

The Role of Social and Behavioral Sciences in Water Resources Planning and Management

Proceedings of an
Engineering Foundation Conference
in conjunction with the
Universities Council on Water Resources

Santa Barbara, California
May 3-8, 1987

Approved for publication by the Water Resources
Planning and Management Division of the
American Society of Civil Engineers

Edited by Duane D. Baumann
and Yacov Y. Haimes



Published by the
American Society of Civil Engineers
345 East 47th Street
New York, New York 10017-2398

1988

HOW FAR ALONG THE LEARNING CURVE IS THE CONTINGENT VALUATION METHOD?

Robert Cameron Mitchell*
Richard T. Carson**

Informed decisions about the provision of water resources require an assessment of their likely costs and benefits. One of the newest and, in many respects, most promising techniques to measure the benefits of water resource programs is the contingent valuation (CV) method. Until the last few years, contingent valuation was regarded as an experimental approach whose validity was yet to be established. Contingent valuation studies were relatively few in number, individually handicrafted, and directed more at testing for bias and demonstrating the methodology's promise than at providing usable benefit estimates. Recently, however, there has been a dramatic shift in the method's status and use. CV surveys directed towards policy needs are now multiplying in almost exponential fashion. CV practitioners are aggressively lowering production costs by substituting mail surveys¹ for the previously dominant in-person administration method. And policy analysts are beginning to use CV findings in estimating the benefits of pure public goods. Certainly this shift is gratifying to CV practitioners as a demonstration that their efforts to legitimize the method have been fruitful. But it raises false and, perhaps, dangerous expectations about the method's current capabilities. Put another way, it presupposes that we are farther along on the learning curve in developing and implementing the CV methodology than is actually the case.

The Learning Curve

Learning curves are conventionally used to describe the increase in productivity which occurs as workers and managers become familiar with a new production process (Yelle, 1979). One important aspect of this phenomenon² is "learning by doing" (Arrow, 1962), which permits improvements in the number and quality of widgets produced per worker for a given unit of time. In making heuristic use of the concept to illuminate the process whereby knowledge is acquired about a benefit measurement methodology, it is necessary to modify the classic learning curve in three respects. First we expand the scope of learning to include pre-production development. Second, we replace product unit cost, which has no clear analogue for a method that aims to value

1. And, to a lesser extent, the more expensive telephone surveys.

2. Other determinants of unit cost improvements are spillovers of learning from other producers (Spence, 1981), the nature of the productive technology, improvements in general technical knowledge, and scale economies (Chenawat, 1985).

* Graduate School of Geography, Clark University, Worcester, MA 01610.

** University of California-San Diego, San Diego, CA 92100.

otherwise nonpriced goods, with knowledge about the method on the vertical axis. Third, we treat the acquisition of knowledge as part of a multistage process, each with its characteristic type and pace of learning. The knowledge obtained about the method at each stage is used to evaluate its effectiveness and validity and to make judgments about whether it is worthwhile to proceed to a higher application stage. Figure 1 summarizes this concept of the learning curve for new methodologies which applies to other nonmarket valuation techniques such as hedonic pricing as well as contingent valuation.

The first stage in the development of a new benefit measurement methodology is its invention and initial application. If the method appears plausible and researchers begin to use it, the second or prototype stage is reached. During this stage basic knowledge is acquired about the method's properties, possibilities, and limitations. As problems are identified, researchers develop and test possible solutions. They also seek evidence of its validity. Once the method's basic feasibility and validity is established, attention turns to learning how it can be implemented in a routine fashion to obtain information useful to policy makers. Stage three, the early implementation (or production) stage, is characterized by a reduced, but still fairly steep slope to the learning curve as information is acquired about problems in applying the method under different circumstances and aggregating the findings to estimate program benefits and how these problems can be addressed. During this stage learning focuses on how systematically to implement the method in such a way that the quality and cost of its output are acceptable. If the overall results are satisfactory (as judged by peers, journal editors, review panels, and policy analyst consumers), the method may reach stage IV on the learning curve. At this stage, the method is regarded as being relatively well understood so attention is directed at making it more cost efficient. During this stage the learning curve gradually flattens until it reaches a plateau.

It is our view that the method's basic validity is sufficiently well established to place us beyond Stage II on the learning curve. But how far beyond? Owing to the inherent difficulty of using surveys to value public goods, the short amount of time that we have been in Stage III, and what we perceive as a reduction in funds for fundamental methodological research on the CV method,³ we believe CV is currently closer to the beginning than to the end of the Stage III segment of the learning curve. If so, we are not yet far enough along the learning curve to encourage novice CV researchers to rely on model questionnaires or to employ lower-cost methods such as telephone and, especially, mail surveys to obtain benefit estimates for policy purposes.

In what follows, we first describe CV and how it achieved its

3. The major source of CV research funds during the previous decade, the Environmental Protection Agency, has tended in recent years to channel its research monies to CV studies that apply the method for policy purposes.

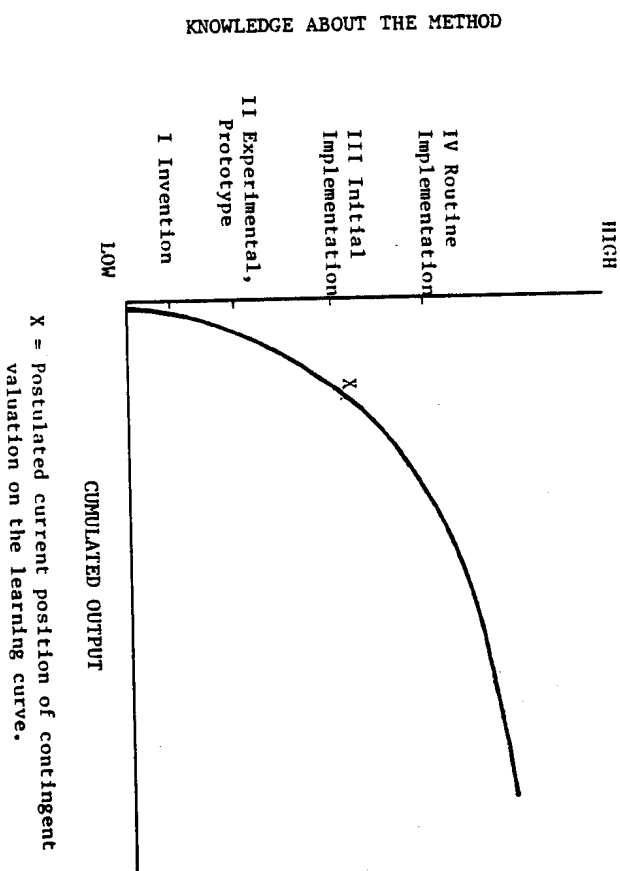


FIGURE 1. LEARNING CURVE FOR NEW METHODOLOGIES

current level of legitimacy as a valid benefit measurement procedure. We then review briefly the nature of sample surveys and the distinctive character of CV surveys to highlight the still-formidable methodological challenge faced by CV practitioners at this stage in the method's development. In the balance of the paper we consider three topics whose treatment in a guide prepared by the Corps of Engineers' illustrates our contention that there are as yet important methodological issues that need to be resolved before the contingent valuation method is ready for routine implementation.

Growth in Acceptance

In contrast to other benefit measurement techniques, which infer the value of nonmarketed goods from consumer behavior in related markets, contingent valuation uses surveys to ask respondents how much, in dollars, they would pay for a carefully specified change in the provision of an amenity. The CV approach, which received its first use in the 1960s (Davis, 1963, 1964), makes it possible to directly measure the quantity desired by economic theory and in a way that permits the valuation of a much wider range of goods, including those not yet provided, than the other methods. If surveys can be trusted -- and for some this is a big "if" -- the method offers economists the opportunity to directly measure benefits for such things as park improvements, national water quality benefits, or reductions in risks from hazardous waste contamination that were previously thought to be virtually unmeasurable. This promise led natural resource and pollution control agencies to fund exploratory research on the method during the 1970s and early 1980s. Since many economists were doubtful about the validity of a method that relies on people's responses to hypothetical questions instead of their behavior in actual markets, research efforts during the experimental-prototype stage (II) were largely devoted to demonstrating that the method could obtain valid estimates for nonmarketed goods.

We have elsewhere (Mitchell and Carson, forthcoming) discussed the findings of this research. One is that the threat of strategic behavior -- that respondents would deliberately distort their WTP amounts to influence the study's outcome -- is much less of a problem in the CV context than economists had originally anticipated. Another is that CV estimates compare favorably with those obtained by other methods. A number of researchers conducted studies to compare the findings of a CV study with the WTP amount estimated from a travel cost or hedonic price study of the same good. According to the most extensive review of the findings of these cross-method comparisons (Cummings, Brookshire, and Schulze, 1986), they support the conclusion that the CV method is reasonably accurate, at least when it is used to value the kinds of amenities that are also accessible to the indirect methods. Heberlein and Bishop (1986) went one step further and conducted a series of imaginative experiments that probed the ability of hypothetical markets to predict the outcome of real markets. In their most recent study, their CV-based estimate of sportsmen's willingness-to-pay for a license to hunt in a Wisconsin state game preserve closely matched the mean amount actually paid by another set of hunters who were given the chance to bid for these licenses in an auction market conducted by Bishop and Heberlein (with the permission of the State). Finally, parallel

theoretical work established that CV data are generated in forms consistent with the theory of welfare change measurement (Randall, Ives and Eastman, 1974; Freeman, 1979; Brookshire, Randall and Stoll, 1980; Just, Hueth, and Schmitz, 1982; Hanemann, 1986; Hoehn and Randall, forthcoming). Taken together, these findings support the validity of the CV method. They do not, of course, guarantee that the method is easy to use nor that every application is equally trustworthy (Mitchell and Carson, 1987).

