CAN RESPONSE ORDER BIAS EVALUATIONS?

GLENN D. ISRAEL and C. L. TAYLOR

University of Florida

ABSTRACT

Response-order effects can bias evaluations. Measurement error is introduced by order effects, and this, in turn, can affect associations between variables. This study attempts to identify types of questions in mail questionnaires which are susceptible to order effects and illustrates the consequences of response-order effects on associations. Data from a Cooperative Extension Service evaluation showed order effects for multiple response and attributive questions but not for single response ones. Aspects of question complexity and social desirability may contribute to order effects. In at least one instance, response-order effects were found to influence associations between items and thus, the substantive conclusions.

CAN RESPONSE-ORDER BIAS EVALUATIONS?

Although much research has been directed at assessing the effects of the order and wording of questions on responses in evaluation surveys, less attention has been directed toward bias which can occur from the order of response alternatives in questions. Response-order effects in closed survey questions are said to occur when the order in which the alternatives are listed influences respondents' choices (Schuman & Presser, 1981). The consequences of response-order effects are two-fold. First, the distribution of responses can be affected, thereby introducing measurement error in variables for evaluation studies. For panel studies, comparability requires the same response sequences be used (Carp, 1974; see also Kraut, Wolfson, & Rothenberg, 1975), but the possibility of measurement error for each instance of data collection remains. Second, response-order effects can influence associations between variables (Kalton, Collins, & Brook, 1978). Measurement error that is introduced by response-order effects can attenuate relationships, resulting in reduced significance and/or the nonrejection of null hypotheses when they should be rejected. As Kalton et al. (1978) note, potential bias due to response order deserves wider attention by evaluators than has previously been the case.

To date, most studies of response-order effects used telephone and face-to-face interviews and little attention has been directed at mail questionnaires (see, for example, Bishop, Hippler, Schwarz, & Strack, 1988; Carp, 1974; Kraut, Wolfson, & Rothenberg, 1975; Krosnick & Alwin, 1987; Schuman & Presser, 1981). However, mail surveys continue to be an important means for gathering evaluation data. The purpose of this paper is to examine the occurrence of response-order effects in a mail survey and to identify which types of questions are susceptible to these effects. The potential for response order to influence substantive conclusions about the effectiveness of a program is illustrated.

ISSUES FOR STUDY

Among the most commonly mentioned response-order effects are primacy and recency effects. A primacy effect is the tendency to select a response near the beginning of a list of alternatives. Likewise, a recency effect is the tendency to select the responses near the end of a list of alternatives. A number of studies using long lists of response choices have found a primacy effect (Becker, 1954; Brook & Upton, 1974; Krosnick & Al-
win, 1987; Mueller, 1970). In a study of radio show preferences, Becker (1954) found that the chances of being selected as one of the five most preferred (of 16 programs on the cardlist) increased as the program type moved from the sixteenth position to the first position on a printed checklist. The voting studies by Brook and Upton (1974) and Mueller (1970) both found that in nonpartisan elections with numerous nominees, candidates whose names appeared near the top of the ballot gained votes as a consequence of that position.

Although primacy effects have tended to be found for printed lists, especially those with numerous response options, recency effects have tended to be found for orally administered questions (Krosnick & Alwin, 1987; Schuman & Presser, 1981). Primacy effects in written lists are said to occur when respondents select the first satisfactory response or responses from the list of alternatives, rather than selecting the optimal response (Krosnick & Alwin, 1987). Respondents who are low in cognitive sophistication are more likely than those high in cognitive sophistication to use the satisfying strategy in selecting responses. Recency effects tend to occur from limitations in short-term memory to recall earlier alternatives when no written reminders are present.

Response-order effects resulting from different response strategies may be minimized in self-administered surveys because respondents are likely to have more time to think about answering each question (Bishop et al., 1988). An optimizing strategy might be used by more respondents to a mail questionnaire and this can yield more reliable data than would be obtained by other data collection methods (e.g., telephone survey or personal interview).

Although self-administered surveys may minimize response-order effects, those that remain can result from a number of factors. The complexity of the question responses and the extent to which opinions are defined has been suggested as reasons for response-order effects (Becker, 1954; Rugg & Cantril, 1944; cf. Schuman & Presser, 1981). There are many aspects to complexity, including the number of alternatives, physical presentation (e.g., one column or multiple columns), and the number of responses to be selected (e.g., single or multiple). Each of these can have an effect on the use of a satisfying or optimizing mode to respond to questions.

