Journal of Environmental Management (1991) 32, 177–190

Evaluation of Mail and In-person Contingent Value Surveys: Results of a Study of Recreational Boaters

Gregory Mannesto

National Marine Fisheries Service, 1 Blackburn Drive, Gloucester, Massachusetts 01930, U.S.A.

and John B. Loomis*

Division of Environmental Studies, University of California, Davis, California 95616, U.S.A.

Received 11 *October* 1989

There is controversy in the literature over the appropriate method for administering the contingent value method (CVM) survey. This research was conducted to determine if there were any differences in responses between the mail and in-person survey methods. Three criteria were used in evaluating the suitability of mail versus in-person surveys: overall response rate; item non-response; and data comparability. Item non-response and the willingness-to-pay function (analysis of covariance regression approach) found more similarities between methods than differences. The mail method had higher individual question response rates than the in-person method for sensitive (income) and CVM (complex, future-oriented) questions. The mail method may be better suited for CVM questions because it allows for contemplation and reduces pressure to an immediate answer. The in-person method may be better suited for recalling past behavior questions. Overall, both methods provided similar results, so it is important to continue to compare the two methods to further establish their validity and replicability.

Keywords: recreation, boating, contingent value surveys

1. Introduction

The contingent valuation method (CVM) is an increasingly recommended technique for valuing non-market natural resources. The method is being widely used by not only university economists, but also by State and Federal agency personnel. One CVM issue that has been much discussed but not empirically investigated deals with the relative performance of mail versus in-person CVM surveys.

Two schools of thought have emerged regarding mail surveys. One school of thought

*To whom correspondence should be addressed.

0301–4797/91/020177 + 14 $03.00/0

is that CVM surveys are complex and require both explanation and respondent motivation to achieve accurate answers and response rates high enough to be validly generalized to the target population (Mitchell and Carson, 1989, p. 109). These authors argue that given the needs for visual aids (such as maps), extensive explanation and respondent motivation that in-person surveys are "the method of choice for most CVM surveys" (Mitchell and Carson, 1989, p. 109).

For many studies, in-person interviews are quite costly, particularly if interviews take place in the respondent's home. However, Mitchell and Carson (1987) argue that not enough is known about a respondent's cognitive strategies or motivations in answering CVM questions to accept widespread use of mail CVM surveys.

Other researchers (Moser and Dunning, 1986), state that the CVM method has progressed to a point where it can be applied in a routine manner through mail surveys. They state that "experience in testing CV (contingent value) questionnaires suggests that it is well suited for self-administered surveys (pIV-2)." On behalf of the U.S. Army Corps of Engineers, they published a guide for the routine use of the CVM in many natural resource policy settings with either mail, telephone, or in-person surveys. Moser and Dunning state that one advantage of mail surveys is the lack of interviewer bias which is a possibility with the in-person surveys. Yet, they have not provided empirical evidence of why mail surveys are appropriate for the CVM. It would appear they have drawn their conclusions with only a partial evaluation of the mail/in-person survey format.

In this paper, we propose some additional criteria with which to compare mail and in-person surveys. These criteria (discussed below) have been given little attention in the current debate about the appropriate CVM survey method. In addition, we provide empirical evidence and statistical tests of Mitchell and Carson's criteria (discussed below) as well as this new criterion using survey data from a recreational boating study in California.

2. Survey evaluation criteria and the total design method

There are three criteria that are useful in evaluating mail versus in-person surveys. The first two are drawn from Mitchell and Carson (1989): overall response rate and data comparability. To this is added a third criterion, item non-response.

The interaction between survey method (e.g. mail or in-person) and both overall response rate and item non-response can be better understood by extending Dillman's total design method (1978) of questionnaire design and implementation down to the individual question level. His "total design method" (TDM) pays close attention to every detail of the survey setting. TDM is based on the theory of social exchange developed by Homans (1961), Blau (1964) and Thibault and Kelly (1959). Social exchange theory states that an individual's response is motivated by the expected results the responses will bring. In a CVM survey, respondents would expect that their answers will be used to help make decisions on the use of the public good described in the survey.

Dillman (1978) describes his theory of respondent behavior in terms of three major components: reward, costs and trust. First, the costs to the respondent of participating in the survey must be minimized. Second, the rewards to the respondent for accurately and completely answering the questions must be maximized. Third, it is important that the respondent feel a sense of trust that the rewards will be delivered. Respondents are less likely to answer a survey if costs are high, rewards are low, and there is no feeling of trust.