The accomplishments during stage II directly led to the method's increasing acceptance by economists and decision makers. One indicator of this was the willingness of editors of leading economics journals to accept articles based on CV findings. The *Journal of Environmental Economics and Management* published its first CV article as early as 1974 and served as a publishing outlet for many of the pioneering CV studies. Other prestigious economics journals in this country eventually followed suit as shown by the publication of CV articles in the *Quarterly Journal of Economics* (Greenley, Walsh, and Young, 1981), the *American Economic Review* (Brookshire et al., 1982), and the *Journal of Political Economy* (Smith and Desvousges, 1987).

Another sign of the method's increasing acceptance was the willingness of the water resource agencies to recognize it as a valid benefit measurement methodology. In 1979, the Water Resources Council (1979) published its newly revised "Principles and Standards for Water and Related Land Resources Planning" in the *Federal Register*. This important document set forth guidelines for federal participation in project evaluation and specified which methods were acceptable for use in determining project benefits. The inclusion of the CV as one of the three recommended methods -- the other two were the travel cost and the unit day value methods -- contributed importantly to the method's perceived legitimacy. Unfortunately, the *Principles and Standards* document, which was modified and expanded slightly in 1983, presumed that contingent valuation was farther along the learning curve than was actually the case in 1979 when it designated the then prevailing CV practice of eliciting values by means of a "bidding game" as the "preferred" elicitation format.⁴ Subsequent research demonstrated that the bidding game format was very vulnerable to starting point bias which occurs when respondents' WTP are influenced by the particular dollar amount the researcher uses to start the bidding process (Boyle, Bishop, and Walsh, 1985; Mitchell and Carson, 1985; Roberts, Thompson, and Pavlyk, 1985). As might be expected, the bidding game format proved difficult to dislodge once it was enshrined as a bureaucratic

4. In this method the respondent is asked whether he or she is willing to pay a predetermined amount, such as \$5.00 a month, for a particular level of provision. If the respondent says yes, the amount is increased by predetermined steps, until the respondent says no. The reverse process occurs if the respondent rejects the opening bid.

prescription.⁵

Not too long after the Federal Register listing approved its use, one of the largest consumers of water benefit research, The Corps of Engineers, began to facilitate the routine use of contingent valuation to measure project benefits. By July 1986, the various Districts and the Corps' Institute of Water Resources had conducted 15-20 CV studies of varying degrees of sophistication⁶ and more were on the way. In the same year, the Institute published A Guide for Using the Contingent Valuation Methodology in Recreation Studies (Hoser and Dunning, 1986; hereafter "Guide") to provide step by step guidance for Corps' funded researchers who wished to use CV. Although this document bore less formal prescriptive authority than the Water Resources Council's "Principles and Standards," it offered much more detailed advice about how to conduct a CV study, advice predicated on the assumption that the method had reached the stage where correct procedures were unproblematic and its implementation was straightforward. The Guide also presumed that the learning curve had progressed to the point where production could be achieved at a lower unit cost per interview through the use of mail survey methods.⁷

The sudden rush of CV studies,⁸ the increasing use of inexpensive and little tested (in the CV context) mail surveys, and the publication of the Corps' Guide, which comes complete with model questionnaires suggest that at least some of the agencies who fund benefit research harbor the belief that the CV method has already attained stage IV on the learning curve. Our concern is that this misapprehension of the method's status may jeopardize its further maturation. The nature of surveys is such that despite the apparent simplicity of writing a questionnaire and soliciting answers there are many potential sources for bias which can, if proper caution is not exercised, lead to large distortions in CV-based benefit estimates. Such distortions, if produced by researchers inexperienced in survey research who assume the method is proven and unproblematic, could mislead policy makers and

5. As recently as 1986, contract administrators were known to insist, over the objections of CV researchers who were bidding on a proposed project, that they employ the "approved" bidding game elicitation format if they wished to win the government contracts.
6. C. Mark Dunning, personal communication, July, 1986.
7. Contingent valuation has also been recognized as an approved method for measuring benefits/damages under the Comprehensive Environmental Response, Compensation, and Liability Act of 1980 (Superfund) according to the final rule promulgated by the Department of the Interior (1986).
8. In addition to the CV studies conducted for the Corps of Engineers, we know of approximately over 100 studies that have used the method (Mitchell and Carson, forthcoming; appendix B), most of them conducted within the past five years.

provide skeptics with grounds to discredit the contingent valuation method before we acquire enough knowledge to know just when, where, and how the method can best be used in a routine fashion.

The Methodological Challenge⁹

An examination of where the contingent valuation method is on the learning curve properly begins with nature of the methodological challenge posed by using surveys to value public goods. Rossi, Wright, and Anderson (1983) define sample surveys as "relatively systematic, standardized approaches to the collection of information...through the questioning of systematically identified samples of individuals." Thirty years ago, Rensis Likert (1951) predicted that survey methodology, then a relatively new research technique, would have an increasingly wide application in all the social sciences. Proof of his prescience is as close as the nearest major journal in social psychology, political science, sociology, or economics where, depending on the discipline, currently 20 to 56 percent of the articles use survey data (Presser, 1984).

The modern survey is the result of two key methodological developments. The first is probability sampling, which enabled survey findings to be accurately projected to larger populations. Sampling itself is centuries old, but sampling based on strict probability principles has a much shorter history. In the United States it was first applied to survey research in the 1930s and came of age after the 1948 presidential election, which most of the major polls had confidently predicted would go to Dewey. Although the pollsters' use of the inappropriate quota sampling method was only one of the factors behind it, the debate provoked the national survey organizations to adopt the area probability methods first developed by the census bureau (Rossi, Wright, and Anderson, 1983). These techniques give each element in the population a known nonzero probability of being selected into the sample, thus making possible the use of statistical inference to project the results to the relevant population.¹⁰ If rigorously implemented, findings based on sample sizes as small as 600 to 1500 people can be representative of the entire United States population (or any other population) with a high degree of confidence.¹¹

The second development on which the modern survey rests is "the art of asking questions" (Payne, 1951). This phraseology nicely captures the largely qualitative nature of this achievement, which is based on a great deal of experience and a relatively small number of controlled

9. The following discussion draws, in part, on Mitchell and Carson (forthcoming).
10. Kish (1965), Sudman (1976), Cochran (1977), Yates (1980) and Frankel (1983) provide useful overviews of sampling theory and practice.
11. The same sample size will be equally effective for large or small populations.

experiments, most of them of recent vintage (e.g., Schuman and Presser, 1981), to test the effects of different ways to word questions. The result is that there is no one "correct" version of a questionnaire. Such diversity should not be taken to imply that anything goes, however. Writing survey questions appears deceptively easy to many people. In reality, it is difficult to convey even relatively simple ideas so they are uniformly and correctly understood by respondents who come from diverse backgrounds, have differing levels of education, and may or may not feel comfortable in the interview situation. Words that seem clear to the researcher may be ambiguous to some or many respondents. Seemingly slight changes in word order may convey unexpected meanings (Payne, 1951). Words have multiple meanings. Fee (1979) found that his respondents held at least nine distinct understandings of the term "energy crisis."

Another aspect of the art of asking questions is determining the best sequence in which to ask the questions. Experience has shown, for example, that questions about the respondent's personal characteristics -- the background questions -- are best left to the end of the questionnaire when the respondent is more relaxed about being interviewed and less likely to take offense at having the interviewer probe into his private life. A sequence of questions in a questionnaire that "works" will flow from one topic and question to another so that the transitions are smooth, the interrelationship between the parts is perceived by the respondent to be logical, and the pace is varied enough to retain the respondent's attention.

The Goal of Data Comparability

Irrespective of how it is administered, a major requirement of a survey is to ensure that the data it obtains are comparable -- that is, the information is gathered in a standardized fashion so that one person's answers can be compared with those given by another person. To this end, survey organizations devote considerable care and resources to pretesting questionnaires and training interviewers. Pretesting (see Converse and Presser, 1986) is the survey equivalent of the test flight. Just as no plane manufacturer would go into production without rigorously testing its latest design, no survey writer would assume that a questionnaire on a new topic, especially if the questionnaire is complex, could be sent directly to the field without careful tryouts under field-like conditions. Even experienced survey practitioners are often surprised when certain questions work better than they had anticipated and others, which they thought were winners, turn out to be fatally ambiguous. Pretests normally consist of an extended period of trial and error experimentation with draft versions of the questionnaire to see which alternative question wordings and orderings work best. If the topic is novel, the pretest process may include preliminary in-depth research, perhaps using focus groups (Desvousges, Smith, Brown and Pate, 1984; Randall et al., 1985; Mitchell and Carson, 1986), to learn how people conceptualize and talk about the topic.

Comparability also imposes demands on how interviewers conduct themselves in surveys. As David Riesman (1958) once observed, the basic task of the interviewer is to "adapt the standardized questionnaire to

the unstandardized respondents." Except for mail surveys, questioning is a social process. Each interaction between an interviewer and a respondent is unique owing to the particular circumstances in which the interview occurs and the personal characteristics of the two participants. In order to "adapt the questionnaire" without distorting or changing it, the interviewer has to motivate the respondent to enter into a special kind of relationship. Schuman and Bradburn describe how interviews differ from ordinary conversations.

The survey interview...is a transaction between two people who are bound by special norms; the interviewer offers no judgment of the respondents' replies and must keep them in strict confidence; respondents have an equivalent obligation to answer each question truthfully and thoughtfully. In ordinary conversation we can ignore inconvenient questions, or give noncommittal or irrelevant answers, or respond by asking our own question. In the survey interview, however such evasions are more difficult. The well-trained interviewer will repeat the question or probe the ambiguous or irrelevant response to obtain a proper answer to the question as worded (1982: 5).

