

- Tourangeau, R. (1992). Attitudes as memory structures: Belief sampling and context effects. In N. Schwarz & S. Sudman (Eds.), *Context effects in social and psychological research*. New York: Springer Verlag.
- Tourangeau, R., & Rasinski, K.A. (1988). Cognitive processes underlying context effects in attitude measurement. *Psychological Bulletin*, 103, 299 - 314.

In R.J. Kopp, W.W. Pommerehne & N. Schwarz (eds.), Determining the value of non-marketed goods.
The Netherlands: Kluwer Academic Publishers (1997).

168
8

**WHAT DO PSYCHOLOGISTS WANT?
CONTINGENT VALUATION AS A
SPECIAL CASE OF ASKING
QUESTIONS**

Baruch Fischhoff

INTRODUCTION

For a hundred years or so, psychologists have been asking people evaluative questions. In order to gain approval for their work, they have had to satisfy peer reviewers in disciplinary journals run primarily by other psychologists. In this world, technical reports carry little credibility. Psychologists' training typically includes graduate and undergraduate courses in research design, culminating in an experimental dissertation. Unless they are high rollers, they stick to methods that are incremental improvements on existing ones, with (what are believed to be) well-understood strengths and weaknesses. Somewhat more innovation is tolerated in creating theoretical accounts. Top journals seldom accept studies motivated primarily by practical, rather than theoretical issues. Great emphasis is placed on anticipating criticisms -- by raising alternative explanations, explicitly citing sources of residual uncertainty, and identifying research needs.

For twenty years or more, economists have been asking people questions in the specific context of contingent valuation studies. The audience for this work has included juries, lawyers, regulators, and fellow economists. Their work is as likely to appear in technical reports as scientific journals. Reviewers tend to be restricted to CV practitioners and professional critics. They may even try to advance private interests rather than public understanding. Combatants' formal training usually is in the standard fare of economics: sophisticated theory and statistical analysis of archival data. Rather than being considered the elite of their profession (a fancy entertained by experimental psychologists), CV researchers are often hounded by traditional economists for daring to trust expressed preferences. Their clients often impose a siege mentality, discouraging the active pursuit and admission of problems.

If this characterization is even roughly correct, then most experimental psychologists and CV investigators have quite different backgrounds and perspectives. The fact that they are ostensibly in the same business, that of asking questions, has encouraged the naive assumption that they can speak directly to one another. However, their interactions have rarely been in settings suited to cultivating mutual understanding: working collaboratively on common problems, with equal status and commitment. (Viscusi [in press] and Schulze et al. [in

press] cite some productive exceptions.) A few conferences have offered the limited opportunities for understanding afforded by formal papers and responses (e.g., Cambridge Economics, 1992; Peterson, Driver & Gregory, 1988). Occasionally, psychologists have been on advisory panels, able to offer comments, but too far from the action to shape the basic design of studies. As a result, they find themselves constantly carping, without the hands-on role needed to "sign off" on results. Sometimes, psychologists are hired to savage CV studies by clients more interested in damaging adversaries than in evaluating natural resources. Radical skepticism makes one few friends, however competently it is executed. These unsatisfactory public interactions are augmented by countless hours of private fuming.

These are not a pleasant or productive circumstances. As a possible step toward bridging this gap, I will describe what I, as a psychologist, see -- and miss -- in conventional CV studies. These observations lead to several recommendations for removing barriers to mutually acceptable research. The second half of the paper confronts psychology's own inherent limits when addressing the challenges facing CV. These limits are revealed in the course of offering an alternative methodology, violating the norms of both professions. Given the social, environmental, and economic stakes riding on the pricing of nonmarket goods, it would be a shame not to pool our respective training and talents. Cooperation might reduce the risk of our becoming pawns in the machinations of those hoping to distort natural resource policy by manipulating science. The spirit of this essay is that of a member of the loyal opposition. I would dearly like CV to work. It is, indeed, too important to fail. However, those stakes make it critical to face problems coming from all quarters. Being an uncritical friend would be a disservice.

From my (limited) viewpoint, CV seems to be at a turning point. Investigators increasingly conduct studies that are conceptualized as experiments. In so doing, they can look for better methods, rather than having to claim to have definitive ones. In this shift, CV goes from an exercise in applied basic social science (adapting existing techniques to applied problems) to an arena for basic applied social science (addressing fundamental issues arising in applied problems). Although this promises to be an exciting period, it may create tensions between those who view CV as a playground, where fascinating topics are pursued, and those who view it as a workshop, where authoritative estimates are produced. Balancing these conflicting objectives requires both good will and a strategic plan. The first half of this paper hopes to clarify some sources of misunderstanding, while the second half offers a possible bridge.

What follows are personal opinions, derived from following CV with some intensity for a decade and more passively for somewhat longer (e.g., Fischhoff, Slovic, & Lichtenstein, 1980).¹ I won't claim to speak for all psychologists, although I suspect that many of these tastes are shared (Ajzen & Peterson, 1988;

Baron, 1994). I won't attempt to document my claims systematically, beyond reference to what I have written already (Fischhoff, 1988, 1990, 1991; Fischhoff & Furby, 1986, 1988; Furby & Fischhoff, 1989).

THE PSYCHOLOGICAL STUDY

The American Psychological Association's *Publication Manual* (1984) sets our standard for reporting research results. Although much of it involves arbitrary rules (e.g., how to format tables, where to place acknowledgments), it also reflects psychology's evolving philosophy of science, regarding the essential elements of a study.² I will organize my comments according to the sections that the *Manual* prescribes for the standard psychological paper, in part because of its comprehensiveness, in part to emphasize the deep-seated conventions shaping our perceptions. The plea for full, consistent reporting by NOAA's Contingent Valuation Panel (Department of Commerce, 1993, at p. 4608) is one indication of how slowly such norms evolve.

Components of Psychological Reports

Introduction

When reading a scientific paper, psychologists expect to see a new hypothesis that attempts to accommodate all existing studies and resolve some uncertainty in them. Ideally, that hypothesis is rooted in existing theory. For example, prospect theory (Kahneman & Tversky, 1979) builds on results in psychophysics and adaptation-level theory in order to explain some choice anomalies. Once a study has passed peer review, the burden of proof shifts to those who would ignore its results.

In CV studies, I see less interest in theory testing, in deference to the practical goal of evaluating particular amenities. I see little awareness of the relevant psychological literature. I see dismissal of studies satisfying our methodological norms, for failing to satisfy the norms of CV investigators. Moreover, those norms seem inconsistent and unstable. At any one time, different CV investigators uphold different inviolate standards. Over time, norms seem to change for unsatisfying reasons (e.g., shifting tastes, producing unduly high estimates). The NOAA panel's attempt to legislate methodology feels like the sort of action that might be undertaken by clinical psychologists, a professional guild that we hold at arm's length (Dawes, 1994). As a result, CV studies seem to be introduced in ways that exclude much of what is dear to us.

Design

For psychologists, valid measures are sensitive to relevant changes in the measured object and insensitive to irrelevant changes (Campbell & Fiske, 1959).

