

Determinants of Stated Willingness to Pay for Public Goods: A Study in the Headline Method

DANIEL KAHNEMAN

Woodrow Wilson School of Public Affairs, Princeton University

ILANA RITOV

Ben-Gurion University, Beer-Sheva, Israel

Abstract

Respondents were shown brief statements (“headlines”) referring to various threats to the environment or to public health, and other public issues. An intervention to deal with each problem was also introduced by a single sentence. Some respondents were asked to indicate their willingness to pay for the interventions by voluntary contributions. Others indicated their opinion of the intervention on a conventional rating scale, rated the personal satisfaction of contributing to it, or rated the importance of the problem. Group averages of these response measures were obtained for a large set of issues. Computed over issues, the rank-order correlations between the different measures were very high, suggesting that group averages of WTP and of other opinion statements are measures of the same public attitudes. Observed preference reversals and violations of monotonicity in contributions are better explained by a concept of attitude than by the notion of economic value that underlies the contingent valuation method. Contributions and purchases do not follow the same logic. Possible implications for the contingent valuation method are discussed.

Key words: contingent valuation, willingness-to-pay, preference reversals

The prospect of litigation over high stakes has caused a remarkable surge of interest in the contingent valuation method (CVM) for the assignment of dollar values to environmental goods. The controversy is sharpest in the context of the measurement of existence, or passive use values. When applied to the use value of goods from which the respondents derive consumption benefits, the contingent valuation method presumably shares the strengths and weaknesses of familiar market research techniques. The situation is quite different when respondents are asked to value public goods that they will not personally use. Respondents have no relevant market experience, and no opportunity to choose between existence goods and their substitutes. Although some temperate voices have been heard (Fischhoff and Furby, 1988; Lazo et al., 1992; McClelland et al., 1992), the controversy about the use of CVM to measure existence value has been unusually intense, and the proponents of the method (e.g., Hanemann, 1992; Randall, 1992) and opponents (e.g., Diamond and Hausman, 1993) share little common ground. It is fair to say that some critics consider CVM thoroughly discredited by conceptual and methodological weaknesses, whereas the practitioners of contingent valuation are often appalled by critical research that they judge to be of poor quality.

One possible reason for these unusually sharp differences may be that the practice of CVM has evolved very rapidly in recent years, sometimes in response to critiques, sometimes in advance of the critics. As a consequence, critical studies can often be dismissed because they do not meet ever-improving standards. Issues of research costs also play an important role. The practitioners of CVM have worked in the context of large litigation projects, and have been subject to few cost constraints: the currently recommended procedure dedicates an entire face-to-face interview to obtain a single yes/no response to one hypothetical referendum question. The cost of duplicating such elaborate procedures for a critical experiment is dauntingly high. Would-be critics have also been hampered by the slow dissemination of information about state-of-the-art studies conducted in the context of litigation. Thus, the critics of contingent valuation have in some occasions been shooting their arrows at a target that is both rapidly moving and partly concealed. The obstacles to credible critical research hinder orderly scientific dialogue about the method.

The present article illustrates a new approach to the study of willingness to pay for public goods, which we call the *headline method*. Our starting point is the observation that statements of willingness to pay for public goods are easy to elicit, even when the goods in question are sketchily described. We seek to model the reader of a newspaper headline that describes a public problem and a proposed intervention, such as: "SHARP DROP IN VISIBILITY RECORDED IN SEVERAL NATIONAL PARKS. Special fund set up to improve air quality in parks." A question about willingness to pay (WTP) into this fund will yield a distribution of WTP that is not obviously different from results obtained in conventional contingent valuations. The headline method makes it relatively easy to study the determinants of WTP, because a respondent can quickly state WTP for 10–20 different issues without succumbing to boredom or exhaustion.

There are large and important differences between the methods used to elicit willingness to pay in the headline study and in contingent valuation. In particular, the respondent in a professionally conducted contingent valuation is given substantial information, both about the problem and about the outcome of the proposed intervention. Some authors would restrict the use of the term "willingness to pay" to studies that meet these standards, which the headline method does not attempt to meet. Given these differences of method, why would findings obtained in the headline method be relevant to contingent valuation? We justify the headline method in terms of a general hypothesis of *process continuity*. Research on human judgment and decision making during the last two decades has generally supported the proposition that the errors and biases that affect quick intuitive judgments and decisions are usually present, though sometimes in modified or moderated form, in judgments that are made more slowly and carefully (Bazerman, 1990; Gilovich, 1992; Kahneman, Slovic, and Tversky, 1982). Process continuity provides a useful first approximation, in spite of some demonstrated differences between effortless, automatic processing of information and the effortful and deliberate variety (Petty and Caccioppo, 1986). In the present context, process continuity suggests substantial overlap between the factors that determine the response to a WTP query in contingent valuation and in the more casual setting of the headline method.

Clearly, the conclusions of a headline study cannot simply be assumed to extend to contingent valuation, because of the multiple procedural differences between the methods.

However, we propose that the findings of a headline study may yield credible hypotheses about contingent valuation, and in some cases shift the burden of proof to the proponents of CVM. Our methodological position is best explained in terms of a legal analogy. We consider it equally unreasonable for the CVM to be considered innocent of all flaws until proven guilty, or to be declared flawed because of some difficulty that affects a less stringent method. The solution is to introduce the Scottish verdict of "Not proven." The demonstration of some undesirable properties of the WTP response in a headline study should change the status of the relevant property of CVM from "presumptively innocent" to "not proven," with an implied requirement for proponents of the method to eliminate the suspicion if they can do so. The present article includes several observations that appear to challenge contingent valuation in this fashion.

Two competing interpretations of willingness to pay for public goods have been advanced in discussions of contingent valuation: we call them the *purchase model* and the *contribution model*. The purchase model is the theoretical foundation of the contingent valuation method. A respondent exposed to a valuation question is supposed to compare two states of the world: the status quo (or the state of affairs that will result if no action is taken), and an alternative state in which a public good is provided and the respondent's wealth is diminished by some amount. The maximal WTP for the good is taken to be the highest cost at which the respondent still prefers the provision of the public good over the status quo. The same model is applied when the survey question is framed as a referendum about a levy earmarked for a particular good: respondents are assumed to support the new tax if and only if it is below their maximal WTP for the good. The valuation of public and private goods is fundamentally similar in this model, which therefore permits the estimation of a contingent market value for public goods. The assumption that public and private goods are valued alike is appealing, but not necessarily true: Green (1992) has observed large systematic differences in the shape of the distribution of WTP for these families of goods.

The contribution model (Kahneman and Knetsch, 1992; Kahneman et al., 1993) assumes that respondents view the provision of some public goods as good causes that need supporting. This interpretation is particularly applicable to goods that have only existence or passive-use value, such as the pristine beauty of sites one never expects to visit. A stated willingness to support a new tax or to make a voluntary monetary contribution to the solution of a public problem is one of many possible indications of the respondent's *attitude* to the problem and to its solution (Ajzen, 1988; Eagly and Chaiken, 1993). Attitude is construed as an evaluative tendency, which can be favorable or unfavorable (Eagly and Chaiken, 1993). A favorable attitude to an object is usually correlated with favorable attitudes to actions that will protect that object from harm, or restore it if it has been harmed. In the context of a contribution model, we interpret stated WTP as an expression of attitude to a public good, using a (possibly arbitrary) dollar scale. Note that by labeling stated WTP a measure of attitude we do not imply that this response is irrelevant to actual behavior in the real world. The relationship between stated attitudes and behavior is a much researched and very complex issue, which we set aside in this article (see Eagly and Chaiken, 1993, chapter 4). Our purpose here is to contrast the psychological rules that govern attitudes to the logic of economic values.

Unlike economic values, attitudes do not generally conform to the rule that “more is better.” An individual may have an equally positive attitude to the Grand Canyon and to the broader notion of National Parks. Another difference between the concepts is that an attitude is assumed to be elicited by an object considered in isolation; a choice between two objects is not necessarily made by picking the object that evokes the more favorable attitude. Thus, the notion of attitude is compatible with “preference reversals,” in which a respondent is willing to pay more for A than for B, but will prefer B over A in a direct choice. Attitudes are also affected by factors that would be considered irrelevant in standard economic analyses, such as the special “dread” associated with some causes of death (Savage, 1993). Finally, an important difference between attitudes and economic values is that contributions that express attitudes may provide intrinsic satisfaction, whereas payments that procure economic values are commonly assumed to be aversive. The contrast between attitudes and economic values helps explain the difference between purchases and contributions. Consider the question “What is the most you would contribute to prevent starvation deaths in Africa this year?” The purchase interpretation of the question is “what is the highest cost to your household at which you prefer total famine relief in Africa to the currently expected level of famine?” This interpretation is far-fetched. The assumptions of a contribution are more reasonable: some famine relief will surely be provided, relief is unlikely to be complete, and any individual’s contribution is a drop in a large bucket. The contribution is an expression of the intensity of the attitude toward the provision of famine relief. Another way of saying the same thing is that an individual who has a favorable attitude to a cause derives utility from contributing to it. The utility of contributing \$10 to famine relief comes from the warm glow of giving (Andreoni, 1990), not from the minuscule increase in the funding of that cause (Margolis, 1982). Thus, a contribution can be viewed as a purchase of moral satisfaction (Kahneman and Knetsch, 1992). To test this notion, Kahneman and Knetsch (1992) asked respondents to indicate how much personal satisfaction they would derive from contributing to different causes. They found that the issues that were rated high on personal satisfaction tended to elicit the highest WTP. The present study will extend these findings.