There are not *many* rewards that the researcher can offer the respondent in the survey setting. The introduction to the survey can explain to the subjects that they are part of a carefully selected sample and that their answers are important in order for the study to be successful (i.e. that their answers would be used in making decisions). It is a reward to be considered part of a select group that can make a difference. Being part of a select sample enhances the possibility of influencing the outcome (Hoehn and Randall, 1987; Randall *et al.*, 1983).

If researchers develop questions that are interesting and entertaining, this can be a reward by itself. By being sociable, the interviewer can motivate the respondent to give complete information. Dijkstra and Vanderzouwen (1987) found that showing interest and understanding, acts as a reward for the respondent in general. Their study illustrated that the higher the level of motivation, the more effort the respondents would put forth to match their responses with their perceptions. They showed that respondents interviewed in a friendly style gave more information and performed more accurately than those interviewed in a formal interviewing style.

The *costs* of participating in the survey can be minimized by careful survey design. Time is the major cost factor for the respondents, therefore, the survey should be as short as possible. The mental effort required to answer the questions is a cost. Mental costs can be reduced by designing the questions to be easy to understand and answer. This is particularly important with a mail survey. Participating in the survey should not require direct monetary costs for the subject. For example, in a mail survey the inclusion of a stamped return envelope would eliminate a cost and increase response rate.

It is important that the researcher build a feeling of *trust* with the respondent. In the personal interview, trust can be built by the use of suitable interviewers. A person who comes across as a professional and shows interest in what the respondent says can build trust.

In the mail survey, the development of an appealing survey and the provision of an incentive in advance can build trust. Dillman stated that incentives were effective in mail surveys not because of their monetary power but because of the fact that they represented a symbol of trust. He found that the size of the incentive does not positively correspond to response rate. It was found that increasing the incentive may actually decrease response rate. It should not be made to appear that the researcher is paying the respondent to complete the survey, as the size of the token is rarely adequate payment.

Another example of an incentive woud be a promise to the respondents that they would receive a copy of the results and that results would be used by goverment officials in deciding resource management or public policy.

Dillman (1978: 12) states:

> "that people attempt to keep their costs below the rewards they expect to receive. Fundamentally then, whether a given behavior occurs is a function of the ratio between the perceived costs of doing that activity and the rewards one expects the other party to provide at a later time."

The expected net benefits to the subject of participating in the survey can be expressed in equation (1) as follows:

$$\text{Expected net benefits} = [(\text{Rewards} \times \text{Trust}) - \text{Cost}] \quad (1)$$

The logic behind the formulation of equation (1) is that the respondent perceives the cost of filling out the survey and these costs are incurred at the time the survey is

completed. Many of the rewards of the survey in terms of contributing to improved policy are uncertain at the time the survey is completed. Trust is based on the premise that the rewards of participation will be forthcoming in the future. Thus, reward is multiplied by the trust variable. Trust is the probability that the rewards stated by the researcher will be delivered.

3. Testable hypotheses

The expected net benefits formula will be used to help formulate hypotheses about different formats (mail and in-person surveys) and differential item non-response to varying question formats (open- versus closed-ended questions). Equation (1) suggests the following hypotheses:

1. The overall response rate will be higher for in-person as compared to mail surveys. The in-person interviewer can minimize the mental costs of completing the survey by explaining questions at the pace needed by the respondent. Personal contact aids the building of trust.
2. Overall, equation (1) would suggest the mail item non-response rate will be higher than the in-person survey.
3. The more complex the question the higher the item non-response rate will be because of the greater mental effort (hence cost) required to answer the question (i.e. simple, closed-ended question should have lower non-response rates than the more complex, open-ended questions).
4. The item non-response rate of both the mail and in-person surveys will show less difference for the simpler questions and more difference for the complex questions. The reason is because complex questions have higher costs.

It should be kept in mind that the in-person interview only raises the "discomfort cost" to a respondent of not answering a question but cannot eliminate item non-response. For questions that require a great deal of thought and reflection about intended behavior (e.g. willingness-to-pay) mail surveys may have lower item non-response, since each person sets their own pace.

To test whether the number of responses and non-responses on each item between the mail and in-person survey is statistically different the Fisher test is used. A 2×2 contingency table is constructed in order to use the Fisher test.

4. Data comparability hypotheses

There are two components of data comparability. The first one stressed by Mitchell and Carson (1989, p. 112) relates to insuring different interviewers do not send different verbal or non-verbal messages to interviewees. There is a potential for the interviewer to influence the respondent by either tone of voice, facial expressions, or modification of survey wording. If this occurs, the net result is that respondents are answering different surveys and the data cannot be compared between interviewers.