This test is easy to apply when relevance is defined independently (e.g., with length or weight); it is much harder with subjective phenomena like values or beliefs. Psychometrics is psychology's philosophy of measurement for coping with subjectivity (Dawes, 1972). It requires the active pursuit of problems, systematically varying factors that should or should not affect measurement. For cognitive psychologists, interested in *how* people think, even measurement failures can be useful if they reveal new behavioral phenomena. Indeed, experimental psychology has been described as a process of converting methodological artifacts into theoretical main effects, worth studying in their own right (McGuire, 1969).

In CV studies, I see a conscientious effort to design the single best evaluation question. All resources are then invested in securing a large sample of responses to it. CV investigators typically rely on impressionistic summaries of pretest interviews with individuals or focus groups. Although they can uncover problems, these techniques create conditions quite different than those of the actual interviews (Merton & Kendall, 1946; Merton, 1987). As a result, there is uncertainty about how actual respondents have interpreted their task.

Stimuli

Psychologists place great importance on mapping the elements of their experimental set-ups onto the concepts in their theories. This means making the theory as explicit as possible, describing its boundary conditions, and excluding irrelevant factors from the task (e.g., giving no hint at the investigator's intent). There is a strong bias toward simplicity, focusing subjects on the investigators' issues and reducing cognitive load. We are, nonetheless, probably too optimistic regarding how well our stimuli are understood and accepted. Even when the words are simple, the underlying concepts and context may be quite foreign to respondents (Fischhoff, 1993).

CV studies necessarily use much richer stimuli, in order to describe complex environmental goods. However, I often cannot see how the specific details were selected. As a result, CV stimuli look quite arbitrary. Seeing the great variation in the features that different CV investigators chose to specify, Lita Furby and I created a framework (Fischhoff & Furby, 1988) for the features that should be addressed in a fully specified evaluation task. The typical CV study omits many of these details, forcing (or allowing) respondents to guess at what was meant. Nonetheless, most CV studies still present more detail than subjects could reasonably absorb in a typical interview. In the absence of comprehensive manipulation checks with actual subjects, I remain skeptical that subjects are answering the intended question. In a study using simplified stimuli, we found that the embedding effect (Kahneman & Knetsch, 1992) decreased when we interpreted responses in terms of the task that subjects reported answering rather than the one we had actually posed (Fischhoff et al., 1993). Without this

manipulation check, we would have been unduly optimistic about our ability to communicate and unduly pessimistic about subjects' ability to respond.

Response mode

Psychologists hope is to get as much information as possible out of subjects without creating (and not just revealing) preferences. The late Clyde Coombs articulated this philosophy in his *A theory of data* (1964). There, he showed techniques for eliciting the most precise preferences that subjects can give (e.g., pick the single best and worst options, pick the three best options). We seldom elicit anything stronger than interval scale responses.

CV investigators share our aversion to reactive methods. However, their tasks seem extraordinarily demanding: monetary evaluation of goods whose precise formulation must be very unfamiliar. That formulation must be understood even if evaluations are expressed in a referendum format. Although ostensibly simple, referenda may obscure the difficulty of a task, by making it easy to respond casually in ways that look sensible. I suspect that subjects often translate CV tasks into something more manageable under the time constraints (e.g., "does this merit a large, medium or small standard contribution -- given that I'm supposed to pay something for it."). Or, they may just flee with "protest responses." Strack and Schwarz (1992) use "implicit collaboration" to describe subjects' attempts to make sense out of unreasonable tasks.

Subjects

Cognitive psychologists tend to be casual about sampling. Because we believe that basic thought processes are widely shared, we are content with convenience samples. Because effects must be fairly large in order to have theoretical interest, relatively small samples will have adequate statistical power (although, left to our own intuitions, we tend to overestimate that power [Cohen, 1962; Tversky & Kahneman, 1971]). When individual differences matter, we assume that correlations are invariant across samples. Our primary worry is truncating the range on a variable, so that there is too little heterogeneity for correlations to emerge.

In CV studies, I see a diligent commitment to recruiting representative samples. This befits determining *what* the population as a whole thinks. Less understandable is concentrating all those resources on the single best question. The resulting precision is misplaced without a trustworthy elicitation methodology. It seems more efficient to develop such a methodology with experiments using small, diverse samples. Insisting on large samples for any CV-related study represents a restraint of inquiry, by disenfranchising investigators who cannot command the needed resources.

Results

Our training typically includes a graduate year of "Statistics for Psychologists" and several supervised experimental projects. One taste that this training inculcates is having a large tool kit of statistical procedures. A second taste is letting the responses speak for themselves, reporting them in something like their original form and avoiding heavily interpretative statistical models. A third is discarding subjects with great reluctance.

In CV studies, I see a problem that we rarely face: many subjects spontaneously rejecting their task or providing unreasonable answers. Deleting such responses makes the effective sample in many CV studies less representative than is intended (and sometimes claimed). It raises doubts about the responses that are retained, some of which may seem reasonable just by chance. CV's reporting norms are improving. However, it is still often unclear what responses mean to those who provide them, nor how respondents' residual confusion should be reflected in data analyses (Hamm et al., 1992). In order to exploit the power of econometric procedures, CV investigators often require stronger assumptions, regarding the orderliness of subjects' decision-making, than most psychologists would be comfortable making.

Discussion

In this section, psychologists pick up the themes laid out in the Introduction, consider alternative explanations for their data, and identify needed future research.

In CV studies, I often see an aggressively defensive attitude, making strong claims, with little self-criticism. Concern is often focused on issues that we tend to ignore, like strategic responses (Orne, 1962; Schwarz, 1994). It is hard for me to distinguish what CV researchers really believe from what their clients need to hear.

Some recommendations

This is, of course, a caricature. CV studies vary widely in their quality, just as we deviate from psychology's standards more than we would like to think or claim. Nonetheless, there seems to be a clash of paradigms here. The fact that some psychologists and economists have found ways to co-author papers suggests reason for hope. Reason for despair can be found in the incentives for unqualified claims and critiques in a litigious and politicized society. Too many people involved with CV cannot take "yes" or "no" for an answer, but must be ruthless proponents or opponents. Perversely, this unreasonable environment has gotten a hearing for psychology (Department of Commerce, 1993), although in a context

that kept psychologists at arms length. Here are some institutional and intellectual suggestions that might help tilt the odds toward the optimists:

Institutional suggestions

Make resources available for basic research on elicitation, without having to evaluate environmental resources that are currently "in play" or conform to current CV norms.

Make psychologists and economists equal partners in allocating resources, directing projects, and reviewing papers -- forcing them to work together.

Encourage cross-disciplinary postdoctoral work, producing scholars who have undergone both apprenticeships.

Include independent peer review in CV projects; publish some reviews, along with responses, as a way of airing differences.

Intellectual suggestions

Accept the legitimacy of studies meeting the standards of each research tradition, if they recognize the standards of the other (e.g., "we sampled haphazardly," "we performed no manipulation checks," "these procedures are not widely accepted within psychology").

