Navigating the Confluence of Two Streams of Social Research: Contingent Valuation and Normative Standards

Robert Manning, Steven Lawson and Lesley Frymier
School of Natural Resources, University of Vermont
Burlington, VT 05405
USA

Abstract

Interdisciplinary research on contingent valuation and normative standards is reviewed and integrated. The objectives of the study are to identify issues of concern to both areas of research and to describe findings from each area of research that might inform the other. Seven theoretical and methodological issues are identified and reviewed. Based on this analysis a series of conditions are described in which contingent valuation and normative standards research are most applicable and effective, and recommendations for future research are developed.

Keywords: contingent valuation, normative standards, literature review, theoretical issues, methodological issues, interdisciplinary research

Introduction

Contingent valuation (CV) and normative standards (NS) represent two streams of social research. CV has become a common approach to measuring the economic value of public goods, such as parks and aesthetic beauty. NS is becoming an increasingly common approach to measuring acceptable conditions in parks and related recreation areas. These two streams of social research have been developed independently — CV research has evolved primarily out of the discipline of economics, while NS research has evolved from the disciplines of sociology and social psychology - but share several theoretical and methodological issues. This paper reviews and synthesizes CV and NS research. The objective of the paper is to identify findings from CV research that may inform NS research and vice-versa. In this way, we hope to contribute to the advancement of more efficient and effective methods of CV and NS research.

CV Research

Economists traditionally rely on observation of the actions of buyers and sellers in a market in order to identify the values individuals place on goods and services. CV was developed as a means of estimating the economic value of goods for which no explicit market exists. CV draws upon consumer theory and the methods of survey research to elicit the values individuals place upon these “non-market” goods (Mitchell and Carson 1989).

Surveys are used by CV researchers to present respondents with a scenario describing a baseline and a hypothetical alternative level of provision of a non-market good or resource. Respondents are asked to state their maximum willingness-to-pay (WTP) or minimum willingness-to-accept compensation (WTA) for the hypothetical change in the quality or quantity of the good described in the scenario. Individual WTP or WTA values are averaged for the sample and then aggregated over the relevant population to estimate the total economic benefits associated with the scenario. Other questions typically included in a CV survey ask respondents about their socioeconomic characteristics (e.g., income, education, gender) and their use of the resource (Mitchell and Carson 1989).

CV was introduced in the 1960s by Robert K. Davis, who used questionnaires to estimate the benefits of outdoor recreation in a Maine backcountry area (Mitchell and Carson, 1989). Since the 1970s, CV has been used by economists to measure the benefits of a wide variety of non-market goods, including outdoor recreation (Walsh, Miller and Gilliam 1983), reductions in morbidity and mortality risk (Rowe and Chestnut 1984; Tolley and Babcock 1986; Brajer, Hall and Rowe 1991), congestion in a wilderness setting (Walsh and Gilliam 1982; Walsh et al. 1983), wildlife populations (Cocheba and Langford 1978), water quality (Gramlich 1977; Mitchell and Carson 1981, 1984; Carson and Mitchell 1993b), and visibility (Rowe, d’Arge and Brookshire 1980; Schulze, Cummings, Brookshire, Thayer, Whitworth and Rahmatian 1983).

Many economists have debated the validity of economic value estimates generated using the CV method. Two developments in environmental regulation have heightened the significance of this debate. In 1986, the Department of the Interior declared that passive or non-use values should be included among the losses parties are responsible for under the Comprehensive Environmental Response, Compensation,
and Liability Act (CERCLA). Non-use values are the benefits individuals derive from a resource without actively using it. Because non-use values are derived from non-market behavior, they cannot be estimated directly from the transactions of buyers and sellers in a market (Arrow, Solow, Portney, Learner, Radner and Schuman 1993).

The second development took place in 1990, when Congress enacted the Oil Pollution Act of 1990 (OPA). The OPA requires responsible parties to provide compensation for damages caused by oil spilled into waters or on shorelines under the jurisdiction of the United States, including non-use values.

In order to include non-use values, the implementation of both CERCLA and OPA required a method to assess the magnitude of natural resource damages not captured from observations of market transactions. The National Oceanic and Atmospheric Administration (NOAA) commissioned a panel of economic experts, co-chaired by Nobel laureates Kenneth Arrow and Robert Solow, to evaluate the use of CV in making natural resource damage assessments including non-use values. The panel concluded that the CV method could be used for such purposes, subject to numerous conditions regarding the design and administration of the survey instrument.

A substantial body of CV research has focused on the survey design and administration issues inherent in the CV method. Some of the topics CVM research has focused on include minimizing the potential for bias in responses (Sutherland and Walsh 1985; Brown and Duffield 1995; Loomis 1996), selecting the appropriate format of the elicitation questions (Kealy and Turner 1993; Boyle and Bishop 1988), and conducting statistical analysis of CV data (Lindsey 1994).

NS Research

Developed in the fields of sociology and social psychology, NS have attracted considerable attention as an organizing framework in recreation research and management. As applied to parks and related areas, norms are generally defined as standards that individuals and groups use for evaluating behavior and social and environmental conditions (Donnelly, Vaske and Shelby 1992; Shelby and Vaske 1991; Vaske, Graefe, Shelby and Heberlein 1986). If visitors have norms concerning relevant aspects of recreation experiences, then such norms can be studied and used to guide management of parks and related areas.

