Available online at www.sciencedirect.com
SCIENCE DIRECT®
ELSEVIER
Journal of Environmental Economics and Management 50 (2005) 1–22
JOURNAL OF ENVIRONMENTAL ECONOMICS AND MANAGEMENT
www.elsevier.com/locate/jeem

Rethinking the scope test as a criterion for validity in contingent valuation

Thomas A. Heberlein[a,*], Matthew A. Wilson[b], Richard C. Bishop[c], Nora Cate Schaeffer[d]

[a]*Department of Rural Sociology, University of Wisconsin-Madison, 348b Agriculture Hall, 1450 Linden Drive, Madison, WI 53706, USA*
[b]*School of Business Administration, and The Gund Institute for Ecological Economics, University of Vermont, USA*
[c]*Department of Agricultural and Applied Economics, University of Wisconsin-Madison, USA*
[d]*Department of Sociology, University of Wisconsin-Madison, USA*

Received 12 July 2002; received in revised form 27 August 2003; accepted 30 September 2004
Available online 2 December 2004

Abstract

The scope test is widely recommended as a way to evaluate the validity of contingent valuation (CV) studies. We measured contingent values of parts and wholes for four different environmental goods and applied the conventional scope test. In addition, the study design allowed examination of scope relationships for individual respondents. We also used social psychological theory to expand the definition of scope to include attitudinal and behavioral scope. Looking at the individual responses showed that conventional economic scope test results, which depend on comparisons of average values, can hide important relationships that are relevant to the validity of CV studies. We showed that, as a validity test, the conventional scope test can lead to false positives and false negatives. We are led to question the efficacy and cost effectiveness of making scope tests a routine part of validity assessment in applied CV studies.

Keywords: Contingent valuation; Validity; Embedding; Scope test; Wolves; Water quality; Biodiversity

*Corresponding author. Department of Rural Sociology, University of Wisconsin-Madison, 1450 Linden Drive, Madison, WI 53706, USA. Fax: +609 262 6022.
E-mail address: taheberl@facstaff.wisc.edu (T.A. Heberlein).

0095-0696/$ - see front matter © 2004 Elsevier Inc. All rights reserved.
doi:10.1016/j.jeem.2004.09.005

1. Introduction

According to accepted economic thinking, more of a desired market good should normally lead to more consumer utility and hence more economic value. Researchers and policy makers have logically assumed that consumers should likewise be sensitive to changes in the size and scope of environmental amenities: more clean lakes should be preferred to fewer lakes, more endangered species should be preferred to less and so on [3,6,9]. Scope sensitivity tests—which involve comparison of mean values from separate samples—are thus widely regarded as one of the best available steps for gauging the validity of contingent valuation (CV) studies. Failure to obtain larger mean values for more of an environmental good is taken as evidence that the study in question lacks validity [22]. Indeed, some have even gone so far as to suggest that because their studies failed to pass scope tests, the CV method itself should be abandoned [11,12,16,24].

In order to better understand the utility of the scope test, we need to better understand the *conditions* that produce scope failure. A review by Carson [5] showed that out of 22 studies where scope sensitivity was investigated, only four failed to pass the scope test outright while two others showed mixed results. That a large majority passed is encouraging, but the failures are nevertheless troubling. Where would the theory of gravity be if one out of five rocks when dropped landed on the ceiling? Theoretical validity would obviously be thrown into serious doubt. But what if the rocks that 'float' up to the ceiling were discovered not really to be rocks at all, but helium balloons cleverly disguised as rocks? Then the theory of gravity would be preserved. The object of this study is to expose and examine the theoretical conditions that might lead gray balloons to be misunderstood for rocks by thoughtful and rational researchers and policy makers.

To increase the transferability of our research findings, rather than look at one environmental good in isolation, we considered four very different environmental goods: (1) water quality—the signature resource in a region well known for its freshwater lakes; (2) Wisconsin's wild wolf population—a species in which respondents had relatively little interest; (3) Indian spear fishing—an emotionally charged issue; and (4) biodiversity—a complex concept of interest to the scientific community. Using Mitchell and Carson's [21] terminology, we evaluated a predefined "part" and "whole" for each of these goods, thus forming the basis for our scope test analysis.[1]

This study offers new and innovative ways to look at the scope test problem. We first followed conventional practice and looked at traditional split sample comparisons of mean values for each environmental part and whole. Using methods from social psychology, we looked underneath conventional scope test results and showed how successes and failures can be consistent with underlying attitudes toward the goods in question. We did this by showing how the concept of scope can be expanded beyond economic scope to encompass attitudinal and behavioral scope. We also looked beyond the aggregates in a within-subjects experiment where each respondent valued both the part and the whole for each of the goods. Individual respondents were then assigned to one of three economic scope categories for each of the goods: (1) those who showed positive scope (i.e., valued the whole more than the part); (2) those who showed negative scope (valuing the part more than the whole); and (3) those who showed no scope by giving the same value (possibly zero) for both part and whole. By comparing results at the aggregate and individual levels and considering relationships between economic, attitudinal, and behavioral

[1]Mitchell and Carson [21] clearly anticipated scope issues when they spoke of the possibility of "part-whole bias."

scope, we were able to look deeper into the issues surrounding scope tests than has been possible heretofore. Finally, we sought additional, if tentative, insights from ex post personal interviews with selected study subjects.

Our results raised difficult questions about the interpretation of conventional scope test results, but they also offered possible answers. To be sure, the failure to pass conventional scope tests may simply indicate problems in CV study design and execution. The problem is that scope test failures can also occur for other reasons, reasons that are quite compatible with fundamental economic reasoning and social psychological theory. Thus, we will argue below that the real problem may not be with the scope test itself, but rather its simplistic interpretation as a pass/fail validity criterion. Failure to show scope must inevitably lead to questions about why—questions that can be complex and expensive to answer. Accordingly, arguments in favor of the utility of scope tests in the assessing CV study validity are weakened.

2. Previous failures to find scope sensitivity

The idea of scope insensitivity first originated with Kahneman [18]. To support his hypothesis, Kahneman [18] presented a graph (p. 191) showing three demand curves derived from a telephone survey of Ontario residents. Each curve was interpreted as representing respondents' demand for one of three *nested* goods: fishing in the lakes of (1) all of Ontario—the 'whole;' (2) the Haliburton region in Ontario—a 'part' of the 'whole;' and (3) the Muskoka region in Ontario—also a 'part' of the whole [18]. All respondents were asked their willingness to pay (WTP) in the form of a tax to maintain the quality of fishing in these three geographically distinct regions. The resulting graph shows three demand functions that are very similar in shape and magnitude. In other words, a theoretical problem is revealed: "people seem to be willing to pay almost as much to clean up one region or any other, and almost as much for any one region as for all Ontario together" [18, p. 191].

In addition to reiterating the original findings from the Ontario experiment, Kahneman and Knetsch [19] provided new empirical data that has since been widely cited as further evidence of scope insensitivity [12]. All respondents were given the same information describing an inclusive package of public services including education, health, police protection, roads, and environmental services. WTP questions were then administered to three subsamples. Respondents in one sample received three WTP questions starting: (1) environmental services, a subset of that inclusive good; (2) improved disaster preparedness; and (3) improved rescue equipment and trained personnel. Respondents in a second sample were only asked two WTP questions: the first question asked about WTP for improved disaster preparedness and a subsequent question asked about WTP for improved rescue equipment and trained personnel. Finally, respondents in a third sample were asked about their WTP to improve the availability of equipment and trained personnel for rescue operations alone. Kahneman and Knetsch [19] observed no statistically significant difference between mean WTP for the first good valued by each of the subsamples. Hence, they concluded that respondents were insensitive to the inclusiveness of the public good being valued and that the magnitude of the good had no discernible effect on WTP.