In-person surveys are more apt to contain answers reflecting a "social desirability bias" (Sudman and Bradburn, 1974). Specifically, are the respondent's answers influenced because he is trying to please the interviewer or make his answers socially acceptable so it will reflect well on the respondent.

Another interpretation of data comparability (useful in this study) relates to comparability of responses between mail and in-person surveys. Even if mail surveys

have much lower overall response rates, are the answers statistically different from the in-person survey?

Wellman *et al.* (1980) much cited study suggested that early respondents and those that responded to subsequent mailings were similar. They suggested that response rates of 70% would represent the population just as well as 100%. Mitchell and Carson (1987, p. 30) take issue with this on conceptual grounds regarding the difference between late respondents and non-respondents. Mitchell and Carson (1987, p. 30) feel that non-respondents are in fact quite different and even simple weighting strategies may be unable to overcome this problem.

In this paper we will statistically test to see if mean responses and willingness-to-pay functions are different between mail and in-person surveys. The effect of experienced versus inexperienced interviewers and mail survey format will be statistically tested using an analysis of covariance approach. The outcome of these statistical tests will allow us to determine if there is data comparability: (1) between interviewers in the in-person survey; and (2) between the in-person survey with a high response rate and a mail survey with a relatively low response rate.

5. Data sources and survey procedures

The sample frame is recreational boaters and anglers in the Sacramento–San Joaquin Delta of California. Two different approaches, in-person and mail-back, were used in order to test the hypotheses about overall response rate, item non-response, and data comparability. The personal interviews were conducted at public boat ramps in the delta. The interviews took place on both morning and afternoon weekdays and weekend periods from 29 August to 9 October 1987. A stratified random sample of public boat ramps was used.

The second technique was a mail-back questionnaire that was placed on vehicles with boat trailers parked at the same public boat ramps where the in-person interviews were conducted that day. At the end of each in-person interview period (1.00 p.m. or 5.00 p.m.) questionnaires were placed on the windshields of the remaining vehicles with boat trailers. Each survey contained a prepaid envelope. This technique has two advantages: (1) the same type of visitors as the in-person survey could be targeted on the same day at low cost; and (2) by leaving surveys on windshields, no personal contact was made. This approach closely matches a mail to the home survey where no personal contact is made.

Using the two different data collection methods allowed for a comparison of the two approaches for a population for which no names and addresses were available ahead of time.

5.1. INTERVIEWING PROCESS

The interviewers were trained in standard interviewer techniques. This involved training as a group so that they could learn from each other. The interviewers took part in the pretest and this allowed them to become familiar with the questionnarie and the sampling schemes which the interviewers were instructed to follow. This gave them a chance to provide input so the questionnaire and the interview process was more understandable. This technique motivated the interviewers to be more enthusiastic about the project. All interviewers were staff members of Environmental Resources Branch of the U.S. Army Corps of Engineers in Sacramento, California.

The research results will be examined to determine the influence of different types of

interviewers. One interviewer was more knowledgeable and experienced in conducting surveys and he trained the other three interviewers who had no previous survey experience. Therefore, the in-person survey has two treatment groups: the experienced interviewer and the less experienced interviewer group.

5.2. DESIGN OF THE SURVEY

The basic design of the survey, question format and question order, followed the format suggested by Dillman (1978). The first set of questions were developed to collect information of the current recreation trip (see Mannesto, 1989, for copy of the survey). The second set of questions collected information on annual trips and expenses. The beginning, easier questions helped prepare the subjects for the CVM questions. The first CVM question asked visitors their willingness-to-pay (WTP) to maintain the current wetland situation. It was an open-ended question that asked them to state their maximum WTP. A payment card was presented to the respondent in the personal interview format to provide a possible range of values, but the interviewer made it clear that the respondent was free to choose any value. The values on the payment card ranged from 0 to $850. The card was arranged in a rectangular matrix form to minimize potential for bias (Moser and Dunning, 1986). The second and third CVM questions asked respondents what their maximum WTP for two alternative levels of restored wetlands (25% and 50% increase in the number of total delta wetlands). The current wetland situation and the two levels of improvement were depicted on a colored map, either displayed by the interviewer or provided as a separate attachment in the mail questionnaire. After the first two WTP questions, a protest question was asked in order to determine which zero bids were valid and which were protests to some aspect of the CVM question. Respondents who take offense to some aspect of the hypothetical market are unlikely to give accurate values to the WTP question. Therefore, in order to increase accuracy of results, people who protest are removed from the WTP analysis.