Perform studies that systematically vary potentially relevant and irrelevant features of CV scenarios, within an independent framework for determining their relevance.

Provide complete descriptions of responses, including reliability estimates, manipulation checks, no-response rates, and protest-response rates.

Accept the possibility that respondents do not know what to think about the specific questions we pose; grapple with helping them to articulate their values and with using incomplete values in decision making.

Interim conclusion

Most of these suggestions are controversial. For example, a National Research Council (1982) report recommended independent peer review of "survey measures of subjective phenomena," only to provoke a passionate dissent from a leading figure in the survey research industry. Some would be hard to implement. Nonetheless, CV seems ripe for wrestling with core issues, rather than dealing with symptomatic conflicts. CV could become an exciting arena for interdisciplinary contact, or remain a frustrating one.

The next section describes a constructive approach to addressing these issues. It is intended as a zero-based approach, unconstrained by existing ones, but drawing on their hard-earned experience. It is unlikely to satisfy any discipline entirely. Economists may be particularly offended by its claim of taking the contingent market metaphor more seriously than they have (i.e., improving on the economics of CV). Psychologists (and survey researchers) might be particularly aghast at its abandoning the conventional commitment to nonreactive measurement.

The exposition begins with the criteria that a method for eliciting environmental valuations should satisfy. The next section outlines the proposed method, in terms of its organizing metaphor and basic methods. The following section elaborates on how it addresses the criteria and how one should validate an implementation. Finally, the proposal is critiqued, in terms of related approaches and the work to be done. Some differences among approaches are empirical questions, which could be resolved by studying the data that each produces. Other differences are matters of principle, regarding what kind of environmental values should be used in policy decisions.

A CONSTRUCTIVE APPROACH TO ELICITING ENVIRONMENTAL EVALUATIONS

Criteria for value elicitation procedures

In order to fill the niche that CV claims, a procedure should meet these seven criteria:

General applicability

A method for eliciting public values should be accessible to any citizen willing to invest the effort. Thus, it cannot require extensive prior knowledge, nor discourage individuals without it.

Well-formulated policy issue

Participants must receive information on all issues that they deem essential. Vague questions may encourage vague responses (e.g., general expressions of sympathy), while leaving too much freedom to interpret what those responses mean. Participants should not have to read between the lines of the questions; readers should not have that option with the answers.

Standardized response mode

In order to be summarized efficiently and objectively, respondents' beliefs must be expressed in a common format. That summary should capture the critical features of subjects' values, including how deeply they are held.

Understanding of science

In order to provide informed consent of the governed for the actions based on their evaluations, respondents must understand the size and predictability of the environmental changes at stake.

Articulated preferences

Once they have understood the issues, participants must identify the responses that reflect their values. If that process normally requires rumination, conversation, and consultation, then the elicitation procedure must provide the functional equivalent.

Validation

Any scientific study must assess the quality of its measures. That includes determining how sensitive responses are to relevant and irrelevant changes in procedure.

Impact clarity

The consumers of research need a bottom line, summarizing what the results mean for their concerns. They must be able to decide whether the estimates are good enough to justify action, and whether alternative procedures might have fared very differently.

None of these criteria seems controversial.³ The proposed approach tries to take each seriously. In some cases, there is much to build on; in other cases, it is largely guesswork. It is called constructivist because it assumes that meaningful responses to novel tasks must be created at the investigators' demand (Burgess, Harrison & Limb, 1988; Fischhoff, 1991; Gregory et al., 1993; Lazo et al., 1992; Schulze et al., in press; Whittington & Davis, 1993).

Some organizing metaphors

Contingent valuation might be seen as having two organizing metaphors. One is the *contingent market*, within which respondents imagine their behavior. The second is the *opinion poll*, within which respondents report their current thinking on some issue. Table 1 shows some examples of people's willingness to respond.

TABLE 1: SOME POLL RESULTS

65% of Americans believe that frozen pizza will never be any good, and there is nothing that science can do about it.
10% of Americans would pay \$5 to see Sen. Orrin Hatch battle a big mean dog on television. 80% would root for the dog.
45% of Americans think that rain doesn't feel as good in real life as it seems to in the movies

(margin of error = 9%)

Source: TV Nation, NBC, July 20, 1994

Our approach accepts the former metaphor, perhaps even more seriously than is usual. However, it rejects the latter, as incompatible with the seven criteria. People cannot have existing opinions on the specific CV scenario that they receive. Even if respondents have thought about the good (e.g., a watershed), the terms of a well-specified proposal are inevitably novel. People cannot instantly absorb its details and convert them into valuations. It may take a while even to conceptualize environmental effects as "goods," with a price tag no less. As a result, a process of learning and value construction is inevitable, making measurement inherently reactive. Participants must change if they are to respond meaningfully. A measurement procedure must face and manage those changes.

Reactive measurement seems compatible with the contingent-market metaphor. Participants in actual markets can provide flash estimates for the worth of novel commercial or financial products. However, they might not want to be held to those answers in any consequential setting. Markets work when people can identify specific actions that serve their general interests (Viscusi & Magat, 1987). A properly functioning market should capture *stable* expressions of participants' preferences, in the sense that nothing changes upon further reflection or additional information. Ill-informed responses represent (contingent or conventional) market failure.

Citizens may vote on a complex ballot measure with similarly little prior thought. These casual votes might be predicted by flash responses to an opinion poll. However, that would mean predicting a failed electoral process. A successful referendum is preceded by extensive doses of such democratic processes as partisan proclamations, academic analyses, conversations with peers, question-and-answer sessions, and time to think about the specific proposal. Indeed, one might argue that the referendum metaphor should be used only by investigators who deliver these accouterments (Barber, 1991; Verba & Nie, 1972).

"Referendum" has typically been appropriated for much more casual presentations in CV studies.

For serious elicitation, a more appropriate metaphor is the citizens commission: A representative sample of citizens is selected to learn about an issue, on behalf of the electorate. Their opinions are interpreted as resembling those that other citizens might adopt, if offered the same opportunity. An elicitation procedure instantiating this metaphor would have to capture the salient features of such a commission, including balanced briefings, partisan testimony, group discussions, and the chance to explain complex opinions (Crosby, Kelly & Schaefer, 1986; Fiorino, 1990; Reich, 1985). This metaphor is adopted here, with one important exception: Evaluations are collected privately, in order to reduce the effects of group processes and provide estimates of individual preferences.

Metaphors matter, not only because of the claims they make, but also because of the guidance they provide. Opinion polls on real issues attempt to simulate the conditions of the ballot box (secret responses, no proximal campaigning, etc.). Although they might ask how respondents they expect their opinions to change by election day, polls would not attempt to simulate that change process. With artificial issues, like those necessarily posed by CV, opinions must evolve (on the specific CV proposal, if not on the underlying issues). The commission metaphor embraces that change. It aspires to simulate an idealized political process, with all the access that participants would want to information, analysis, discussion, and rumination. It replaces the polling norm of "tell them nothing" with an attempt to "give them everything." It seeks the thoughtful public values needed for environmental policy. It attempts to confront and exploit the reactivity that is anathema to most psychology and survey research.