Three additional studies commonly cited as supporting the claim of scope insensitivity were carried out by Desvousges et al. [11], Diamond et al. [13], Diamond and Hausman [12], and

Schkade and Payne [24]. First, to test whether the CV method was sensitive to scope variations, Diamond et al. [13] elicited WTP estimates to avoid a 1 percent annual commercial timber harvest in several different wilderness areas throughout the Western United States. Diamond et al. [13] tested the hypothesis that WTP would vary by the *size* of the wilderness area being protected. Using the same split-sample design, the three areas were the Selway Bitteroot Wilderness (1.3 million acres), the Bob Marshall (1.0 million acres), and Washakie (.7 million acres). No significant difference was found between WTP estimates for the three areas. Several other treatments involving wilderness areas also failed to find scope sensitivity. The authors concluded that in general, "whatever CV surveys may be measuring, they are not measuring consumers' economic preferences over environmental amenities" [13, p. 61].

Desvousges et al. [11] investigated the sensitivity of WTP to prevent (a) 2000, (b) 20,000, or (c) 200,000 birds from being killed in oil holding ponds in the Central Flyway. The CV survey involved a self-administered questionnaire conducted with respondents in Atlanta shopping malls. Respondents in three different subsamples were asked their WTP to prevent the deaths of either 2000, 20,000, or 200,000 birds. The resulting means for the three treatments, \$80, \$78, and \$88, were not statistically different, leading the authors to conclude that, "WTP estimates of non-use values do not satisfy simple validity and reliability requirements…current methods for estimating nonuse values are neither valid nor reliable for damage-assessment purposes" [11, p. 93]. This study was replicated by Schkade and Payne [24] using cognitive think-aloud interviews and concluded with similar results.

3. Rethinking scope sensitivity using social psychological theory

Differing conceptions of "value" have long been recognized and debated in both economics [1,2,20] and social psychology [17,23,25]. Traditionally, economic theory has defined 'value' in a rather strict and narrow behavioral sense [15,20]. In this context, the economic value of a commodity is no more and no less than the amount of money a person is willing to give up to get the commodity or the amount the person requires as compensation for loss of the commodity. On the other hand, in terms of attitude theory, WTP as an elicited contingent value is best conceived as a behavioral intention—an expression of *willingness* on the part of survey respondents to engage in a behavior (paying or receiving money) [2,17]. A CV value is *not* an observable behavior like buying or selling commodities in the marketplace; rather it is an expressed intention to make a purchase should the opportunity be available.

Attitude theory suggests that behavioral intentions are influenced by the affective and cognitive dimensions of a person's "attitude" towards a commodity [32,33]. On the one hand, 'cognitions' involve the thoughts and knowledge that people might have about an environmental commodity. For example, a survey respondent might say, "I know a lot about air quality" or "I think a lot about the air that I breathe." Cognitions are generally conceptualized as *information, knowledge or beliefs*, where beliefs are understood to be the associations or linkages that people establish between the attitude object and various objective attributes [14]. The affective dimension of attitudes deals with emotions. An example might be a statement such as, "I like the air quality in my neighborhood."

Scope tests have traditionally looked at the behavioral intention domain of attitudes. But attitude theory would suggest that we might also want to look at 'affective scope'—liking the whole more than the part—and 'cognitive scope'—knowing more and thinking more about the whole than the part. Where affective and cognitive scope are present, we anticipate that respondents would tend to exhibit the virtues of a rational consumer and express a higher WTP (behavioral intention) for more of an environmental amenity than less. Conversely, a respondent may show something akin to 'negative affective scope'—liking the part more than the whole—or 'negative cognitive scope'—knowing more and thinking more about the part than the whole. In this case, attitude–behavior theory suggests that a higher WTP (behavioral intention) will be expressed for the part of the environmental good rather than the whole. In this manner, scope insensitivity can be placed in a richer theoretical context; one that appreciates both the attitudinal and situational characteristics that might reasonably lead the individual to show scope insensitivity.

4. Study design

4.1. Four environmental goods

Based on 27 pre-survey face-to-face interviews with randomly selected property owners in Vilas and Oneida Counties in northern Wisconsin (the "Lakeland Area"), we identified four salient environmental goods for study. *Water quality* in lakes was selected because it is a concrete object of vital economic and social importance to residents of the Lakeland Area. The part was the well-known Minocqua Chain of lakes in the geographical center of the study area. The whole was all of the lakes in Vilas and Oneida Counties, including the Minocqua Chain. There are 991 named lakes in the two counties. While only four of these lakes are part of the Minocqua Chain, two of the four are among only a handful of lakes in the entire region that exceed 1000 acres in area.

The second environmental good chosen for study was wolves, or more precisely, *wolf populations* in northern Wisconsin. The current population is about 200 wolves [29]. At the time of our research, the Wisconsin Department of Natural Resources was developing a wolf management plan and the question of wolf populations in northern Wisconsin was being publicly debated. Wolf populations were numerically nested: the whole was 800 wolves in northern Wisconsin and the part was 300.

Policy makers would sometimes like to have estimates of economic values for complex scientific concepts like carbon sequestration. In our developmental interviews, we asked people what they knew or thought about *biodiversity*. What we found was that most Lakeland property owners thought of biodiversity as a straightforward proxy for the concept of "nature." For example, while deer in northern Wisconsin are overabundant and biologists complain that their presence actually reduces the number and distribution of plant species, most residents we interviewed felt that deer actually represented biodiversity. The whole was protecting biodiversity in all of northern Wisconsin while the part was protecting biodiversity only in Vilas and Oneida Counties.

In the 1980s, local Indian tribes won state and federal court cases that reestablished their right to hunt, fish and gather off reservation. Tribal fishers began to exercise their rights to harvest game fish with spears during the spawning season in April, before the regular sport fishing season

opened. This created great controversy and confrontations at boat landings that bordered on race riots. Although the research team went into the field thinking that the conflict was long over, the topic came up repeatedly in our developmental interviews with respondents expressing very strong feelings about the issue. In an effort to have one environmental good where perhaps rationality was overwhelmed by emotion, we selected the restriction or elimination of *Chippewa Indian spear fishing* and nested the good in a manner similar to water quality with the part represented by spearing in the Minocqua Chain and the whole represented by spearing in all of the lakes in Vilas and Oneida Countries.

4.2. Mail and telephone survey

Information about knowledge, interest, and satisfaction with the wholes and the parts of the four environmental goods was obtained using a 19-page mailed questionnaire.[2] The inside cover of the survey booklet had a map of the two counties where respondents could circle the lakes they had experience with and a color map showing the state with northern Wisconsin and Vilas and Oneida Counties highlighted so people could see what we meant by the whole and the part areas.