6. Results

6.1. HYPOTHESIS 1: IN-PERSON AND MAIL SURVEY OVERALL RESPONSE RATE

Personal interviews were completed for 241 visitors out of a possible 248 attempts, a 97% completion rate, during the interview period from August to October, 1987. The mail-back questionnaires were placed on 645 vehicles with boat trailers at public access sites in the Sacramento–San Joaquin Delta. They were returned by 155, a response rate of 24%. The 24% response rate is reasonable (but not desirable) considering no incentives were given for returning the questionnaire and it was not possible to send a follow-up post card or a second mailing of the survey to increase response rate as is standard with the Dillman (1978) repeat mailing approach. Corps of Engineers's legal counsel stated the survey had to be kept anonymous so no personal addresses could be recorded for follow-up mailings. Even with follow-up mailings, the overall response rate would have likely been below the in-person survey (Loomis, 1987). Nonetheless, the in-person survey did have a substantially higher overall response rate.

6.2. COMPARISON OF ITEM NON-RESPONSE BETWEEN THE TWO METHODS

There were only five questions answered by all respondents in both the mail and in-

TABLE 1. Comparison of response (R) and non-response (NR) between survey methods

	Mail			In-person			
	NR	R	NR (%)	NR	R	NR (%)	One-sided significance
Simple, open-ended questions							
Total	12	1326	0·9	10	2359	0·4	0·059
Complex (past behavior), open-ended questions							
Total	73	181	28·7	30	387	7·2	0·000†
Complex (future behavior), open-ended questions							
Add fishing	16	72	18·2	40	121	33·1	0·148
WTP current	5	161	3·0	18	238	7·0	0·056
WTP improve	14	152	8·4	48	206	23·3	0·001‡
Total	35	385	8·3	106	565	15·8	0·000‡
Simple, closed-ended questions							
Total	8	744	1·1	44	1236	3·4	0·001‡
Complex, closed-ended questions							
Total	108	478	18·4	96	833	10·3	0·000†
Embarrassing, closed-ended questions							
Total	19	144	11·7	27	229	10·5	0·445
Overall results							
Total	255	3258	7·3	313	5609	5·3	0·0001†

† Significant difference: mail survey has a higher non-response rate.
‡ Significant difference: in-person survey has a higher non-response rate.

person surveys. They were: "number of miles driven"; "group size"; "hours recreated"; "species fished for"; and "number of fish caught". It is interesting to note that all of these items were at the beginning of the survey, where interest and motivation is still high and the cost of answering questions is still low. This finding supports Bradburn's (1983) finding of a fatigue effect over the course of an interview. Also, they were all simple questions to answer which involved little mental effort (cost).

Table 1 compares the number of non-responses to the different types of questions in both surveys to determine if there are any significant differences. The potential mail respondents who did not return the survey are not included in the comparison and this represents a significant groups of non-responses (76%). The results will be interpreted with this omission in mind so that the limits of the results are clearly understood.

The overall results show that the mail survey had an item non-response rate of 7·8% for returned questionnaires while the in-person survey had a 5·6% item non-response rate. This was expected based on equation (1) and supports hypothesis (2) that mail surveys would generally have higher item non-response. For the complex, open-ended questions, such as willingness-to-pay or future recreation, the mail survey had a statistically lower item non-response rate than the in-person. It may be these questions are hard to quickly answer and a person would rather move on in the interview. In the mail survey they have time to reflect and answer at their own pace.

The analysis of item non-response on the WTP for improving wetland question (Table 2) shows that there is a significant difference between experienced and less

TABLE 2. Comparison of the three groups on WTP responses

	Non-response	Response	NR (%)
Experienced interviewer	20	126	13·7
Less experienced interviewer	28	79	26·2
Mail	14	152	8·4

experienced interviewers (Fisher exact = 0·0045). These results support DeLamater's (1982) findings that technically competent interviewers may be able to reduce non-response. There was a significant difference between less experienced interviewers and mail respondents (Fisher exact = 0·0022), but there was not a significant difference between respondents to the experienced interviewer and mail respondents (Fisher exact = 0·426).

The questions with the highest percentage of non-response in both the mail and in-person survey were: "recreation benefit" (44%); "income" (11%); and "know better ways to improve fishing" (38%). All of these high non-response items (except income) require a more complex thought process to answer than most of the other questions in the survey. These relatively high item non-response rates are consistent with equation (1) and hypothesis 3. Respondent behavior would predict that respondents could minimize their costs by skipping complex items.