It is precisely at the point of probing and handling respondent queries that comparability can be lost unless the interviewer rigorously follows instructions not to offer any information or explanations other than those described in the handbook for the study.¹²

The Problem of Response Effects

Despite fifty years of experience with surveys, survey methodologists agree that our knowledge about how respondents answer survey questions is still relatively primitive (Bishop, 1981; Schuman and Presser, 1981; Dijkstra and van der Zouwen, 1982; Bradburn, 1983; Jabine, et al., 1984; Turner and Martin, 1984). We know a great deal about how surveys can go wrong and much less about how to keep them from going wrong. Fortunately, there is currently considerable interest in the latter topic¹³ with the result that our understanding of response effects is growing more rapidly than ever before. Nevertheless, the current state-of-the-art in survey research is sufficiently problematic that two of the leading academic survey methodologists (Kalon and Schuman, 1982), advise against taking the marginal distributions of

12. The Research Triangle Institute's (1979) Field Interviewers General Manual offers an informative overview of the interviewer's role and training.

13. See, for example, the reports, prepared under auspices of the National Research Council, of the Panel on Survey Measurement of Subjective Phenomena (Turner and Martin, 1984) and the Advanced Research Seminar on Cognitive Aspects of Survey Methodology (Jabine, et al., 1984) and the research program currently being undertaken by Roger Tourangeau and his colleagues (Tourangeau et al., 1985).

answers to non-factual questions "too seriously" (Kalton and Schuman, 1982), advocating instead that researchers concentrate their attention on some form of correlational analysis. According to this view, the mean value of scales measuring concern about economic growth and environmental attitudes may fluctuate from survey to survey, but the relationship between the two variables is likely to be stable and meaningful. Since the distribution of the WTP amounts is the primary focus of attention in CV surveys, this observation deserves notice as a warning against methodological hubris.

Response effects are a type of nonsampling error that can distort a survey's results. They occur when one or more characteristics of the question, the respondent, the interviewer, and the context in which the interview takes place unduly affect response behavior. The many dimensions involved make the opportunities for such effects legion. For example, when asked a question, respondents may misunderstand it or their answer may be influenced by the order in which it appears or by the interviewer's manner. One team of hardy researchers (van der Zouwen and Dijkstra, 1982) arrayed causal factors by various types of survey outcomes and determined that an adequate nonsampling errors model needs to take into account at least 300 potentially important interaction effects. The current state of response effect theory is suggested by the fact that "theory" does not appear in the index of their book, *Response Behavior in the Survey-Interview* (1982), which is devoted to examining the 104 bivariate propositions they developed about respondent behavior.

Why is the survey interview vulnerable to response effects? The answer lies in the social nature, motivational mysteries, and cognitive capacity of human beings. In-person and telephone interviews are social situations and humans respond to these situations in complex ways -- part idiosyncratic, part programmed by learned rules of behavior or norms, and part in reaction to the prevailing norms. The ideal respondent would be motivated to devote as much time and energy as required and to answer every question truthfully. The actual respondent's motivation is likely to depart from this ideal in ways that fluctuate according to his current obligations, his sense of what the interviewer wants to hear, his current degree of self-confidence, his self-image, and his reaction to the interviewer.

As for cognitive capacity, survey researchers have long recognized that surveys must adjust to human frailties. These frailties include the difficulty some people have in understanding seemingly simple questions and instructions and the difficulty many respondents have in accurately recalling even relatively recent and concrete events such as a visit to the doctor a week before. More recently, survey researchers have become interested in the findings of cognitive psychologists (e.g., Kahneman, Slovic, and Tversky, 1982; Nisbett and Ross, 1980) who, on the basis of their laboratory experiments, have suggested that humans tend to use certain rules of thumb (heuristics) in ways that can affect their responses to survey questions. For example, the "availability" heuristic (Tversky and Kahneman, 1974) holds that people will tend to give the answer that is most immediately available to them in memory when asked to give a verbal report of any kind. Applied to surveys,

this implies that "the behavior itself of answering questions may act as a 'primer' which makes some cognitions more accessible or salient than others" (Bishop, 1981). The "anchoring" heuristic is another tendency identified by the cognitive psychologists. In this case, people make estimates by starting from an initial value, which may be suggested by the formulation of the problem, that then is adjusted to yield the final answer (Tversky and Kahneman, 1974). For example, when subjects were asked to estimate the frequency in death from each of 40 different causes, their estimates tended to be higher or lower depending on whether they were first told that about 50,000 people die annually in motor vehicle accidents or that about 1,000 people a year die from electrocution (Slovic, Fischhoff, and Lichtenstein, 1982).¹⁴

CV Surveys Compared with Conventional Surveys

Surveys are diverse in what they try to measure and the uses to which they are put. Having called attention to the general methodological challenge facing the survey researcher, we now turn to the special problems posed by CV surveys. CV surveys are a particularly demanding application of the survey methodology. This is not, we hasten to add, because the types of questions asked in CV surveys are fundamentally different from those asked in "standard" attitude surveys. Table 1 lists the types of items surveys commonly measure. Even the *raison d'être* of the CV survey, the use of survey questions to predict consumer behavior, has its counterpart in the attempts of political pollsters to predict elections and the efforts of market researchers to predict how new products will fare in the marketplace. Nor are CV surveys unique in posing hypothetical situations to their respondents. Similar "what-if" questions are used in other settings to divine the prospects of potential candidates or to discover which characteristics of prospective products will most endear them to consumers. "If Ted Kennedy was the Democratic candidate in the 1988 presidential election and George Bush was the Republican candidate, which one would you vote for?" "If the new toothpaste I described to you was green in color would this make you more likely to try it, less likely, or wouldn't it affect your decision one way or another?"

14. Some people look to these cognitive research findings as a possible basis for a theory of response effects (Bishop, 1981; Jabine, et al., 1984). Others, however, are skeptical about the potential payoff from this source. The survey methodologists Schuman and Presser (1981: 313), for example, found that cognitive research findings offered surprisingly little theoretical guidance for the solution of the question wording and "context effect" problems they studied in their split-sample survey experiments. They attribute this situation to the great difference between controlled laboratory experimentation and "the encounters with ordinary people that characterize surveys" (Schuman and Presser, 1981:313). This is a perceptive comment which may be augmented by the observation that the cognitive research enterprise is devoted to finding factors that create differences in respondent outcomes, whereas survey methodologists endeavor to identify ways to avoid or to minimize differences caused by such factors.

COMPARISON OF INFORMATION OBTAINED BY GENERAL
AND CONTINGENT VALUATION SURVEYS

TABLE 1

Type	General Examples	CV Examples
Self reports of behavior	Current employment status; Voting behavior	Recreational participation; Consumption of bottled water
Demographic Information	Sex; Education; Income	Sex; Education; Income
Personal knowledge	Names of respondent's Senators; Awareness of the accident at Three Mile Island	Awareness of risks to drinking water quality from PCB's
Opinion attitudes	Satisfaction with job; Concern about crime; "Are we spending too much, too little or right amount for education in this country?"	Concern about air pollution; confidence in government
Expected Future behavior	Voting intentions; Expected future purchases of consumer goods	Willingness to pay for specified amenities; voting intentions in a hypothetical referendum

Source: Mitchell and Carson (forthcoming, chapter 4).

The reasons why valid CV surveys are even more difficult to conduct than regular surveys lies in: (1) the novelty of the situation most CV surveys pose to the respondent, (2) the need to construct a market in which to sell the good and, (3) the effort required for a respondent to arrive at a meaningful answer to many WTP questions.

CV surveys ask respondents to make a judgment that is novel to most people. While respondents are generally familiar with the idea of making decisions about referenda propositions or expressing opinions about whether or not they will buy a car this year, the request to set a dollar value on what they would be willing to pay for a hypothetical change in a public or quasi-private good which is described in some detail is unfamiliar to most people. The novelty of the valuation request is enhanced when respondents lack direct experience with the good being valued, such as a park they have never visited. In our experience, professional survey researchers who encounter a complex CV instrument for the first time tend to assume that these types of judgments are beyond the capacity of respondents in a survey setting and are surprised at the quality of the valuation data which can be obtained.

Respondents in CV surveys value the good in a specific setting. Survey researchers have long recognized that respondents' answers to general questions are often poor predictors of how they would respond to specific applications of the general issue. In a survey taken shortly after World War II (Cottrell and Eberhart, 1948: 8), large majorities supported the idea that the U.S. government should do everything it could to stimulate world trade, yet opposed a specific plan to extend large credits to Great Britain for this purpose. Because they create a hypothetical market in the good, CV surveys are required to be even more specific than this. They must, in Randall, Hoehn, and Brookshire's (1983: 637) words, confront "the respondent with a well-defined situation and elicit(ing) a circumstantial choice contingent upon the occurrence of the posted situation." The "posted situation" typically includes such factors as the current level of the amenity's provision, the amount of its increase or decrease the respondent is to value, how this will be provided, how the respondent will pay for it, and who else will pay for it. The amount of description required to construct a scenario which presents a plausible market and describes the good in adequate detail is sometimes considerable.

15. Pure public and quasi-private goods are primarily distinguished by whether or not access to the good can be limited. A pure public good is a good for which potential access is available to all members of a collectivity and it is impractical or impossible to exclude people from consuming the good (such as air visibility). A quasi-private good, in our usage, is one which every person in a collectivity has the right to use, but where the collective may grant differential access, often for a fee and on an equal basis (such as a recreational area). Note that fee typically does not cover the full cost of providing the amenity. See Mitchell and Carson (forthcoming, chapter 3) for a further discussion of this distinction.

This stands in sharp contrast with conventional surveys which offer comparatively few details about situations the respondent is asked to value. Consider the following public opinion question:

Now, I'd like to find out how worried or concerned you are about a number of problems I'm going to mention: a great deal, a fair amount, not very much, or not at all....

..How worried or concerned are you about the presence of toxic chemicals such as pesticides or PCBs in the environment?

Cleaning up our waterways and reducing water pollution?