Procedural overview

Table 2 outlines the method. It begins with a recruitment procedure that explains the study, introduces the issue, and solicits initial responses on a questionnaire designed to get participants thinking. A confirming phone call allows for clarifying questions. Group sessions provide the opportunity to present and discuss the issues in detail. (The number of sessions depends on the complexity of the issues.) Only private responses are analyzed.

TABLE 2: STANDARD TASK SEQUENCE	
Recruitment	Letter of Invitation (introducing issue and procedure) Letter of Acceptance (including brief questionnaire) Phone Confirmation (and clarification)
Initial Group Session	Introductions (clarifying questions, no discussion) Procedure CV Transaction Underlying Science (natural, social) Preliminary Evaluations (assisted interview) Group Presentation and Mediated Discussion Summary Evaluations (proposal, procedure)
Additional Group Session(s)	Procedural Discussion Clarifying Questions Opening Evaluation Group Presentation and Mediated Discussion Summary Evaluations (proposal, procedure)
Follow-up (sub-samples, with/without intervening events)	

The initial meeting introduces each element of the procedure, including the CV scenario, the relevant science, the response modes, and the organizing metaphor -- so that participants will know what is wanted of them. After this introduction, respondents complete each evaluation form (in private), in order to clarify the task and elicit initial impressions. In keeping with the metaphor, these are treated as tentative preliminary views, subject to change as subjects have time to learn, talk, and think. The remainder of the session is devoted to introducing and discussing the issues. It concludes with the first "serious" evaluations. The following

sessions involve more of the same: presenting information, clarifying uncertainties, discussing implications, and eliciting evaluations in private.

Participants evaluate the procedure as well as the issues. They perform manipulation checks, rate fairness, comment on the interactions, and request changes. These evaluations guide the process, in addition to characterizing its quality. Because moderators respond to requests and problems, different groups may receive nominally different material, in order to achieve functionally equivalent experiences. This adaptive strategy takes seriously the "conversational" approach to survey design, which assumes that respondents apply the norms of everyday conversation to interpreting survey questions (Dietz, 1987; Strack & Schwarz, 1992). With unfamiliar topics and heterogeneous audiences, no one wording may be interpreted similarly (and appropriately) by all respondents (Quadrel, Fischhoff & Palmgren, 1994). As a result, getting the message across requires actual conversation.

For most respondents, the elicitation process ends with the final session. Some are interviewed later in order to assess the stability of responses. Additional interviews might be conducted later, if events occur that should change evaluations (e.g., the release of significant scientific findings).

A Work plan

Table 3 lists the tasks involved in implementing this approach and, arguably, any approach to value elicitation that addresses the seven criteria. Each task raises basic research issues, some of which are discussed below, organized by the criterion that each is most closely attached to achieving.

TABLE 3: INSTRUMENTS AND MATERIALS (by primary criterion)	
<i>General Applicability</i>	Sampling Frame and Size
<i>Well-Formulated Policy Issue</i>	Specification of CV Scenario
<i>Standardized Response Mode</i>	Proposal Evaluation Protocol Procedure Evaluation Protocol
<i>Understanding of Science</i>	Explanation of CV Scenario Explanation of Underlying Science
<i>Articulated Preferences</i>	Supporting Materials Moderator Guidelines

<i>Validation</i>	Manipulation Checks Validity Tests Reliability Tests
<i>Impact Clarity</i>	Summaries for Policy Makers

General applicability

Sampling frame and size. A procedure that assesses public values should include a representative sample of the citizenry. That could mean random population sampling with aggressive call backs. However, those recruitment efforts would reduce the resources available for interacting with respondents. The value of random sampling is further limited by frame omissions, refusals to participate, and protest responses. The method proposed here should cost more per subject because of its intensity, but less per subject because of its group administration. It also makes a larger capital investment in stimulus development. Having fewer subjects would be justified by having better responses. Response rates might also improve if a more involving procedure increased cooperation. The tradeoff between sample size and response precision deserves some basic research attention.

At this developmental stage, we are selecting participants at random from the membership rolls of diverse organizations (churches, clubs, etc.). This allows assembling people in groups with similar backgrounds, a setting where opinions are often shaped. Responses could be summarized as representing known groups (and not just atomized citizens), or extrapolated to populations of people with similar characteristics. At present, our main concern is establishing a method that is widely applicable. Alternative approaches to representativeness are another topic for future research.

Well-formulated policy issue

Some years ago, Lita Furby and I reviewed thirty-odd studies eliciting evaluations of atmospheric visibility. We were struck by the great variability in the features that different investigators chose to specify. Concluding that there was little consensus on what constituted a "market," we created a framework for specifying "transactions," within which a *payment* is exchanged for a *good*, within some *social context* (Fischhoff & Furby, 1988). It attempts to identify all issues that respondents might wish to consider.

Because respondents may guess at missing features, our framework also identifies the opportunities to read between the lines of a CV scenario. It does not, however, show how to plug the gaps. In some cases, knowledge of the

substantive issues will tell us what to say. In others, we may lack answers to participants' questions (e.g., who else will pay? what else am I going to be asked to pay for? who will implement the promised change? what if I want to contribute time, rather than money? how will my responses be used? what are others saying? how would I pay?). Where we cannot answer authoritatively, it may be tempting to remain silent. We may even have normative reasons why a feature should not matter. However, saying nothing about a feature need not prevent respondents from making assumptions on their own. Indeed, they may feel compelled to do so, in order to render the task meaningful. Deciding how to fill in the blanks is another research challenge.⁴

Standardized response mode

Proposal evaluation protocol. Complex, novel proposals may evoke complex, novel opinions. People might be pressured to summarize those feelings in a simple way. However, that summary captures but a portion of their opinions, leaving observers to guess at their full meaning -- just as election results leave room to speculate about the meaning of voters' choices, not to mention their abstentions.⁵

The ambiguity of election results appeals to politicians, leaving them freedom to interpret the mandate that they have gained or lost. It is not, however, very helpful for determining future actions. That requires supplementary research, like detailed exit polls. Whatever format is used, responsible use of CV results requires understanding what respondents meant. Given the latitude that CV investigators have in specifying their scenarios, one needs to know how alternative tasks would have affected responses. To this end, our subjects evaluate the proposals from a variety of perspectives, including how strong their preferences are, how various changes in the proposal would affect them, how fair the proposal seems, and how their responses would change if the science changed.

Procedure evaluation protocol. Interviews are a special kind of social interaction. As a result, the procedure itself may color participants' responses. For example, "protest responses" may represent objections to the proposals or to the opportunities to understand them (Hamm et al., 1992; Schkade & Payne, 1994). In order to clarify sources of random error, our respondents evaluate statements like "I think that I know enough about the issue to make an informed decision about it," "I found the explanation confusing," "I think that my views will be taken seriously by policy makers," and "I think that my position is well represented by my responses." In order to clarify sources of bias, respondents evaluate statements such as "I felt pressure to answer the question in a particular way," "The proposal was not stated honestly," and "The environmental effects were exaggerated." Such questions might both explain puzzling results and reveal problems with seemingly orderly ones (e.g., cases where respondents adopt a mechanical response strategy, in order to get through a frustrating task).