Two weeks after we received the completed questionnaire, respondents were contacted by phone to elicit their WTP. The interviews were conducted by a professional survey research organization, the University of Wisconsin Survey Center. The first telephone interview asked respondents what they would be willing to pay for all four goods, two parts and two wholes. The question-order sequence of environmental goods was randomly assigned. Furthermore, during the first telephone interviews (referred to as the Time 1 interview), each respondent was randomly assigned either the part or the whole for each good (300 wolves or 800 wolves; water quality in the Minocqua Chain or all lakes in Vilas and Oneida Counties; etc.). Two weeks after their first completed telephone interview, respondents were contacted for a second interview (the Time 2 interview), which obtained their WTP for the remaining four part-whole complements.[3]

The CV questions were posed in an open-ended fashion. While we recognized that dichotomous choice or referendum questions are preferred by many researchers—indeed we routinely use that approach ourselves [4]—we felt that open-ended questions would give us a much richer data set in this case. In particular, meaningful results from comparisons of part and whole values at the individual level would be much more difficult using dichotomous choice responses. Obviously, using the open-ended approach required the assumption that associated biases, such as those associated with lack of incentive compatibility, would be neutral with respect to scope issues.

[2]A copy of the questionnaire can be found at http://www.aere.org/journal/index.html.

[3]In the design and administration of the Computer Assisted Telephone Survey, all part-whole question orders were completely randomized between time 1 and time 2 surveys for all respondents. Similarly, the order in which each respondent was asked about each environmental good was completely randomized for each time 1 and time 2 survey contact. To test the effectiveness of randomization, a series of ANOVA means tests were run to compare the part-whole results for each good (part first or whole first) and the sequence in which each good was asked (first, second, third or fourth). None of the comparisons yielded statistically significant results, revealing that there was no statistically discernable instrument bias in the experimental design.

4.3. Procedures and response rate

We conducted a mixed-mode mail and telephone panel survey of $N = 1500$ private property owners, identified via property tax records in Vilas and Oneida Counties. The final sample size for the mail questionnaire was 1435 cases—65 questionnaires were undeliverable or the individual was deceased or no longer living at the specified address. The survey was mailed out in the fall of 1998. A total of 124 respondents either refused participation by returning their mail questionnaire with a note stating that they did not want to participate or by telling the interviewer that they did not wish to continue with the study during the reminder telephone calls. By January 25, 1999, the final number of completed mail surveys was measured $n = 876$, with an overall mail response rate of $876/1435 = 61$ percent.

For the telephone interviews, response rates and sample dispositions were calculated both separately and cumulatively for Time 1 interviews and Time 2 interviews. As anticipated, attrition rates and observed non-response errors were substantially higher for Time 1 telephone interviews than Time 2 interviews, with 70 cases refusing outright to participate in the first telephone interview and 120 determined to be "not available" by the UW Survey Center staff after repeated calls. The final within-mode response rate for telephone Time 1 was $686/876 = 78$ percent with a cumulative response rate of $686/1435 = 48$ percent. In the case of Time 2 interviews, attrition rates and non-response declined to only $n = 29$ refusals with only $n = 49$ cases determined to be not available. The final within-wave response rate for Time 2 was therefore $617/676 = 90$ percent with a cumulative response rate of $617/1435 = 43$ percent. Hereafter, the final sample of $n = 617$ cases described above will serve as the primary data used in this study.

4.4. Independent variables: affect, cognition, recreational experience, and ownership

In the initial three pages of the mailed survey, respondents were asked to report where they lived in the Lakeland Area, if it was a seasonal residence, and how much time they spent there. Whether they owned property on the Minocqua Chain or on other lakes in the Lakeland Area served as data on ownership in the analysis that follows. They were also asked to report whether and where they participate in water-related outdoor sports (fishing, boating, etc.). A recreation experience scale from 1 to 5 was constructed according to the number of recreational activities in which the respondent engaged on both the Minocqua Chain and lakes in the Lakeland Area as a whole.

The next 14 pages of the questionnaire measured attitudes toward each of the four environmental goods. Two affective variables were measured: the respondent's rating of each good on a scale from 1 (extremely bad) to 7 (extremely good) and the respondent's satisfaction with the resource on a scale from 1 (extremely dissatisfied) to 7 (extremely satisfied). Similarly, two cognitive dimensions were investigated for each good: the respondent's knowledge of the good on a scale from 1 (almost nothing) to 5 (I am an expert) and how often the respondent thinks about the good on a scale from 1 (never) to 6 (everyday).

The last 3 pages of the survey measured respondent's attitudes toward payment and standard social and economic variables.

4.5. Measures of willingness to pay

Box 1 illustrates the format we used in the CV questions.

The issues associated with payment vehicle choice have been much discussed in the literature [21]. Plausibility to respondents must often be traded off against theoretical qualms and vagueness. In our case, these issues were compounded by our desire to deal with four goods in one survey. Furthermore, our developmental interviews showed that many respondents were averse to conventional CV payment vehicles like property taxes when it came to issues like reducing the amount of spear fishing in lakes, yet were amenable to the idea of using taxes to increase the number of wolves in the area. At various points in our developmental interviews, respondents spontaneously brought up the idea of voluntary contributions to protect and increase biodiversity or mentioned things like higher construction and building permit fees to protect water quality in the lakes. We also feared that switching payment vehicles would have been distracting to respondents.

Box 1

Sample willingness to pay question

One way to raise money would be for people to pay into *a public trust fund* that would be set aside by the State of Wisconsin to increase the number of wolves from 200 to 800…

…At this time, we don't know how you might be asked to pay into the trust fund by the State of Wisconsin, but we do know that payments to the fund would take place on a one-time basis and money could be collected in one of the following four ways:

- You might pay directly to the trust fund through a one-time voluntary donation.
- Your property taxes might increase on a one-time basis, affecting you directly through your tax bill or indirectly through the rent on your residence.
- If you are a Wisconsin resident, your state income taxes might increase on a one-time basis
- Or, you may pay directly through one-time government charges and fees on things like new housing construction, well drilling, septic system, and other permits.

Now, suppose that the number of wolves in Wisconsin could be increased from 200 to 800 if enough money were raised by the Trust Fund.

If you were given the one-time opportunity to pay money to the Trust Fund, what is the most money you would be willing to pay to ensure that the number of wolves in Wisconsin is increased to 800?

Hence, we used the same set of four vehicles in all the CV questions. A one-time payment to an "environmental trust fund" was selected as the best payment strategy [26,27]. We offered respondents a range of four payment alternatives that would be directed into the public trust fund: (1) a one-time voluntary donation, (2) a one-time levy on property taxes, (3) a one-time levy on state income taxes, or (4) one-time charges or fees for things like new housing and construction permits. Our multiple vehicles satisfied our need to have the same vehicles for all the goods. Furthermore, we think it is likely that describing four plausible vehicles as possibilities tended to reduce tendencies toward vehicle bias. It does not, for example, focus so much attention on the objectionable aspects of property taxes or opportunities to free ride when programs are funded through donations.

5. Findings

5.1. Conventional scope tests

We used only Time 1 interviews at this point in the study. This made our approach conform to the normal split sample design of conventional scope tests. As Table 1 below shows, mean WTP values to maintain water quality and prevent spear fishing passed the conventional economic scope test. Importantly, respondents also showed cognitive, affective and behavioral scope sensitivity to the differences between these goods. Respondents tended to say that they knew more about and thought more about water quality and spear fishing in all lakes than in the chain, and that they liked the water quality and disliked spearing more in all the lakes than in the chain. They also recreated more in all the lakes than the chain, one of our measures of behavioral scope. And a larger share of the respondents owned or had owned property on lakes outside the Minocqua Chain, a second measure of behavioral scope.