The disposal of industrial chemical wastes that are hazardous? (Mitchell, 1980)

It covers three environmental amenities. The description of each is very brief, there is no attempt to describe different levels of provision, and the only response requested of the respondent is a simple choice between four levels of concern on an ordinal response scale. In comparison, a CV scenario for just one of these goods, reduced risk from toxic chemical contamination of a water resource, would need to identify the chemicals, their uses, consequences, and current presence in the environment. It would also describe several levels of contamination in enough detail so the respondents could understand the changes they are asked to value in dollars, and provide information about how these changes would occur and how the respondent would pay for them. Compared with the minute or two it would take to ask the conventional survey question on this topic, a CV version would require anywhere from ten to thirty minutes of the interviewer's time.

Finally, CV surveys tend to require a greater effort from the respondent than most conventional surveys. This is particularly the case for unfamiliar and complex topics such as CO₂-induced climatic change and genetic engineering. Few respondents bring well realized values for such amenities to the CV interview, yet they are asked to pay attention to the (sometimes lengthy) description of the market, search their preferences, take their income constraint and possible substitutes into account, and determine a dollar amount which represents the most they would pay for each level of the good the survey attempts to value. Fischhoff, Slovic, and Lichtenstein (1980) point out that people's opinions about such goods are likely to be "labile" or subject to change because they may not have thought through the implications of their views, they may have contradictory values which they bring to bear on the situation, or they may vacillate between incompatible, but strongly held, positions. In situations such as these, respondents will be tempted to minimize the effort by resorting to strategies which ease the decision burden such as giving an off-the-cuff answer or one suggested by an aspect of the scenario which is not intended to convey value.

Even when people are asked to value amenities with which they are personally familiar, such as the types of recreational facilities considered in the Corps of Engineers' Guide, the researcher cannot assume that the respondents will readily grasp the meaning of the

scenario as the researchers intend it to be understood nor that they will immediately comprehend the request to express the dollar value they personally (or their household) hold for the amenity under the stated conditions.

The measurement problems faced by CV surveys become especially important when their intended use is taken into consideration. Outside of market surveys and political polling for candidates, most opinion surveys are not directly used in making decisions or, if they are so used, a high level of precision is not assumed.¹⁶ The avowed purpose of CV surveys, on the other hand, is to obtain benefit estimates for benefit/cost decisions where an efficiency criterion is employed to allocate resources to their highest valued uses. In order to fulfill this mission, the WTP amounts obtained in a particular study are typically aggregated across the relevant population and presented as an authoritative representation of the good's benefits. In such an enterprise, questions about the quality of the data, their validity, and the precision of the estimates will be raised by policymakers and, as has already happened in several instances, by lawyers and judges.¹⁷ This places a special burden on CV studies to use the best available methodology. Unfortunately, at this stage in the method's development, there is no one best methodology for each case that can be taken off the shelf. The researcher needs to be aware of recent methodological developments and, in many cases, to test new approaches or the application of old approaches to new situations in order to arrive at credible benefit estimates.

Three Issues in the Use of CV Surveys

We have argued that the CV method has the potential to obtain valid WTP estimates for nonmarketed goods, but that the method on which it rests, survey research, requires careful attention to the problems of conveying meaning to diverse respondents and motivating them to undertake the effort necessary to arrive at a value for the amenity. A successful CV scenario must be understandable by the respondent as

16. Exceptions to this generalization exist, such as the Census Bureau's employment and cost-of-living surveys, the results of which may trigger important decisions, but these are large, routine, data gathering efforts which measure relatively well defined, previously experienced behavior. Smith (1986) notes, however, that even these questions require a fair amount of judgment on the respondents' part.

17. In 1986, benefit estimates based on CV surveys figured in an administrative hearing between the Northern states Power Co. and the Minnesota Pollution Control Agency (Welle, 1985; Carson et al., 1986) and in several court cases brought against corporations by the Colorado Department of Law under the Comprehensive Environmental Response, Compensation and Liability Act of 1980 (known as variously as the CERCLA or superfund legislation) (Department of the Interior, 1986).

intended by the researcher, and be perceived as plausible, and meaningful. If these goals are not met, respondents may "guess" at a value and give unreliable WTP amounts, or they may fail to respond and give WTP amounts that are influenced by one or more features of the survey instrument, or they may give responses, such as protest zeros, which do not mean what they appear to mean.¹⁸

The complexity of using surveys to value water resource amenities at this stage on the method's learning curve may be illustrated by considering three of the several important topics on which the Corps' Guide prematurely assumes methodological consensus. As mentioned, the Guide is intended to offer Corps planners practical assistance in conducting CV surveys to measure recreation benefits. Its treatment of instrument design consists of a very brief chapter on the subject augmented by an appendix containing several model questionnaires. The three topics we will consider are the relationship between bias and information, the measurement of nonuse benefits, and the special problems posed by the use of mail surveys in CV studies.¹⁹ In each case the Guide's recommendations (or implied recommendations) may result in biased benefit estimates. Our discussion will assume that the aim of a CV study is to value the relatively easy-to-understand recreational amenities covered by the Guide such as beaches, boat-ramps, boat marinas, and dredged channels.

The Relationship Between Bias and Information

To the extent that an instrument measures the concept under investigation (is unbiased, in the terminology of statisticians and economists) it is valid (Bohrstedt 1983). The absence of systematic error is implied by:

$$E(RWTP_j - TWTP_j) = 0, \quad -j \quad (1)$$

Systematic errors are perhaps the more serious threat to the accurate measurement of respondents' WTP amounts, and are more difficult to assess and adjust for, than are many types of random errors.²⁰ Unlike random error, which is amenable to assessment by sampling and test theory, there is no applicable body of theory by which validity can be assessed (Garmines and Zeller, 1979; Bradburn, 1982) because we lack an

18. See Mitchell and Carson (forthcoming) for an extended discussion of these issues including a typology of potential biases in CV surveys.

19. For a more detailed discussion of these topics see Mitchell and Carson (forthcoming). Mitchell and Carson (1987) discusses the issue of how the validity of individual CV studies may be assessed.

20. In a recent review of the general survey research methodology literature, George Bishop concluded that the magnitude of systematic error "greatly exceeds that stemming from random sampling or non-sampling error" (1981:591)

explanatory model of the cognitive processes which underlie respondents' verbal self-reports (Bishop, 1981:591). In these circumstances, the prevention of systematic error necessarily has an *ad hoc* character about it, although survey researchers have developed rules of thumb, based on experience and a growing body of survey experiments, which serve to minimize bias.

The question of bias is complicated in CV surveys by the general absence of a measurable true WTP value for public goods which can be used to assess the validity of a given study. This means that bias must be inferred from our partial understanding of respondent behavior -- if you ask the question this way, people will be likely to distort their answers -- or from evidence in the survey which shows that changing the wording of the scenario in ways that are not expected to affect the WTP amounts in fact does so. "Not expected" is a key phrase here, because such differences may in fact be legitimate contingent effects.

Until recently there has been some confusion in the CV literature on this point which is reflected in the Guide's discussion of the biases that should be avoided by the designer of a CV survey (Moser and Dunning, 1986: chapter 1). The Guide appears to assume that only the nature and amount of the amenity being valued should influence the WTP amounts; all other scenario components, such as the payment vehicle and amount of provision, should be neutral in effect. Thus, while it advises the researcher to choose a payment vehicle that is as realistic and familiar as possible -- such as admission passes to a beach -- it cautions that this vehicle should not influence the respondent, otherwise payment vehicle bias will result (Moser and Dunning, 1986: chapter 3). There is ample evidence that this degree of scenario neutrality is difficult to achieve (see Mitchell and Carson, forthcoming, for a review) and a real question as to whether it is an appropriate goal. Others (Arrow, 1986; Kahneman, 1986; and Randall, 1986) now argue that important conditions of a scenario, such as the payment vehicle, should be expected to affect the WTP amounts because respondents in a CV study do not value levels of provision of an amenity in the abstract, they value a policy which includes the conditions under which it will be provided, and the way the public is likely to be asked to pay for it.²¹ One important implication of this latter position is that without further evidence it cannot be assumed that the findings of a particular CV study can be generalized to other settings unless:

- (a) the other setting matches the scenario employed in the study, or
- (b) the researcher can show empirically²² that the WTP amount in the original study is insensitive to the scenario elements which do not apply. It would be misleading, for example, to use a benefit estimate based on a scenario which implies that the proposed recreation area the respondents are asked to value has no nearby substitutes to value a

21. The notion that a public good does not have a value independent of its method of financing goes back to at least Wicksell (1976) and is fully consistent with economic theory.

22. By incorporating a test in the study design.

recreation site for which nearby substitutes are available.

The Guide also describes "hypothetical bias" as if there were a unique systematic distortion (bias) attributable to the use of hypothetical markets or situations. What hypothetically may induce unless the survey researcher is careful, is either reduced reliability (increased statistical variance) owing to respondents offering casual, unthought out answers, or vulnerability to one or another bias, such as starting point bias, where the uncertain respondent is vulnerable to perceived cues from elements of the survey instrument.

The Measurement of Nonuse Benefits

The Guide advises CV researchers to first measure use values and then ask questions "designed to determine the option and existence values for the recreational opportunity." These can be separate questions or they can be "related to the previously elicited user value by asking if the respondent is willing to pay X percent more, in addition to the use charge, to ensure that the recreational site is available in the future." While we believe CV surveys are capable of measuring benefits that include a nonuse dimension, we are less optimistic about their ability to obtain meaningful estimates of separate component values in the fashion recommended at this point in the Guide. In the particular case of option value, there are theoretical grounds to argue that even if such an approach were workable, it would be inappropriate to measure option value in this fashion.