Application of norms in outdoor recreation is most fully described by Shelby and Heberlein (1986) and Vaske et al. (1986). These applications rely heavily upon the work of Jackson (1965), who developed a methodology — return potential curves — to measure norms. Typically, respondents are asked to evaluate the acceptability of a range of social and/or resource conditions. Data on the personal norms of individuals can then be aggregated to test for the existence of social norms, or the degree to which norms are shared across groups.

NS research in outdoor recreation has focused largely on the issue of crowding (e.g., Heberlein, Alfano, and Ervin 1986; Patterson and Hammitt 1990; Shelby 1981a; Vaske et al. 1986; Whittaker and Shelby 1988; Williams, Roggenbuck, and Bange 1991), but also has been expanded to include other social and ecological variables, including ecological impacts at wilderness campsites (Shelby, Vaske and Harris 1988), wildlife management practices (Vaske and Donnelly 1988), and minimum stream flows (Shelby and Whittaker 1995). As research on NS in recreation has matured, attention has turned to a variety of conceptual and methodological issues, including the theoretical basis of norms and their application to outdoor recreation (Heywood 1996a, 1996b; Noe 1992; Roggenbuck, Williams, Bange and Dean 1991; Williams et al. 1991; Shelby and Vaske, 1991; Manning, Lime and Hof 1996b); alternative evaluative measures (Manning, Valliere and Jacobi 1997); visual representation of social and environmental conditions (Manning, Lime, Hof and Freimund 1995a; Hof, Hammitt, Rees, Belnap, Poe, Lime and Manning 1994; Manning, Lime and McMonagle 1995b; Manning et al. 1996b; Manning, Lime, Freimund and Pitt 1996a); congruence between norms and behavior (Hammitt and Ruttim 1995; Patterson and Hammitt 1990; Lewis, Lime and Anderson 1996; Manning et al. 1996a; Manning et al. 1996b; Vaske et al. 1986); and question formats designed to reduce respondent burden (Shelby 1981b; Vaske et al. 1986; Manning et al. 1997).

Common Issues

From a broad conceptual standpoint, both CV and NS research are concerned with uncovering information about people's preferences with respect to the provision of various resources (e.g., the economic value of visiting a national park, or the acceptable number of other park visitors). Moreover, from a similarly broad methodological standpoint, both areas of study rely primarily on survey research (i.e., they elicit from respondents a dollar value of visiting a park or the acceptability of encountering selected numbers of other park visitors). The broad similarities between CV and NS are evident upon review of the literature in these two areas of research. The theoretical foundations underlying each area of research, and their practical application are confronted by similar issues and challenges. Issues common to both areas of research are illustrated in Table 1 and are described below.
on voting behavior. An individual’s vote is considered to reveal their true preferences despite the fact that voters often make spontaneous decisions and respond to some relatively unfamiliar items on the ballot (Arrow et al. 1993). Limited research also suggests that most respondents are reasonably confident of the validity of their answers. Schkade and Payne (1994), for example, conducted a verbal protocol analysis, allowing respondents to self-assess how they answered a series of WTP questions. While 20 percent of respondents reported they merely guessed, most felt that they had carefully weighed the value of the resource under study and that they had considered their personal budget constraints before answering.

Familiarity with the Resource

Critics of the CV method also argue that “respondents do not understand what it is they are being asked to value” (Arrow et al. 1993, 4603). According to critics, the WTP responses provided by individuals who do not understand or are unfamiliar with the good being valued are not meaningful (Arrow et al. 1993).

Several CV studies have been conducted to investigate the influence of respondents’ familiarity with the good being valued on WTP estimates. Lindsay, Halstead, Tupper and Vaske (1992) found WTP to protect coastal beaches from erosion was higher for individuals who were familiar with coastal laws. In addition, Kealy and Turner (1993) found that WTP estimates are more robust for a familiar private good (chocolate candy) than a less familiar public good (aquatic ecosystem protection) when several question formats are used in the CV survey. On the other hand, Boyle, Welsh and Bishop (1993) found no significant difference between experienced and inexperienced whitewater boaters with respect to their WTP for different flow levels on the Colorado River. Mitchell and Carson (1989) conclude that when confronted by an unfamiliar, hypothetical situation, respondents tend to construct meaning based on previous experience and to arrive at an opinion that does reflect their true tastes and preferences.

Suggestions for Implementation

CV researchers have made efforts to increase the reliability of CV responses by attempting to create realistic decision frameworks that respondents are familiar with. For example, Walsh, Loomis and Gillman (1984) asked Colorado residents to indicate how much they would be willing to pay into a special fund used exclusively for protecting wilderness. The fund was selected as the payment vehicle because of its similarity to the nongame wildlife preservation fund Colorado residents can elect to contribute to on their state income tax return.

Some authors have suggested that, given uncertainty over the ability of respondents to provide meaningful responses to CV questions, researchers should include a “no vote” response option. Providing a “no vote” option would allow respondents who do not have well-informed responses to so indicate (Arrow et al. 1993: Fischhoff 1991).