When it comes to wolves, on the other hand, Table 1 indicates that respondents failed to show sensitivity to the scope. People on average were no more willing to pay for 800 than they were for 300 wolves. They failed to show cognitive or affective scope as well, actually expressing negative cognitive and affective scope. In other words, it appears that our respondents knew more and thought more about 300 wolves than 800 and they tended on average to be more satisfied with 300 than 800 and to rate 300 higher. This pattern of responses to the attitude questions is important and we will return to it later.

Table 1 further shows that biodiversity failed the conventional scope test. In fact, on average, respondents valued the part more than the whole. They were willing to pay significantly more to maintain biodiversity in their local area than they were in the whole region. Concurrently, respondents also showed negative cognitive and affective scope in that they consistently expressed more positive attitudes toward protecting biodiversity in the two counties surrounding their property than toward doing so in the entire region.

5.2. Individual scope

The majority of studies cited previously as having failed to show scope were based on aggregate, split-sample experimental designs. That is, survey respondents got to express their

Table 1
Aggregate WTP and attitude scope for four environmental goods

	Part	Whole	Difference	Positive scope
Water quality	Minocqua Chain	All lakes		
Willingness to pay	$107	$260	**+$153**	Yes
Know about	2.23	2.59	**+.36**	Yes
Think about	2.82	3.44	**+1.62**	Yes
Satisfaction	3.42	3.73	**+.31**	Yes
Rating	4.97	5.50	**+.33**	Yes
Recreation experience	2.50	4.23	**+1.73**	Yes
Ownership	.14	.58	**+.44**	Yes
Spear fishing	Minocqua Chain	All lakes		
Willingness to pay	$47	$102	**+$55**	Yes
Know about	3.23	3.47	**+.24**	Yes
Think about	2.84	2.95	**+.11**	Yes
Satisfaction	2.25	2.13	**−.12**	Yes[a]
Rating	2.26	2.13	**−.13**	Yes[a]
Recreation experience	2.50	4.23	**+1.73**	Yes
Ownership	.14	.58	**+.44**	Yes
Wolves	300 Wolves	800 Wolves		
Willingness to pay	$42	$40	−$2 ns	No
Know about	1.97	1.71	**−.26**	No
Think about	2.34	1.94	**−.40**	No
Satisfaction	4.44	3.48	**−.96**	No
Rating	4.49	3.46	**−1.03**	No
Biodiversity	2 Counties	N. Wisconsin		
Willingness to pay	$173	$125	**−$48**	No
Know about	2.42	2.29	**−.13**	No
Think about	2.96	2.79	**−.17**	No
Satisfaction	5.09	5.03	**−.06**	No
Rating	5.11	5.04	**−.07**	No

Values in bold represent $p < .05$.

[a]The affect and satisfaction questions measure how satisfied the respondent is with the "current level" of off reservation spear fishing by Chippewa. The CV question asks the respondent how much money he or she would be willing to pay to halt the current level of off reservation spear fishing. Affective scope implies that the respondent is willing to pay more to reduce spear fishing where they are more dissatisfied with the current level. Thus, the negative difference score represents positive affective scope.

WTP for only a whole or a part, but not both. Thus the analyses were restricted to statistical averages compared across sample groups. But as we know, human behavior is complex and much is hidden behind averages. By asking respondents to express their WTP for *both* the whole and the part in separate telephone interviews, we are able to explore scope sensitivity across each of the four goods for each individual respondent. As previously noted, to avoid possible survey design biases, both the order that respondents were asked to value the part or the whole and the order of the four goods was completely randomized during survey administration and follow-up analysis showed no statistically significant effect from either part-whole or good sequence.

Table 2
Percent showing economic scope sensitivity

	Positive scope (%)[a]	No scope (%)[b]	Negative scope (%)
Water quality	42	40 (17)	18
Wolves	18	60 (42)	22
Biodiversity	30	48 (19)	22
Spear fishing	20	63 (41)	17

[a]Positive scope here represents conventional economic scope—higher value for the whole than the part.
[b]Values in brackets are the percent of total that said zero for the whole and the part.

In Table 2, which reports results for economic scope, each respondent was classed as expressing positive scope, negative scope, or "no scope." Notice that no scope is used in a special way here. Those in the no-scope category gave the same value for the part and the whole. In many cases, they responded with zero for both part and whole, but others in the no-scope group responded with the same positive values for a given good for part and whole.

At first glance, negative scope may seem objectionable on theoretical grounds. How can more of a desired good elicit a lower WTP value and be based on a fully informed, rational judgment? To start, note that we would not expect negative scope for private goods even if non-satiation were violated. Too much of a good—i.e., enough to yield negative marginal utility—would simply cause the rational consumer to reject unacceptable units above the satiation level even if the good is free. But with environmental *public* goods like wolf populations, consumers have no choice about how much they "consume." If increasing the number of wolves in the population at the margin would cause enough negative marginal utility to accumulate—say because study subjects feared the effects of the larger numbers of wolves on deer or livestock—then values could decline enough to cause negative scope. Of course, this argument is harder to make for environmental amenities like water quality and biodiversity. Marginal utility may approach zero for such goods as supplies approach pristine levels, but likely would not turn absolutely negative. However, as we will show below, there may still be good psychological reasons for a respondent to express values with negative scope, reasons that would not necessarily invalidate observed contingent values.

As the results in Table 2 show, *a majority of respondents failed to show positive economic scope for each of the four environmental goods*. For *water quality*, which passed the scope test in the aggregate analysis, only 42 percent of individual respondents showed positive scope—that is, only 2 out of 5 respondents expressed WTP more to clean up all of the lakes in the Lakeland Area than the four lakes in the Minocqua Chain. For Indian spear fishing, which also passed the split-sample aggregate economic scope test, only 1 in 5 respondents revealed positive scope sensitivity at the individual level. The public thought spear fishing was important and had strong feelings about it, but nevertheless 46 percent said they would pay nothing to stop spear fishing on either the chain or all of the lakes.

For wolves—a good about which most respondents had little interest and no strong feelings—60 percent showed no scope. It is noteworthy that fully 42 percent of all respondents said they would pay nothing for 300 or 800 wolves. Biodiversity, which showed negative economic scope in the aggregate, only had 22 percent of the people showing negative economic scope when we consider the individual data.

5.3. Affective scope

To show positive affective scope an individual had to express more positive feelings about the environmental whole than the part. For negative affective scope, the opposite had to hold. We use as an example here the ranking of the various items on a 7-point attitude scale where 7 was "extremely good" and 1 was "extremely bad". To quantify affective scope we subtracted the rankings for each environmental part from its counterpart whole. For example, if a respondent thought 300 wolves was somewhat good (a ranking of 5) and 800 wolves somewhat bad (a ranking of 3) he or she would be assigned a negative value (ranking −2) and would be classified as showing negative affective scope. Conversely, if a respondent thought water quality in the Minocqua Chain was "very good" (a ranking of 6) and the quality in all the lakes was "very good" (a ranking of 6) then he or she would be assigned a zero and would be classified as showing no affective scope. The results presented in Table 3 were generated using this ranking system.