The contribution model makes sense of the *embedding effect*, the observation that WTP sometimes fails to increase with the size of the public good that is to be provided. Consider an individual who would state equal WTP to eliminate famine in Ethiopia or in the whole of Africa. A purchase model implies the unreasonable inference that this individual attaches no value to famine relief in the rest of Africa, if only Ethiopia is saved. The more plausible inference from a contribution model is that the individual has equally favorable attitudes to the two activities of famine relief, when considering them separately. Thus, the contribution model draws significant support from demonstrations of embedding effects, such as the early finding by Kahneman and Knetsch of almost equal WTP for a cleanup of all Ontario or in one region of it (Kahneman, 1986; see also Desvousges et al., 1992; Fischhoff et al., 1993; Kahneman and Knetsch, 1992; Kemp and Maxwell, 1993). We shall report a new variant of an embedding effect.

A contribution model and a purchase model assign different weights to two elements normally included in WTP questions: the *problem* (e.g., a threat to a species), and the *intervention* proposed to alleviate it (e.g., the establishment of refuges). In contrast to

everyday purchases, where customers normally insist on fairly detailed information about the goods they buy, contributors to good causes are usually content with the general knowledge that something will be done. The impetus for charitable giving is the urgency of a problem, not the perfection of a solution. Accordingly, we expect participants in WTP surveys to show little sensitivity to interventions. The present study examines this issue in two different ways: (i) by direct manipulations of the scope of interventions, and (ii) by correlating WTP for an intervention with judgments of the importance of the problem, made with no reference to any particular intervention.

The main role of the concept of attitude is to organize a set of correlated measures. In standard studies of attitudes the relevant correlations are computed over individuals; the covariance to be explained is the covariance of individual differences. In the context of contingent valuation, however, the correlation of interest is over issues, not over individuals. The task of contingent valuation is to provide a useful measure that discriminates issues by the average value that the public attaches to them, and the question to be answered here is whether different measures order issues in the same way. A central goal of the present study was to establish whether issues are ranked alike by a set of alternative indications of public attitude, including WTP, political support for governmental intervention, ratings of the importance of the issue, and ratings of the personal satisfaction of participating in the solution. We examine these correlations over a large and diverse set of public issues, including threats to the environment and to public health.

The present study in the headline method differs from contingent valuation in multiple respects, including the selection of the sample. Our purpose in this study was not to estimate the distribution of willingness to pay or other responses in the general population. Our objective was to look for particular patterns in comparisons across response measures, across issues, and across proposed solutions to the same problem. A convenience sample was considered adequate for this limited purpose.

1. Method

Data were collected during the summer of 1991 from 1441 volunteers, recruited among visitors at the Exploratorium, a science museum in San Francisco. A sign was posted describing the study as a brief survey of attitudes about the environment. Only volunteers who stated that they had filed a tax return for the preceding year were invited to participate. Respondents were paid \$2 and informed that a contribution would be made to the museum for each completed questionnaire. Most respondents completed their questionnaire in 10–15 minutes. Questionnaires were anonymous, but respondents were asked to report their age, sex, education, and state of residence. The sample included respondents between the ages of 17 to 80; mean age was 34.1. There were 46.4% female, 49% male, and 4.6% did not report their sex. Residents of California made up 80% of the group of respondents. The median educational level was “some years of college.” A set of 37 issues were selected, representing threats to animal species, threats to vegetation, major ecological issues, threats to public health, and miscellaneous issues of public goods. Each issue was described by a brief headline, as in the following example: “FERTILITY LOSS DUE TO POLLUTION THREATENS SEVERAL SPECIES OF

REPTILES ON THE MEXICAN COAST.” For some of the issues we constructed several alternative versions, to provide an experimental manipulation. For example, the word “REPTILES” in the preceding headline was replaced in other versions by “LIZARDS” and by “SEA TURTLES.” Including the alternative versions, there were a total of 53 different headlines.

Each headline was paired with one or more interventions. For example, the intervention associated with the three versions of the reptile problem was stated as follows: “Intervention: support program to increase fertility by hormonal treatment.” For some of the problems, the intervention was experimentally manipulated across alternative versions. For example, the headline “MANY INNER CITY HOMES HAVE LEAD-BASED WALL PAINT. LEAD POISONING AFFECTS THE CENTRAL NERVOUS SYSTEM” was associated with four different interventions in separate questionnaires. A complete list of headlines and interventions is given in the Appendix, which also associates a brief descriptive phrase with each problem. These phrases are used to identify the problems in section 2.

The general introduction to the questionnaires stated that “This study is concerned with people’s attitudes about environmental and related public issues. Information about the goals of the survey will be provided after you complete the questionnaire.” The instructions further stated that “In order to understand your attitudes about environmental issues, a series of problems will be presented. Most are real, some are fictitious, all are possible.” From that point, different instructions were given in four types of questionnaire, devoted to the four dependent measures of the present study: *WTP*, *Support* (for government intervention), personal *Satisfaction*, and *Importance*. Respondents were warned in advance that they would be asked about several environmental and other public issues, and were instructed to consider each issue separately.

WTP. “For each problem, a possible intervention is briefly described. This intervention would be in addition to the normal activities of federal and state governments in that area. It would be funded by voluntary contributions. The state income tax form in some states includes a section in which taxpayers can make voluntary, tax-deductible contributions to various activities. In considering each of the following problems, we ask you to respond as you would do if you encountered it as an entry in such a tax form, and had to determine whether or not you want to make a voluntary contribution, and how large it would be.”

Support. “For each problem, a possible intervention by state or federal government is briefly described. This intervention would be in addition to the normal activities of federal and state governments in that area. We want to know whether you would support these extra interventions, which would impose additional demands on public funds.”

The response scale for the Support measure ranged from “0 ___ No, I would not support intervention, if it involves any extra costs” to “4 ___ Yes, I would support intervention, whatever the costs.”

Satisfaction. “For each problem, a possible intervention is briefly described. This intervention would be in addition to the normal activities of federal and state governments

in that area. It would be funded by voluntary contributions. For each problem, please answer the following question: Consider the possibility of making a voluntary contribution of money to the proposed intervention. How much personal satisfaction would you expect to get from making such a contribution?"

The response scale ranged from "0 ___ No personal satisfaction at all" to "4 ___ A great deal of personal satisfaction." This question was asked about only 19 of the problems.

Importance. "Many, but not all of these problems have to do with environmental issues. Please consider each problem as one of the multitude of projects on which public money and attention could be spent. Try to assess the importance of the proposed intervention relative to all other projects for which public action is required. Rate each project on the scale shown below.

- 0 ___ does not deserve attention
- 1 ___ significant but not important
- 2 ___ less important than most other problems
- 3 ___ moderately important
- 4 ___ important
- 5 ___ very important
- 6 ___ one of the most important

For each problem, we are interested in your intuitive, immediate response as to its importance. Since our interest is in your first reaction, please respond as quickly as you can. In addition, we would like to emphasize that there are no right or wrong answers."

There were a total of 19 different questionnaires: 7 for WTP, 7 for the Support measure, 2 for Satisfaction, and 3 for Importance. The questionnaires for the first three variables included 9-11 questions; the Importance questionnaires, which did not mention interventions, included 17-19 headlines. For WTP and Importance two versions of each questionnaire were constructed, with different question orders. Most questionnaires ended with two choice questions. Two problems and associated interventions were introduced, and the respondent was asked to choose between them, assuming that budget constraints made this choice necessary.

2. Results and discussion

2.1. Overview

The results of the study will be presented in the following sequence. In section 2.2 we describe the distribution of WTP and examine order effects. Section 2.3 presents the

mean responses to 37 issues and provides an account of the variance of the different response measures (table 1). Section 2.4 presents the correlations between the dependent measures, over the set of 37 issues (table 2). The next sections are concerned with particular experiments. Section 2.6 reports a new variant of the embedding effect, in which WTP and measures of attitude vary non-monotonically with the inclusiveness of a problem (table 3). Sections 2.7 (table 4) and 2.8 (table 5) respectively present results on the effects of varying the stated severity of a problem, and of attributing it to a human or to a natural cause. Section 2.9 (tables 6–9) summarizes several experiments that manipulated the intervention proposed for a given problem. Finally, section 2.10 (tables 10–11) examines preference reversals, in which the ordering of issues by WTP does not correspond to the preferences indicated in a direct choice.

2.2. Preliminary analyses

We begin with some descriptive statistics on the distribution of WTP responses. Overall, about half of our sample ($N = 786$) answered WTP questions, offering a total of 8194 responses. There were 53% positive responses, a rather lower value than is commonly observed in contingent valuations, but perhaps not surprising because many of the issues we presented were selected to appear rather trivial. The grand mean of WTP was \$21.5 and the median was \$1. The mean and median of positive responses were respectively \$40.20 and \$10. There were 1.6% of responses in excess of \$100. Winsorizing the data by adjusting all these responses to \$100 reduced the grand mean of WTP to \$12 (a reduction of 44%), but did not affect the median WTP for any of the questions. Unless otherwise described, all subsequent analyses are based on Winsorized data.

Over individuals, the proportion of positive responses ranged from 0% to 100%, with a mean of 53% and a median of 55%. The correlation between the proportion of positive responses and the mean value of these responses was .07. Thus, the propensity to make positive contributions and the size of these contributions appear to be essentially independent characteristics of respondents.

The issue of order effects arises in any questionnaire in which respondents answer multiple questions of the same type. We considered a gradual decline of responses the most plausible general hypothesis about an order effect (Tolley and Randall, 1983). To test this hypothesis, we computed, within the responses of each individual, a correlation between the response and question number.¹ The mean correlations across all questionnaires were $-.001$ for WTP ($N = 730$), $-.003$ for political support ($N = 401$), $-.02$ for ratings of importance ($N = 145$), and $-.06$ for moral satisfaction ($N = 103$). Only the correlation for moral satisfaction differed significantly from zero ($p = .02$). We conclude that respondents faced with multiple WTP questions in a headline study can follow an instruction to treat the problems as independent. The assumption of independence between successive WTP questions greatly simplifies the analysis.