6.3. SUMMARY OF ITEM NON-RESPONSE BETWEEN METHODS

Hypothesis (3) would predict that the more complex the question the more cost to the respondent and the more likely he would skip the question in order to increase his total net benefits of participating in the survey. The combined results of the two surveys reflect this postulate extremely well. The results also support Hippler and Schwarz's (1987) finding that closed formats will reduce uncertainty (a cost to the respondents) and result in higher response rates. Table 3 summarizes these results.

The simple, open-ended questions had a combined mail and in-person item non-response rate of less than 1% and the simple, closed-ended questions also had a very low non-response rate—2·6%. The complex, closed-ended questions had a non-response rate of 13·5%. The complex (future behavior), open-ended questions (12·9%) and embarrassing, closed-ended questions (11%) had very similar non-response rates to the complex, closed-ended questions (13·5%). The group of questions with the highest non-response rate at 15·4% was the past behavior, open-ended questions. There appears to be a real cost of recalling past behavior that is even higher than thinking about the future.

TABLE 3. Summary of item non-response rates in different types of questions

	Simple open (%)	Simple closed (%)	Complex closed (%)	Complex future open (%)	Complex past open (%)	Embarrassing closed (%)
In-person	0·4	3·4	10·3	15·8	7·2	10·5
Mail	0·9	1·1	18·4	8·3	28·7	11·7
All	0·6	2·6	13·5	12·9	15·4	11·0

There was not a significant difference in responses to simple, open-ended questions between mail and in-person respondents. This is consistent with hypothesis (4). Simpler questions involve less cost in answering than complex questions, therefore, less significant difference is expected between the two methods on simpler questions.

An interesting point is that past behavior questions had a significantly higher non-response rate for the mail survey than the future behavior questions, while the future behavior questions had significantly higher item non-response than the past behavior questions in the in-person survey (opposite results). This demonstrates that the mail survey may be better for thinking about future events (where they have time to contemplate) while in-person surveys may be better for recently past (last recreation trip) behavior. These results support Bradburn and Sudman's (1979) finding that memory was the most important factor in influencing response rate. The implication for future-oriented CVM questions is that mail surveys may be better than in-person because of the ability to contemplate future behavior without time pressures.

6.4. COMPARISON OF MEANS BETWEEN MAIL, EXPERIENCED INTERVIEWER AND LESS EXPERIENCED INTERVIEWER RESPONDENTS

There were significant differences between the three groups in WTP for the current wetland situation and for improvements. Tables 4 and 5 summarize these results.

There was not a significant difference between the less experienced interviewers and mail surveys in either WTP question (t-test, two-sided, $P=0{\cdot}949$ for difference on current wetland WTP and $P=0{\cdot}652$ for improved wetland WTP question). Less experienced interviewer respondents behave more like mail survey respondents than experienced interviewer respondents.

TABLE 4. Comparison on the WTP for the current wetland situation

Group	Mean	Standard error	Sample size
Experienced interviewer	69·80	8·48	122
Less experienced interviewer	37·12	9·88	74
Mail	37·85	5·84	109

Anova F value = 5·77.
Significance level = 0·0035.

TABLE 5. Comparison on the WTP for the improved wetland situation

Group	Mean	Standard error	Sample size
Experienced interviewer	59·27	6·81	116
Less experienced interviewer	39·47	12·51	58
Mail	33·14	6·34	102

Anova F value = 3·49.
Significance level = 0·032.

TABLE 6. Number of protestors for mail, experienced, and less experienced interviewers

	Protests	Non-protests	Protests (%)
Mail	52	114	31·3
Less experienced interviewers	27	80	25·2
Experienced interviewer	14	132	9·6

There were significant differences between the experienced interviewer and the less experienced interviewer on the current WTP for wetlands question (t-test, two-sided, $P=0{\cdot}012$). It may be that the experienced interviewer obtains the higher bids because he understands the logic and purpose behind the WTP question. Consequently, he can guide the respondent to understand the survey and provide more thoughtful answers. Perhaps the respondents to the experienced interviewer are giving a higher bid than they would normally because they feel the experienced interviewer's enthusiasm in the interview, and bid higher as a result of this social pressure. Therefore, factors other than training and experience may be responsible for the higher bids. In terms of means, the mail and inexperienced interviewer groups are equal.