The findings in Table 3 in many ways mirror those shown above for economic scope. A majority of survey respondents did not report ranking the whole higher than the part for any of the environmental goods. In fact, for three of the four goods—water quality, spear fishing, and biodiversity—a large majority of respondents reported ranking the part and whole equally. With respect to positive and negative scope, the four goods showed remarkable variation in this attitude measure. For wolves, more than 60 percent of respondents said that they ranked the part higher than the whole. Only in the case of water quality did more than 10 percent of survey respondents rank the whole higher than the part with 33 percent giving higher rankings to water quality in the regional lakes rather than the chain of lakes.

Our other affective scope variable —which asked for satisfaction levels on a 1–7 scale—showed very similar patterns and is not presented here.

5.4. Cognitive scope

Cognitive scope was assessed in the same manner as affective scope. Table 4 shows results for the question that asked how often they thought about the part and the whole. On a six-point scale respondents could answer from "every day" (6 points) to "never" (1 point). Those respondents who reported thinking more about the environmental whole than the part were thus considered to have shown positive cognitive scope and those who had thought more about the part than the whole showed negative cognitive scope. Those who gave equal ratings to both had zero scope.

For none of the four goods did a majority of respondents show positive cognitive scope. On the contrary, majorities reported that they thought about the whole and the part equally often. Water

Table 3
Percent showing affective scope for four goods

	Positive affective scope (%)	No affective scope (%)	Negative affective scope (%)
Water quality	32.6	59.1	8.3
Wolves	8.2	29.2	62.6
Spear fishing	3.0	84.2	12.8
Biodiversity	12.0	81.0	7.0

Table 4
Percent showing cognitive scope for four goods

	Positive cognitive scope (%)	No cognitive scope (%)	Negative cognitive scope (%)
Water quality	40.6	51.8	7.6
Wolves	1.8	67.8	30.5
Spear fishing	17.5	80.5	2.0
Biodiversity	1.2	82.7	16.1

quality had the highest percentages with positive cognitive scope. For wolves nearly one out of three thought about 300 wolves more often than about 800 wolves. Similar patterns emerged for our other cognitive scope measure, a question probing how much respondents knew about the attitude objects.

5.5. Individual data: economic scope and attitudinal scope

Based on social psychological theory, we expected that those respondents with positive affective and cognitive scope would also be more likely to show positive economic scope; and conversely for those showing negative affective and cognitive scope.

To estimate the joint effects of cognitive and affective scope sensitivity on economic scope we used multinomial logistic regression. The dependent variable was a three-way categorical economic scope variable: negative scope, zero scope, and positive scope. As is often done with attitudinal measures, the two affective scope variables (rating of the part and whole and satisfaction with the part and whole) were combined to form one independent variable for affective scope. This was done by adding the two scores for the part and subtracting that from the combined score for the whole. Hence a positive value reflects positive affective scope and a negative value the opposite. A combined cognitive scope variable was constructed in the same way and reflects how much each respondent reported knowing about and how often she or he thought about the part and the whole.[4] Additional regressions using the four disaggregated affective and cognitive scope measures as independent variables did not lead us to question any of the conclusions drawn below.

While this effort was not as successful as we had hoped, it did produce some interesting results. Overall, the analysis clearly showed that attitudinal scope did play a significant role in explaining economic scope for both water quality and wolves (see Table 5). Those respondents who were willing to pay more to maintain water quality in all of the lakes also said they thought more about and knew about the water quality in all of the lakes than in the chain (see Table 6). Affective scope appeared to play no role in economic scope for lake water quality. Put another way, those respondents who were more satisfied with and rated the water quality higher in all of the lakes compared to the chain were not willing to pay more to maintain quality in the whole.

For wolves, the story is quite different. Those respondents who said they knew more or thought more about 800 than 300 wolves did not show any difference in their economic scope sensitivity.

[4]Intercorrelations of the constructed affective and cognitive scope variables were as follows: for water quality, .394 ($p<.001$); for wolves, −.047 (ns); for spear fishing, −.017 (ns); and for biodiversity .158, ($p<.001$).

Table 5
Significance of attitudinal scope multinomial regression models

	Water quality	Wolves	Spear fishing	Biodiversity
−2 log likelihood	40.24	24.74	33.35	30.652
χ^2	24.60	16.01	4.89	5.33
Significance	**.001**	**.003**	.298	.255
Pseudo R^2	.058	.034	.011	.011

Table 6
Effects of attitudinal scope on economic scope from multinomial regression

		Water quality		Wolves		Spear fishing		Biodiversity	
		B	SE	*B*	SE	*B*	SE	*B*	SE
Positive scope	Affective scope	.185	.227	**1.40**[*]	.365	.970	.550	.331	.357
	Cognitive scope	**.745**[*]	.226	−1.01	1.05	.082	.202	−1.63	1.082
Negative scope	Affective scope	.114	.282	.468	.432	1.12	.619	.183	.391
	Cognitive scope	−.360	.293	−.160	.668	.117	.251	.337	.609

[*]Values in bold $p<.05$. *B* is the regression coefficient and SE denotes the estimated standard error.

But those respondents, with positive affective scope for wolves, were also more likely to express economic scope sensitivity. Therefore, in the case of wolves, it appears that affective scope is a good predictor for economic scope. Cognitive scope played no role in accounting for economic scope sensitivity for wolves.

For spear fishing and biodiversity there were no significant relationships between cognitive or affective scope and economic scope sensitivity. Moreover, it appears that neither affective nor cognitive scope variables accounted for observed negative scope in any observed cases.

5.6. Behavioral scope

Like economics, social psychology makes a clear distinction between overt behavior and attitudes. Behavior can be observed, although for convenience we most often ask people to report on their behavior on surveys rather than observing it directly. Attitudes, however, must always be inferred from self-reported responses or from interpretations of behavior. Focusing on behavior alone, independent of the underlying attitudes, we considered the role of behavioral scope for lake-related goods, water quality and spear fishing. Suppose, for example, a respondent uses the lakes in all of all the Lakeland Area for fishing, swimming, and boating more than he or she uses the lakes of the Minocqua Chain. This can be considered "positive behavioral scope." Conversely, if the respondent uses the lakes associated with the chain more than the lakes in the two counties we can consider this "negative behavioral scope".

Similarly, we can measure behavioral scope in terms of where each respondent owns property. If they own property on the Minocqua Chain but not on the other lakes in Vilas and Oneida

Table 7
Percent showing behavioral scope for recreation and ownership

	Positive behavioral scope (%)	No behavioral scope (%)	Negative behavioral scope (%)
Recreation scope	63.9	25.6	8.6
Ownership scope	54.5	35.4	10.0

Table 8
Behavioral and economic scope means across two environmental goods

	Economic scope			
	Positive	Zero	Negative	Significance
Water quality				
Recreation scope	.73	.50	.39	$p<.001$
Ownership scope	.55	.45	.29	$p<.004$
Spear fishing				
Recreation scope	.63	.57	.48	ns
Ownership scope	.42	.45	.41	ns

Counties, they were classified as having negative behavioral scope. This indicates that they had a greater material interest in the part than the whole. If they owned property on lakes outside the Minocqua Chain but not on the chain we classify them as having positive scope. Those without lake property and those with property on both the Minocqua Chain and other Lakeland Area lakes were classified as showing no ownership scope.

As Table 7 shows, most respondents showed positive behavioral scope in that they were more likely to own property outside the chain and fish and swim more outside the chain.