2.3. Responses to issues

Table 1 presents some summary statistics of the different response measures, for 37 issues. The issues are categorized by content. Stars indicate issues for which several different versions were used; the values shown in the table are weighted means of the responses to all versions. Differences among versions will be discussed later. The total number of respondents answering each question ranged from 104 to 116 for the WTP measure, and from 48 to 60 for all other measures. The number of respondents is correspondingly larger for starred issues.

In addition to the mean of Winsorized WTP, two other statistics of the WTP measure are shown for each of the 37 issues: the percentage of positive (non-zero) responses, denoted $\%(WTP)$, and the mean of individually normalized positive contributions, denoted $N(WTP)$. To obtain $N(WTP)$, the WTP responses of each individual were divided by the mean of the positive contributions of that individual. Because the distribution of positive WTP tends to be positively skewed within the data of individual respondents, 70% of these values were strictly lower than 1. The mean of the normalized positive contributions was then computed for each issue. This value is larger than 1 for issues that generally attracted higher-than-average contributions, when they attracted any at all.

The restricted range of mean values is a striking aspect of the WTP results. The means of uncorrected WTP range from a low of \$2.24 for a threat to Spanish moss to a high of \$118.40 for solar energy research to combat global warming, but this variability was largely due to a few extreme contributions. After Winsorizing, the means ranged from \$2.24 for a threat to Spanish moss up to \$23.96 for a response to massive burning and clear-cutting in the rain forests of South America. The individuals who made a positive response to the preservation of Spanish moss (25% did), indicated a willingness to pay half of the mean of all their positive contributions for that insignificant cause. Over $\frac{2}{3}$ of the mean Winsorized WTP shown in table 1 fall between \$5 and \$16. This lack of sensitivity to the importance of issues presents a serious problem, because it seems highly implausible that a ratio of 10:1 (or even 100:1) provides an adequate range for the value of public goods. Note that there is no particular reason to expect greater variability of responses in contingent valuation, where respondents face a single issue, than in the present headline method. An issue considered in isolation is likely to gain in importance because it has been singled out. In contrast, the exposure to many different headlines should have encouraged our respondents to increase the variability of their responses. The limited range of WTP in the headline method raises a question that should be answered for CVM as well: Does the WTP response have the sensitivity required of a measure of the value of public goods? D'Arge (1989) raised this question some years ago, and concluded that estimates of WTP cluster in an implausibly narrow range.

The bottom rows of table 1 summarize an account of the variance of the different variables. Two-way analyses of variance (Issues X Respondents, both treated as random effects) were conducted for each of the 19 separate questionnaires used in the study (7 for WTP, 7 for political support, 3 for importance ratings, and 2 for moral satisfaction). For each of these data sets we computed the percentages of variance associated with (i) differences in the mean responses to the different issues, (ii) individual differences in mean responses, and (iii) a residual term that combines Issue X Respondent interactions

Table 1. Responses to all issues

	WTP	% (WTP)	N (WTP)	Support	Importance	Satisfaction
<i>Animal Species</i>						
American elk	7.69	59%	.82	2.07	5.00	2.51
Ferret	9.55	44%	.66	1.66	4.92	1.62
Elephants	16.28	74%	.98	2.64	5.88	2.96
Marine life	13.08	59%	.96	2.41	6.16	2.92
Kangaroo rats	6.33	34%	.64	1.65	4.19	—
Florida panther	6.81	45%	.72	1.83	5.35	—
Falcon shell	11.21	60%	.76	2.13	5.86	2.38
Dolphins*	12.57	56%	1.02	1.92	5.60	—
Australian mammals*	8.42	51%	.82	1.70	4.98	2.45
Coastal reptiles*	3.60	24%	.61	1.27	5.30	1.67
Wildlife*	13.30	65%	.94	2.36	5.63	2.45
Birds*	8.91	54%	.77	2.18	5.63	—
Spotted owl*	14.55	63%	.99	2.12	5.33	—
<i>Plant species</i>						
Coral reefs	12.03	60%	.99	2.47	5.66	2.61
Mushroom	3.63	25%	.55	1.36	4.51	1.56
Pine disease	5.76	46%	.69	2.24	4.96	2.18
Spanish moss	2.24	25%	.52	1.02	4.39	—
Pine trees*	9.79	54%	.70	1.71	4.65	—
<i>Ecological damage</i>						
Wetlands	11.49	58%	.90	2.29	5.47	—
CO ₂ in 3rd world	5.44	41%	.90	2.03	5.53	—
CO ₂ —oil burning	19.88	63%	1.26	2.68	6.19	—
Automobile pollution	18.57	54%	1.41	2.36	5.88	2.63
Burning rain forest	23.96	79%	1.28	2.84	6.17	—
Visibility in Parks*	9.56	47%	.94	2.02	4.65	—
Toxic waste dumps*	20.67	72%	1.26	2.86	6.36	2.95
Shrinking rain forest*	16.93	66%	1.20	2.73	6.22	—
Toxic spills*	9.98	43%	1.03	2.44	6.56	—
Solid waste*	15.64	67%	1.07	2.61	6.35	—
<i>Miscellaneous goods</i>						
Historic buildings	5.67	47%	.78	1.74	5.13	2.27
Lighthouses	6.44	34%	.60	1.04	2.96	—
Earthquake safety	8.40	37%	.78	1.81	5.23	—
<i>Public health</i>						
Power lines leukemia	17.42	69%	1.11	2.56	6.20	2.57
Increase in anemia*	15.27	68%	.97	2.25	5.57	—
Skin cancer in farmers	12.18	56%	.93	2.43	5.22	2.31
Increase in myeloma*	12.91	62%	1.00	2.23	5.58	2.29
AIDS in Africa*	10.96	45%	1.03	1.99	5.89	—
Lead paint poisoning*	13.84	62%	.95	2.43	6.06	2.71
<i>Variance account (%)</i>						
Problems	4.1	—	—	8.3	15.8	12.3
Respondents	46.4	—	—	46.9	25.9	39.4
Error	49.5	—	—	44.8	58.3	48.3

*Different versions of the problem were used. Mean satisfaction rating based on one or two versions only.

and error of measurement. The entries in the table are weighted means of these percentage estimates, over all the questionnaires in which each response measure was collected. (The variance account was not computed for the % (WTP) and N (WTP) measures). The most striking result of this analysis is that the proportion of systematic variance associated with differences among problems is far smaller for WTP than for the other response measures of the present study (see also Kahneman et al., 1993). Because the purpose of all these response scales is to identify systematic differences among the issues, the overall conclusion is that the WTP measure appears to be psychometrically inferior to ratings of both political support and problem importance.

2.4. Correlational analyses

A measurement instrument, whether it be a thermometer or a contingent valuation survey, is only useful to the extent that it discriminates objects in a relevant domain. We have already seen that the WTP measure is less efficient in discriminating issues than some of the other response measures collected in this study. Next, we ask whether the different measures order issues in the same way. To answer this question we computed rank-correlations between the mean responses to the set of issues shown in table 1. The results are shown in table 2. The correlations that involve moral satisfaction are based on rankings of 19 issues; other coefficients are based on 37 issues. The entries in table 2 are correlations between ranked mean responses obtained from different samples; they are not estimates of the within-subject correlations that would be obtained if every respondent provided all the responses to each of the issues. Some view this as a serious problem (Nickerson, 1994), but we do not. There are two legitimate interpretations of the correlations of table 2. Most conservatively, the correlations can be taken for what they are: indicators of the agreement between the rankings of issues by the average responses of separate samples answering different questions. This interpretation is adequate in the context of the contingent valuation method, which deals in such averages. Alternatively, the correlations can be taken to describe an idealized subject, as is often done in psychological theorizing.

Table 2. Rank correlations between response measures

	WTP	% (WTP)	N (WTP)	Support	Importance	Satisfaction
WTP	.87					
% (WTP)	.88	.88				
N (WTP)	.87	.72	.90			
Support	.84	.82	.81	.85		
Importance	.76	.66	.83	.84	.88	
Satisfaction	.84	.80	.88	.87	.85	.90

Because correlations between variables are constrained by the reliability of their measurement, we used a bootstrapping procedure to estimate the sampling variability of the ranking of issues by each response measure. New samples were constructed by randomly drawing from the responses to each questionnaire, with replacement, a sample as large as the original group of respondents for that questionnaire. Values corresponding to those shown in table 1 were computed in the data set constructed in this manner, and the issues were ranked by the computed means. This procedure was repeated 1001 times. The rank correlation between the means of each variable in immediately successive samples was computed. The diagonal entries in table 2 are the means of 1,000 correlations for each variable. The results indicate adequate reliability of the orderings. Unlike a previous study in which we carried out the same analysis (Kahneman et al., 1993), the reliability of WTP rankings is quite high, in spite of the fact that the proportion of systematic variance is low for that variable. The larger sample size we used for the WTP variable was apparently sufficient to yield stable rankings of issues. We note that the Winsorizing procedure improves the reliability of the measure: without Winsorizing, the average intercorrelation of mean WTP across samples was .79; after Winsorizing it was .87.

Following the example of McFadden and Leonard (1992) our analysis separates two components of WTP: the proportion of positive contributions, and their size, which in this study is normalized to the mean positive contribution of each individual. The correlation between $\%(WTP)$ and $N(WTP)$ over the 37 issues is only .72, because of a few large discrepancies. The three largest rank differences between the two measures occurred on problems where refusals to contribute were relatively common, but where the mean of positive contributions was high. For example, the issue of toxic spills ranks first in table 1 by its importance, and is tied for 7th rank by $N(WTP)$, but it ranks 30th among the 37 issues by $\%(WTP)$, the percentage of positive responses. Two other issues for which $\%(WTP)$ was low and $N(WTP)$ was high were the development of electric automobiles and the treatment of AIDS in Africa. In all these cases, there is some reason to question whether American taxpayers are "responsible" for finding a solution, and we suspect that a large proportion of zero responses were really protest bids where respondents refused to pay for a public good that they valued. Unfortunately, the design of the study did not allow a more detailed analysis of such protest bids.