The NOAA panel indicated that a CV survey must include an accurate and thorough description of the good or service being valued. In addition, the NOAA panel recommended the use of pre-tests and focus groups to determine whether the information provided to respondents is adequate for answering the CV question. CV studies that follow these recommendations are likely to enhance the reliability of WTP responses.

Existence of Norms

NS researchers have begun to address the issue of whether survey questions about normative standards can be answered meaningfully. First, the theoretical foundations of norms have been reexamined (Heywood 1996a, 1996b; Noe 1992; Roggenbuck et al. 1991; Williams et al. 1991; Shelby and Vaske 1991; Manning et al. 1996a). As noted earlier, normative theory has been borrowed from the disciplines of sociology and social psychology. Within these disciplines, norms are characterized by several distinguishing features, including the fact that they are obligatory, they are enforced by sanctions, they guide behavior, and they are shared by social groups. Application of normative theory to NS research in outdoor recreation has adopted a more expansive view of norms suggesting that (1) recreation often involves emerging norms for which a strong sense of obligation and sanction has yet to fully evolve; (2) recreation-related norms can apply to social and resource conditions as well as behavior because such conditions are often a function of individual behavior; (3) recreation-related norms often regulate collective rather than individual behavior, and (4) research has documented some degree of consensus regarding a number of recreation-related norms (Shelby and Vaske 1991).

Second, empirical findings of NS studies are suggestive of the extent to which norms may exist. Most studies have found that visitors to recreation areas are able to respond to NS questions and that, as noted earlier, these studies have addressed a variety of social and resource conditions. The extent to which there is agreement or consensus about such norms is less certain (see, for example, Roggenbuck et al. 1991; Shelby and Vaske 1991; Williams et al. 1991). While there are a number of ways to measure consensus, there is no broad agreement about the degree of consensus needed to establish normative standards. Moreover, the degree of consensus is affected by a number of intervening variables.
These issues are considered in greater depth later in this paper.

"No Vote" Option

Third, as in CV research, NS researchers have considered the advisability of incorporating a "no vote" option for respondents (Roggenbuck et al. 1991; Hall and Shelby 1996). This option would allow respondents to indicate that the impact under consideration is important, but that they can't specify a maximum amount of impact acceptable. This would allow respondents who do not have well-informed opinions to so indicate. Initial studies suggest that a minority of respondents will select this option (Hall and Shelby 1996). However, the advisability of this strategy is not universally acknowledged. One study suggests that respondents who choose this option are similar to those who report a norm, and that this response option may simply be an easy way for some respondents to avoid a potentially burdensome question (Hall and Shelby 1996).

Finally, in keeping with the theoretical foundations of norms, some researchers have suggested that NS questions be formulated with a more direct emphasis on notions such as obligation (e.g., the maximum amount of impact that should be allowed), internal or informal sanctions (e.g., a sense of personal responsibility), and the extent to which they are thought to be shared by others (e.g., what "others" are believed to think) (Heywood 1996a; Roggenbuck et al. 1991).

Influence of Information on Responses

In both CV and NS surveys a tension exists between providing adequate information and overloading respondents with such information. Past research suggests the amount and type of information provided can influence responses obtained in both CV and NS research. In the CV literature, this issue is often termed "information bias." Varying amounts of information on the ecological and social services of wetlands were found to influence WTP for wetland protection; the more services described, the higher the WTP estimates (Bergstrom, Stoll and Randall 1990). Aszen, Brown and Rosenthal (1996) concluded that the nature of the information provided can profoundly affect WTP estimates, and that subtle contextual cues can seriously bias these estimates, especially when the good being valued has low personal relevance to respondents. Information on the WTP of other respondents has been found to influence individual WTP responses (Rowe et al. 1980). Early in the development of CV, it was assumed that only the nature of the good and the amount of the amenity being valued should influence WTP; all other information (such as the payment vehicle) should be neutral (Rowe et al. 1980). But Arrow (1986), Kahneman (1986) and Randall (1986) have argued that important conditions of the scenario should be expected to affect WTP amounts. In their view, respondents are valuing a policy that includes the amenity under certain conditions (Mitchell and Carson 1989).

The role of information on NS has been explored in the context of examining narrative versus visual approaches to resource description. Traditionally, the resource under study is described to the respondents in a brief narrative. For example, respondents may be asked to consider a situation in which they are hiking a wilderness trail and encounter five other groups at a scenic attraction. Alternatively, respondents could be presented with a picture or visual simulation of the situation (Hof et al. 1994; Manning, Lime, Hof and Freimund 1995a; Manning et al. 1995b; Manning et al. 1996a; Manning et al. 1996b). Initial research on this issue suggests that visual presentations of normative scenarios may result in higher crowding norms (Manning et al. 1996a). Respondents may cognitively "process" some people in the visual representation at a subconscious level because they are perceived to be "like" the respondent and therefore do not substantially contribute to perceived crowding. In contrast, narrative descriptions call explicit attention to all people "encountered."