Order effects also can occur from social desirability, the tendency of some respondents to be chronic “yeasayers” or “nay-sayers” on Likert-type items, and the nature of the question (see Carp, 1974). A social desirability effect can occur in questions asking about attribution of program effects by listing the sponsoring agency in a prominent position (i.e., first or last). For example, Cooperative Extension Service surveys which ask where the respondent obtained educational information might induce a larger number of respondents to select “Extension” from the list of alternatives if that response was listed first, simply because that organization sent the survey. Previous contact between respondents and Extension personnel can foster the development of an exchange relationship (Homans, 1961) in which support for the organization is expected from its clientele in return for the provision of educational information.

The nature of the question also can influence whether or not response-order effects might occur. In a survey of elderly persons on transportation issues, Carp (1974) found response-order effects occurred for evaluative questions consistently, but infrequently for factual questions. An extension of this finding would suggest that questions which include response alternatives that are ambiguous can produce “wording effects” (cf. Bishop et al., 1988) which influence the appearance of response-order effects.

This study examines the influence of response order on single response questions, multiple response questions, and attributive questions to see which types are likely to show order differences. The effect of the number of alternatives listed (from three categories for several questions to nine for another) and whether single or multiple categories are selected for a single question is examined to assess the effects of question complexity. In addition, questions about the attribution of services form a basis to examine response-order bias from social desirability.

**METHODS**

The source of data for this study is a survey conducted by the Florida Cooperative Extension Service in 1986 to evaluate the impact of educational programs concerning the production and management of beef cattle. County Extension Agents obtained the data from producers using mail questionnaires. From their county mailing lists, agents identified 2458 beef producers in nine counties in the Greater Tampa Bay Area, Florida. A random sample of 344 was selected. Using mailing techniques from the “Total Design Method” (Dillman, 1978), 168 usable questionnaires were obtained. This represents at 48.8% response rate. The resulting error rate for a sample of this size is ±7.3% within a 95% confidence interval.

Two survey forms were utilized for response-order experiments. The forms were identical in format and question order. The only differences were the order of response options for 10 questions on the instrument. At
least one question was located on each page of the four-page questionnaire. The forms were randomly assigned to potential respondents by flipping a coin. A total of 76 respondents returned form “A” and 92 respondents returned form “B.” One of the 10 questions was eliminated from the study because too many respondents were screened out by the immediately preceding question. The specific questions for the remaining nine items are presented in the accompanying tables. Five of the questions solicit a single answer and four ask for multiple responses, with one of each type also being an ascriptive question.

The analysis compares the distributions of responses from the two forms using the chi-square statistic. A single chi-square test was applied for questions with mutually exclusive response categories. For questions in which multiple responses could be selected, chi-square was calculated for each category to test the hypothesis of equal proportions.

RESULTS

Of the nine questions about beef management and production practices tested, two suggest that the order of the alternatives influences the distribution of responses. The results for the single response questions are presented first, followed by the multiple response questions, and then the attributive questions.

Single Response Questions

Producers were asked about their plans for their beef operation over the next 5 years. As shown in Table 1, no significant difference is apparent between the two listings of response alternatives. For each of the four alternatives, differences between the first order of listing and the second are 1.6 percentage points or less.

The results from Table 2 for the question in which producers evaluate the importance of making a profit are similar to the first. Producers were asked how important it was for their beef operation to make a profit. Nonsignificant order effects are apparent. Percentage differences are 2.5 or less.

A third single response question was employed when producers were asked about the adequacy of their winter forage program. As shown in Table 3, no significant difference was apparent when the order of the three alternatives was changed, although the percentage who responded “yes” increased 10.8 points when the order of the response was changed from first to second in the list.

The last single response question included in the Beef-Forage survey asked producers to identify the single most important problem facing the beef cattle industry. The results are shown in Table 4. The difference between the distribution of the two listings of responses is not significant. The largest percentage difference is 6.5 points for the “diet/health concern” option, followed by 5.3 percentage points for “price rancher receives.” Neither this question nor the other three single response questions were found to have response-order effects.

Multiple Response Questions

Beef producers also were asked for information about the methods that they used for parasite control (see Table 5). This question was more complex for two reasons. First, the item involving the alternative response

---

1Although the sample size for the two forms differed, the response rate was nearly identical (47.5% for form A and 50.0% for form B). Respondents were found to differ on the number of years in the cattle business, but not on the size of the operation (in acres), herd size, farm structure, off-farm employment, and whether the farm is a major source of income. Logistic regression analysis (not shown here) indicates that differences in longevity in the cattle business between producers receiving form A and those with form B had no statistically significant effect on the findings reported later in the paper.
TABLE 3
RESPONSE-ORDER EFFECTS ON ADEQUACY OF WINTER FEEDING PROGRAMS OF BEEF PRODUCERS

The question read: Do you feel your winter feeding program is adequate for greatest net returns?

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Order of Alternative</th>
<th>A. Yes</th>
<th>B. No</th>
<th>C. Uncertain</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABC</td>
<td>39.2%</td>
<td>50.0%</td>
<td>22.6</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>CAB</td>
<td>33.8%</td>
<td>27.4%</td>
<td>27.0</td>
<td>100.0</td>
</tr>
</tbody>
</table>

chi-square = 1.864, df = 2, n.s.