The most striking observation about table 2 is that the correlations between the different responses to issues are almost uniformly very high. Especially noteworthy is the close agreement ($\rho = 0.84$) between the ranking of issues by personal willingness to contribute (WTP) and by a conventional measure of political support for government action. The correlation is not substantially lower than the reliabilities of the separate rankings by the two variables. The high correlation is compatible with the simple null hypothesis that the two measures yield orderings that are not systematically different. The concordance of importance with the other measures is also of special interest, because, unlike the other measures in table 2, importance was judged by subjects who only saw the problem headline by itself, with no mention of any particular intervention. The high correlations of Importance with WTP and Support reinforce the claim that both responses mainly reflect a response to the problem.

The findings summarized in table 2 support the hypothesis that the average WTP for interventions to solve public problems is first and foremost a measure (and not the most efficient measure) of respondents' average attitude to the problem. Although the different ways of tapping the attitude evoke slightly different responses to the same issue, the similarity of the measures is far more impressive than their differences. It is appropriate to conclude that WTP, at least in the context of a headline study, is best viewed as an indicator of the intensity of attitudes to public problems, measured on a dollar scale which may well be arbitrary (see also Kahneman et al., 1993). The conclusion that WTP is one of many essentially interchangeable expressions of attitude suggests the possibility of using more reliable attitude measures to assess the value of public goods. We return to this suggestion at the conclusion of the paper.

2.5. *Experimental studies of the determinants of WTP*

The preceding sections described the general pattern of responses to headlines concerning damage or threat to various public goods. We now turn to several small experiments in which we used the headline method to examine particular hypotheses about the determinants of WTP.

The experiments reported below were generally set up in a factorial design, with the same experimental manipulation applied to several different problems. The appropriate analysis for these data is analysis of variance, with both factors treated as fixed effects. However, the application of ANOVA to WTP data presents special problems, because of the two peculiarities of the distribution of WTP responses: a concentration of zero responses, and a heavy right tail. Fortunately, the impact of deviation from normality on the interpretation of ANOVA results can be assessed by using the parameter D , defined for each data set as $k_4/(nk_2^2)$, where k_m is the sample m -cumulant.³ In particular, k_2 is the sample variance, and k_4 is a combination of the first four sample moments; k_4 , and hence D , is 0 for a normal random variable. The parameter δ is defined as $(1/(1 - D))$.

Box and Watson (1962) showed that if under normality assumptions the F -statistic has an $F_{m,k}$ distribution under the null hypotheses, then in the present design it will have an $F_{m\delta,k\delta}$ distribution when the normality assumption is valid (see also Seber, 1977, pp. 149-151). Computing δ for the Winsorized WTP data used in our analyses, we find that δ is below 1.029 in all the results reported in tables 3-11 below, and below 1.013 in all but one case (the exception is table 8). For $N(\text{WTP})$ analyses, δ is in all cases below 1.004. Hence the analyses of both measures can be regarded as robust. We have also observed that the robustness of the analysis depends on the Winsorizing procedure: we conducted several analyses without this transformation, and observed δ values ranging from 1.19 to 8.61.

2.6. *Embedding of problems: violations of monotonicity*

The embedding effect is generally considered one of the main objections to the validity of the contingent valuation method (Contingent Valuation Panel, 1993). Two variants of

embedding effects have been distinguished. *Perfect embedding* occurs when WTP fails to increase with the scope of the good mentioned in the question. When the quantity of a single good is varied, perfect embedding is simply another name for a non-increasing utility function. Several examples of perfect, or almost-perfect, embedding have been documented. In an early demonstration, Kahneman and Knetsch found that residents of Toronto were willing to pay little more to clean up all the lakes of Ontario than they were willing to pay to clean up the lakes in separate regions of that province (Kahneman, 1986). Kahneman and Knetsch (1992) reported very similar WTP for improvements in the provision of goods that varied greatly in scope. Desvousges et al. (1992) found that WTP to save birds endangered by open oil ponds hardly varied when the number of birds expected to be saved varied from 2,000 to 200,000. Fischhoff et al. (1993) observed similar WTP for preserving 110 or 10,000 acres of wetland in New Jersey, though the latter intervention was somewhat more attractive in a direct comparison to another public good (protecting the Chesapeake Bay watershed). Diamond et al. (1992) reported that WTP increased significantly, but only by a small amount, when the number of wildlife refuges to be built was varied between 1 and 57.

Perfect embedding entails another result, which has been labelled *regular embedding*: WTP for a good is much higher when it is evaluated on its own than when it is an allocated part of the contribution made for a more inclusive good. Because the standard theory of contingent valuation does not provide any reason to expect a difference between the two ways of assessing WTP for a good, the finding of regular embedding suggests a serious indeterminacy in CVM. Responses to this threat have varied. Some proponents of CVM have focused on weaknesses in demonstrations of embedding (Smith, 1992). Others who have acknowledged the fact of embedding have suggested a charitable interpretation for the effect and procedures to measure it (McClelland et al., 1992), or offered a rationale for using the unembedded value in the context of damage assessment (Randall and Hoehn, 1992).

In the present study we draw a new distinction between two types of embedding: *problem embedding*, where the scope of the problem is varied, and *solution embedding*, where the scope of the intervention is varied. Our analysis of WTP in terms of a contribution suggests somewhat different hypotheses for the two types of embedding: (i) the idea that WTP reflects the importance of the issue implies that variations in the scope of a problem may influence WTP, but not necessarily in monotonic fashion: a specific threat can appear either more or less urgent than a more inclusive threat; (ii) the contribution model entails generally low sensitivity to the scope of the solution, except perhaps when the proposed intervention is grossly inadequate.

To examine the first hypothesis, we constructed two inclusive problems: "A RISE IN THE INCIDENCE OF MULTIPLE MYELOMA" and "FERTILITY LOSS DUE TO POLLUTION THREATENS REPTILES ON THE MEXICAN COAST," and two subsets of each problem, which we label favored and inferior. The favored subset was chosen to elicit a particularly strong response: we expected a rising incidence of myeloma among the young to be seen as a greater threat than a rising incidence of myeloma, and we expected turtles to be more important than reptiles. The inferior subset was selected to be relatively unimpressive: we thought that a rising incidence of myeloma among the old would be viewed as a mild threat

and that lizards would be less important than reptiles. The three versions of each problem were included in different questionnaires, answered by different respondents. The results are shown in table 3.

An ANOVA was conducted for each variable, using problem (myeloma or reptiles) as one factor, and set type (inclusive, favored or inferior) as the other factor. The bottom part of table 3 presents an account of the proportion of the variance associated with the two factors and with their interaction. Our main prediction for this experiment was that the favored subset of the problem (myeloma among the young, threat to turtles) would elicit the strongest responses. The predicted pattern of results was found for all variables, and was statistically significant in planned comparisons for WTP, for % (WTP), and also for the Support and Importance ratings.

The present study illustrates the potential for violations of monotonicity in WTP for public goods: more is not always more valuable, even when satiation is not an issue. These findings are challenging, because it is not possible to take a WTP estimate of the value of a good seriously if WTP for a more inclusive good could be lower, or WTP for a subset higher, and because the possibilities for violations of monotonicity are endless. Thus, saving 20,000 birds from dying in oil spills appears to be a well-defined goal, although the manner of their death is not fully specified. However, WTP to save 2,000 birds from dying of convulsions after an oil spill could well be higher. As this example illustrates, it is not easy to tell when a description of a good is sufficiently detailed to preclude violations of monotonicity.

2.7. Severity of problem

We have proposed the hypothesis that WTP generally reflects the attitude to the problem, rather than to the solution offered for it. This hypothesis suggests that respondents will be alert to indications of the importance and urgency of problems. Qualifiers such as

Table 3. Violations of inclusion

	WTP	% (WTP)	N (WTP)	Support	Importance
Myeloma	12.80	63%	1.00	2.11	5.48
Young	17.35	76%	1.08	2.77	5.92
Elderly	8.55	47%	.87	1.80	5.37
Reptiles	3.48	23%	.56	1.17	5.35
Turtles	4.97	34%	.65	1.54	5.48
Lizards	2.27	15%	.58	1.09	5.08
<i>Variance account (%)</i>					
Problem	2.3 **	**	7.4 **	6.1 **	1.5
Category	1.7 **	**	0.7	4.9 **	0.6
P × C	0.5	NS	0.2	0.8	0.2
N	677	677	293	337	292

**significant at .01 level

“large,” “major,” and “disastrous” are common clues to the importance of problems, and we expected respondents to be sensitive to them. To test this prediction, we constructed two variants of each of three problems, using adjectives to indicate the severity of a threat or damage. The two versions of each problem were included in different questionnaires. Table 4 displays the means of all response measures and the results of separate ANOVA’s for these variables.

The most important conclusion to be drawn from the variance account in table 4 is that WTP is drastically less sensitive than the two other response measures. The main effect of problem was significant for WTP, but it accounted for much less variance than in either Support or Importance. The effect of the manipulation of severity was not significant for WTP, though it was for the proportion of positive responses. We have no explanation for the odd result that for one of the problems (diminished visibility in National Parks) the WTP measure showed no difference at all between a “mild” and a “severe” drop of visibility, though the Support and Importance ratings showed the expected effect.