Understanding of science

CV investigators have often worked hard to convey their scenarios. For example, simulating different levels of atmospheric visibility was once a major methodological pursuit. Concern over embedding has increased the attention paid to conveying quantitative features. In this volume, Navrud (in press), Schulze et al. (in press), and Schwarz (1994) all discuss presentation issues. Our approach builds on this work, and on basic research in scientific communication:

Explanation of CV scenario. An obvious challenge is the complexity of CV scenarios, which would only get worse if they were specified completely. Lazo et al. (1992) propose simplifying scenarios by asking subjects how relevant particular facts are, then confirming these judgments by deleting irrelevant facts in scenarios presented to additional subjects. They report promising results (i.e., similar evaluations with and without those facts). A residual risk is artificial convergence, should subjects give roughly the same value to all scenarios, regardless of their details.

Quadrel, Fischhoff and Palmgren (1994) offer a related procedure, differing in two ways: (a) Subjects think aloud as they read proposed experimental materials, in order to avoid the potential problems with retrospective introspections (Ericsson & Simon, 1984; Nisbett & Wilson, 1977). (b) Confirmatory studies test whether people show the same sensitivity (and insensitivity) in structured tasks as emerged in the think-aloud tasks (Fischhoff, 1992; Stokol & Altman, 1986). A minimal test of efficacy is manipulation checks, asking subjects (in actual studies and not just pretests) to report how they interpreted their task.

Explanation of underlying science. One recurrent result in communication research is that people have difficulty processing information that does not correspond to their "mental model" of the situation. They may even reject factual assertions that do not make sense. Thus, successful communication requires knowing the preconceptions that people bring to a task and presenting new information in that context. In effect, respondents need to be taught some science and not just given quantitative end states (diSessa, 1993; Gentner & Stevens, 1983). That education might involve economics, as well as natural science (Furnham, 1989; Green, in press; Voss, Tyler & Yengo, 1983).

For topics that occur in many studies, it might pay for CV to create a science of lay understanding (e.g., How do people estimate the personal effects of "higher prices for goods and services?" or How do they assess the feasibility of referenda on designated taxes for environmental goods? What loopholes do they imagine in the contract offered by a CV transaction?). Specific goods (e.g., groundwater pollution, ozone-related health effects) may require dedicated studies. The methods are quite inexpensive, considering the stakes (e.g., Bostrom et al., 1992;

Fischhoff et al., 1993; Jungermann, Schutz, & Thuring, 1988; Leventhal & Cameron, 1987; Schulze et al., in press).

Articulated preferences

Supporting Materials. Articulated preferences are ones that remain stable in the face of additional time to think, additional information from neutral sources, and additional perspectives from partisan ones. They are immune to framing effects, often observed in confrontations with unfamiliar tasks (Hogarth, 1982; Tversky & Kahneman, 1981). The order in which issues arise might affect the path by which preferences evolve, but not their final state. As a result, it is critical to present a balanced set of evaluative perspectives as well as scientific facts (saying, in effect, here's a way that you could think about this problem). One test of balance is asking partisans whether we have captured their views.⁶ Finally, we apply a simple test suggested by the late Cathie Marsh, asking subjects who they think produced the materials.

Moderator Guidelines. The interactive character of our method reflects a belief that the norms of conversation (Schwarz, 1994) require actual conversation. That is, however diligent the pretests, no single wording can address the diverse perspectives that respondents bring to a CV study. They need individual help to understand the issues and develop positions. We have chosen to provide this help in group sessions, both for efficiency and because complex information is often processed socially. Following the underlying metaphor, our moderators are expected to suppress their own opinions, and encourage the expression of diverse views. Their success is monitored by reviewing session tapes and asking participants.

Validation

Types of validity. A measurement procedure's success is often seen as having three components: (a) *face validity*, does it include the critical features of the concepts that it represents?; (b) *predictive validity*, does it predict behavior in some designated situation (e.g., an election)?; and (c) *construct validity*, do responses fit a theoretically derived pattern?

Ordinarily, face validity is a minimal condition. The need to define a complex transaction makes it a particularly demanding test for CV tasks. Previous sections describe our procedures for ensuring that the right things get said (well-formulated policy issue) and heard (understanding of science). Process evaluations and manipulation checks measure our success at conveying the intended context and content. They involve (a) asking respondents to interpret the task in front of them, (b) eliciting their memories for tasks that they have just completed, and (c) posing inferential questions regarding the implications of the literally stated task (Kintsch, 1986; Reder, 1985).

Predictive validity is relevant for CV tasks that have clear real-world analogues. CV tasks have often done fairly well in such tests -- when they involve familiar goods and settings. There is nothing in the world to predict, however, for the typical CV scenario (Mitchell & Carson, 1989; Smith & Osborne, in press; Viscusi, in press). Our procedure should predict the result of a comparable public debate.

Construct validity means showing sensitivity to meaningful changes in conditions and insensitivity to non-meaningful ones. Our framework (Fischhoff & Furby, 1988) distinguishes between *substantive* features, which subjects can choose to value or neglect (e.g., visibility, species preservation), and *formal* features, determining the amount of the substantive features (e.g., the probability that the promised environmental change will be as described or that the promised payment will be collected). CV researchers need independent indicators of relevance, followed by empirical tests of sensitivity. The most extensive tests of sensitivity to a formal property involve scope, with mixed results (Baron, 1994; Carson & Mitchell, 1993; Green, Kahneman & Kunreuther, 1994; Schulze et al., in press). Sensitivity tests should also be conducted for other formal properties, and for substantive features of demonstrated relevance or irrelevance. A systematic research plan is needed for identifying and prioritizing such tests.

Reliability. The limiting case for testing insensitivity to irrelevant changes is test-retest reliability. The obvious obstacle to such tests is that subjects will remember their previous responses (but be interpreted as having generated them consistently). Increasing the time period between tests decreases the role of memory, while increasing the risk that intervening events will change beliefs. We use an intermediate time period as a compromise, and also test secondary judgments (e.g., confidence, evaluation of alternative proposals), which should be less memorable. Reliability is necessary, but not sufficient for validity; people might routinely pay the same flat fee for poorly understood goods that investigators force upon them.

Impact clarity

Summaries for Policy Makers. The proposed method creates "thick" descriptions of respondents' beliefs, showing what they thought at several junctures about many issues. Some readers will want to know it all, in order to get an encompassing view of public perceptions. Other readers will focus on particular responses (e.g., final best-guess estimates, first-cut confidence, robust protests). Data need to be summarized coherently for each such audience. Readers who are unfamiliar with a methodology need epistemological guidance as well, telling them how solid the method is. The NOAA review panel provided a narrative summary for some aspects of conventional CV studies. Funtowicz & Ravetz (1990) reviewed and proposed some more formal summaries.