Inherent Biases

As with all applications of survey research, CV and NS studies are subject to numerous forms of potential bias. "Hypothetical bias" concerns the degree to which responses to survey questions may not reflect actual behavior. CV researchers have explored this issue both theoretically and empirically. In its most fundamental form, hypothetical bias is a manifestation of the sometimes weak and inconsistent relationships documented between verbal measures of attitude and observation of actual behavior. This suggests that hypothetical bias may be a substantial problem in WTP research (Fishbein and Ajzen 1975). Mitchell and Carson (1989) argue that the key problem in CV is the novelty of valuing a public good — that respondents have a varying degree of familiarity with the good and how they currently pay for its provision. In an empirical test for hypothetical bias, Neill, Cummings, Ganderton, Harrison, and McGuckin (1994) found hypothetical values of respondents to be significantly higher than real economic commitments.

There is a significant amount of research that suggests the relationship between attitudes and behavior is not weak and inconsistent. In fact a number of studies argue that there is a positive relationship between attitudes and behavior (Cannary and Seibold 1983; Schuman and Johnson 1976; Stouffer and Lumsdaile 1949; Brannon, Cyphers, Hess,
Hesselbarb, Keane, Schuman, Vaccaro and Wright 1973; Kelly and Miner 1974; Vinokur-Kaplan 1978). Additionally, Bishop and Heberlein (1986) argue that attitudes encompass intended behavior. Perhaps the findings of these studies explain why, in a summary of studies comparing hypothetical CV markets and markets simulated by using real money, Mitchell and Carson (1989) found that various validity assessments are generally favorable to CV's potential for measuring valid WTP amounts.

NS researchers have recently begun to address the potential for hypothetical bias through study of "norm congruence," or the degree to which respondent behavior and/or evaluation of conditions corresponds to previously reported normative standards. While study findings are not uniform, they generally suggest that hypothetical bias is not a substantial problem (Hammit and Ruthn 1995; Lewis et al. 1996; Manning et al. 1996a, 1996b; Williams et al. 1991; Patterson and Hammit 1990). For example, respondents who report encountering more groups of hikers than their personal norm tend to report higher levels of perceived crowding than do respondents who report seeing fewer groups of hikers than their personal norms. In addition, those respondents who report encountering more groups of hikers than their personal norm are more likely to report adopting some action to avoid such encounters.

Several other sources of bias may affect results of CV and NS research, though these issues are addressed exclusively in the CV literature. "Social desirability bias," or "compliance bias" concerns the degree to which respondents might be influenced by perceived social norms or cues from the context or administration of the survey (Snyder and Swann 1976; Harris, Driver and McLaughlin 1989). For example, a questionnaire addressing the value of free-flowing rivers may implicitly or explicitly imply that free-flowing rivers are important, thereby influencing respondents. Or interviewers may unknowingly impart subtle cues to respondents.

"Strategic bias" may occur if respondents wish to intentionally influence study findings in one direction or another (Fishoff, 1991). For example, respondents may deliberately understate their WTP if they feel they may be asked to actually pay for a resource based on study findings. Or they may overstate their WTP if they want more of the resource, but feel it is unlikely they will be asked to actually pay.

"Starting point bias" is a more technical, methodological issue concerning the degree to which an initial value proposed in a study may ultimately influence WTP. Starting point bias is an issue in the case of CV studies that use multiple-bound dichotomous choice questions. The respondent is asked to indicate with a "yes" or "no" response, whether they would be willing to pay a proposed dollar amount for the resource. The next question poses a higher dollar value for those individuals who selected a "yes" response to the previous question, and a lower dollar value to those individuals who responded "no." Starting point bias is a concern if the initial dollar amount proposed effects the magnitude of WTP estimated in the study (e.g., larger starting bids result in larger WTP values than lower starting bids). Several studies have explored this issue, but findings are inconclusive (Rowe et al. 1980; Desvousges, Smith and McGivney 1983; Thayer 1981).

A final type of potential bias in CV studies is termed "embedding bias." Some economists argue that the consistency of economic theory, CV responses should vary depending upon the magnitude of the environmental amenity. Boyle, Desvousges, Johnson, Dunford and Hudson (1994) found no significant differences in WTP of nonusers to prevent 2000, 20,000, or 200,000 migratory waterfowl deaths in the Central Flyway of the United States. Kahneman and Knetsch (1992) presented empirical evidence that values derived from CV for any one public good may be somewhat arbitrary because value estimates may vary depending on whether the good is valued by itself or as part of some broader package. Loomis, Lockwood and DeLacy (1993) tested for embedding effects in contingent valuation of forest protection and concluded that while there was some evidence of embedding bias, the effects were less when respondents were clearly informed of the regional context of the good being valued. Carson and Mitchell (1993a, 1995), however, based on evidence from the CV literature, and empirical evidence from their own study, concluded that respondents are sensitive to the scope of the good being valued.

Mitchell and Carson (1989) cite several studies that identify a similar issue to embedding bias called the "adding up effect" — that WTP values for two goods elicited individually might be significantly different than the WTP for the two goods valued together. Mitchell and Carson (1989) argue that individually measured WTP cannot be added without some over-counting effect. Instead, respondents tend to value each good sequentially as if it were a small, incremental (marginal) addition to the existing set of environmental amenities they enjoy, instead of valuing each good separately. Mitchell and Carson (1989) argue that this method of valuing additional units of a good is consistent with economic theory, but Randall and Hoehn (1996) state that although economic theory anticipates embedding, it is possible that CV exacerbates and even amplifies these effects. The adding up effect remains a matter of contention in contemporary CV research.