2.8. *Natural vs. human causes*

The next experiment was concerned with the influence of the cause of the damage or threat to an environmental good. Economic logic suggests that, for public as well as for private goods, the manner in which a good may be lost should not affect the value of its continued existence, but psychological research has shown that the causes of harms do affect values. For example, respondents favor higher compensation for injury if the harm

Table 4. Manipulations of severity of problem

	WTP	% (WTP)	N (WTP)	Support	Importance
Pine trees					
Minor threat	7.52	51%	.61	1.53	4.19
Major threat	11.99	58%	.77	1.89	5.17
Anemia					
Mild increase	13.39	59%	.88	1.93	5.29
Sharp increase	17.20	77%	1.05	2.58	5.86
Visibility in Parks					
Mild drop	9.56	46%	1.02	1.85	4.31
Sharp drop	9.57	47%	.85	2.18	5.00
<i>Variance account (%)</i>					
Problem	1.5**	**	5.7**	3.2**	6.1**
Severity	0.4	*	0.3	3.4**	4.7**
Problem × severity	0.2	NS	2.1*	0.3	0.3
N	659	659	372	348	296

*significant at .05 level

**significant at .01 level

was caused by human action than by nature, even when the added compensation can have no deterrence effect (Baron, 1993; Baron and Ritov, 1993). It is therefore of interest to examine whether the stated cause of an environmental problem affects WTP for its prevention or amelioration.

In an earlier study we showed that subjects' responses to various threats or actual damages depends to some extent on their causes. We called this an outrage effect (Kahneman et al., 1993). Thus, for example, people were willing to pay more to save seabirds from oil spills than from a new epidemic disease. Table 5 presents the result of a partial replication of that study in a between-subject design.² Two of the three problems were included in the earlier study (Kahneman et al., 1993). The dolphin problem was constructed for the present replication.

The results of Table 5 confirm the finding of a modest effect of outrage on WTP (Kahneman et al., 1993), though this effect appears to be smaller in the between-subject design of the present study than in the within-subject design that we employed earlier.

2.9. Tests of sensitivity to solutions

The results summarized in tables 3–5 indicated some sensitivity to relatively subtle variations in the formulation of the problem. However, we have argued that variations in the scope of proposed solutions are less likely to be noticed, and that the personal satisfaction of contributing is not generally related to the scope of the intervention (Kahneman and Knetsch, 1992). We do not suggest that respondents ignore proposed solutions altogether; an intervention that is obviously feeble or inept will surely be rejected. Our

Table 5. Natural vs. human cause

	WTP	% (WTP)	N (WTP)	Support
Skin cancer				
Natural	11.40	51%	.88	2.38
Human	12.97	61%	.98	2.47
Dolphins				
Natural	6.35	44%	.86	1.72
Human	18.85	68%	1.12	2.14
Australian mammals				
Natural	8.35	46%	.79	1.57
Human	8.49	55%	.86	1.83
<i>Variance account (%)</i>				
Problem	0.8	NS	1.3	5.4**
Cause	1.3**	**	1.5*	1.0
Interaction	1.7**	NS	0.5	0.3

*significant at .05 level

**significant at .01 level

hypothesis is that support for interventions mainly reflects the importance of the problem, and that plausible reactions of widely varying scope may consequently be supported to about the same degree. The present section collects several small experiments that examined this issue.

In the first experiment we varied both the scope of the problem and the completeness of proposed solutions. Three versions of each of three problems were constructed. For example, one version of the rain forest problem mentioned a threat to the rain forest in both Brazil and Ecuador, and proposed an intervention in both countries. Another version mentioned the threat in both countries but proposed an intervention for Ecuador alone. Finally, the third version restricted the stated problem to Ecuador, and matched the proposed intervention to the problem. The different versions of each problem were included in different questionnaires. The results are shown in table 6.

The middle problem in each of the triads of issues in table 6 describes an intervention that is clearly restricted to only part of the problem. We supposed that an obviously incomplete solution might be relatively unattractive, but planned comparisons failed to support this hypothesis; the predicted effect was not significant for any of the response measures listed in the table. Another planned test compared the broader intervention (the top item in each triad) to the two other versions of each problem. One measure showed a significant effect in the predicted direction: $N(\text{WTP})$ was significantly higher for the broader intervention.

Table 6. Completeness of solution

	WTP	% (WTP)	$N(\text{WTP})$	Support	Importance
Shrinking rain forest					
BE-BE	15.61	64%	1.28	2.80	6.31
BE	14.98	62%	1.16	2.68	
E-E	19.96	71%	1.16	2.70	6.13
Birds					
PFE-PFE	9.44	57%	.81	2.32	5.85
PFE-F	7.96	56%	.75	1.98	
F-F	9.34	49%	.75	2.26	5.41
AIDS					
KZU-KZU	16.92	53%	1.12	2.09	6.04
KZU-Z	8.98	48%	.94	2.10	
Z-Z	6.55	34%	.91	1.78	5.75
Variance account (%)					
Problem	1.9**	**	8.0**	5.4**	2.9*
Category	0.3	NS	0.9	0.3	1.1
$P \times C$	1.1*	NS	0.2	0.6	0.1
N	982	982	542	530	306

*significant at .05 level

**significant at .01 level

A further attempt to test the sensitivity to the scope of an intervention is summarized in table 7. The four questions referred to speeding up the cleanup of old toxic waste dumps. Two of the versions stated that "20 OLD TOXIC WASTE DUMPS HAVE BEEN IDENTIFIED AS SOURCES OF POSSIBLE TOXIC LEAKAGE INTO THE WATER." Two other versions did not specify the number of leaking dumps. The proposed intervention was cleanup of either 4 or 20 dumps. The results were analyzed in a 2×2 ANOVA, testing the same hypotheses as in the previous study: (i) highlighting the incompleteness of a solution could reduce its appeal; this would produce an interaction effect in the ANOVA; (ii) the size of the intervention could have a main effect on the WTP and Support measures. Neither of these hypotheses was supported.

An exception to this litany of negative results was found in a set of questions concerning the response to toxic spills, summarized in table 8. The experiment was patterned after a study of solution embedding reported by Kahneman and Knetsch (1992). It was concerned with a comparison of two interventions that differ greatly in their scope: "set up a fund to improve emergency response to chemical accidents" or "institute special training programs for emergency personnel." Kahneman and Knetsch (1992) found that WTP for a similar pair of interventions did not differ significantly, but the results of table 8 show a significant difference in the proportion of positive responses ($p = .03$) and a marginally significant difference in Support ($p = .054$).

Table 9 summarizes the results of another study of solution embedding. The headline describing the problem was: "MANY INNER CITY HOMES HAVE LEAD-BASED WALL PAINT. LEAD POISONING AFFECTS THE CENTRAL NERVOUS SYSTEM." Different groups of subjects saw this headline paired with different proposed interventions, including replacement of the paint, blood tests for screening, and free medical care for lead poisoning. Two of the proposed policies were described as conjunctions of interventions, and two of the constituent interventions were also evaluated on their own.

The salient result of table 9 is that WTP for a conjunction of two interventions (replacing lead paint and providing medical tests to possible victims) was actually lower than for

Table 7. Specificity of problem and solution

	WTP	% (WTP)	N (WTP)	Support
Dumps, unspecified				
Cleanup 20	23.17	72%	1.32	2.92
Cleanup 4	18.47	66%	1.31	2.85
20 Dumps				
Cleanup 20	20.00	80%	1.25	2.89
Cleanup 4	21.00	72%	1.18	2.77
Variance account (%)				
Problem	0.1	NS	0.4	0.07
Solution scope	0.3	NS	0.0	0.04
P \times S	0.2	NS	0.1	0.01

Table 8. Solutions to toxic spills

	WTP	% (WTP)	N (WTP)	Support
Emergency fund	12.43	49%	1.08	2.67
Training	7.26	34%	0.96	2.22
<i>Variance account (%)</i>				
Solution	1.5	*	1.1	3.1
N	215	215	90	118

*significant at .05 level

Table 9. Solutions to the lead problem

	WTP	% (WTP)	N (WTP)	Support
Replace	20.58	70%	1.07	2.48
Test	11.33	59%	0.96	2.22
Replace and test	8.14	59%	0.81	2.43
Replace and care	14.75	60%	0.93	2.62
<i>Variance account (%)</i>				
Intervention	3.11**	NS	2.91*	1.21
N	438	438	271	232

*significant at .05 level

**significant at .01 level

either of the two interventions alone. A planned comparison yielded a significant result, but the direction was opposite to predictions, both for WTP and for the normalized mean of positive contributions, $N(\text{WTP})$. Because no parallel result was found for the Support measure, we suspect that the strong violation of monotonicity in WTP could be a statistical fluke. The conservative conclusion from table 9 is that adding a valued component to an intervention does not necessarily add to its value, contrary to standard choice theory.

We report more briefly two other experiments that yielded negative results. One was an attempt to use numbers to define the scope of an intervention. The problem was the threat to the spotted owl, and the proposed intervention was to protect some part (ranging from 10% to all) of the untouched old-growth groves. The responses to the different versions did not differ statistically for any of the measures. Another small experiment compared two reactions to inadequacies of the capacity available for the treatment of solid waste. One aimed to increase the treatment and disposal of waste; the other, which we thought more radical, proposed reduction of solid waste at the source. Again, the mean responses to these two interventions were not statistically different, for any of the measures we examined.

The experiments summarized in this section generally supported the hypothesis that WTP for public goods does not conform to the logic of economic value, at least in the

context of a headline study. The results suggest that respondents made only crude discriminations among problems, and that they were even less sensitive to the statement of solutions. In general, we believe that the effective range of stated WTP is radically inadequate for a measure of the value of public goods. We also suspect that the range of estimates is narrow because the willingness to contribute to insignificant causes is too high; our hypothesis is that a contingent valuation question is likely to elicit substantial WTP even for trivial causes. It is important for proponents of contingent valuation to demonstrate that their methods can overcome this possibly critical weakness.

2.10. Preference reversals

Standard choice theory requires that an individual who is willing to pay more to acquire A than to acquire B should also prefer to obtain (or to retain) A rather than B in a direct choice. In contrast, the psychological analysis of attitudes focuses on responses to objects considered in isolation; it is not part of attitude theory that choices are necessarily made by selecting the option that elicits the most favorable attitude. Thus, the idea that WTP is a measure of attitude allows goods to be ordered differently by WTP and by direct choice. Indeed, preference reversals between pricing and choice have been documented repeatedly in recent years. Preference reversals are identified by significant differences between the proportions of respondents who appear to favor a particular option, depending on the response used. Systematic preference reversals between pricing and choice do not only imply that these responses differ, but also that the sign of the difference is predictable.