We begin with the latter point. Recent work on the concept of option value indicates that it is important to distinguish ex ante and ex post perspectives in thinking about the concept.²³ Consider two time periods where what the state of the world will be in period 2 is unknown in period 1. Option price (OP) is defined as the ex ante (period 1) state independent willingness-to-pay for a specified change in the level of the public good in question. Option price is usually considered to be the relevant welfare measure in benefit/cost analysis, because governments must decide what actions to carry out in period 1 and

23. Since Weisbrod's (1964) seminal work on uncertainty, which made specific reference to the possible irreversible consequences of the destruction of a national park, scholars have considered uncertainty in a benefit-cost analysis from two distinct perspectives. The first, which we consider here, is the "timeless" or option value approach (Cicchetti and Freeman, 1971; Krutilla, et al., 1972). From this perspective, option value is the price that people will pay for a contract which guarantees them the opportunity to purchase a good for a specified price at a specified point in the future. The second perspective is known as the "time-sequenced" or quasi-option value approach (Arrow and Fisher, 1974; Henry, 1974) where quasi-option value is regarded as the risk premium people will pay to delay an activity which, if undertaken, might foreclose making a better informed decision at a later time.

citizens' payments for the change are rarely dependent on what state of the world actually materializes in period 2. In contrast, expected consumer surplus (ECS) is an ex post measure.²⁴ It is found by summing over the probability of each state of the world times the expected consumer surplus in that state of the world as if that state were known to have occurred in period 2.²⁵ The problem with trying to measure option value as possible future use in a CV survey comes from the fact that what CV surveys almost always offer for sale is certain provision of the amenity for possible use in the future at a specified price.²⁶ Instead of an option to be able to buy the amenity in the future at some fixed price. As Smith (forthcoming) puts it, one is really selling access to public goods, not use. Contingent contracts are rarely relevant in the context of valuing public goods.

We now turn to the use of follow-up questions to measure other types of existence values such as valuing the good because others will be able to use it (vicarious consumption) or because providing it is a good thing in itself (inherent value). Our view of respondent behavior in the CV setting is that when people are asked to value an amenity, such as the provision of a boat ramp or setting aside prime forest lands for use as wilderness, they do so by making a holistic judgment. Instead of going through a mental process where they separately value each of the relevant benefit categories (such as use and existence values) before combining them in their mind to arrive at a total value, respondents arrive at a global judgment about what the amenity is worth to them based on the conditions described in the scenario. This judgment reflects the total configuration of benefits -- use and nonuse -- they believe will accrue to them from the amenity. Consider a consumer analogy. While consumers could probably place a dollar value

24. We want to clearly distinguish between ex ante and ex post states of the world (is it going to rain/it is raining), and ex ante expected utility from undertaking an activity and ex post actual utility gained from undertaking the activity. Only the states-of-the-world use is relevant to our discussion and the option value literature. Differences in utility under the expected/actual utility approach may be much larger, but are rarely considered in economics because of the impossibility of providing contingent contracts against such differences unless they are due solely to the realized state of the world.

25. In practice expected consumer surplus is measured only in the state of the world that did occur.

26. We can envision cases of quasi-private goods where one might pay something now to join a pool of people who had the right to buy a public good in the future at some fixed amount. The CV study by Brookshire, Eubanks, and Randall (1983) of hunters willingness to pay for a grizzly bear or bighorn sheep stamp which would entitle them, on payment of a fixed amount, to participate in a drawing for licenses to shoot these endangered species at a given time in the future, comes close to invoking such a situation.

on a given car, if asked to do so, they would not arrive at this value by adding up in their minds how much they would be willing to pay individually for the vehicles' styling, horsepower, the comfort of its seats, its trunk, and the prestige associated with its ownership.

This view of the valuation decision leads us to be skeptical of simple attempts to ask respondents to separately value the benefit categories for a given amenity identified by economists. William James (1890 in Fischer, 1970:209) labeled the "psychologists' fallacy" the error of assuming that a person who has a given psychic experience is conscious of it with the same degree of insight as an observing psychologist. In a similar vein, the error of assuming that respondents are aware of what motivates their value judgments with the degree of precision desired by the economist researcher may be called the "fallacy of motivational precision." It is difficult enough for respondents to place a dollar value on the global desirability of a nonmarketed good, but to ask them to first restrict their WTP amount to the use dimension and then, in subsequent questions, to say how much they are willing to pay just to know that others, but not themselves, can use it (vicarious consumption), is even more difficult. Faced with this type of follow-up question, many respondents will dutifully comply with the interviewer's request and offer a dollar amount, but there is a danger that the amounts are based on a desire to be helpful or not to appear stupid.

The Guide is not insensitive to the fallacy of motivation precision, but it fails to give it sufficient weight despite the fact that invalid existence values, even if small on a per household basis, have the potential to bias the presumed benefits of a project owing to the large number of households involved. If some measure of existence value is required, which approaches offer the most promise of avoiding the fallacy of motivational precision? In what follows, we briefly examine four possible measurement strategies. The first three are based on subjective judgments made by the respondents whereas strategy IV uses self-reported information about the respondents' use of a given amenity to infer a lower bound on aggregate existence value. Strategies I and II are particularly prone to the fallacy of motivational precision.

In strategy I, each type of value is individually described to the respondents who are directly asked how much this dimension of the amenity is worth to them. If separate values for direct use, indirect use, and existence were obtained in this way, the total WTP amount for the good could theoretically be obtained by adding the values. Although this strategy has the advantage of apparent simplicity, it poses a potential for invalid or meaningless answers because respondents may mistakenly include more than one value dimension (use plus existence) when they give their initial WTP amount. When used with care to value recreational facilities which are excludable, however, and this is the type of amenity described in the Guide's model questionnaire, strategy I may be suitable for measuring (only) use value through the use of an entrance fee payment vehicle.

Strategy II, the decomposition strategy, involves asking respondents to separate a previously obtained total WTP amount into one

or more benefit categories. For example, each respondent might be asked: "Of this amount (total WTP amount), how much of it would you pay for (description of category or subcategory)." This approach is preferable to the previous approach because it first obtains a presumably valid total WTP amount before attempting to break it into potentially invalid component values. Obtaining the total WTP amount first also helps respondents grasp the idea that the component values are a subset of the overall value. To operationalize this definition, it is first necessary to measure a respondent's total value for the amenity. This WTP amount will include both the use and existence value dimension. The respondent has for the good. The existence component could, at least in theory, be ascertained by asking the respondent to say how much he would be willing to pay for it if, for some reason, his household could not use or have access to it. To the extent that the new contingency was plausible to the respondent this approach might avoid the motivational precision fallacy.

The experience of CV researchers who have attempted to use these strategies underscores the methodological difficulties associated with their use. Employing strategy I, one set of researchers (Greenley, Walsh and Young, 1982) asked their respondents to take their chances of future recreational use into account and say how much they would pay to postpone a deterioration in quality for a river basin in Colorado. They then asked how much they would pay just to know that clean water exists at a given level. "If it were certain you would not use the South Platte additional separate questions to measure bequest value for future generations and (quasi) option value. Although Greenley, Walsh and Young believe their respondents were able to comprehend their questions so as to give meaningful answers, the values they obtained for the three non-use values are sufficiently similar²⁷ to suggest that the respondents may not have been able to adequately differentiate these dimensions in their minds (Mitchell and Carson, 1985).

Walsh, Sanders, and Loomis (1985) used strategy II in a study of the benefits of protecting up to 15 wild and scenic rivers in Colorado. After obtaining the respondents' total value for this resource they asked them to give the proportion they would assign to the following types of benefits: use, "an insurance premium..." to guarantee your choice of recreation use of these rivers in the future, "bequest, and inherent. When the value of these benefits were elicited in this fashion, the mean WTP amounts were much more diverse,²⁸ suggesting that presumably spurious repetition was avoided. Whether respondents could

27. For their water fee vehicle the mean value in dollars per month for all resident households in the South Platte River Basin is \$1.16, \$1.23 and \$0.90 for the three types of benefits (Greenley, Walsh and Young, 1982).

28. On a per household basis, they obtained \$19, \$16, \$28, and \$36 for the four types of benefits for the "fifteen most valuable rivers" (Walsh, Sanders, and Loomis, 1985: 72).

meaningfully distinguish the four categories is another matter.

Strategy III involves posing two or more scenarios to separate subsamples or, possibly, to the same respondents,²⁹ which differ only with respect to the benefit component which the researcher desires to measure. The difference between the total WTP amounts for the two scenarios yields an estimate of the desired quality such as a natural area being available or unavailable for recreational purposes. This strategy avoids the fallacy of motivational precision because it only asks the respondents to give a total value for a given scenario. To the extent that the different scenarios are plausible and only invoke the intended value dimension, this strategy offers a potentially useful, though expensive, approach. A variant of strategy III, and one which is mentioned in the Guide, would be to create a single scenario where the only plausible motive for valuing the amenity is a specific benefit class or category.³⁰

The last strategy, IV, is also mentioned in the Guide. It avoids the fallacy of motivational precision because it does not rely on respondents' subjective assessments of the benefit components at all. Instead it uses reported use and/or anticipated future use of the good to indirectly estimate existence value. On the basis of self-reports the respondents are divided into those who use the amenity and those who do not. The WTP amounts are measured in the conventional manner. The WTP amounts given by the nonusers for the amenity are treated as a relatively pure expression of nonuse or existence value whereas the users' WTP amounts include some combination of use and implicit values. Since no defensible external criteria are available to determine which portion of the users' WTP amounts should be assigned to the existence category, estimating existence values by this approach will result in a lower bound for existence value (Fisher and Rauber, 1984). Despite its methodological limitations this strategy has provided suggestive evidence about the magnitude of existence values.

The Use of Mail Surveys

The third area where matters are more complex in the conduct of CV surveys than indicated by the Guide is the use of mail surveys in CV studies. CV surveys have been conducted using the three principal methods of administration: in-person, telephone and mail surveys. Cost aside, the in-person survey, where the interviewer conducts the interview

29. This would be much less expensive than two surveys, but raises questions about contamination across scenarios. Can respondents value a second scenario without being influenced by the content of the first scenario?