A final issue that has been investigated in the CV literature is whether respondents answer all WTP questions with the same response. Economic theory suggests that WTP
responses should differ significantly when respondents are faced with WTP choices involving different objects (Smith 1996). Smith (1996) evaluated two WTP questions - one to expand a popular flower planting program and a second to facilitate the use of recycled tires in making asphalt for highways. Smith found significant differences between estimated WTP for the two plans, suggesting that respondents can discriminate between questions for different goods when answering CV survey questions.

Influence of Question Format on Responses

CV and NS research rely on survey methods. Thus, issues of question format are pertinent. A principal issue concerns attempts to reduce respondent burden: how can questions be asked so they are easier or less time-consuming to answer? The choice of question format has a significant influence on the complexity or amount of time involved in responding to survey questions. Question formats, or "elicitation techniques", as they are often referred to, have evolved as researchers learn more about the strengths and weaknesses of such techniques. Mitchell and Carson (1989) identify nine elicitation methods, categorized by whether actual maximum WTP is obtained and whether a single WTP question or a series of questions is asked.

The most widely used elicitation methods have been open-ended (direct question), and dichotomous choice (yes or no). In open-ended questions, respondents are asked to state a maximum dollar amount they would be willing to pay for the hypothetical scenario. In dichotomous choice (closed-ended) questions, respondents are told how much each individual would have to pay if the hypothetical scenario is adopted and then asked to vote "yes" or "no". The dichotomous choice method is so named because only two responses ("yes" and "no") are available (Arrow et al. 1993).

Several studies have examined the WTP values obtained using both open-ended and dichotomous choice formats (e.g., Kealy and Turner 1993; Loomis 1990; Loomis, Brown, Lucero and Peterson 1997; Randall, Hoehn and Brookshire 1983). While there is no clear consensus among researchers about which question format is more valid, several studies suggest that close-ended questions may yield higher WTP values than open-ended questions.

Both the open-ended and dichotomous choice question formats have advantages and disadvantages (Loomis 1990). Using open-ended questions is a more direct measure of maximum WTP. However, the open-ended format is more burdensome to respondents as it requires them to offer a dollar amount with little or no assistance. As a consequence, there tends to be an unacceptably large number of nonresponses and zero bids from individuals with actual WTP values greater than zero (Desvousges, Smith and Mcgivern 1983).

The dichotomous choice approach has the disadvantage of requiring more sophisticated statistical techniques in order to derive WTP indirectly from responses to the CV question (Boyle and Bishop 1988). In addition, responses to dichotomous choice questions only provide a bound on individuals' actual WTP, and therefore a large sample size is needed to get an acceptable level of statistical precision. However, efficiency can be increased if the dichotomous choice question is followed up by further single dichotomous choice questions with the dollar values being revised upward or downward depending on the respondent's initial answer (Carson, Hanneman and Mitchell 1986).

The NOAA panel report provides a strong recommendation in favor of the dichotomous choice format of the CV question (Arrow et al. 1993). The first argument the NOAA panel presents against the use of open-ended CV questions is related to the "familiarity" issue discussed in the previous section of this paper. According to the NOAA panel, the open-ended format "lacks realism since respondents are rarely asked — to place a dollar value on a particular good" (Arrow et al. 1993, 4606). Because respondents are unfamiliar with this task of pricing goods, the reliability of their responses is questionable. In addition, the open-ended format provides respondents with an opportunity to overstate or understate their true WTP in an attempt to influence the outcome. The dichotomous choice format, on the other hand, provides a more familiar and realistic situation for respondents since — "refers to the provision of public goods are not uncommon in real life" (Arrow et al. 1993, 4606). In addition, the close-ended format eliminates the opportunity for strategic bidding.

An analogous issue in the NS literature concerns "long" — sometimes called the "repetitive item format" (Shelby 1981b; Vaske, Graefe and Dempster 1982) — versus "short" question formats. Early applications of NS research employed the long question format by asking respondents to evaluate a range of resource or social conditions. For example, studies of crowding norms might have asked respondents to evaluate the acceptability of seeing 0, 5, 10, 15, and 20 other groups while hiking a trail. To reduce respondent burden, it has become common to employ a short, open-ended version of NS questions where respondents are asked to simply state the maximum amount of impact (or number of other groups, in the above example) they feel is acceptable. Only one study has explored the comparability of these two question formats (Manning et al. 1997). The long or close-ended question format was found to yield somewhat higher norms than the short or open-ended question format.