Tversky and his colleagues have identified several different mechanisms that produce preference reversals (Tversky, Sattath, and Slovic, 1988; Tversky, Slovic, and Kahneman, 1990). Most relevant in the present context is a prominence principle: in general, the most important (prominent) attributes of options have greater weight in choice than in pricing. For example, the prominence principle explains why personal safety looms larger in choices between options that vary in both safety and cost than in the pricing of such options (Magat, Viscusi, and Huber, 1988). Safety is more prominent than money, and it is accordingly assigned greater weight in choice than in pricing. Irwin et al. (1993) have also invoked the prominence principle to explain a preference reversal that they observed in the valuation of public goods: their respondents frequently indicated a preference for a reduction of pollution over an increment in the quality of a private good such as a VCR, even when they assigned higher WTP to the upgrade of the private good. The authors explain that "... choice involves arguments and there are many powerful, even noble, arguments in favor of one's placing high personal value on improved environmental quality" (p. 8).

We hypothesized that "existence goods" such as inaccessible lakes and unfamiliar species are less prominent than the health and safety of humans, and that this difference in prominence would cause predictable preference reversals. To test this hypothesis we selected 12 pairs of goods. In 7 of these pairs one item involved passive use, the other health or safety. In 3 other pairs both items involved health or safety, and in the two

remaining pairs both items involved existence value. There was little ambiguity in the classification of issues, except for the burning of rain forests in South America, which we classified as a high-prominence problem because of its widely publicized connection to the threat of massive ecological change. A page requesting two choices between pairs of interventions was included at the end of most questionnaires in the study. The issues included in the choice items did not appear elsewhere in the same questionnaire. The following example illustrates the format of the choice question:

“It sometimes happens that budget constraints force a choice between two desirable projects. One has to be given up, at least for now, so that the other can go forward. If you had to choose one of the following interventions, which would you choose?”

Problem: Skin cancer from sun exposure is common among farm workers.

Intervention: **support free medical checkups for threatened groups.**

Table 10. Choice and WTP: Human safety vs. passive use

	Choice	WTP	Chi-square
Replace paint	154	47	
Elephants	57	40	10.08 $p < .001$
ties		28	
Skin cancer	72	33	
Australian mammals	31	36	8.47 $p < .01$
ties		43	
Myeloma	60	20	
Coral reefs	50	46	9.78 $p < .01$
ties		46	
Skin cancer	42	33	
Dolphins	16	40	9.77 $p < .01$
ties		40	
Earthquake safety	29	19	
Spotted owl	29	58	9.26 $p < .01$
ties		37	
Toxic spills	38	22	
Sharp visibility	18	33	8.67 $p < .01$
ties		49	
Carbon dioxide	42	29	
Mild visibility	15	32	8.40 $p < .01$
ties		51	

Table 11. Choice vs WTP: Problems matched in prominence

	Choice	WTP	Chi-square
Myeloma	12	23	
Burning rain forest ties	45	53 40	1.42 NS
Global warming	44	51	
Toxic spills ties	13	13 39	.11 NS
Reduce waste	40	43	
Toxic spills ties	19	27 44	.56 NS
Pine disease	43	47	
Spanish moss ties	17	21 45	.09 NS
Sharp visibility	35	22	
Florida panther ties	37	27 40	.16 NS

Problem: Several Australian mammal species nearly wiped out by hunters.

Intervention: **contribute to worldwide fund to provide safe breeding areas for these species.**

Which would you choose (check one):
medical checkups ___ safe breeding areas ___”

The results are summarized in tables 10 and 11, which present the numbers of respondents indicating a preference for each item in WTP and in direct choice, respectively for items differing in prominence (table 10) and for items of similar prominence (table 11). A chi-square test was computed for each problem, to test the hypothesis that the distribution of support for the two options is the same for WTP and for choice (ties in WTP were excluded from this analysis). The results are clear. A significant difference in preferences was found in every problem involving a choice between existence value and health or safety. No significant reversals of preference were found in choices between items within each of these broad categories. If choice is viewed as the more fundamental operation, these results imply that the WTP question systematically overestimates the value of existence goods. Here again, we must repeat the now familiar caveat that the preference reversals were observed in the context of a headline method. We also ask again whether there is any obvious reason for preference reversals to be eliminated by the procedures of contingent valuation.

3. General discussion

This article has addressed a methodological dilemma and a set of substantive issues. The methodological problem is that the current procedures of contingent valuation are too elaborate and expensive to be adopted in experimental critiques of the method. The headline method is a partial response to this quandary. We have shown that it is possible to elicit WTP responses by sketchy headline scenarios. The WTP values elicited by headlines are in the range of values observed in contingent valuations for issues of comparable importance, and they show some of the familiar characteristics of CV results, notably a heavy right tail in the distribution of responses. We do not claim that every finding in this study would apply as well to orthodox contingent valuation. However, we have advanced a proposition of *process continuity*, which suggests that findings in the headline method merit the status of plausible hypotheses about contingent valuation. In our view, the designers of future contingent valuation surveys should accept the burden of showing that their procedures are adequately sensitive to the scope of problems and of interventions, that they avoid embedding effects, and that they are robust to preference reversals (for a similar view on the issue of burden of proof, see Contingent Valuation Panel, 1993). Perhaps most important, it appears essential to demonstrate that contingent valuation properly discriminates significant causes from trivial ones. In the data of table 1, the gap between the defense of Spanish moss or mushroom varieties and major problems of ecology or health was simply too small.

The substantive issues with which we were concerned arise from two competing views of the WTP response in contingent valuation. Contingent valuation is based on a purchase model, in which the respondent determines how much a public good is "worth" to her household, by conjoining the provision of the good with a decrement of wealth. The alternative is a contribution model, in which respondents treat public issues as problems about which something must be done, in a spirit of donation rather than acquisition. The present study used voluntary contribution as payment vehicle, and its results conform closely to the predictions of the contribution model. Could the setting of conventional contingent valuation induce respondents to adopt a purchase model? We cannot be certain that the answer is negative, but see no psychological reason to believe that the procedures of contingent valuation would profoundly transform respondents' approach to the WTP question.

Willingness to contribute is one aspect of the psychological notion of attitude, and our main substantive hypothesis is that WTP is a measure of attitude on a scale of (hypothetical) dollars. The concepts of attitude and economic value are constrained by different logics. For example, an assumption of extensionality (Arrow, 1982) or invariance (Tversky and Kahneman, 1986) is essential to the notion of economic value: preferences are taken to be robust to inconsequential variations in the framing of options, or in elicitation procedures. In contrast, the psychological analysis of attitudes assumes that measures of attitudes are highly susceptible to framing and that different measures of attitude are imperfectly correlated. Furthermore, attitudes are not necessarily constrained by monotonicity. The looser logic of attitude allows people to feel more strongly about turtles than about reptiles, or about 2,000 birds that died of convulsions in an oil spill than about 20,000 birds that died in an oil spill. It is important to note,

however, that the difference between economic value and attitude is not a difference between order and chaos. Although much weaker than invariance, the hypothesis that alternative indices of attitudes will be highly correlated still imposes strong constraints on data. Furthermore, the violations of invariance are themselves orderly: Framing effects, preference reversals and violations of monotonicity are systematic and predictable observations.

The main finding of the present headline study was that issues were ordered in much the same way by WTP and by more conventional measures of attitude. Furthermore, the psychometric properties of the WTP measure appeared to be inferior to those of other measures of attitude that we examined. The proportion of problem-related variance was two to four times larger for ratings of political support, importance and personal satisfaction than for WTP. The difference is easily explained: conventional attitude measures use a bounded scale whereas the WTP scale is both unbounded and unfamiliar. This psychometric deficiency of WTP is likely to extend to contingent valuation as well, because CV procedures do not usually provide help in the difficult task of expressing an attitude to an unfamiliar good in units of money.

Willingness to pay was assessed in the present study by an open-ended question about the maximal amount that a household would enter as a voluntary contribution on a tax form. This type of open-ended question has been displaced in recent applications of CVM by a referendum technique (Carson et al., 1992), which has also been endorsed by the Contingent Valuation Panel (1993). In the new technique respondents are asked how they would vote on a proposition to provide a specific public good, at a cost of \$ X to the household. The value of X is varied for different respondents, and the parameters of an assumed underlying distribution of WTP are estimated.

An important virtue of the referendum question is that it orients the respondent to a realistic question about a possible voting choice; this is a persuasive solution to the problem of payment vehicle. Nevertheless, adoption of the referendum format is unlikely to dispel all the doubts that have been raised about the validity of WTP. Referendum questions elicit attitudes just as open-ended WTP questions do, and there is little reason to expect the troubling phenomena of embedding and preference reversal to disappear in this format. Note that the accuracy with which surveys predict actual referenda is not the issue here: a real referendum should also be rejected as a measure of economic value, if its outcome is demonstrably susceptible to effects of framing, or to the inclusion of another proposition on the ballot. The deeper problem is that WTP may not satisfy the consistency requirements of an adequate measure of economic value, for the purposes of benefit-cost analysis or damage assessment. Further research will tell, of course, but we do not expect that the referendum technique will solve the problems of contingent valuation.