30. WTP for the second scenario (the absence of the benefit class) is implicitly assumed to be a zero.

in the respondent's dwelling, is the method of choice for most CV situations. The physical presence of the interviewer offers the greatest opportunity to motivate the respondent to cooperate fully with a complex or extended interview and the interviewers can be expected to probe unclear responses, use skip patterns where the choice of followup question depends on the respondent's answer to previous question, or to tailor the interview to the individual respondent's needs.³¹ In-person interviews lend themselves to the use of various types of visual aids or "display cards" to help convey complex ideas or bodies of information, and they support missing data techniques. Finally, since refusals typically occur before the potential respondent learns that the purpose of the survey is to value a particular good, the in-person methodology minimizes sample selection bias. This bias occurs when those who are interviewed differ systematically in the value they hold for the good from those who refuse to be interviewed.

The major drawback of in-person interviews is the expense involved in selecting the sample and properly administering the questionnaire. Telephone surveys are somewhat less expensive than in-person interviews because they circumvent the time-consuming process of locating and interviewing a respondent at home and also because their length is usually limited to around fifteen minutes. Mail surveys, however, are the cheapest type of survey to administer on a per completed questionnaire basis, especially when the sample can be based on an up-to-date mailing list. This and the fact that they also offer the researcher the opportunity to use charts, pictures, and maps, has made them increasingly popular among CV researchers and their funding agencies. The Guide, for example, strongly supports the use of mail surveys for measuring recreational amenities with the proviso that the researcher use state-of-the-art techniques -- such as followup letters, inducements, sponsor identification and the like -- which have been shown to enhance response rates. It does not specify what response rate would be "acceptable" except to suggest that a 10-15 percent return is very low and that samples based on response rates under 50 percent are likely to under or over represent important categories of people in the population. When response rates are under 50 percent the Guide recommends that the sample be adjusted to correct for sample nonresponse bias. This adjustment ensures that categories of respondents, such as those in the various income groups, contribute to the WTP estimate in proportion to their distribution in the population from which the sample was drawn.

There are two additional drawbacks to the mail survey method which the Guide does not address but which are likely to pose significant problems for the validity of mail CV studies. The first problem is the mail survey's reliance on the ability of the respondent to read and understand the description given in the scenario. The reading level of

31. Accepted survey practice forbids interviewers from providing ad hoc explanations or answers to respondent questions, a point unfortunately not mentioned in the Guide which recommends the use of nonprofessional interviewers.

many adults is surprisingly low. According to the National Assessment of Educational Progress, which is based on a careful study of literacy among a national sample 3,600 young adults between the ages of 21 and 25, 6 percent of those who were interviewed (in-person) were unable to read a short sports story in a newspaper. Twenty percent could not read as well as the average eighth grader, 37 percent could not present the main argument in a newspaper column, and only 43 percent could use a street map (Kirsch and Jungblut, 1986). These data actually underestimate adult reading comprehension problems since the young adults interviewed for that study have a higher level of education than comparable cohorts of older age groups. The implications are clear. Unless the scenario in a mail questionnaire is very short and simple, or the respondent is reasonably well educated and also highly motivated, there is a good chance that the respondent will miss important details or misinterpret one or more aspects of the scenario. For example, comprehension problems could easily lead some respondents to miss a scenario's instructions to consider only use in the use segment of a scenario as discussed above even when the wording appears to the researcher to be invulnerable to misinterpretation.

Sample selection bias, the second drawback, results from the self-administered character of mail surveys. Respondents to in-person and telephone surveys typically make their decision to participate before they learn very much about the nature of the interview. In contrast, those who receive mail surveys are able to become familiar with the subject matter of the questionnaire by looking it over before deciding whether or not to go to the effort of filling it out and returning it. This introduces the possibility that those who respond will hold different values for the amenity than the nonrespondents. It is likely that those who have a stronger interest in an amenity (and usually a higher value for it) will be more inclined to go to the trouble to fill out the questionnaire than those with similar demographic characteristics who are less interested and whose WTP amounts, if they were obtained, would tend to be lower. To the extent that this occurs we have sample selection bias.

Sample selection bias has the potential to seriously distort the WTP estimates and cannot be corrected by the standard weighting or imputation procedures recommended by the Guide for sample nonresponse bias. Sample nonresponse bias, as noted earlier, is caused by the differential propensities for population subgroups to respond to surveys of any kind. The adjustment procedure the Guide recommends to correct for sample nonresponse bias is valid only if there is random nonresponse within identifiable categories of people so that those in, say, a low income group who return the mail survey hold the same values for the amenity as their low income counterparts who normally throw all mail surveys in the waste basket. This assumption makes it possible to use the values given by those who respond as a surrogate for the amounts demographically similar nonrespondents would have offered if they had returned the questionnaire. We cannot use this procedure to compensate for sample selection bias because we cannot say for sure why the

questionnaires were not returned.³² The dilemma, in the words of Anderson, Basilevsky, and Hum (1983) is that "...in the case where bias is least problematic (random nonresponse within identifiable groups), we can fix it. For the case where bias is a serious threat (nonrandom nonresponse within identifiable groups), the standard methods are not appropriate."

The only currently defensible way to address the bias introduced by the sample selection problem in mail surveys (if a lower bound on WTP is desired) is the procedure used by Bishop and Boyle (1985) in their study of the economic value of an Illinois beach nature preserve where they assigned a \$0 value to each nonrespondent. This practice yields a lower bound estimate of the amenity's value (presuming that their study is otherwise accurate) as it is based on the conservative assumption that every single nonrespondent (excluding those with incorrect addresses) to their mail survey (about 30 percent of their sample) was not willing to pay anything for the preserve. If they had proceeded under the same assumption that those who returned their questionnaires held the same views about the amenity as the nonrespondents, it is likely that their study would have substantially overestimated the amenity's value.

Conclusion

We proposed an application of the learning curve concept to new benefit methodologies in this paper and applied the model to the contingent valuation methodology. As knowledge accumulates about a method, researchers proceed along a learning curve whose end point, presuming that no insoluble problems are encountered along the way, is a stage where the method can be implemented in a routine fashion. Our analysis leads us to conclude that although the CV method has passed beyond the experimental prototype stage it is not understood well enough to have reached stage IV or the routine application stage.

We paid particular attention to what survey researchers have learned about the art of asking questions: their wealth of experience

32. Some CV researchers have argued that nonresponse bias is not likely to be significant on the basis of the findings of a study conducted by Wellman, et al. (1980). The Wellman et al. study compared early and late respondents to a mail nonCV outdoor recreation survey, which achieved a 70 percent response rate. The authors argued, on the basis of apparent similarities between these groups on a number of characteristics, that "time, effort, and dollars spent in intensive followups to increase recreation survey response rates" might better be expended on other phases of the research process. This finding is an insufficient basis to assume random nonresponse as Wellman et al. did not study the 30 percent of their sample who failed to respond to their survey. There are no grounds for believing that late respondents to mail surveys such as theirs are a valid surrogate for the nonrespondents and a priori (see above) and empirical (Anderson, Basilevsky, and Hum, 1983: 479-80) evidence to the contrary.

has revealed many subtle problems, but also sensible solutions. CV surveys are less-travelled terrain, with fewer solutions precisely when they are most needed because of the nature of the subject matter and the uses to which CV studies are put. The principal challenge facing the CV researcher is to make the scenario sufficiently understandable, plausible, and meaningful to respondents so that they can and will give valid and reliable values despite their lack of experience with one or more of the scenario's dimensions. The difficulty of writing CV scenarios which accurately communicate the intended meaning to respondents with varying levels of education, life experiences and interest in the topic is often underestimated by researchers with little experience in survey research. Unless respondents understand all the components of a scenario in the way that the researcher intends them to be understood, there is no assurance that those surveyed will properly value the good. Even if an instrument is understandable, the market it portrays must also be plausible. Respondents are unlikely to take seriously relationships or situations which do not seem credible, such as the use of an electric utility bill payment vehicle for a study of risk reductions from toxic waste dump sites. Finally, unless respondents are able to relate the scenario to their personal knowledge and experience in such a way that the market is genuinely meaningful to them, they will not be motivated to expend the effort necessary to determine their personal value for the good. Part of the challenge in conducting a CV study is to get the respondents to accept making a tradeoff between money and the public good which maintains the same level of utility. Economists usually assume that asking people to make this tradeoff is unproblematic whereas research conducted by social psychologists suggests that people have difficulty with tradeoffs (Abelson and Levi, 1985: 287).

These considerations underlie our view that we still have much of importance to learn about the CV method and that it is vulnerable to misuse. This view runs counter to what we perceive as a tendency among some researchers and agencies to treat the method as proven and relatively cheap, at least when used to value project recreational amenities. The Corps of Engineers' Guide for Using the Contingent Value Methodology in Recreation Studies (Hoser and Dunning, 1986) illustrates this tendency. We offered three illustrations of where, in our opinion, the Guide prematurely assumes a level of knowledge about the CV method that is inappropriate at this point on its learning curve. These are: the degree to which scenario elements should or should not influence WTP amount outcomes; how or whether to measure option and existence values; and the advisability of using mail surveys for CV studies. In each case our knowledge is growing; in no case can foolproof instructions currently be provided to novice CV researchers.