Our method will eventually include guidelines for creating summaries. Its implementation would have to be evaluated empirically, using directed readings -- in which individuals from each target audience read materials aloud, commenting on the conclusions that they derive and the confusion that they experience (Schriver, 1989).

CONCLUSION

In some senses, this proposal is quite consistent with conventional CV practice. It could, for example, be viewed as a way to achieve Carson and Mitchell's criteria for "a valid CV response," namely, that respondents "(i) clearly understand the characteristics of the good they are being asked to value; (ii) find the CV scenario elements related to the good's provision plausible; and (iii) answer the CV questions in a deliberate and meaningful manner." (1993, p. 1267)

However, the analysis opening this chapter suggests a clash of paradigms, in which people using the same words are still speaking different languages.⁷ Inappropriately assuming common meaning can exacerbate conflicts, by obscuring the existence and source of differences. The acrimony of the debates over CV, the personal recriminations, the differing evidentiary standards -- all suggest disciplines that are solving different problems.

Comparison with alternatives

In that spirit, I will conclude by accentuating some differences between this approach to measuring environmental values and others. These differences are partly empirical questions and partly matters of principle.

Public debate

An obvious source of insight into the public's values is listening to what they say. In principle, all citizens can take part in public debates, formulating issues in their own terms and saying what they want. In practice, though, citizens may know too little to participate effectively, may posture for rhetorical purposes, may fear airing unpopular views, may speak off the (CV) topic, and may express themselves imprecisely. Our proposed approach attempts to create and capture the best of such discourse, while overcoming its weaknesses.

Public hearings

Properly conducted hearings allow participants to study in advance, to organize their thoughts, and to set them down in a formal record. Cross-examination and discussion can clarify ambiguities. Thinking may even evolve, as participants hear other opinions and dwell on the issues. However, comments may address

different issues and use different modes of expression, making them hard to summarize. Here, too, strategic responses and heated exchanges can obscure underlying beliefs. Those who speak may represent the population poorly, being skewed toward those who can pay (or get paid) to appear, develop testimony, etc. The present approach attempts to capture the focused and evolutionary character of hearings, with better specified tasks and broader participation.

Focus groups

Every city has firms that pay citizens to join group discussions on specific topics, under the direction of moderators trained to encourage contributions and manage conflict. Such groups can uncover unexpected views on an issue, as can open-ended individual interviews (Chi, in press; Merton & Kendall, 1946; Schriver, 1989). They are less suited to systematically presenting information, airing views, and eliciting individuals' values. Focus groups are usually analyzed impressionistically, as befits their heuristic value (Merton, 1987). The present approach attempts to capture the learning potential of well-moderated groups, but with more systematic sampling, presentation, and elicitation. It elicits values privately.

Opinion polls

Surveys are the obvious method for systematic research into public opinion, and have often shown the public's willingness to sacrifice for the environment. They elicit views in relative privacy. However, they share many limits of public debates. Respondents may be ill-informed about the issues in general and about the specific proposal. If so, they must choose between certain disenfranchisement (by refusing to answer) and some risk of misinterpretation (by answering without understanding). Moreover, what they can say is restricted to a few structured response options, with little chance for elaboration. Survey researchers' commitment to using a single formulation for the entire population means keeping questions simple -- and still leaving some respondents behind. The present approach aspires to elicit any citizen's private views, but abandons the constraints that require simple, uniform, and uninformative questions.

Contingent valuation

Over time, CV research has increasingly freed itself from the constraints it inherited from survey research. CV studies present much more elaborate information and ask much more complicated questions than almost any surveys. In order to do so, CV researchers have developed increasingly involved procedures for preparing, testing, and presenting material. However, these changes seem incremental, without an overall approach to the paradigm shift that they involve. By consciously adopting an alternative philosophy, our approach is forced to face some issues that do not seem to have emerged spontaneously. One

such issue is how to create a fully specified scenario, not just with a lot of detail, but with all the detail that respondents need. A second issue is how to convey these details and help subjects articulate their responses. A third is how to assess the success that we, and subjects, have had. The fourth is how to extract the policy implications of the inevitable imperfections. A fifth issue is how to conceptualize and manage the actual elicitation process.

Multi-attribute utility theory (MAUT)

Looking at some of the same issues, Gregory et al. (1993) proposed using multiattribute utility theory to elicit environmental values. They draw on a well-developed technical literature, with techniques for decomposing diverse situations, in order to reduce decision-makers' cognitive load (Keeney, 1993; Keeney & Raiffa, 1976). MAUT accepts the need to help people understand issues and articulate responses. However, its implementation here would face questions for which MAUT lacks complete answers (e.g., framing effects, information selection, failures of introspection, strategic responses). It is unclear if the benefits of decomposition in familiar domains extend to such unfamiliar ones (Fischhoff, 1980; Dawes, 1988).

As a CV substitute, MAUT eliminates the market metaphor, addressing just the environmental good. It assumes that respondents will view that good in isolation, while ignoring the other issues raised in our transaction framework. Although Gregory et al. (1993) call MAUT a "constructive" approach, that seems to be in the sense of assembling overall evaluations from component judgments rather than allowing the evolution of emergent values. As a result, its attitude to reactivity is unclear. Psychology as a whole is still grappling with how to handle situations in which people might not know what they want (Fischhoff, 1991; Fischhoff et al., 1980; Hogarth, 1982; Tversky & Kahneman, 1981).

A Two-fold strategy

The elicitation side of CV has been an exercise in applied basic social science, using existing techniques in unfamiliar settings. The limits to those techniques have prompted CV investigators to innovate continuously. However, there may be limits to ad hoc changes. Our procedure offers a comprehensive alternative view. It is one version of what psychologists might have created had they been given a chunk of the CV budget. Its elaboration shows the need for basic applied research, examining new fundamental issues highlighted by this application. Some studies of "embedding" have this character. Failure to distinguish between basic applied and applied basic research (Baddeley, 1979) has contributed to the acrimony of this debate (Cambridge Economics, 1992; Harrison, 1993; Nickerson, 1993; Smith, 1992), as has neglecting the clash of paradigms.

The economic side of CV has attracted much more basic applied research, looking at such issues as how to identify strategic responses and how to extract continuous estimates from discrete choices (Kristrom, in press; McConnell, 1990; Mitchell & Carson, 1989; Viscusi, in press). From my perspective, these challenges seem to be addressed with greater equanimity than those coming from psychologists. Perhaps the threats are less serious; perhaps economists present their critiques less caustically; perhaps a member of a profession will not rock the boat too vigorously. Our framework (Fischhoff & Furby, 1988) is, in a sense, an unconventional exercise in economic research, asking how to create a well-specified contingent market. We are, in effect, asking what a market is, perhaps not a question that an economist would raise or a psychologist should answer -- alone. Raising these questions might prompt the sort of mutually respectful relations that their solution requires.