There are no significant differences in income considering the three groups at one time. Also, there was no difference between groups analysed two at a time (experienced versus less experienced, $P=0{\cdot}338$ (two-sided); less experienced versus mail, $P=0{\cdot}087$; and experienced versus mail, $P=0{\cdot}386$). This may mean that each group was selected from the same population.

6.5. COMPARISON OF PROTESTERS AND NON-PROTESTERS

One would also expect that there will be significant differences between mail, experienced interviewer and less experienced interviewers in the number of respondents protesting the WTP question. The results are presented in Table 6.

The Chi-square statistic with 2 degrees of freedom is 22·02. There is a significant difference between the groups (significance level, $P=0{\cdot}00$). Testing two groups at a time, there is not a significant difference between the mail and less experienced interviewer ($P=0{\cdot}17$). But, there is a significant difference in mail and experienced interviewer respondents ($P=0{\cdot}00$) and also there is a significant difference between less experienced and experienced respondents ($P=0{\cdot}00$) in number of protest questions. A possible explanation is that the experienced interviewer is better at answering concerns the respondent may have about the nature of, and reason for, the CVM question. Thus, this should lead to a reduction in the number of protesters to the WTP question. However, it is also possible that the experienced interviewer's enthusiasm may have generated a social desirability bias toward giving positive (non-zero) willingness-to-pay bids which are not considered protests.

There were no differences between mail and less experienced interviewer respondents which may mean the less experienced interviewer respondents act more similar to a group that has no interviewer than a group that has an experienced interviewer. The complexity of the WTP question can lead to misunderstanding and since either a less experienced interviewer is present or nobody is present to explain it to the mail respondents, the respondents in these situations are more likely to skip the question instead of investing additional time and mental cost to figure out the question. This additional cost to the mail and less experienced interviewer respondents answering the WTP question would increase the number of protesters.

6.6. DIFFERENCES IN WPT FUNCTIONS BETWEEN METHODS

The null hypothesis regarding data comparability states that the WTP function of the experienced interviewer respondents is not significantly different from the WTP function of either the less experienced interviewer respondents or the mail respondents. A regression model was developed to test this hypothesis based on one dependent variable (WTP to improve wetlands) and six independent variables. The following independent variables were used in the model: age; miles driven; group size; recreation benefit; quality rating; and conservation rating. The analysis of covariance (dummy variable) approach is used to compare the three groups.

The recreation benefit variable's inclusion in the model drastically reduced the number of cases that could be analysed because many respondents did not answer this question (it had the highest non-response rate of any question—44%). A change was made in the recreation benefit variable so the number of cases in the overall analysis could be increased. First, a new variable was created that was assigned a value of one if the case was missing and zero if the recreation benefit variable contained data. If this variable has a significant t-value ($t > 1{\cdot}96$) this indicates the missing data cases are different from the cases with data. The second change was the insertion of the average value of the recreation benefit variable in all the cases containing missing data.

The analysis of covariance analysis to test for difference in the WTP function will be implemented using dummy variables. The advantage of the analysis of covariance over the Chow test is that it allows for identification of the specific variables which are different rather than accepting or rejecting only the entire equation as the Chow test does. The two dummy variables, $D = 1$ for mail respondents and $D = 0$ for all other cases and $K = 1$ for experienced interviewer respondents and $K = 0$ for all other cases are included along with the six independent variables and the missing data variable for the recreation benefit variable. Equation (2) is expressed as:

$$\begin{aligned} \mathrm{WTP} = {} & B_0 + B_1 D + B_2 T + B_3 G + B_4 R + B_5 Q + B_6 C + B_7 A + B_8 M \\ & + B_9 K + B_{10} DT + B_{11} DG + B_{12} DR + B_{13} DQ + B_{14} DC + B_{15} DA + B_{16} DM + \\ & B_{17} KT + B_{18} KG + B_{19} KR + B_{20} KQ + B_{21} KC + B_{22} KA + e \end{aligned} \tag{2}$$

WTP, log of willingness-to-pay to improve wetlands; B_0, constant; D, dummy variable 1, 1 = mail respondents, 0 = all other cases; K, dummy variable 2, 1 = experienced interviewer respondents, 0 = all other cases; T, log of the number of miles respondents drive one way to the site; G, log of the number of people in the respondent's group; R, log of recreation benefits: the amount of money a respondent would spend for his last recreation experience above actual expenses; Q, log of the respondent's rating of the quality of their recreational experience; C, log of the respondent's rating on a conservation/development scale; A, log of the respondent's age; M, missing data variable; e, error term.
$DT, DG, DR, DQ, DC, DA, DM$, the interaction variables for dummy variable No. 1.
$KT, KG, KR, KQ, KC, KA, KM$, the interaction variables for dummy variable No. 2.