"Forcing" the CV method by prematurely applying it instead of letting it mature naturally is the understandable outcome when natural resource agencies need to justify projects and have limited funds to conduct the necessary research. Unfortunately, forcing poses a threat to the method's further progress along the learning curve. For one thing, poorly conducted studies or studies based on erroneous assumptions may produce embarrassingly implausible estimates which could result in bad policy or serve to discredit the method. For another, a

view that the method is well understood is likely to divert scarce agency funds from badly needed methodological research to bare-bones field applications. At the present time basic research on the contingent valuation method is needed to study such things as the role of perceived property rights in influencing WTP responses; how respondents consider the perceived uncertainty of supply when they value amenities; the size of the substitution elasticities between public goods; the role of visual stimuli in CV surveys; and the statistical techniques for analyzing discrete choice CV responses, identifying outliers, and imputing missing values to name just a few examples. Field applications need to be funded with sufficient generosity so that researchers can conduct thorough pretests and, in the final survey, use samples that are large enough to conduct split-sample experiments to test for possible instrument effects. CV surveys can be done on the cheap with apparent success as people can be found who will answer even meaningless questions. But skimping on questionnaire development costs, pretesting, interviewer quality and training, or sample size can easily lead to very sizable benefit misestimates. The new trend toward conducting mail surveys is particularly disquieting in this regard, as their use for CV purposes has received little formal study.

All benefit estimation methods are vulnerable to error, including those based on actual market prices. The history of the contingent valuation method supports the notion that it is generally possible to overcome problems when they have been identified or at least to minimize their effects and still obtain useful data. Our intention in pointing out the difficulties involved in using survey research to measure benefits is not to counsel despair, but to promote methodological sophistication about the contingent valuation method's present place on the learning curve.

References

- Abelson, Robert P. and Ariel Levi, 1985. "Decision Making and Decision Theory," in Gardner Lindzey and Elliot Aronson (eds.), The Handbook of Social Psychology, 3rd ed., Vol. 1, Newbery Award Records, New York.
- Anderson, Andy B., Alexander Basilevsky, and Derek P.J. Hum, 1983. "Missing Data," in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research, Academic Press, New York.
- Arrow, Kenneth J., 1962. "The Economic Implications of Learning by Doing," Review of Economic Studies, Vol. 29, pp. 155-173.
- Arrow, Kenneth J., 1986. "Comments," in Ronald G. Cummings et al., Valuing Environmental Goods, Rowman and Allanheld, Totowa, NJ.
- Bishop, George F., 1981. "Survey Research," Chapter 21 in Dan D. Nimmo and Keith R. Sanders (eds.), Handbook of Political Communication, Sage, Beverly Hills and London.

- Bishop, Richard C. and Kevin J. Boyle, 1985. "The Economic Value of Illinois Beach State Nature Preserve," Report to the Illinois Department of Conservation, Heberlein, Baumgartner Research Services, Madison, WI.
- Bloomgarden, Kathy, 1982. "Toward Responsible Growth: Economic and Environmental Concern in the Balance," The Continental Group, Inc., Stamford, CT.
- Bohnstedt, George W., 1983. "Measurement," in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research, Academic Press, New York.
- Boyle, Kevin J., Richard C. Bishop, and Michael P. Welsh, 1985. "Starting Point Bias in Contingent Valuation Surveys," Land Economics, Vol. 61, pp. 188-194.
- Bradburn, Norman M., 1982. "Question-Wording Effects in Surveys," in Robin M. Hogarth (ed.), Question Framing and Response Consistency, Jossey-Bass, San Francisco.
- Bradburn, Norman M., 1983. "Measurement: Theory and Techniques," in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research, Academic Press, New York.
- Brookshire, David S., Larry S. Eubanks, and Alan Randall, 1983. "Estimating Option Price and Existence Values for Wildlife Resources," Land Economics, Vol. 59, No. 1, pp. 1-15.
- Brookshire, David S., Alan Randall and John R. Stoll, 1980. "Valuing Increments and Decrements in Natural Resource Service Flows," American Journal of Agricultural Economics, Vol. 62, No. 3, pp. 478-488.
- Brookshire, David S., Mark A. Thayer, William P. Schulze, and Ralph C. d'Arge, 1982. "Valuing Public Goods: A Comparison of Survey and Hedonic Approaches," The American Economic Review, Vol. 72, No. 1, pp. 165-176.
- Carmine, Edward G. and Richard A. Zeller, 1979. Reliability and Validity Assessment, Sage Publications, Beverly Hills, CA.
- Carson, Richard T., Theodore Graham-Tomasl, Charles F. Rund, and William W. Wade, 1986. "Problems with the NPCA Contingent Valuation Survey Instrument and Survey Results: Exhibit 324," Report prepared for Northern States Power by Dames and Moore, San Francisco, CA.
- Chavass, Jean-Paul, Richard C. Bishop, and Kathleen Segerson, 1986. "Ex Ante Consumer Welfare Evaluation in Cost-Benefit Analysis," Journal of Environmental Economics and Management, Vol. 13, No. 3, pp. 255-268.
- Cochran, William G., 1977. Sampling Techniques, 3rd ed., Wiley, New York.

- Converse, Jean M. and Stanley Presser, 1986. Survey Questions: Handcrafting the Standardized Questionnaire, Sage Publications, Beverly Hills, CA.
- Cottrell, Leonard S., Jr. and Sylvia Eberhart, 1948. American Opinion on World Affairs in the Atomic Age, Princeton University Press, Princeton, NJ.
- Cummings, Ronald G., David S. Brookshire, and William D. Schulze (eds.), 1986. Valuing Environmental Goods: A State of the Arts Assessment of the Contingent Method, Rowman and Allanheld, Totowa, NJ.
- Department of the Interior, 1986. "Final Rule for Natural Resource Damage Assessments under the Comprehensive Environmental Response Compensation and Liability Act of 1980 ("CERCLA"), "Federal Register (August 1), Vol. 51, No. 148, pp. 27674-27753.
- Devousges, William H., V. Kerry Smith, Diane H. Brown, and D. Kirk Pate, 1984. "The Role of Focus Groups in Designing a Contingent Valuation Survey to Measure the Benefits of Hazardous Waste Management Regulations," Draft Report to the U.S. Environmental Protection Agency, Research Triangle Institute, Research Triangle Park, NC.
- Dijkstra, Wil and Johannes van der Zouwen, 1982. "Introduction," in Dijkstra, Wil and J. van der Zouwen (ed.), Response Behavior in the Survey-Interview, Academic Press, New York.
- Fee, J., 1979. "Symbols and Attitudes: How People Think About Politics," unpublished Ph.D. dissertation, University of Chicago.
- Fischer, David Hackett, 1970. Historians' Fallacies, Harper & Row, New York.
- Fisher, Ann and Robert Raucher, 1984. "Intrinsic Benefits of Improved Water Quality: Conceptual and Empirical Perspectives," in Advances in Applied Economics, V. Kerry Smith (ed.), JAI Press, Greenwich, CT.
- Fischhoff, Baruch, Paul Slovic, and Sarah Lichtenstein, 1980. "Knowing What You Want: Measuring Labile Values," in Thomas S. Wallsten (ed.), Cognitive Processes in Choice and Decision Behavior, Lawrence Erlbaum Associates, Publishers, Hillsdale, NJ.
- Frankel, Martin, 1983. "Sampling Theory," Chapter 2 in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research, Academic Press, New York.
- Freeman, A. Myrick III, 1979. The Benefits of Environmental Improvement in Theory and Practice, Johns Hopkins University Press for Resources for the Future, Baltimore, MD.
- Freeman, A. Myrick III., 1984. "The Size and Sign of Option Value," Land Economics, Vol. 60, pp. 1-13.

- Chenavat, Pankaj, 1985. "Building Strategy on the Experience Curve," Harvard Business Review, Vol. 85, pp. 143-149.
- Greenley, Douglas A., Richard G. Walsh, and Robert A. Young, 1981. "Option Value: Empirical Evidence from a Case Study of Recreation and Water Quality," Quarterly Journal of Economics, Vol. 96, No. 4, pp. 657-672.
- Greenley, Douglas A., Richard G. Walsh, and Robert A. Young, 1982. Economic Benefits of Improved Water Quality: Public Perceptions of Option and Preservation Values, Westview Press, Boulder, CO.
- Hanemann, W. Michael, 1986. "Willingness to Pay and Willingness to Accept: How Much Can They Differ?" draft manuscript, Department of Agricultural and Resource Economics, University of California, Berkeley.
- Heberlein, Thomas A. and Richard C. Bishop, 1986. "Assessing the Validity of Contingent Valuation: Three Field Experiments," The Science of the Total Environment, Vol. 56, pp. 99-107.
- Hoehn, John P. and Alan Randall. Forthcoming. "A Satisfactory Benefit-Cost Indicator for Contingent Valuation," Journal of Environmental Economics and Management.
- Jabline, Thomas B., Miron L. Straif, Judith M. Tanur, and Roger Tourangeau (eds.), 1984. Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines, National Research Council, National Academy Press, Washington, DC.
- Just, Richard E., Darrell L. Huech, and Andrew Schmitz, 1982. Applied Welfare Economics and Public Policy, Prentice Hall, Englewood Cliffs, NJ.
- Kahneman, Daniel, 1986. "Comments," in Ronald G. Cummings et al., Valuing Environmental Goods, Rowman and Allanheld, Totawa, NJ.
- Kahneman, Daniel, Paul Slovic, and Amos Tversky (eds.), 1982. Judgement Under Uncertainty: Heuristics and Biases, Cambridge University Press, New York.
- Kalton, Graham and Howard Schuman, 1982. "The Effect of the Question on Survey Responses: A Review," The Journal of the Royal Statistical Society, Series A, Vol. 145, pp. 42-73.
- Kirsch, Irwin S. and Ann Jungelut, 1986. Literacy: Profiles of America's Young Adults, National Assessment of Education Progress Report No. 16-PI-02, Educational Testing Service, Princeton, NJ.
- Kish, L., 1965. Survey Sampling, John Wiley, New York.