The average behavioral scope scores are presented in Table 8 for each kind of economic scope. ANOVA results reveal that those respondents who showed positive economic scope were also more likely to show recreation behavior and ownership scope only for water quality. Those who showed positive economic scope for spear fishing were not significantly more likely to show positive recreation or ownership scope.

6. Post survey interviews

Thirty retrospective interviews were conducted with respondents who had completed all three waves of the survey. Respondents were purposively sampled in order to ensure representation from all of the scope conditions (negative scope, positive scope, and no scope) across the four environmental goods. Respondent debriefing interviews were framed as a question–answer process about the survey in order to gain a better understanding of how individual respondents interpreted the questions that we asked them [10,30]. Retrospective think-aloud protocols were used so that each respondent would ''think aloud'' as they read through their answers to the survey questionnaire, thus verbalizing the contents of their personal memories about why they

answered each question in the manner that they did [28]. The interviews were conducted in person by the authors of this paper and were tape recorded and yielded several hundred pages of transcriptions. The items below represent a very small selection but will give the reader an idea of the thought processes that underlay the quantitative findings.

6.1. Water quality

A respondent who showed negative scope explained: "I'm familiar with Lake Minocqua and to some extent Tomahawk Lake [lakes of the Minoqua Chain]…And, I'm less familiar with other lakes in Vilas and Oneida County although I'm familiar with the large flowages, Rainbow Flowage." This fits with our quantitative analysis: cognitive scope and economic scope worked together as expected. If you know more about the part than the whole you will pay more for the part and the whole.

In a majority of cases, the quantitative analysis showed that respondents did not differentiate between the whole and the part—they said they thought about the part and whole equally and liked the whole and part equally. A respondent who was willing to pay $100 for the chain and $100 for all lakes explained: "I believe in that. I could give $100 toward it…yes I felt good about that program and this was above and beyond my taxes. $100 is just sort of in my head as that would be, it, in other words the $100 had nothing to do with how much it would pay toward cleaning up a lake. It simply would be our family's maximum amount that we would ever put into a program outside of environmental things we're already involved in." Equal economic values followed this lack of differentiation in attitudes.

Because the part is embedded in the whole these sorts of responses may be a bit suspect from an economic perspective, but they are certainly understandable from a psychological point of view. If the criterion for validity is how much people would actually pay, it is certainly plausible that such reasoning would come into play. We do not see an argument for declaring such responses to be invalid.

6.2. Wolves

Our quantitative analysis revealed that many people felt more positive toward 300 than 800 wolves and that affective scope was associated with economic scope. When asked why, respondents explained themselves in ways that were consistent with the data. As one respondent who expressed negative economic scope explained: "I rated 300 wolves as very good and 800 wolves as good. And the difference is because of the potential for the very kind of public dissatisfaction that we're seeing right now…Oh yeah, it [800 wolves] would be a problem for the wolves and it would be a problem for the wolf advocates and so on…So therefore, I would be concerned about whether 800 wolves would cause problems."

Another respondent who said they would pay $200 for 300 wolves and $100 for 800 wolves held similar beliefs: "I think 800 is too many where it could become a problem for them and possibly a people problem with, or problems with us…. Thinking that 300 is probably closer to the reasonable number than 800, I would be more willing to support the 300."

In the formal survey, over four out of 10 respondents said they would pay nothing for 300 wolves and nothing for 800 wolves. In the words of one respondent: "I think we could get along without them [wolves] very well, which we have…I don't think they're very important, I don't

think they're very unimportant. Nature kind of takes care of itself." Another told us, "I would rather see them do other things with tax money than increase the wolf population. 800 no, no, I don't think we should have 800 wolves anyway and I wouldn't pay penny one to get 800 wolves." In other words zero means zero. This was not a protest bid, nor did this person value wolves but was expressing an unwillingness to pay because he did not like the vehicle. He simply did not value wolves. We were quite convinced from our interviews that these zeros indicated that people did not get utility from wolves and would not be likely to pay more than they said they would. Likewise those that were willing to pay more for fewer wolves, had carefully thought out the issues and we feel really would pay more for fewer.

6.3. Spear fishing

Forty-six percent of all respondents were willing to pay nothing to stop spear fishing even though they tended to have strong negative feelings about it, and clearly would be better off if the Indians did not spear fish. Respondents often felt such payments would not be effective or appropriate: "No, you won't buy those people off. [pause] You know this isn't the first thing on the spearing and fishing—probably say maybe 30 years ago they did come up with a reservation fishing license, and they have one now but it didn't go over at all… You're not going to change anything until you change people." For spear fishing, we came away with the strong impression that feelings toward Indians and the politics of sharing resources were more important than cognitive and affective scope about the resource.

6.4. Biodiversity

In the formal survey, over 80 percent of all respondents failed to show cognitive and affective scope for biodiversity, and nearly fifty percent gave the same value for the whole and the part. One respondent who expressed WTP of $500 to protect biodiversity in Vilas and Onieda Counties and in all of Northern Wisconsin was not thinking about wholes and parts. "Well, I suppose I would associate it [biodiversity] with…just looking out in the backyard, you know, looking at the bird feeders and the deer feeders and ferns and the elms and the oaks making a place interesting, making it um I don't know, ecologically uh balanced is what I think about…In fact, this morning I saw a mother [deer] and two fawns out here…it was wonderful, and the hummingbirds were flying at the feeder at the same time, so here I'm looking at all these beautiful, wonderful things that make me feel just great, it's a spiritual connection to me." When it came to biodiversity the public was not thinking like scientists, and nature near their home was worth more to them than nature somewhere else in the north. Biodiversity failed to pass the scope test but the lack of validity of the CV measures was obvious from our interviews both prior to the survey and after the survey, and dividing the sample to show the lack of scope was an expensive and largely unnecessary exercise.

7. Conclusions

In the beginning of this paper we said that we were searching for conditions that lead to scope failure as a way of assessing the utility of the scope test. Most studies in the literature do show

scope. But sometimes they do not and this can be troubling. Our findings lead us to the following conclusions and recommendations.

Conclusion 1: *When respondents know more about the part, like the part more, and have more experience with the part they are likely to assign higher economic values to the part than the whole* (see Table 1). Scope failure may simply signal that the wholes and parts are different goods for the respondents. This was most clearly shown for wolves. People on average tended to like the part more than the whole and to know and think more about the part than the whole and 60 percent of the respondents preferred 300 over 800 wolves (Table 3). Water quality passed the scope test, but there were many who were willing to pay more for the part than the whole because they knew more about the part, liked the part more, and used the part more. Had this group been larger for water quality, it could easily have led to a scope failure or even negative scope for the aggregate. Likewise for biodiversity, on average, people had more positive attitudes toward the part than the whole and gave higher values for the part than the whole.

Conclusion 2: *Contingent values are more likely to be valid (whether they show scope or not) when respondents have knowledge about, experience with, and well-formed attitudes toward the good and the whole and the part*. Water quality showed scope and we believe based on our interviewers and the survey data that the measures have substantial validity—that is people would actually pay roughly what they said they would pay. They knew where the lakes were and what current water quality was like, and they knew more about the whole than the part. Likewise we believe the values given for wolves have substantial validity, even though they did not show scope, because people knew what wolves are and could and did differentiate the attributes of the whole and the part. They were concerned that 800 wolves could be a problem. Also many did not value wolves and their zero bids were not protest bids, but rather reflected disinterest in or antipathy toward wolves. We do not believe that the values we got for biodiversity were valid because respondents were confused about what biodiversity was nor did they differentiate very clearly the whole from the part.