The regression equation, which has 22 variables, includes all six of the independent variables plus all the interaction terms (each dummy variable times each independent variable to test to see if the slopes of the variable change with method).

The results of this full interaction equation produced two significant interaction variables (Table 7). There was a significant difference in slope between the experienced interviewer respondents and the less experienced group on the quality of experience

TABLE 7. Regression analysis, full interaction equation

Variable	Coefficient	Standard error	*t*-Statistic
Intercept	12·79	7·04	1·81
Miles	0·50	0·31	1·58
Group size	−0·80	0·73	−1·09
Recreation benefit	−1·41	0·86	−1·65
Quality of experience	−0·18	0·29	−0·63
Conservation-developer	−0·63	0·82	−0·77
Age	−2·21	1·62	−1·36
Missing data variable	1·08	2·95	0·37
Dummy variable 1	−6·21	8·47	−0·73
Dummy variable 2	−6·35	8·07	−0·79
D_1/miles	−0·39	0·42	−0·93
D_1/group size	−1·06	0·97	−1·10
D_1/recreation benefit	3·22*	0·96	3·34
D_1/quality of experience	0·23	0·72	0·32
D_1/conservation developer	1·00	1·03	0·97
D_1/age	−0·43	2·01	−0·21
D_1/missing data variable	−0·56	3·07	−0·18
D_2/miles	−0·54	0·40	−1·35
D_2/group size	0·01	0·95	0·01
D_2/recreation benefit	0·63	1·08	0·59
D_2/quality of experience	1·07*	0·42	2·53
D_2/conservation-developer	0·38	0·95	0·40
D_2/age	1·10	1·86	0·59

* Indicates coefficients significant at $P<0·05$.

variable. The other significant variable was the recreation benefit variable which represents a difference in slope between mail and all other cases. In the strictest statistical sense, we must reject perfect data comparability between the three treatments. The analysis of covariance on the willingness-to-pay function indicates there is a statistically significant difference for one variable between experienced and less experienced interviewers and for another variable for mail versus in-person interviewer groups. Nonetheless, the significant differences between the group are minimal. Only two out of 22 variables showed significant differences which is only 9%.

7. Conclusions

The major objective of this research was to determine if there were any differences in responses between types of questions and between mail and in-person respondents. Overall, the in-person survey had a much higher survey completion rate than the mail survey. Without the normal follow-ups conducted in most surveys, the mail survey received a 24% response rate, while the in-person was 97%. These differences must be kept in mind when comparing the mail and in-person results that follow.

There was a trend in the data from simple to complex questions: the more complex the question the lower the item response rate. All of the findings are consistent with our adaption of Dillman's theory to the individual question level. It predicts that as motivation decreases or the benefit of answering questions decreases or the cost of answering questions increases, item non-response rate should increase. The implication of these findings to survey research is to have short questionnaires with salient, easily understood questions.

A major finding was that it appears certain types of questions are better suited (have a higher response rate) for each method: the mail survey was better for more complex (future) behavior questions and the in-person was better with complex (past) behavior questions. Mail surveys received a 20% higher response rate on future versus past behavior questions (8·3% item non-response versus 28·7% item non-response for in-person). The results show it may be easier for the mail respondents to answer future questions than to recall past events. Many types of CVM questions are complex (future) behavior questions which involve contemplation so the ability to remove the time constraint from the in-person surveys could make the mail survey better for CVM questions than the in-person survey.

The results show the experienced interviewer respondents are significantly higher in their willingness-to-pay bids for the current wetlands situation than both the less experienced interviewer respondents and mail survey respondents. The less experienced interviewer group behaved more like the mail respondents than the experienced interviewer group. The higher bids on the experienced interviewer group may be due to a social desirability bias exerted by the more enthusiastic experienced interviewer.

The mail respondents bid significantly lower on the WTP to improve wetlands. People may bid less because they have more time to think about their answers in the mail survey or perhaps it is due to lack of social desirability influence of an interviewer. Thus, using the mail survey may result in lower CVM values than the in-person survey.

The data showed that an experienced interviewer is an important variable in reducing the number of protests to the CVM question. The reduction in number of protesters may be because the experienced interviewer understands and can better explain the questions and their importance to the success of the study, thereby reducing mental cost to respondents answering the question.