- Libert, Rensis, 1951. "The Sample Interview Survey as a Tool of Research and Policy Formation," in D. Lerner and D. Laswell (eds.), The Policy Sciences, Stanford University Press, Stanford, CA.
- Mitchell, Robert Cameron, 1979. "Silent Spring/Solid Majorities," Public Opinion, Vol. 2, No. 4, pp. 16-20, 55.
- Mitchell, Robert Cameron, 1980. Public Opinion on Environmental Issues: Results of a National Opinion Survey, President's Council on Environmental Quality, Washington, DC.
- Mitchell, Robert Cameron and Richard T. Carson, 1985. "Comment on Option Value: Empirical Evidence from a Case Study of Recreation and Water Quality," Quarterly Journal of Economics, Vol. 100, No. 1, pp. 291-294.
- Mitchell, Robert Cameron and Richard T. Carson, 1986. "Property Rights, Protest, and the Siting of Hazardous Waste Facilities," American Economic Review, Vol. 76, No. 2, pp. 285-290.
- Mitchell, Robert Cameron and Richard T. Carson, 1987. "Evaluating the Validity of Contingent Valuation Studies," Discussion Paper 92-87-06, Resources for the Future.
- Mitchell, Robert Cameron and Richard T. Carson. Forthcoming. Using Surveys to Value Public Goods: The Contingent Valuation Method, Resources for the Future, Washington, DC.
- Moser, David A., and C. Mark Dunning, 1986. A Guide for Using the Contingent Value Methodology in Recreation Studies, National Economic Development Procedures Manual -- Recreation, Volume II, IWR Report 86-R-5, Institute for Water Resources, Fort Belvoir, VA.
- Misbett, Richard and Lee Ross, 1980. Human Inference: Strategies and Shortcomings of Social Judgement, Prentice-Hall, Englewood Cliffs, NJ.
- Payne, Stanley L., 1951. The Art of Asking Questions, Princeton University Press, Princeton, NJ.
- Plummer, Mark L. and Richard C. Hartman, 1986. "Option Value: A General Approach," Economic Inquiry, Vol. 24, pp. 455-471.
- Presser, Stanley, 1984. "The Use of Survey Data in Basic Research in the Social Sciences," Chapter 3 in Charles F. Turner and Elizabeth Martin (eds.), Surveying Subjective Phenomena, Vol. 2, Russell Sage Foundation, New York.
- Randall, Alan, 1986. "Human Preferences, Economics, and the Preservation of Species," in Bryan G. Norton (ed.), The Preservation of Species, Princeton University Press, Princeton, NJ.

- Randall, Alan, Glenn C. Blomquist, John P. Hoehn, and John R. Stoll, 1985. "National Aggregate Benefits of Air and Water Pollution Control," Interim Report to the U.S. Environmental Protection Agency, Washington, DC.
- Randall, Alan, John P. Hoehn, and David S. Brookshire, 1983. "Contingent Valuation Surveys For Evaluating Environmental Assets," Natural Resources Journal, Vol. 23, pp. 635-648.
- Randall, Alan, Berry C. Ives and Clyde Fastman, 1974. "Ridding Games for Valuation of Aesthetic Environmental Improvements," Journal of Environmental Economics and Management, Vol. 1, pp. 132-149.
- Research Triangle Institute, 1979. "Field Interviewer's General Manual," Research Triangle Institute, Research Triangle Park, NC.
- Riesman, David, 1958. "Some Observations on the Interviewing in the Teacher Apprehension Study," pp. 266-370 in Paul Lazarsfeld and Wagner Thielens, Jr. (eds.), The Academic Mind: Social Scientists in a Time of Crisis, Free Press, Glencoe, IL.
- Roberts, Kenneth J., Mark E. Thompson, and Perry W. Pawlyk, 1985. "Contingent Valuation of Recreational Diving at Petroleum Rigs, Gulf of Mexico," Transactions of the American Fisheries Society, Vol. 114, No. 2, pp. 214-219.
- Rossi, Peter H., James D. Wright, and Andy B. Anderson, 1983. "Sample Surveys: History, Current Practice, and Future Prospects," in Peter H. Rossi, James D. Wright, and Andy B. Anderson (eds.), Handbook of Survey Research, Academic Press, New York.
- Rowe, Robert D., Ralph C. d'Arge, and David S. Brookshire, 1980. "An Experiment on the Economic Value of Visibility," Journal of Environmental Economics and Management, Vol. 7, pp. 1-19.
- Schuman, Howard and Stanley Presser, 1981. Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context, Academic Press, New York.
- Slovic, Paul, Baruch Fischhoff, and Sarah Lichtenstein, 1982. "Response Mode, Framing, and Information-Processing Effects in Risk Assessment," in Robin M. Hogarth (ed.), Question Framing and Response Consistency, Jossey-Bass, San Francisco.
- Smith, V. Kerry, 1984. "A Bound for Option Value," Land Economics, Vol. 60, No. 3, pp. 292-296.
- Smith, V. Kerry, 1986. "Intrinsic Values in Benefit Cost Analysis," draft manuscript, Vanderbilt University, Department of Economics, Nashville, TN.

- Smith, V. Kerry. Forthcoming. "Uncertainty, Benefit-Cost Analysis, and the Treatment of Option Value," Journal of Environmental Economics and Management.
- Smith, V. Kerry and William H. Desvousges, 1987. "An Empirical Analysis of the Economic Value of Risk Changes," Journal of Political Economy, Vol. 95, No. 1, pp. 89-114.
- Spence, A. Michael, 1981. "The Learning Curve and Competition," Bell Journal of Economics, Vol. 12, pp. 49-70.
- Sudman, Seymour, 1976. Applied Sampling, Academic Press, New York.
- Sudman, Seymour and Norman M. Bradburn, 1982. Asking Questions: A Practical Guide to Questionnaire Design, Jossey-Bass, San Francisco.
- Tourangeau, Roger, Kenneth A. Rasinski, Robert P. Abelson, Roy D'Andrade, and Norman Bradburn, 1985. "Cognitive Aspects of Survey Responding: Attitudes and Explanations," Preliminary Report on Pilot Studies, National Opinion Research Center, Chicago.
- Turner, Charles F. and Elizabeth Martin (eds.), 1984. Surveying Subjective Phenomena, Vol. 1, Russell Sage, New York.
- Tversky, Amos and Daniel Kahneman, 1974. "Judgement Under Uncertainty: Heuristics and Biases," Science, Vol. 185, pp. 1124-1131.
- Van der Zouwen, Johannes and Wil Dijkstra, 1982. "Conclusions," in Dijkstra, Wil and J. van der Zouwen (eds.), Response Behavior in the Survey-Interview, Academic Press, New York.
- Walsh, Richard G., Larry D. Sanders, and John B. Loomis, 1985. Wild and Scenic River Economics: Recreation Use and Preservation Values, Report to the American Wilderness Alliance, Department of Agriculture and Natural Resource Economics, Colorado State University.
- Water Resources Council, 1979. "Procedures for Evaluation of National Economic Development (NED): Benefits and Costs in Water Resources Planning (Level C), Final Rule," Federal Register, December 14, 1979, Vol. 44, No. 242, pp. 72,892-72,977.
- Welle, Patrick G., 1985. "Potential Economic Impacts of Acid Rain in Minnesota: The Minnesota Acid Rain Survey," prepared for the Minnesota Pollution Control Agency.
- Wellman, J.D., E.G. Hawk, J.W. Roggenbuck, and G.J. Buhoff, 1980. "Mail Questionnaire Surveys and the Reluctant Respondent: An Empirical Examination of Differences Between Early and Late Respondents," Journal of Leisure Research, Vol. 12, No. 2, pp. 164-172.

Wicksell, K., 1967. "A New Principle of Just Taxation," in R. A. Musgrave and A. T. Peacock, Classics in the Theory of Public Finance, St. Martin's, New York.

Yates, F., 1980. Sampling Methods for Censuses and Surveys, 4th ed., Griffin, London.

Yelle, Louis F., 1979. "The Learning Curve: Historical Review and Comprehensive Survey," Decision Sciences, Vol. 10, pp. 302-328.

ALTERNATIVE DISPUTE RESOLUTION IN WATER RESOURCES MANAGEMENT

C. Mark Dunning*

Introduction

In recent years a variety of techniques have been advocated to resolve natural resource conflicts. These approaches have collectively been termed "Alternative Dispute Resolution" methods, or ADR for short. They include mediation, conciliation, and collaborative problem solving. While there are differences in the approaches they share common characteristics: they are voluntary, they encourage face-to-face interaction of parties involved in the dispute, they seek a consensus among the parties, and they are often assisted by an impartial third party.

A substantial body of academic as well as popularized pieces on ADR have appeared (see, for example, Fisher and Ury, 1981; Moore, 1986; Bacow and Wheeler, 1984; Wehr, 1979; Susskind and Weinstein, 1980). ADR approaches have been portrayed as offering the ability to generate "quality decisions" -- that is, decisions which are fair to all, are durable, take community interests into account, are efficient and improve or at least do not damage the relationships among parties (Fisher and Ury, 1981:4). The implicit, or sometimes explicit, message is that ADR methods are superior to other approaches for dealing with conflict.

ADR approaches have been used in water resources management issues. A recent study has recorded 17 instances of ADR approaches in water resources conflicts from the early 1970s through 1984 (Bingham, 1986:xviii). ADR approaches are actively being promoted within water resources agencies in training courses (Corps of Engineers, 1986) and have also been endorsed by executives within agencies.

1. "I am personally committed to ADR, since I am convinced that for many problems it can provide a less expensive, less resource intensive alternative to costly, time consuming litigation....ADR is a proactive alternative to the defensive posture often imposed on the Government by litigation. It is an option which can produce better management decisions and policies, and increase public trust and faith in the overall system of Government." Remarks by Mr. Les Edelman, Chief Counsel of Corps of Engineers at Second Worldwide USACE Legal Services Conference, 16-21 November 1986, San Antonio, TX.

* Sociologist, Research Division, Corps of Engineers, Institute for Water Resources, Fort Belvoir, Virginia. The views presented in this paper are the author's and do not necessarily reflect those of the U.S. Army Corps of Engineers.