Human Ecology Review, Vol. 6, No. 1, 1999
Effect of Evaluative Measure on Validity of Responses

In CV and NS studies respondents are asked to make evaluative judgements. For example, respondents to a CV survey may be asked to evaluate the economic value of a day hike, while respondents to an NS survey may be asked to evaluate the acceptability of different levels of crowding along a hiking trail. However, alternative evaluative measures are possible in both areas of study, and may lead to conflicting estimates of value or normative standards. In CV research, WTP and WTA are the two most commonly used evaluative measures of economic welfare. Economic theory provides guidelines about which measure is most appropriate for a particular scenario. The guidelines are based on assumptions about the study population’s property rights with respect to the good being valued. For example, a CV survey may present a hypothetical scenario involving a program to reduce the concentration of contaminants in a local river. If the researcher assigns the study population property rights to the status quo condition of the river, then the correct measure of economic welfare is the maximum amount of money individuals are willing to pay for the improved river condition. However, if the researcher assigns the population property rights to the improved condition of the river, then the correct measure of welfare associated with the program is the minimum amount of money individuals are willing to accept as compensation to forgo the river improvement program (Freeman 1993).

According to economic theory, the two measures of welfare, WTP and WTA, should be similar when empirically estimated for the same good. However, WTA estimates are typically three to five times greater than WTP values in CV studies eliciting both measures (Adamowicz, Bhardwaj, and Macnab 1993). A number of theories have been proposed to explain the divergence between WTP and WTA estimates. According to the "endowment effect" explanation, individual behavior is driven in part by an aversion to loss. Individuals will act on their loss aversion by demanding more in compensation for the removal of a good from their possession than they would pay to obtain the good in the first place.

A second explanation for the tendency of WTA estimates to exceed WTP estimates, the "substitution effect," was first introduced by Hanneman (1991). Hanneman suggested that the magnitude of the divergence between WTA and WTP is dependent on the number and quality of substitutes available for the good in question. According to Hanneman, the fewer alternatives available to the individual, the greater the divergence between WTA and WTP. Hanneman’s theory is based on the notion that if consumers are able to get the same or similar services from an available private good, there will not be much difference between WTP and WTA for a public good. However, if the public good in question has few substitutes (e.g., Yosemite National Park), "WTP could equal the individual’s entire (finite) income, while WTA could be infinite" (Hanneman 1991, 635-636).

Boyce, Brown, McClelland, Peterson and Schulze (1992) suggest another explanation for WTA estimates exceeding WTP. They assert that there is greater moral responsibility involved in not giving up certain public goods than in paying to obtain them. For instance, individuals may feel a greater moral obligation to refuse compensation in exchange for allowing water quality to deteriorate than they do to pay to enhance water quality.

Substantial research has been conducted to examine the validity of the above explanations for discrepancies between WTP and WTA values (e.g., Morrison 1997; Adamowicz et al. 1993; Boyce et al. 1992), however, the findings have been inconclusive. Thus far, CV researchers have responded to the divergence between WTP and WTA by relying primarily on the more conservative WTP measure to calculate welfare estimates. The NOAA panel report included the use of the WTP format as one of their guidelines for designing a CV survey. In addition to providing more conservative welfare estimates, the WTP format is favored because it provides a more realistic and familiar decision framework for survey respondents (Arrow et al. 1993).

NS studies have only recently begun to examine the validity and appropriateness of alternative evaluative measures. The traditional evaluative measure in NS research has been "acceptability." However, the use of other evaluative measures are possible, including preference, tolerance, and more purely normative notions of what conditions should be. Initial tests of measuring preference, tolerance, and attitudes about what should be suggest that they result in estimates of normative standards that differ significantly from those measured by asking about acceptability (Manning et al. 1997; Shelby and Whittaker 1995). However, there is yet no basis on which to favor one evaluative measure over another; each may offer different insights to managers and policy makers.

Effect of Making Response Implications Salient

The validity of responses to CV and NS questions is based on the assumption that respondents are at least generally aware of the personal and/or management implications of the information they provide in such studies. However, this assumption is generally untested. In CV research, the NOAA panel specifically recommended that the implications of any response should be described in detail. More specifically, the CV scenario description or the CV question should
include a reminder to respondents that they have a limited budget and that electing to spend money for the public good means a reduction in other kinds of goods that can be purchased (Arrow et al. 1993). This recommendation is in keeping with economic theory, which suggests that the availability of substitutes for a good and the individual’s income are important determinants of an individual’s willingness to pay for the good.

Several studies have examined the effect of including reminders about the availability of substitutes and individuals’ budget constraints on CV estimates of WTP. Loomis, Gonzalez-Caban and Gregory (1994) administered two versions of a CV survey, one that included a reminder about substitute goods and budget limitations, and one that did not include these reminders. The researchers found no significant difference between mean WTP to prevent old-growth forest fires estimated from the two survey instruments. However, several studies have found that WTP estimates are significantly lower when respondents are reminded of the availability of substitute goods (e.g., Whitehead and Blomquist 1991; Cummings, Ganderton and McGuckin 1994).

NS research has only recently begun to focus on ensuring that respondents consider the management implications of their responses to NS questions. In NS research, reminders about the management implications of responses have been primarily based on previous research findings. For example, if studies of crowding norms at parks and recreation areas indicate little tolerance for high use levels, then it is likely that management actions will be taken to limit public access, and this will affect a respondents’ ability to use these areas. Initial research suggests that explicit knowledge of such implications can have a significant affect on the crowding norms reported (Manning et al. 1997).

How Scale of Analysis Affects Results

Both CV and NS research use surveys to uncover information about individuals’ tastes and preferences, with respect to public goods. The individual values or standards are then aggregated by multiplying the average response by the total number of individuals in the study population.