Without reviewing the details of each study here, we believe that the failure to find scope in the widely quoted research by Kahneman [17], Kahneman and Knetsch [18], Diamond et al. [12] and Desvouges et al. [10] reflect situations where people were not able to differentiate the part from the whole either because they had little experience with and understanding of the parts and the wholes or they had little information about the part and the whole.

It is interesting that when confronted with scope failures economists seem to take such failures as *prima facie* evidence of bad questionnaire design. There are certainly poorly designed surveys being done, and some of the scope failures in the cited studies may reflect basic errors in survey design, but our study demonstrates that scope failures can easily occur for other reasons. Nor do we agree completely with the almost automatic response of most CV researchers that scope issues can nearly always be resolved by giving respondents more information. Admittedly, in our research, a more elaborate scenario with detailed descriptions of what biodiversity is and how it would change might have succeeded in obtaining more valid measures, and perhaps even estimates that passed a scope test. However, while we favor making questionnaires clear, one must be careful about depending too much on information provided in the survey. Values constructed on the spot that are not firmly rooted in the attitudes, beliefs, and experiences of respondents may be unstable and overly sensitive to the information presented.

Conclusion 3: *There is more to validity than simply showing scope.* Estimates of WTP to stop Indian spear fishing passed the scope test, but we have serious reservations about whether these numbers reflected the utility that residents would get if the Indians stopped spear fishing the lakes. Attitudes toward spear fishing were much stronger and more extreme than the attitudes toward the three other goods [31] and we expected this to lead to higher values than we obtained. Furthermore, neither affective nor cognitive scope predicted economic scope, which we would take to mean that other, potentially biasing, forces were affecting the values expressed. In addition, four out of 10 respondents said they would pay nothing if the Indians stopped spearing in all or some of the lakes. This was surprising given the strong feelings so many respondents expressed about this subject. The same percentage gave zero values to wolves which people showed much less interest in and much less extreme attitudes. From our interviews it was clear that issues outside of our survey dealing with negative attitudes toward Indians and a strong resistance to giving Indians money was leading people to give CV responses which were lower than the benefits they would derive from a cessation of spearing. Thus, we would conclude that our spear fishing values were badly biased downward by protest bids including protest zeroes. Note that this was true even though the aggregate scope test was passed. Furthermore, we do not believe that this is the sort of problem that could easily be overcome by redesigning the survey. Additional information would not have been likely to change these protest bids. Our judgment is that a reduction in spear fishing is one of those goods that cannot be very well assessed by CV techniques. In this case, positive signals from the conventional scope tests were very likely misleading about the validity of the values derived. Passing the conventional scope test was very probably a false positive.

Thus of the four scope goods we examined two passed the scope test and two did not. For wolves we got a false negative, failure to pass scope when additional evidence from the attitudinal surveys and interviews suggested the CV measures were valid. Estimates of biodiversity failed to pass the scope test and our judgments largely independent of this failure led us to believe that these measures were invalid. Only in one of the four cases was positive scope consistent with other evidence that suggested the measures were valid.

Our investigations also uncovered some aspects of the scope concept that need further consideration. A majority was not willing to pay more for the whole than the part for *any* of our goods. This finding deserves more research. Some of this individual scope failure can be explained by the zero–zero responses, which indicate that respondents placed no value on the good (part or whole). But what does the scope test mean if even for goods that people understand and value (like water quality on the lakes they swim in and fish), nearly 60 percent of the individuals failed to show scope? Should future studies find similar results, what does this imply for the logic of scope tests?

We close by offering some final thoughts on the future usefulness of scope tests.

8. Implications for the future use of scope tests

We would suggest that more thought be given to when, in the research process, scope tests are performed. In current research, the final versions of the main survey and the scope survey or surveys are typically conducted simultaneously. If their studies pass scope tests, researchers

breathe a sigh of relief and report the success as support for the validity of their results. If scope tests fail, not much can be done, since the main survey has already been completed. Little information may be available to understand why failure occurred and to judge whether failure really implies that the value estimates are invalid. And even if it is determined that flaws in the survey instrument led to failure, it is too late to do anything to correct the problem. If scope tests are to be performed, it would make more sense to us to do them as part of the survey development process using pilot surveys. Then scope tests could be integrated into the instrument development process including the qualitative research that we have shown to be essential to the interpretation of scope test results. Integrating scope tests with the qualitative research could help to identify and fix problems in the main questionnaire before it is finalized.

Of course, the problem with this recommendation is that such pilot surveys with sample sizes sufficient to test for scope are likely to be beyond the means of most studies. This leads us to wonder whether scope tests, as currently conducted at the final stage of data gathering, are worth what they cost. To be credible, a scope test requires a separate sample of a size sufficient to support needed statistical tests. The scope instrument must be modified and administered and the results of the scope test analyzed. Then, as we have seen, it will be necessary to go deeper in trying to identify false positives and false negatives. Devoting sufficient resources to accomplish all this involves some real tradeoffs.

Suppose for example that we have been commissioned by a government agency to do a CV study of a proposal to increase Wisconsin's wolf population from 200 to 300 animals. Accepted practice would call for qualitative research to develop a survey instrument, pretesting it, and then administering it to a sample of adequate size to support estimation of WTP for the proposed change. Should we also draw a separate sample and administer a separate survey to value 800 animals in order to evaluate the validity of our main survey instrument? What would we really learn from such a scope test that would not have already been apparent from the focus groups and other qualitative research? Alternative uses of scarce time and research dollars to improve the validity of our results might include additional qualitative research to improve the CV scenario, valuation questions, and other parts of the survey. Additional pretests of the main instrument might produce valuable insights and improvements. Larger sample sizes for the main survey might allow us to stratify the sample in order to better assess how well the instrument is working with different groups such as wolf enthusiasts, those who have suffered from wolf damage to livestock, and the general public with its limited knowledge of and interest in wolves. More might be spent to increase the survey response rate. In a case like this, we might even setup a donation-based simulated market following procedures develop by Champ and her co-authors [7,8]. How many resources should be diverted from such uses in order to do a scope test and show that the value of 800 wolves—a number that lacks policy relevance—is not statistically different from the value of 300 wolves?

What role should scope tests play in the future of CV? We fear that testing for scope will continue. The scope test is widely known. It has a close link to economic theory. It was recommended by the NOAA Panel [22] and some leading economists. And it is a fairly sure way of enhancing the credibility of one's study, since most CV studies pass scope tests. We would conclude, however, that this much emphasis on scope tests is misplaced. Our results show that scope tests failures can occur for reasons that are quite consistent with psychological and even economic theory. We conclude that, by itself, failure to pass a scope test is neither a necessary nor

sufficient condition to invalidate any given CV study. And certainly the scope test failures reported so far in the literature do not undermine the validity of the entire CV method.

Acknowledgements

This research was supported by NSF Grant SBR 9525533. The fieldwork was conducted at the University of Wisconsin Kemp Natural Resources Station, and the Trout Lake Station. We would like to thank Richard Stedman and Brad Jorgensen for their assistance with conceptualization and data collection. We would like to thank Tom Steele and the Staff of Kemp Natural Resources Station as well as all of our patient and helpful respondents. We would also like to thank the editor and two anonymous reviewers for their careful reading and thoughtful recommendations on previous drafts.