We have not claimed that people do not value existence goods—of course they do. Our question has been whether existence value is better described in terms of a purchase or a contribution model, as an economic value or as an attitude. *How* the public values existence goods has implications for policy. The dollar amounts that people describe themselves as willing to pay cannot serve as a measure of economic value if WTP is mainly an indication of attitude to the problem, as we have argued. On the other hand, there is a rationale for using public attitudes as an input in policy or litigation, without

pretending that these attitudes satisfy the definition of economic values. We next sketch one possible way of doing so.⁴

A scale of value for environmental damages should be developed by constructing a small number of hypothetical scenarios, covering a broad range of scope and importance. There would be two criteria for inclusion of scenarios in the scale: (i) high consensus in the attitudes of the public to each scenario, and in the ranking of the scenarios; and (ii) a hope of reaching professional and political consensus on appropriate dollar values for these hypothetical problems. We do not address the question of how that consensus would be reached, beyond noting that many different approaches should probably be used. Public attitudes would be one input into this process, but probably not the only one. We expect that the informed judgment of experts would bring in relevant considerations that lay judgment is prone to neglect, such as the scope and duration of the damage. The objective of the scaling effort would be to provide a mapping from attitudes and other relevant factors to dollar values for a particular class of events, such as ecological threats or damages.

Once a scale is established, real issues that arise from time to time could be valued by surveys of attitudes, including explicit comparisons to the scenarios of the original scale. The measures of attitude used in this comparison would be chosen by psychometric criteria: measures of judged importance and political support would probably be used in preference to WTP. A dollar value would be assigned to the target issue by reference to the standard scale. One advantage of this proposal is that the difficult conceptual and methodological problems of anchoring the dollar value of public goods in the preferences and opinions of the citizenry would be addressed just once, in the process of constructing the initial scale linking monetary value to attitude. Clearly, professional and political consensus is more likely to be achieved in dealing with hypothetical questions constructed for the purpose than in evaluating real goods in the context of litigation.

We do not believe that the public interest is best served by anchoring policy in an erroneous model of value. In particular, it may be unreasonable to demand that the dollar values of existence goods be “measured,” if WTP responses do not satisfy the theory of economic value that justifies the measurement. The values that will be used in regulation and in litigation should be negotiated in a process that is informed by all relevant expertise and invokes no assumptions contrary to fact. This is a highly controversial position, of course, and we close on a note of caution, by reminding the reader that the findings of violations of economic logic on which the present argument rests have been observed in a headline method, which is not the same as the method of contingent valuation.

Acknowledgments

This study was supported by a Chevron grant for risk assessment research, administered by the University of California. We thank the San Francisco Exploratorium for facilitating the research, and Don Green and Gary McClelland for their useful comments on versions of

the manuscript. Ya'acov Ritov provided valuable statistical guidance. We are grateful to Karen E. Jacowitz for her assistance in several phases of the research. Our greatest debt is to Paul Grant, for his tireless perfectionism in coordinating a large and complex study.

Notes

1. Individuals who made the same response to all questions were dropped from this analysis.
2. Because of an error in the design, importance judgments for the two versions of each problem were made by the same subjects. To prevent confusion, these results are omitted from the table.
3. We are grateful to Ya'acov Ritov, who suggested this analysis.
4. The following discussion elaborates a communication addressed to a member of the Contingent Valuation Panel in October 1992, developing some notions discussed earlier with Jack Knetsch.

APPENDIX: Problems and proposed interventions¹

Threats to animal species

American elk

THE AMERICAN ELK IS AN ENDANGERED SPECIES IN THE ROCKY MOUNTAINS

Intervention: Provide refuges for the American elk

Ferret

THE BLACK-FOOTED FERRET, A WEASEL-LIKE MAMMAL, IS AN ENDANGERED SPECIES ON THE GREAT PLAINS

Intervention: Breed the black-footed ferret in captivity

Elephants

ELEPHANTS IN ASIA AND AFRICA ARE THREATENED WITH EXTINCTION BECAUSE OF HUNTING BY IVORY POACHERS

Intervention: Join international efforts to provide protection for elephants and to control ivory trade

Marine life

THE PRODUCTION OF BLEACHED PAPER EMITS DIOXIN INTO RIVERS, CAUSING DEFORMITIES AND REDUCED REPRODUCTION IN MARINE LIFE, WITHOUT ENDANGERING HUMANS

Intervention: Provide funds for development of alternative bleaching techniques

1. Alternative versions are shown in square brackets.

Kangaroo rats

THE GIANT KANGAROO RAT POPULATION IN CALIFORNIA ARE DECLINING STEADILY, AS THE OPEN GRASSLANDS SHRINK

Intervention: Increase protected areas of open grassland

Florida panther

THE FLORIDA PANTHER IS THREATENED WITH EXTINCTION DUE TO LOSS OF HABITAT

Intervention: Breed the Florida panther in captivity

Falcon shell

PESTICIDE CONTAMINATION CAUSES THINNING OF THE SHELLS OF FALCON EGGS. THE CHICKS HATCH TOO EARLY, AND DIE. THE SPECIES IS THREATENED

Intervention: Rear falcon population in protected environment

Dolphins

DOLPHINS IN THE MEDITERRANEAN SEA ARE THREATENED BY POLLUTION [THREATENED BY A NEW VIRUS]¹

Intervention: Contribute to international fund to save the Mediterranean Dolphin

Australian mammals

SEVERAL AUSTRALIAN MAMMAL SPECIES NEARLY WIPED OUT BY HUNTERS [BY PREDATORS]

Intervention: Contribute to worldwide fund to provide safe breeding areas for these species

Coastal reptiles

FERTILITY LOSS DUE TO POLLUTION THREATENS SEVERAL SPECIES OF REPTILES [LIZARDS] [SEA TURTLES] ON THE MEXICAN COAST

Intervention: Support program to increase fertility by hormonal treatment

Wildlife

POPULATIONS OF SEVERAL WILDLIFE SPECIES [BLACK BEARS] [SPECIES OF LARGE MAMMALS] HAVE BEEN DEPLETED IN SEVERAL REGIONS OF THE WESTERN STATES, BECAUSE OF HUNTING AND ECOLOGICAL CHANGES

Intervention: Set up and maintain wildlife [bear] [large mammals] refuges in regions where populations have been depleted

Birds

THE PEREGRINE FALCON [WHITE PELICAN, PEREGRINE FALCON, AND BALD EAGLE] IS THREATENED BY POLLUTION

Intervention: Support special program to protect the Peregrine falcon [endangered birds]

Spotted owl

HABITAT FOR THE SPOTTED OWL IS SEVERELY REDUCED BY TIMBER HARVESTING OF OLD GROWTH GROVES

Intervention: Increase the protected area of old growth timberland in the Pacific northwest [to include 10% of all untouched groves] [to include 60% of all untouched groves] [to include all untouched groves]

Threats to plant species*Coral reefs*

CORAL REEFS IN FLORIDA IMPERILED BY OIL SPILLS AND RIVER POLLUTION

Intervention: Create marine sanctuaries and divert tanker traffic

Mushroom

SEVERAL VARIETIES OF MUSHROOM IN CALIFORNIA NATIONAL FORESTS ARE ALMOST EXTINCT BECAUSE THEY ARE HARVESTED AND SOLD AT HIGH PRICES

Intervention: Provide additional inspection to discourage illegal harvesting of mushroom in national forests

Pine disease

NEW DISEASES THREATEN PINE TREES IN SEVERAL WESTERN STATES

Intervention: Increase federal programs to combat pine disease

Spanish moss

SPANISH MOSS, THE GRAY PLANT THAT GROWS ON TREE LIMBS IN THE SOUTHERN STATES, IS DEPLETED IN MANY AREAS DUE TO POLLUTION RELATED DISEASES

Intervention: Plant healthy Spanish moss on barren trees

Pine trees

NEW DISEASE PRESENTS A MAJOR THREAT [A MINOR THREAT] TO PINE TREES IN SEVERAL WESTERN STATES

Intervention: Increase federal programs to combat disease

Ecological damage*Wetlands*

WETLANDS IN CALIFORNIA AND NEVADA, CRITICAL HABITAT TO THE PACIFIC FLYAWAY WATERFOWL, ARE SHRINKING DUE TO DROUGHT AND IRRIGATION PROJECTS

Intervention: Increase water flow into the wetland areas

CO₂ in 3rd world

THIRD WORLD NATIONS ARE UNABLE TO PARTICIPATE IN INTERNATIONAL EFFORT TO REDUCE CARBON DIOXIDE EMISSION

Intervention: Establish funding mechanisms to help poorer nations respond to climate change

CO₂—oil burning

CARBON DIOXIDE EMITTED IN THE PROCESS OF OIL BURNING IS SUSPECTED TO BE ONE OF THE MAJOR CAUSES OF GLOBAL WARMING

Intervention: Increase research on solar energy technology

Automobile pollution

PROGRESS IN CONTROLLING AUTOMOBILE POLLUTION IS SLOWED BY LIMITED RESOURCES

Intervention: Provide active support for development of electrical engines for cars

Burning rain forest

MASSIVE BURNING AND CLEAR-CUTTING THREATEN THE RAIN FOREST IN SOUTH AMERICA

Intervention: Join international efforts to assist South American governments in saving the rain forest

Visibility in Parks

MILD DROP [SHARP DROP] IN VISIBILITY RECORDED IN SEVERAL NATIONAL PARKS

Intervention: Support special fund to improve air quality in parks

Toxic waste dumps

LIMITED RESOURCES CONSTRAIN THE CLEANUP OF OLD TOXIC WASTE DUMPS [CALIFORNIA'S 20 OLD TOXIC WASTE DUMPS WHICH HAVE BEEN IDENTIFIED AS SOURCES OF POSSIBLE TOXIC LEAKAGE INTO THE WATER]

Intervention: Speed up the cleanup of the 20 worst sites in your state [4 worst sites in your state] [all 20 old dumps] [4 worst sites]

Shrinking rain forest

THE RAIN FOREST IS SHRINKING IN ECUADOR [BRAZIL AND ECUADOR] BECAUSE OF HEAVY LOGGING AND BURNING

Intervention: Contribute to international efforts focused on maintaining Ecuador [Brazil and Ecuador] rain forests

Toxic spills

INCREASING INCIDENCE OF ACCIDENTS INVOLVING SPILLS OF TOXIC CHEMICALS

Intervention: Institute special training programs for emergency personnel [set up a fund to improve emergency response to chemical accidents]