TABLE 4
RESPONSE-ORDER EFFECTS ON PERCEIVED PROBLEMS FACING THE BEEF CATTLE INDUSTRY

The question read: What do you feel is the single most important problem facing the beef cattle industry?

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Order of Alternative</th>
<th>A. Retail price of beef</th>
<th>B. Price rancher receives</th>
<th>C. Demand for beef</th>
<th>D. Environmental issues</th>
<th>E. Animal welfare issues</th>
<th>F. Governmental regulation</th>
<th>G. Diet/health concerns</th>
<th>H. Zoning issues</th>
<th>I. Other</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABCDEFGHIA</td>
<td>6.0%</td>
<td>47.8</td>
<td>25.4</td>
<td>0.0</td>
<td>0.0</td>
<td>10.5</td>
<td>6.0</td>
<td>1.5</td>
<td>3.0</td>
</tr>
<tr>
<td></td>
<td>CBADGFEA</td>
<td>1.3%</td>
<td>42.5</td>
<td>22.5</td>
<td>1.3</td>
<td>0.0</td>
<td>10.0</td>
<td>12.5</td>
<td>0.0</td>
<td>10.0</td>
</tr>
<tr>
<td></td>
<td>Ghedfcbai</td>
<td>6.9%</td>
<td>42.5</td>
<td>22.5</td>
<td>1.3</td>
<td>0.0</td>
<td>10.5</td>
<td>12.5</td>
<td>0.0</td>
<td>10.0</td>
</tr>
</tbody>
</table>

chi-square = 9.048, df = 7, n.s.

TABLE 5
RESPONSE-ORDER EFFECTS ON PARASITE CONTROL PRACTICES USED BY BEEF PRODUCERS

Respondents who control for external parasites were asked: If yes, which of the following do you use? (check all that apply)

<table>
<thead>
<tr>
<th>Alternatives</th>
<th>Order of Alternatives</th>
<th>A. Ear tags</th>
<th>B. Pour on</th>
<th>C. Medicated feed</th>
<th>D. Other</th>
<th>E. Back rubber</th>
<th>F. Dust bag</th>
<th>G. Spray</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ABCDEFG</td>
<td>25.0%</td>
<td>30.3</td>
<td>7.9</td>
<td>11.8</td>
<td>9.2</td>
<td>25.0</td>
<td>55.8</td>
</tr>
<tr>
<td></td>
<td>CBADGFE</td>
<td>18.0%</td>
<td>36.0</td>
<td>7.9</td>
<td>9.0</td>
<td>13.5</td>
<td>31.5</td>
<td>61.8</td>
</tr>
</tbody>
</table>

chi-square = 1.210, df = 1, .271

* Responses E, F, and G were listed in the second of two columns.
* Percentage selecting that response.

As a follow-up question to how beef producers felt about the adequacy of their forage program (shown earlier in Table 3), they were asked to identify all the types of forages that were used on their farm. Respondents were instructed to check all that applied. Changing the order in which the alternatives were presented resulted in significant differences (p < .05) for one of the six responses (see Table 6). When "native range" was moved from the first position to the fourth, the percentage of respondents who selected that item decreased from 69.7% to 30.4%. There appears to be a primacy effect associated with this response alternative. Although "deferred grazing" remained in the same position on both forms (second in the list of alternatives), the percentage of producers who selected that item increased from 36.8% to 47.8%. This increase is apparently a result of replacing a high response item "native range" with a low response item "silage" in the immediately prior position. Differences among the remaining items are not significant, although the percentage selecting "winter pasture" increased 6.9 percentage points when its position was moved from fifth to third. In the second listing, "winter pasture" was listed prior to "native range" and apparently benefited from this relationship in the same way that "deferred grazing" did. Of note is the stability in the percentage selecting "hay," with 84.2% for the third position (form A) and 79.4%
for the fifth position (form B). Clearly, order has an effect on the overall pattern of responses for this question.

The distribution of responses to another multiple response question is shown in Table 7. Beef producers were asked to identify which protein supplements to their forage were used during the winter months. No significant differences were found in the percentage selecting individual items, although there is a difference of 12.1 percentage points for one of the six alternatives. There does not appear to be a primacy effect on responses to this question, in contrast to that found for winter forage in Table 6.

### Attributive Questions

The first of two attribution questions asked who performed soil tests for beef producers. The University of Florida is listed among the three alternatives (see Table 8). Since the Florida Cooperative Extension Service is housed at the University of Florida, a social acceptability effect might be anticipated. However, only a 3.0% advantage for the first position relative to the third is found for the “University of Florida” response. The presence of response-order bias due to social desirability is not supported by the responses to this question.