Several aspects of CV research confound the process of moving from individual responses to total social values. One challenge CV researchers face is determining the extent of the relevant study population for the non-market good they wish to analyze. Frequently, political boundaries are used to define the extent of the market or study population for CV studies (e.g., Desvousges et al. 1983; Lindsey 1994; Loomis 1987). However, the CV researcher risks underestimating the true value of a good if the market boundaries selected are too restrictive, and risks overstating economic benefits if the population over which values are aggregated includes distant individuals who are indifferent to the good (Loomis 1996).

Sutherland and Walsh (1985) estimated a relationship between WTP to protect water quality in a Montana lake and distance of the individuals’ households from the lake. The relationship between WTP and distance was used to estimate total WTP for Montana and for a larger market including 7 states and 3 Canadian provinces. The results indicate that aggregate WTP values can be highly sensitive to the definition of the market for the good (Sutherland and Walsh 1985, 209). Loomis (1996) found that when the extent of the market for public goods with national significance (e.g., salmon in the Pacific Northwest or dam removal in national parks) is limited to the local area rather than including the entire nation, estimates will measure as little as 3 percent of the actual economic benefits.

The decisions CV researchers make about what measure of average individual WTP to use to extrapolate to the total population can also significantly influence welfare estimates. Several issues must be considered before generalizing individual values to the population. First a decision must be made whether to use the mean or median of individuals’ WTP to multiply by the total number of individuals in the population. The mean individual WTP is usually used in CV studies, however, a small number of unusually high values can significantly influence the mean WTP of the good in question. The alternative is to use the median WTP, which tends to limit the influence of a few outlying bids, and is likely to yield more conservative WTP estimates (Lindsey 1994).

Secondly, the CV researcher has to develop a decision rule for handling protest zero bids and outliers. An individual’s response of zero to a CV question is defined as a protest zero bid when the respondent’s true WTP is greater than zero (Lindsey 1994). Outliers are responses to the CV question that seem unlikely to be true given their magnitude. If protest bids and outliers are included in calculating average individual WTP, the results will provide inaccurate information about the true economic value of the good in question.

In order to determine whether an individual who provides a zero bid in a CV survey is providing a true value or a protest bid, the NOAA panel recommended that CV questions be followed up with a question asking respondents to explain their answer (Arrow et al. 1993). Explanations of zero bids that are typically classified as protest bids include “object to the question,” “opposed to new taxes,” “not enough information,” and “industry should pay.”

A variety of decision rules have been used in the CV literature to eliminate outliers. For example, researchers have eliminated bids greater than a selected number of standard deviations from the mean or bids greater than a selected per-
centage of the individual’s income, or by using statistical
diagnostics to isolate extreme bids (Lindsey 1994, 125).

NS research tends to favor median values over means.
However, in NS research there is more concern over the issue
of central tendency because social norms are defined as some
level of agreement about appropriate conditions or behaviors.
The extent of agreement is technically referred to as “crysta-
лизация” and is analyzed through several measures of central
tendency, including standard deviations, coefficients of vari-
ation, semi-interquartile ranges, and inspection of frequency
distributions.

In NS research, personal norms are aggregated to derive
broader social norms for larger populations. However, recent
research suggests that a more direct societal unit of analysis,
or at least a stronger societal context, may have some applic-
ability to both CV and NS research. That is, researchers may
wish to allow or even encourage respondents to consider
broad societal perspectives and obligations when formulating
answers to CV and NS questions. As noted earlier, this is
especially important in keeping with the underlying principles of norm-
ative theory, which suggest widely understood and accept-
ced social obligation. In this context, the operative question
in NS research shifts from what is personally acceptable, prefer-
able, tolerable, etc., to what conditions should be maintained
by a broader society. In CV research, the question shifts from
what would the individual be willing to pay to what society
should pay.

Unlike CV research, NS research often aggregates indi-
vidual responses into a graphic form of reporting. “Norm
curves” (or “impact acceptability curves”) are graphs which
illustrate the relationship between a range of resource condi-
tions and aggregate evaluative ratings of these conditions. An
example of a norm curve is shown in Figure 1. While study
data are usually reported in tabular format as well, these
graphs provide an alternative way of reporting study findings
that adds important insights into research results. For exam-
ple, a social norm curve can illustrate the range of acceptable
conditions, the preferred or optimal condition, the minimum
acceptable condition, and “norm intensity” (a measure of the
relevance or salience of the issue under study). The range of
acceptable conditions includes all impact conditions which
receive aggregate evaluative ratings above the neutral line on
the evaluative scale of the graph. The preferred or optimum
condition is the impact condition that receives the highest
aggregate evaluative rating. The minimum acceptable condi-
tion is the impact condition denoted by the point at which the
norm curve crosses the neutral point on the evaluative scale.
Norm intensity is the distance from the norm curve to the
neutral line. Norm curves can also readily indicate when
there may be bi- or multi-model norms in which case com-
monly used measures of central tendency such as means or
medians may be misleading (Shelby and Whittaker 1995;
Whittaker and Shelby 1988).