ENDNOTES

- 1 My initial interest was in the choice anomalies (e.g., starting point bias) that CV investigators had documented, with more systematic samples (and larger budgets) than psychologists had ever mustered. My naive plan was to let the economists collect the data, from which we would then make science. The same data could serve quite different roles for economists trying to get the bugs out of a method and psychologists trying to account for response patterns (Fischhoff, 1991).
- 2 Other sections deal with matters of principle, such as how to avoid sexism in writing (gracefully), how to attribute credit fairly to research collaborators, and how to reassure readers regarding the protection of human subjects and the archiving of research materials.
- 3 Gregory, Lichtenstein & Slovic (1993) offer another set of criteria for this general niche. Generally speaking, I believe that the present set subsumes four of their criteria (accommodate the multi-dimensionality of value, exclude irrelevancies, separate facts from values, ask the right question). However, I reject Gregory et al.'s fifth criterion, minimize response refusals, believing that subjects may have principled objections to a task; capturing those concerns is part of a full description of results. Accepting the legitimacy of such expressions of discomfort should reduce any implicit pressure to suppress them.
- 4 Fischhoff (1993) discusses the analogous problems in experimental psychology. In a sense, these problems are much more severe than with CV, because so much more is typically left unsaid and there is so much less real-world context to provide guidance.

- 5 The failure of California's "Big Green" initiative provides an example of the ambiguous message of a clear electoral majority. The vote told little about which of its many provisions were critical, much less how well they were understood. Equally inscrutable is the June 1994 failure of a California initiative combining environmental protection and crime control, but appearing on the ballot along with three initiatives for earthquake-related spending.
- 6 We have used this procedure in creating risk communications for controversial fields (e.g., the health effects of electromagnetic fields) (Morgan, 1991).
- 7 Witness Carson and Mitchell's distinction between "a WTP question" and a "CV response" (1993, p. 1267), sacrificing one important term to the opposition in order to preserve a distinct meaning for a more critical one.

REFERENCES

- Ajzen, I. & Peterson, G.L. (1988). Contingent value measurement: The price of everything and the value of nothing? In B. Driver, G. Peterson, & R. Gregory (Eds.), *Evaluating amenity resources*. New York: Venture.
- American Psychological Association (1984). *Publication manual* (second edition). Washington, D.C.: American Psychological Association.
- Baddeley, A. (1979). Applied cognitive and cognitive applied research. In L.G. Nilsson (Ed.), *Perspectives on memory research*. Hillsdale, NJ: Lawrence Erlbaum.
- Barber, B. (1991). *Strong democracy*. Berkeley, CA: University of California Press.
- Baron, J. (1994). Rationality and invariance: Response to Schuman. Unpublished manuscript. Presented at DOE/EPA Workshop on Using Contingent Valuation to Measure Non-Market Values.
- Bostrom, A., Fischhoff, B. & Morgan, G. (1992). Characterizing mental models of hazardous processes: A methodology and an application to radon. *Journal of Social Issues*, 48(4), 85-100.
- Burgess, J., Harrison, C.M. & Limb, M. (1988). People, parks and the urban green. *Urban Studies*, 25, 455-473.
- Cambridge Economics. (1992). *Contingent valuation: A critical assessment*. Cambridge, MA: Author.

- Campbell, D.T. & Fiske, R.W. (1959). Convergent and discriminant validity by the multitrait-multimethod matrix. *Psychological Bulletin*, 56, 81-105.
- Carson, R.T., & Mitchell, R.C. (1993). The issue of scope in contingent valuation studies. *American Journal of Agricultural Economics*, 75, 1263-1267.
- Chi, M. (in press). Analyzing the content of verbal data to represent knowledge: A practical guide. *Journal of the Learning Sciences*.
- Cohen, J. (1962). The statistical power of abnormal social psychological research. *Journal of Abnormal and Social Psychology*, 65, 145-153.
- Coombs, C. (1964). *A theory of data*. New York: Wiley.
- Crosby, N., Kelly, J.M. & Schaefer, P. (1986). Citizen review panels. *Public Administration Review*, 46, 170-178.
- Dawes, R.M. (1972). *Fundamentals of attitude measurement*. New York: Wiley.
- Dawes, R.M. (1988). *Rational choice in an uncertain world*. San Diego, CA: Harcourt Brace Jovanovich.
- Dawes, R.M. (1994). *House of cards*. New York: Random House.
- Department of Commerce (1993). Report of the NOAA Panel on Contingent Valuation. *Federal Register*, 58(10), 4602-4614.
- Dietz, T. (1987). Theory and method in social impact assessment. *Sociological Inquiry*, 57, 54-69.
- diSessa, A. (1993). Toward an epistemology of physics. *Cognition and Instruction*, 10, 105-225.
- Ericsson, A. & Simon, H.A. (1984). *Verbal reports as data*. Cambridge, MA: MIT Press.
- Fiorino, D.J. (1990). Citizen participation and environmental risk: A survey of institutional mechanisms. *Science, Technology & Human Values*, 15, 226-243.
- Fischhoff, B. (1980). Clinical decision analysis. *Operations Research* 28, 28-43.

- Fischhoff, B. (1988). Specifying value measurements. In B. Driver, G. Peterson, and R. Gregory (Eds.), *Evaluating amenity resources* (pp. 107-116). New York: Venture.
- Fischhoff, B. (1990). What's it worth to you? A review of *Using surveys to value public goods: The contingent valuation method* by R.C. Mitchell & R.T. Carson. *Public Opinion Quarterly*, 54, 286-294.
- Fischhoff, B. (1991). Value elicitation: Is there anything in there? *American Psychologist*, 46, 835-847.
- Fischhoff, B. (1992). Giving advice: Decision theory perspectives on sexual assault. *American Psychologist*, 47, 577-588.
- Fischhoff, B. (1993). Transaction analysis: A framework and an application to insurance decisions. *Journal of Risk and Uncertainty*, 7, 53-69.
- Fischhoff, B. & Furby, L. (1986). *A review and critique of Tolley, Randall et al. "Establishing and valuing the effects of improved visibility in the Eastern United States"*. ERI Technical Report 86-8. Eugene, OR: Eugene Research Institute.
- Fischhoff, B. & Furby, L. (1988). Measuring values: A conceptual framework for interpreting transactions. *Journal of Risk and Uncertainty*, 1 147-184.
- Fischhoff, B., Quadrel, M.J., Kamlet, M., Loewenstein, G., Dawes, R., Fischbeck, P., Klepper, S., Leland, J., & Stroh, P. (1993). Embedding effects: Stimulus representation and response modes. *Journal of Risk and Uncertainty*, 6, 211-234.
- Fischhoff, B., Slovic, P. & Lichtenstein, S. (1980). Knowing what you want: Measuring labile values. In T. Wallsten (Ed.), *Cognitive processes in choice and decision behavior* (pp. 117-141). Hillsdale, NJ: Erlbaum.
- Funtowicz, S. & Ravetz, J. (1990). *Uncertainty and quality in science for policy*. London: Kluwer.
- Furby, L. & Fischhoff, B. (1989). *Specifying subjective evaluations. A critique of Dickie et al.'s interpretation of their contingent valuation results for reduced minor health symptoms*. USEPA Cooperative Agreement #CR814655-01-0. Eugene, OR: Eugene Research Institute.
- Furnham, A. (1989). *Intuitive theories*. London: Plenum.
- Gentner, D. & Stevens, A.L. (Eds.) (1983). *Mental models*. Hillsdale, NJ: Erlbaum.

