruency issues retested using Rineman's achievement climate notion. *Journal of Social Behavior and Personality* 2: 63-78.
- Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. 1996. Mail surveys for election forecasting? An evaluation of the Columbus Dispatch poll. *Public Opinion Quarterly* 60: 181-227.
- Visser, P. S., Krosnick, J. A., Marquette, J., & Curtin, M. 2000) Improving election forecasting: Allocation of undecided respondents, identification of likely voters, and response order effects. In P. Lavrakas & M. Traugott (Eds.), *Election polls, the news media, and democracy*. New York, NY: Chatham House.
- Wainright, D.M. 2003. More 'con' than 'joint': Problems with the application of conjoint analysis to participatory healthcare decision making. *Critical Public Health* 13: 373-380.
- Wason, P. 1961. Responses to affirmative and negative binary statements. *British Journal of Psychology* 52: 133-142.
- Weisberg, H.F., A.A. Haynes, and J.A. Krosnick. 1995. Social group polarization in 1992. In *Democracy's feast: Elections in America*, ed. H.F. Weisberg, 241-249. Chatham, NJ: Chatham House.
- Weisberg, H.F., J.A. Krosnick, and B.D. Bowen. 1996. *An introduction to survey research, polling, and data analysis*. Thousand Oaks, CA: Sage.
- Welch, W. W. & Wallberg, H. J. 1972. Pretest and sensitization effects in curriculum evaluation. *American Educational Research Journal* 7: 605-614.
- Wesman, A.G. 1946. The usefulness of correctly spelled words in a spelling test. *Journal of Educational Psychology* 37: 242-246.
- Wikman, A., and B. Warneryd. 1990. Measurement errors in survey questions: Explaining response variability. *Social Indicators Research* 22: 199-212.
- Willson, V. L. & Putnam, R. R. 1982. A meta-analysis of pretest sensitization effects in experimental design. *American Educational Research Journal* 19: 249-258.
- Winkler, J.D., Kanouse, D.E., and Ware, J.E. 1982. Controlling for acquiescence

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

response set in scale development. *Journal of Applied Psychology* 67: 555-561.

Wiseman, F. 1972. Methodological bias in public opinion surveys. *Public Opinion Quarterly* 36: 105-108.

Zabel, J.E. 1998. An analysis of attrition in the panel study of income dynamics and the survey of income and program participation with an application to a model of labor market behavior. *Journal of Human Resources* 33: 479-506.

Yalch, R.F. 1976. Pre-election interview effects on voter turnout. *Public Opinion Quarterly* 40, 331-336.

Zabel, J.E. 1998. An analysis of attrition in the panel study of income dynamics and the survey of income and program participation with an application to a model of labor market behavior. *Journal of Human Resources* 33: 479-506.

The descriptions of these methods and approaches and of their utility for ecological valuation at EPA do not represent the consensus views of the SAB Committee on Valuing the Protection of Ecological Systems and Services, nor have they been reviewed and approved by the chartered Science Advisory Board. They are offered to extend and elaborate the very brief descriptions provided in chapter 4 of the draft SAB Report on *Valuing the Protection of Ecological systems and Services* and to encourage further deliberation within EPA and the broader scientific community about how to meet the need for an integrated and expanded approach for valuing the protection of ecological systems and services.

Endnotes

¹ The use of surveys has also been growing in the private sector and the academic world (Presser, 1984; Saris et al., 2003), which likely reflects that: (1) surveys are now capable of generating much more interesting data, via implementation of multifactorial experimental designs and complex measurement procedures; (2) cross-national comparisons are of increasing interest; and (3) social scientists want to collect data on more heterogeneous and representative samples. There is also substantial evidence that the quality of optimally-collected survey data are generally quite high. For example, in the Monthly Survey of Consumer Attitudes and Behavior, a representative national sample of American adults has been asked each month what they expect to happen to the unemployment and inflation rates in the future. Their aggregated answers have predicted later changes in actual unemployment and inflation remarkably well (correlations of .80 and .90, respectively, between 1970 and 1995).

² Presenting a seven-point bipolar rating scale is easy to do visually but is more challenging to do aurally. Such scales can be presented in sequences of two questions that ask, first, whether the respondent is on one side of the midpoint or the other or at the midpoint (e.g., "Do you like bananas, dislike them, or neither like nor dislike them?"). Then, a follow-up question can ask how far from the midpoint the respondents are who settle on one side or the other (e.g., "Do you like bananas a lot or just a little?"). This branching approach takes less time to administer than offering the single seven-point scale, and measurement reliability and validity are higher as well (Krosnick and Berent, 1993).

³ A common set of rating scale labels assesses the extent of agreement with an assertion: strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, strongly disagree (Likert, 1932). Yet a great deal of research shows that these response choices are problematic because of acquiescence response bias, whereby some people are inclined to agree with any assertion, regardless of its content (see, e.g., Couch and Keniston, 1960; Jackson, 1967; Schuman and Presser, 1981), which may distort the results of substantive investigations (e.g., Jackman, 1973; Winkler et al., 1982). Although it might seem that the damage done by acquiescence can be minimized by measuring a construct with a large set of items, half of them making assertions opposite to the other half, doing so requires extensive pretesting, is cumbersome to implement, is cognitively burdensome for respondents, and frequently involves asking respondents their agreement with assertions containing the word "not" or some other such negation, which increases both measurement error and respondent fatigue (e.g., Eifermann, 1961; Wason, 1961). Acquiescers also presumably end up at the midpoint of the resulting measurement dimension, which is probably not where most belong on substantive grounds. Most importantly, answering an agree/disagree question always involves first answering a comparable rating question in one's mind. For example, respondents who are asked their agreement with the assertion "I am not a friendly person" must first decide how friendly they are and then translate that conclusion into the appropriate selection. It would be simpler and more direct to ask respondents how friendly they are on a scale from "extremely friendly" to "not friendly at all." Every agree/disagree question implicitly requires a respondent to make a mental rating of an object on the construct of interest, so asking about that dimension is simpler, more direct, and less burdensome. Not surprisingly, then, the reliability and validity of rating scales that do so are higher than those of agree/disagree rating scales (e.g., Ebel, 1982; Mirowsky and Ross, 1991; Ruch and DeGraff, 1926; Wesman, 1946).