Solid waste

MANY DISPOSAL FACILITIES FOR SOLID WASTE ARE REACHING THE LIMIT OF THEIR CAPACITY

Intervention: Increase programs for treatment and disposal of solid waste [for reduction of solid waste at the source]

Miscellaneous public goods*Historic buildings*

OLD PUBLIC BUILDINGS OF HISTORICAL INTEREST DETERIORATING DUE TO POLLUTION AND INADEQUATE CLEANING

Intervention: Set up fund to clean and waterproof affected buildings

Lighthouses

HISTORICAL LIGHTHOUSES ON BOTH COASTS HAVE BEEN SEVERELY DAMAGED BY HARSH WEATHER CONDITIONS

Intervention: Set up fund to repair damage and provide protective treatment

Earthquake safety

A THIRD OF THE MASONRY BUILDINGS IN OLDER SECTIONS OF CITIES ARE INADEQUATELY REINFORCED FOR EARTHQUAKE SAFETY

Intervention: Set up a program that offers incentives for building owners to strengthen unsafe buildings

Public health issues*Power lines and leukemia*

SOME RESEARCHERS HAVE FOUND UNUSUALLY HIGH INCIDENCE OF LEUKEMIA AMONG CHILDREN WHO LIVE NEAR POWER LINES

Intervention: Provide funds to accelerate research on possible links between leukemia and electromagnetic fields

Increase in anemia

MILD INCREASE [SHARP INCREASE] IN INCIDENCE OF ANEMIA AMONG CHILDREN IN SOME AREAS DUE TO INSUFFICIENT VITAMIN INTAKE

Intervention: Set up a program that encourages children to consume more fruits and vegetables as part of their regular diet

Skin cancer in farmers

SKIN CANCER FROM PESTICIDES [FROM SUN EXPOSURE] IS SUSPECTED AMONG FARM WORKERS

Intervention: Support medical checkups for threatened groups

Increase in myeloma

A RISE IN THE INCIDENCE OF MULTIPLE MYELOMA (A TYPE OF MARROW CANCER) [AMONG THE ELDERLY (75 AND OVER)] [AMONG YOUNG PEOPLE]

Intervention: Increase research on multiple myeloma

AIDS in Africa

AIDS SPREADING RAPIDLY IN ZAIRE [KENYA, ZAIRE, AND UGANDA]

Intervention: Support public health measures in Zaire [in the three countries]

Lead paint poisoning

MANY INNER CITY HOMES HAVE LEAD-BASED WALL PAINT. LEAD POISONING AFFECTS THE CENTRAL NERVOUS SYSTEM

Intervention: Set up program to replace lead-based paint in low-income neighborhoods [provide free blood tests for lead poisoning in low-income neighborhoods] [replace lead-based paint and provide free blood test] [provide free blood tests and free medical care for lead poisoning]

References

- Andreoni, James. (1990). "Impure Altruism and Donations to Public Goods. A Theory of Warm-Glow Giving," *Economic Journal* 100, 464-477.
- Ajzen, Icek. (1988). *Attitudes, Personality and Behavior*. Chicago: Dorsey.
- Arrow (1982). "Risk Perception in Psychology and in Economics," *Economic Inquiry* 20, 1-9.
- Baron, Jonathan. (1993). "Heuristics and Biases in Equity Judgments: A Utilitarian Approach." In B.A. Mellers and J. Baron (eds.), *Psychological Perspectives on Justice: Theory and Applications*. New York: Cambridge University Press, pp. 109-137.
- Baron, Jonathan, and Ilana Ritov. (1993). "Intuitions About Penalties and Compensation in the Context of Tort Law," *Journal of Risk and Uncertainty*, 7, 17-33.
- Bazerman, Max H. (1990). *Judgment in Managerial Decision Making*, 2nd ed. New York: Wiley.
- Box, G.E.P., and G.S. Watson. (1962). "Robustness to Non-Normality of Regression Tests," *Biometrika* 49, 93-106.
- Carson, Richard T. et al. (1992). *A Contingent Valuation Study of Lost Passive Use Values Resulting from the Exxon Valdez Oil Spill*. A report for the Attorney General of the State of Alaska, November 1992.
- Contingent Valuation Panel. (1993). "Report of the NOAA Panel on Contingent Valuation," *Federal Register* 58 (10), January 15, 1993.
- D'Arge, Ralph C. (1989). "A Practical Guide to Economic Valuation of the Environment." In *Thirty-Fourth Annual Rocky Mountain Mineral Law Institute Proceedings*. New York: Matthew Bender and Co.
- Desvousges, William H. et al. (1992). "Measuring Non-Use Damages Using Contingent Valuation: An Experimental Evaluation of Accuracy," *Research Triangle Institute Monograph* 92-1.
- Diamond, Peter A., and Jerry A. Hausman. (1993). "On Contingent Valuation Measurement of Non-Use Values." In J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*. Amsterdam: North Holland Press.
- Diamond, Peter A. et al. (1993). "Does Contingent Valuation Measure Preferences? Experimental Evidence." In J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*. Amsterdam: North Holland Press.
- Eagly, Alice H., and Shelly Chaiken. (1993). *The Psychology of Attitudes*. Fort Worth: Harcourt Brace Jovanovich.
- Fischhoff, Baruch, and Lita Furby. (1988). "Measuring Values: A Conceptual Framework for Interpreting Transactions with Special Reference to Contingent Valuation of Visibility," *Journal of Risk and Uncertainty* 1, 147-184.

- Fischhoff, B. et al. (1993). "Embedding Effects: Stimulus Representation and Response Mode," *Journal of Risk and Uncertainty*, 6 (3), 211–234.
- Gilovich, Thomas. (1991). *How We Know What Isn't So: The Fallibility of Human Reasoning in Everyday Life*. New York: Free Press.
- Green, Donald P. (1992). "The Price of Elasticity of Mass Preferences," *American Political Science Review* 86 (1), 128–148.
- Hanemann, W. Michael. (1992). "Comments Concerning Non-Use Values and The Contingent Valuation Method." Submitted to the NOAA panel on the CVM, July 22, 1992.
- Irwin, Julie R. et al. (1993). "Preference Reversals and the Measurement of Environmental Values," *Journal of Risk and Uncertainty* 6, 5–18.
- Kahneman, Daniel, and Jack L. Knetsch. (1992). "Valuing Public Goods: The Purchase of Moral Satisfaction," *Journal of Environmental Economics and Management* 22, 57–70.
- Kahneman, Daniel et al. (1993). "Stated Willingness to Pay for Public Goods: A Psychological Perspective," *Psychological Science* 4, 310–315.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky. (1982). *Judgment under Uncertainty: Heuristics and Biases*. New York: Cambridge University Press.
- Kemp, Michael A., and Christopher Maxwell. (1993). "Exploring a budget context for contingent valuation estimates." In J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*. Amsterdam: North Holland Press.
- Lazo, Jeffrey K. et al. (1992). "Can Contingent Valuation Measure Non-Use Values?" *American Journal of Agricultural Economics* (December), 1126–1132.
- Magat, Wesley A., W. Kip Viscusi, and John Huber. (1988). "Paired Comparison and Contingent Valuation Approaches to Morbidity Risk Valuation," *Journal of Environmental Economics and Management* 15, 395–411.
- Margolis, H. (1982). *Selfishness, Altruism and Rationality*. New York: Cambridge University Press.
- McClelland, Gary H. et al. (1992). *Methods of Measuring Non-Use Values: A Contingent Valuation Study of Groundwater Cleanup*. Center for Economic Analysis, University of Colorado.
- McFadden, Daniel, and Gregory K. Leonard. (1993). "Issues in the Contingent Valuation of Environmental Goods." In J.A. Hausman (ed.), *Contingent Valuation: A Critical Assessment*. Amsterdam: North Holland Press.
- Nickerson, Carol A.E. (1994). "Does Willingness-to-Pay Reflect the Purchase of Moral Satisfaction? A Reconsideration of Kahneman and Knetsch," *Journal of Environmental Economics and Management*, in press.
- Petty, Richard E., and John T. Cacioppo. (1986). "The Elaboration Likelihood Model of Persuasion." In L. Berkowitz (ed.), *Advances in Experimental Social Psychology* 19, 123–205. San Diego: Academic Press.
- Randall, Alan, and John P. Hoehn. (1992). "Embedding Effects in Contingent Valuation: Implications for Natural Resource Damage Assessment." Staff Paper No. 92–14, Department of Agricultural Economics, Michigan State University.
- Savage, Ian. (1993). "An Empirical Investigation on the Effect of Psychological Perceptions on the Willingness-To-Pay to Reduce Risk," *Journal of Risk and Uncertainty* 6, (1), 75–90.
- Schkade, David A., and John W. Payne. "How People Respond to Contingent Valuation Questions: A Verbal Protocol Analysis," *Journal of Environmental Economics and Management*, in press.
- Seber, G.A.F. (1977). *Linear Regression Analysis*. New York: John Wiley & Sons.
- Smith, V. Kerry. (1992). "Arbitrary Values, Good Causes and Premature Verdicts," *Journal of Environmental Economics and Management* 22, 71–89.
- Tolley, George S., and Alan Randall. (1983). *Establishing and Valuing the Effects of Improved Visibility in the Eastern United States*. Interim report to the U.S. Environmental Protection Agency.
- Tversky, Amos, and Daniel Kahneman. (1986). "Rational Choice and the Framing of Decisions," *Journal of Business* 59, S251–S278.
- Tversky, Amos, Shmuel Sattath, and Paul Slovic. (1988). "Contingent Weighting in Judgment and Choice," *Psychological Review* 95, 371–384.
- Tversky, Amos, Paul Slovic, and Daniel Kahneman. (1990). "The Causes of Preference Reversal," *American Economic Review* 80, 204–217.