- Green, D.P., Kahneman, D. & Kunreuther, H. (1994). How scope and method of public funding affects willingness to pay for public goods. *Public Opinion Quarterly*, 58, 49-67.
- Gregory, R., Lichtenstein, S., & Slovic, P. (1993) Valuing environmental resources: A constructive approach. *Journal of Risk and Uncertainty*, 7, 177-197.
- Hamm, R.M., Bennett, L., Howe, C., & Wunderlich, K. (1992). Contingent valuation: Stability and meaning of "not willing to pay" responses. Unpublished manuscript. Institute of Behavioral Science. University of Colorado, Boulder.
- Harrison, G.W. (1993). Valuing public goods with the contingent valuation method: A critique Kahenman & Knetsch. *Journal of Environmental Economics and Management*, 23, 248-257.
- Hogarth, R. (Ed.) (1982). *Question framing and response consistency*. San Francisco: Jossey-Bass.
- Jungermann, H., Schutz, H. & Thuring, J. (1988). Mental models in risk assessment: Informing people about drugs. *Risk Analysis*, 8, 147-155.
- Kahneman, D. & Knetsch, J. (1992). Valuing public goods in the purchase of moral satisfaction. *Journal of Environmental Economics and Management*, 22, 52-70.
- Kahneman, D. & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica*, 47, 263-281.
- Keeney, R.L. (1993). *Value-focused thinking: A path to creative decision making*. Cambridge, MA: Harvard University Press.
- Keeney, R.L. & Raiffa, H. (1976). *Decisions with multiple objectives*. New York: Wiley.
- Kintsch, W. (1986). Learning from text. *Cognition and Instruction*, 3, 87-108.
- Kristrom, B. (in press). Practical problems in contingent valuation. this volume.
- Lazo, J.K., Schulze, W.D., McClelland, & Doyle, J.K. (1992). Can contingent valuation measure nonuse values? *American Journal of Agricultural Economics*, 75, 1126-1132.

- Leventhal, H. & Cameron, L. (1987). Behavioral theories and the problem of compliance. *Patient Education and Counseling*, 10, 117-138.
- McConnell, K.E. (1990). Models for referendum data; The structure of discrete choice models for contingent valuation. *Journal of Environmental Economics and Management*, 18, 19-34.
- McGuire, W. (1969). Suspiciousness of experimenter's intent. In R. Rosenthal & R.L. Rosnow (Eds.), *Artifact in behavioral research*. New York: Academic Press.
- Merton, R.K. (1987). The focussed interview and focus groups. *Public Opinion Quarterly*, 51, 541-557.
- Merton, R.K. & Kendall, P.I. (1946). The focussed interview. *American Journal of Sociology*, 51, 541-557.
- Mitchell, R.C. & Carson, R.T. (1989). *Using surveys to value public goods: The contingent valuation method*. Washington, DC: Resources for the Future.
- Morgan, G. (1991). *Health risks of electromagnetic fields*. Department of Engineering and Public Policy: Carnegie Mellon University.
- National Research Council. (1982). *Survey measure of subjective phenomena*. Washington, DC: The Council.
- Navrud, S. (in press). Does the presentation of information matter in contingent valuation studies -- experimenting with video. This volume.
- Nickerson, C.A.E. (1993). Valuing public goods: A comment on Harrison's critique of Kahneman and Knetsch. *Journal of Environmental Economics and Management*, 25, 93-102.
- Nisbett, R.E. & Wilson, T.D. (1977). Telling more than we know: Verbal reports on metnal processes. *Psychological Review*, 84, 231-259.
- Orne, M.T. (1962). On the social psychology of the psychological experiment. *American Psychologist*, 17, 776-783.
- Peterson, G.L., Driver, B.L., & Gregory, R. (Eds.), (1988). *Amenity resource valuation: Integrating economics with other disciplines*. State College, PA: Venture.
- Quadrel, M.J., Fischhoff, B. & Palmgren, C. (1994). Adolescents' definitions of risky events. manuscript under editorial review.

- Reder, L. (1985). Techniques available to author, teacher, and reader to improve retention of the main ideas of a chapter. In S.F. Chipman, J.W. Segal & R. Glaser (Eds.), *Thinking and learning skills* (vol. 2) (pp. 37-64). Hillsdale, NJ: Erlbaum
- Reich, R.B. (1985). Public administration and public deliberation. *Yale Law Review*, 94, 1617-1641.
- Schkade, D. & Payne, J.W. (1994). How people respond to contingent valuation questions: A verbal protocol analysis of willingness to pay for an environmental regulation. *Journal of Environmental Economics and Management*, 26, 88-109.
- Schrivver, K.A. (1989). Evaluating text quality: The continuum from text-focused to reader-focused methods. *IEEE Transactions on Professional Communication*, 32, 238-255.
- Schulze, W., McClelland, G., Waldman, D., & Lazo, J. (in press). Information, context, and bias in contingent valuation. This volume.
- Schwarz, N. (1994). Judgment in social context. In L. Berkowitz (Ed.), *Advances in experimental social psychology*, vol 26. San Diego: Academic Press.
- Smith, V.K. (1992). Arbitrary values, good causes, and premature verdicts. *Journal of Environmental Economics and Management*, 22, 71-89.
- Smith, V.K. & Osborne, L. (in press). Do contingent valuation estimates pass a "scope" test?: A preliminary metaanalysis. This volume.
- Stokols, D. & Altman, I. (Eds.) (1986). *Handbook of environmental psychology* (2 vols). New York: Wiley.
- Strack, F., & Schwarz, N. (1992). Communicative influences in standardized question situations. In G.R. Semin & K. Fiedler (Eds.), *Language interaction and social cognition*. London: Sage.
- Turner, C.F. & Martin, E. (Eds.) (1984). *Surveying subjective phenomena*. New York: Sage.
- Tversky, A. & Kahneman, D. (1971). Belief in the "law of small numbers." *Psychological Bulletin*, 76, 105-110.
- Tversky, A. & Kahneman, D. (1981). The framing of decisions and the psychology of choice. *Science*, 211, 453-458.

- Verba, S. & Nie, N.H. (1972). *Participation in America*. Chicago: University of Chicago Press.
- Viscusi, W.K. (in press). Economic and psychological aspects of valuing risk reduction. this volume.
- Viscusi, W.K., & Magat, W.A. (1987). *Learning about risk: Consumer and worker responses to hazard information*. Cambridge, MA: Harvard University Press.
- Voss, J.F., Tyler, S.W. & Yengo, L.A. (1983). Individual differences in the solving of social science problems. *Individual Differences in Cognition*, 1, 205-232.
- Whittington, D. & Davis, D. (1993). *A review and assessment of techniques for Systematic Client Consultation*. Department of Environmental Sciences & Engineering, University of North Carolina at Chapel Hill.