A test of the WTP function (model) for any differences between survey methods was performed using the analysis of covariance (dummy variable) approach. The results showed that the null hypothesis that there would be no differences in WTP function between groups could be rejected. There were significant differences between the mail and the other two groups on the recreation benefit variable. Additionally, there were significant differences between the experienced and less experienced interviewer groups on the quality of experience variable. The conclusion is that there are differences between methods and differences between groups. The full interaction model had only two variables out of 22 variables in the equation that showed a significant difference, thus only 9% of the variables are statistically different. This shows there are more similarities between the groups than differences. Therefore, since there are many similarities between the methods and the mail survey is a cost-effective method of collecting information it is important to continue CVM mail survey research in the future.

One lesson to be learned from this experiment is that it may be premature to unconditionally recommend mail surveys as Moser and Dunning do. However, it is equally unnecessary to reject mail CVM surveys in favor of in-person CVM surveys as Mitchell and Carson do. If the nature of the CVM WTP question relates to valuation of recently experienced events, then in-person surveys may be best. If, on the other hand, the WTP scenarios, are about valuing future scenarios, the mail survey format which allows the respondent to reflect and answer at their own pace may be desirable. Certainly several more research comparisons are needed before anyone can be more conclusive on which type of survey method is most appropriate for different contingent valuation method orientations.

We would like to thank Sy Schwartz and Jim Cramer, University of California, Davis, for their suggestions on all phases of this research. They are not implicated in any way for what follows.

References

Blau, P. M. (1964). *Exchange and Power in Social Life*. New York: John Wiley and Sons.

Bradburn, N. M. (1983). Response effects. In *Handbook of Survey Research* (Rossi, P. H., ed.). New York: Academic Press.

Bradburn, N. M. and Sudman, S. (1979). *Improving Interview Method and Questionnaire Design*. San Francisco: Jossey-Bass Publishers.

DeLamater, J. (1982). Response effects of question content. In *Response Behavior in the Survey Interview* (Dijkstra, W. and Vanderzouwen, J., eds.). London: Academic Press.

Dijkstra, W. and Vanderzouwen, J. (1982). *Response Behavior in the Survey Interview*. London: Academic Press.

Dillman, D. (1978). *Mail and Telephone Survey: the Total Design Method*. New York: Wiley.

Hippler, H. J., Schwartz, N. and Sudman, S. (eds) (1987). *Social Information Processing and Survey Methodology*. New York: Springer-Verlag.

Hoehn, J. H. and Randell, A. (1987). A satisfactory benefit cost indicator from contingent valuation. *Journal of Environmental Economics and Management* **14**, 226–247.

Homans, G. C. (1961). *Social Behavior: Its Elementary Forms*. New York: Harcourt, Brace, and World.

Loomis, J. B. (1987). Expanding contingent value estimates to aggregate benefit estimate: current practices and proposed solutions. *Land Economics* **63**, 396–402.

Loomis, J. B. (1989). Test–retest reliability of the contingent valuation method: a comparison of general population and visitor responses. *American Journal of Agricultural Economics* **71**, 76–84.

Mannesto, G. M. (1989). Comparative Evaluation of Respondent Behavior in Mail and In-Person Contingent Value Method Surveys. PhD Dissertation, University of California/Davis.

Mitchell, R. C. and Carson, R. T. (1987). *How Far Along The Learning Curve Is The Contingent Valuation Method?* Discussion Paper QE 87-07, April, Resources for the Future.

Mitchell, R. C. and Carson, R. T. (1989). *Using Surveys to Value Public Goods: Contingent Valuation Method*. Washington D.C.: Resources for the Future.

Moser, D. and Dunning, C. M. (1986). *National Economic Development Procedures Manual—Recreation—Volume II—A Guide for Using the Contingent Value Method in Recreation Studies*. Fort Belvoir, Virginia: U.S. Corp of Engineers, IWR Report 86-R-5.

Randall, A., Hoehn, J. P. and Brookshire, D. S. (1983). Contingent valuation surveys for evaluating environmental assets. *Natural Resources Journal* **23**, 635–648.

Sudman, S. and Bradburn, N. M. (1974). *Response Effects in Surveys: a Review and Synthesis*. Chicago: Aldine.

Thibault, J. W. and Kelley, H. H. (1959). *The Social Psychology of Groups*. New York: John Wiley and Sons.

Wellman, J. D., Hawk, E. G., Roggenbuck, J. W. and Buhyoff, G. J. (1980). Mailed questionnaire surveys and the reluctant respondent: an empirical examination of differences between early and late respondents. *Journal of Leisure Research* **12**, 164–172.