References

[1] H.J. Aaron, Distinguished lecture on economics and government: public policy, values, and consciousness, J. Econ. Perspect. 8 (2) (1994) 3–21.
[2] I. Ajzen, G.W. Peterson, Contingent value measurement: the price of everything and the value of nothing?, in: G.L. Peterson, B.L. Driver, R. Gregory (Eds.), Amenity Resource Valuation: Integrating Economics with Other Disciplines, Venture Publishing, State College, PA, 1988, pp. 65–75.
[3] R.C. Bishop, P.A. Champ, T.C. Brown, D.W. McCollum, Measuring non-use values: theory and empirical applications, in: R.J. Kopp, W.W. Pommerehne, N. Schwarz (Eds.), Determining the Value of Non-marketed Goods, Kluwer Academic Publishers, Boston, 1997, pp. 59–81.
[4] R.C. Bishop, T.A. Heberlein, Measuring values of extramarket goods: are indirect measures biased?, Amer. J. Agr. Econ. 61 (1979) 926–930.
[5] R.T. Carson, Contingent valuation surveys and tests of insensitivity to scope, in: R.J. Kopp, W.W. Pommerehne, N. Schwarz (Eds.), Determining the Value of Non-marketed Goods, Kluwer Academic Publishers, Boston, 1997, pp. 127–163.
[6] R.T. Carson, R.C. Mitchell, Sequencing and nesting in contingent valuation surveys, J. Environ. Econ. Manage. 28 (1995) 155–175.
[7] P.A. Champ, R.C. Bishop, T.C. Brown, D.W. McCollum, Using donation mechanisms to value nonuse benefits from public goods, J. Environ. Econ. Manage. 33 (1997) 151–162.
[8] P.A. Champ, R.C. Bishop, Donation payment mechanisms and contingent valuation: an empirical study of hypothetical bias, Environ. Resource Econ. 19 (2001) 383–402.
[9] R.G. Cummings, D.S. Brookshire, W.D. Schulze (Eds.), Valuing Environmental Goods: An Assessment of the Contingent Valuation Method, Rowman and Allenheld, Totowa, NJ, 1986.
[10] T.J. DeMaio, J.M. Rothgeb, Cognitive interviewing techniques: in the lab and in the field, in: N. Schwarz, S. Sudman (Eds.), Answering Questions: Methodology for Determining Cognitive and Communicative Processes in Survey Research, Jossey-Bass Publishers, San Francisco, 1996, pp. 177–195.
[11] W.H. Desvousges, F.R. Johnson, R.W. Dunford, K.J. Boyle, S.P. Hudson, K.N. Wilson, Measuring natural resource damages with contingent valuation: tests of validity and reliability, in: J.A. Hausman (Ed.), Contingent Valuation: A Critical Assessment, North-Holland, Amsterdam, 1993, pp. 91–159.
[12] P.A. Diamond, J.A. Hausman, Contingent valuation: is some number better than no number?, J. Econ. Perspect. 8 (4) (1994) 45–64.
[13] P.A. Diamond, J.A. Hausman, G.K. Leonard, M.A. Denning, Does contingent valuation measure preferences?, in: J.A. Hausman (Ed.), Experimental Evidences, Contingent Valuation a Critical Assessment, Elsevier, Amsterdam, 1993.

[14] M. Fishbein, I. Ajzen, Belief, Attitude, Intention, and Behavior: An Introduction to Theory and Research, Addison-Wesley, Reading, MA, 1975.
[15] M. Freeman, The measurement of environmental and resource values, Resources for the Future, Washington, DC, 1993.
[16] J.A. Hausman (Ed.), Contingent Valuation: A Critical Assessment, Elsevier, Amsterdam, 1993.
[17] T.A. Heberlein, Economics and social psychology in amenity valuation, in: B.L.D.G.L. Peterson, R. Gregory (Eds.), Amenity Resource Valuation, Venture Publishing, State College, PA, 1988, pp. 235–244.
[18] D. Kahneman, Comments by Professor Daniel Kahneman, in: R.G. Cummings, D.S. Brookshire, W.D. Schulze (Eds.), Valuing Environmental Goods: A State of the art Assessment of the Contingent Valuation Method, Rowman and Allanheld, New Jersey, 1986, pp. 185–194.
[19] D. Kahneman, J.L. Knetsch, Valuing public goods: the purchase of moral satisfaction, J. Environ. Econ. Manage. 22 (1992) 57–70.
[20] P. Milgrom, Is sympathy an economic value? Philosophy, economics and the contingent valuation method, in: J.A. Hausman (Ed.), Contingent Valuation a Critical Assessment, Elsevier, Amsterdam, 1993.
[21] R.C. Mitchell, R.T. Carson, Using Surveys to Value Public Goods: the Contingent Valuation Method, Resources for the Future, Washington, DC, 1989.
[22] National Oceanic and Atmospheric Administration, Report of the NOAA panel on contingent valuation, Fed. Register 58 (1993) 4602–4614.
[23] G.L. Peterson, B.L. Driver, R. Gregory (Eds.), Amenity resource valuation, integrating economics with other disciplines, Venture Publishing Inc., State College, PA, 1988.
[24] D.A. Schkade, J.W. Payne, How people respond to contingent valuation questions: a verbal protocol analysis of willingness to pay for an environmental regulation, J. Environ. Econ. Manage. 26 (1994) 88–109.
[25] N. Schwarz, Cognition, communication, and survey measurement: some implications for contingent valuation surveys, in: R.J. Kopp, W.W. Pommerehene, N. Schwarz (Eds.), Determining the Value of Non-marketed Goods, Kluwer Academic Press, Boston, 1997, pp. 165–188.
[26] C.L. Spash, N. Hanley, Preferences, information, and biodiversity preservation, Ecolog. Econ. 12 (1995) 191–208.
[27] T.H. Stevens, J. Echeverria, R.J. Glass, T. Hager, Measuring the existence value of wildlife: what do CVM estimates really show?, Land Econ. 67 (4) (1991) 390–400.
[28] S. Sudman, N.M. Bradburn, N. Schwarz (Eds.), Thinking About Answers: The Application of Cognitive Processes to Survey Methodology, Jossey-Bass Publishers, San Francisco, 1996.
[29] R.P. Thiel, The Timber Wolf in Wisconsin, University of Wisconsin Press, Madison, 1993.
[30] G.B. Willis, T.J. DeMaio, B. Harris-Kojetin, Is the bandwagon headed to the methodological promised land? Evaluating the validity of cognitive interviewing techniques, in: M.G. Sirken, D.J. Herrmann, S. Schechter, N. Schwarz, J.M. Tanur, R. Tourangeau (Eds.), Cognition and Survey Research, Wiley, New York, 1999, pp. 133–153.
[31] M.A. Wilson, Rethinking scope sensitivity and contingent valuation surveys: strong environmental attitudes and contingent economic values, Ph.D. Dissertation, University of Wisconsin-Madison, 2000.
[32] R.B. Zajonc, P. Pietromonaco, J. Bargh, Independence and interaction of affect and cognition, in: M.S. Clark, S. Fiske (Eds.), Affect and Cognition: The Seventeenth Annual Carnegie Symposium on Cognition, Lawrence Erlbaum Associates, Hilsdale, NJ, 1982.
[33] R.B. Zajonc, Feeling and thinking: preferences need no inference, Amer. Psychol. 32 (2) (1980) 151–175.