The second of the attribution questions asked respondents where or to whom they looked for information on beef management and/or production problems. In this question, the response option “Extension agent/specialist” is a potential source of social acceptability effects and, thus, response-order bias. As shown in Table 9, there is a significant difference. The item of central interest, “Extension agent/specialist,” shows a significance \(p = .026\) primacy effect in the predicted direction. When the position of this option is moved from third to first, the percentage of beef producers who select that response increases from 41.3% to 58.7%. This suggests that a conservative approach be used in positioning response options for questions in which social desirability effects might occur.

### The Influence of Response-Order Effects on Associations

In the previous section, order effects on the distribution of responses for individual questions have been examined. But as issue of greater importance is that of how response-order effects for one variable can result in different relationships with other variables, depending on the order in which response options are listed. The potential to reach substantively different conclusions solely as a result of response order is at issue here.

Given the above analysis, the relationship between using information\(^2\) in selecting replacement herd bulls

---

\(^2\)Information includes weaning weight, blood lines, and other relevant data.
The findings suggest that response-order effects can be an important source of bias in evaluation studies using mail questionnaires. The order of response alternatives was found to affect both the distribution of responses for some items and the associations between these items and others. Reasons for response-order effects were supported by the findings in some cases but not in others. The idea of social desirability appears to warrant further investigation with one of the two questions involving two columns of alternatives. The data indicates that Extension might be perceived to be ineffective in getting farmers to use information in selecting bulls on the basis of being listed as the first response alternative but might be viewed as ineffective when listed third among sources of information. Clearly, there is an order effect that is biasing this relationship.

As shown in Table 10, the order in which the various sources of information are listed appears to create differences in the relationship between using information in selecting replacement bulls and looking to Extension for information. When Extension is listed third among sources of information, there is essentially no relationship between looking to Extension for information and the use of information in selecting replacement bulls. In contrast, looking to Extension for information shows a strong positive effect on using information to select bulls when Extension is listed first among the alternatives. The data indicates that Extension might be perceived to be effective in getting farmers to use information in selecting bulls on the basis of being listed as the first response alternative but might be viewed as ineffective when listed third among sources of information. Clearly, there is an order effect that is biasing this relationship.

### CONCLUSIONS AND DISCUSSION

The findings suggest that response-order effects can be an important source of bias in evaluation studies using mail questionnaires. The order of response alternatives was found to affect both the distribution of responses for some items and the associations between these items and others. Reasons for response-order effects were supported by the findings in some cases but not in others. The idea of social desirability appears to warrant further investigation with one of the two questions involving two columns of alternatives. The data indicates that Extension might be perceived to be ineffective in getting farmers to use information in selecting replacement bulls on the basis of being listed as the first response alternative but might be viewed as ineffective when listed third among sources of information. Clearly, there is an order effect that is biasing this relationship.

Evidence about the effect of complexity on response order also is mixed. Although all the questions were relatively short, no order effects were evident for the question which involved two columns of alternatives. Likewise, the number of alternatives in the list shows no relationship with the presence or absence of order effects. On the other hand, while items in which respondents were to select a single category showed no order effects, items which allowed multiple categories to be selected showed response-order effects in two of four experiments.

One question which demonstrated a clear order effect (for ages used during the winter months, Table 6) suggests that ambiguity in the meaning of the response options to this question may be a cause of bias. If the term “native range” can be applied to a variety of pasture types by respondents, then the order in which this option is presented might influence the probability of selection. The order effect due to the wording of response categories can be as serious as that from the wording of questions (cf. Bishop et al., 1988). Thus, response alternatives need to be clear and matched with terms familiar to potential respondents.

The examination of order effects on associations serves to illustrate the practical application of response order effects and the need for evaluators to concern themselves with this issue. It would seem that we are in danger of incorrectly estimating the impacts of programs by ignoring potential order effects while conducting evaluation studies. As Schuman and Presser (1981, p. 77) note, “greater understanding of these fun-
Can Response Order Bias Evaluations? 371

damental issues is urgently needed if attitude surveys [or evaluations] are to avoid tripping on their own arti-
facts.” In situations where response-order effects are suspected, Payne (1951) recommended using two or more forms in which the response options are alter-
nated. But, the cost of producing multiple forms and bookkeeping efforts can become prohibitive for evalu-
ators who have limited budgets. However, efforts to re-
duce response-order effects can enhance the validity of evaluation studies.

REFERENCES

BECKER, S. L. (1954). Why an order effect. Public Opinion Quar-
terly, 18, 271–278.


DILLMAN, D. A. (1978). Mail and telephone surveys: The total de-


SCHUMAN, H., & PRESSER, S. (1981). Questions and answers in attitude surveys: Experiments on question form, wording, and con-