Conclusions and Recommendations

The above discussion indicates that there are a number
of theoretical and methodological issues common to CV and
NS research. In some cases, both areas of research may have
addressed the same (or a very similar) issue, and research
findings can be mutually reinforcing. In other cases, research
in one area may suggest and address an issue that is applica-
table to the other area. We hope the comparative interdisci-
plinary review and synthesis of both bodies of research liter-
ature provided here is effective at informing theoretical and
methodological concerns related to CV and NS.

Based on the literature review and synthesis, we make
the following recommendations:

1. Both CV and NS research should further develop and
adopt tests of validity. In particular, self-assessment proto-
cols developed in CV research should be further refined and
applied and should be adapted for use in NS research.
Moreover, “no vote”?“don’t know”?“don’t care” response
options should be further tested to determine their effect on
CV and NS values.

2. Both CV and NS research should further explore the
issue of hypothetical bias. In both cases, this research should
examine the degree to which WTP and NS values relate to
corresponding measures of behavior. In CV research, this
should focus on the degree to which WTP values relate to
actual economic behavior. In NS research, this should focus
on the degree to which NS values relate to evaluations of
actual recreation conditions and associated behaviors.

3. Both CV and NS research should further explore the
tradeoffs between providing too much and too little informa-
tion to respondents. Pretesting of survey instruments may be
especially useful in addressing this issue. Special consideration should be given to exploring the effectiveness of visually-based information where this might be applicable.

4. Both CV and NS research should further explore the potential effects of social desirability/compliance bias, strategic bias, starting point bias, and embedding bias. Social desirability and compliance bias should be addressed through careful attention to question construction to help ensure objectivity or neutrality. Careful training of interviewers is also necessary when surveys are administered personally or by telephone. CV researchers should test for strategic bias to the extent possible and remove outliers and protest bidders to reduce the strategic effect. Starting point bias should be addressed in CV research by varying the initial bids presented to respondents and empirically testing for potential starting point effects. NS research should test the extent to which starting point bias may influence study findings.

5. Both CV and NS research should further test the effects of alternative question formats. In CV research, this effort should focus on the effects of open-ended, and single and multiple bounded close-ended elicitation techniques on WTP values. In NS research, this effort should focus on the effects of long and short versions of normative questions on the NS values derived.

6. Both CV and NS research should continue to explore alternative evaluative dimensions. Generally, WTP is to be favored over WTA compensation in CV research because it results in more conservative estimates. "Acceptability" is the traditional evaluative dimension used in NS research, but other evaluative dimensions, including preference, tolerance, acceptability to others, and management action, may result in alternative NS estimates. Incorporation of multiple evaluative dimensions in both CV and NS research may be wise as a method to enrich resulting information to planners, managers, and policy makers.

7. Both CV and NS research should incorporate potential implications of study questions and findings more directly into research designs. In CV research, respondents should be reminded explicitly of their personal budget constraints and the availability of alternative resources. In NS research, respondents should be informed of potential management actions based on study findings. However, researchers should be careful that such information does not lead to potential strategic bias.

8. CV research should explore the usefulness of graphic approaches to reporting study findings where applicable.

9. NS research should further explore the applicability of alternative measures of central tendency. The issue of consensus is especially pertinent to NS research and more experimentation and agreement about such measures is needed. CV research should continue to investigate the influence of using different measures of average WTP (mean and median) on aggregate WTP estimates. In addition, CV studies that use both mean and median WTP to calculate total social values are likely to provide planners, managers, and policy makers with a richer set of information.

Endnote

1. Email address: rmanning@neture.smnu.vmu.edu

References


8 July 1999

Dr. Robert Manning
School of Natural Resources
Burlington, VT 05405

Dear Dr. Manning,

Here is a copy of the issue of Human Ecology Review that includes your paper. In a continuing effort to increase the journal's visibility, we would appreciate your asking your library to subscribe to the journal, if you have not already done so. We have enclosed a copy of the library recommendation form for that purpose.

In addition, the editorial office would like to send a copy of this issue of the journal to two of your colleagues who will be interested in your paper. You may email the names and addresses to us at: HERreview@osf1.gmu.edu.

Thank you again for giving Human Ecology Review the opportunity to serve as an outlet for your work.

Sincerely,

Linda Kalof
Editor
Library Recommendation Form for Human Ecology Review

(please photocopy)

To recommend this journal to the appropriate subject specialist in your library, please complete this form and return it to:

Melville Coté
Executive Director, SHE
105 Eden Street
College of The Atlantic
Bar Harbor, Maine 04609

We will send a sample copy of the journal to the librarian along with this recommendation. The institutional rate for the journal is $50. Single issues may be purchased for $25/issue.

I RECOMMEND THAT THE LIBRARY SUBSCRIBE TO: Human Ecology Review

SEND SAMPLE COPY TO:

Librarian: ____________________________________________

Library: ____________________________________________

College/University/Institution: ________________________________

Mailing Address: __________________________________________

City, State, & Zipcode: ______________________________________

Telephone Number: ______________________________________

RECOMMENDED BY:

Name/Title: ____________________________________________

Department/School: ______________________________________

College/University/Institution: ________________________________

Mailing Address: __________________________________________

City, State, & Zipcode: ______________________________________

Telephone Number: ______________________________________

Signature: _____________________________________________