AN EXPERIMENTAL STUDY OF THE EFFECTS OF LEVEL OF PARTICIPATION ON PROXY REPORTS OF VACATION PLANNING

Barbara Bickart, University of Florida
Johnny Blair, University of Maryland
Seymour Sudman, University of Illinois-Urbana
Geeta Menon, New York University
Barbara Bickart, 205 Matherly Hall, University of Florida, Gainesville, FL 32611

KEY WORDS: Proxy Reporting, Questionnaire Design, Cognition, Response Effects

One way to reduce costs in surveys, both in terms of time and money, is by using one person to report about other household members (proxy reporting). Often, the female head-of-household will serve as the informant, as she is more likely to be at home and/or to answer the telephone. Prior research investigating the accuracy of proxy reporting has produced mixed results (see Moore 1988). The acceptability of proxy reports appears to vary as a function of survey topic (e.g., reports of crime victimization and health care), type of judgment required (subjective versus factual), and knowledge about the other (Moore 1988; Mathiowetz and Groves 1985; Sudman et al. 1991). As Moore (1988) points out, many of the comparisons reported in the literature are difficult to interpret because proxies have been used only when the respondent is unavailable. Thus, reporting and sample biases are intermixed. Further, in most of these studies a validation source for the behaviors and events of interest is not available.

We are currently engaged in a project funded by the National Science Foundation, which looks at the processes respondents use in constructing proxy reports. In earlier research, we examined the effects of level of participation, discussion and confidence on the convergence between reports (Sudman, et al. in press), the strategies used by proxy respondents to answer behavioral frequency questions (Bickart, et al. 1990), proxy respondents' use of anchor and adjustment strategies in answering questions (Bickart, et al. in press; Menon et al., 1990), and gender differences in proxy reporting (Bickart, et al. 1991). Our past research, however, has not dealt directly with the accuracy issue. In this paper, we report the results of a pilot study designed to assess the accuracy of proxy reporting as a function of experimentally manipulated factors. Then, we discuss the implications of this research for survey design and the relationship between memory about others and reporting.

RESEARCH QUESTIONS

The key reason for conducting this research was to address the limitations of using convergence between self and proxy reports as an indication of the quality of proxy reporting. Convergence may be a more appropriate indication of quality for attitudinal items than for factual or behavioral items (cf. Corfman 1991). For behavioral items, both the self and proxy reports could be inaccurate. When this occurs, high convergence would not be meaningful. Ideally, both reports should be compared to a validation source. A key objective of this research was to develop a methodology which allowed us to compare proxy reports to actual behavior, while at the same time manipulating variables of theoretical interest. In this pilot study, we looked at two issues related to peoples' memory about their partners' behavior and how these issues affect the accuracy of proxy reports.

First, we were interested in comparing the accuracy between proxy and self reports for questions of varying specificity. Because knowledge about one's self is greater than knowledge about others, we expected self reports to be more accurate than proxy reports. Further, for self reports we expected reporting to be more accurate for more specific questions. The same may not be true for proxy reports. It is possible that general knowledge structures about one's partner may facilitate inferences made about general behaviors. Information about specific behaviors may not be available, however, impeding one's ability to report accurately about their partner's specific behavior.

Second, we were interested in determining the conditions under which proxy reports are most likely to be accurate. Our previous research indicated that higher knowledge, or levels of participation in the behavior or event, was related to greater convergence between self and proxy reports (Sudman, et al., in press). In this experiment, we see if the accuracy of reports also varies by level of participation. Specifically, we hypothesize that the knowledge a respondent has about his/her partner's

behavior should be positively related to the accuracy of proxy reporting.

METHODOLOGY

The design was developed based on three criteria. First, we wanted to simulate a real-world situation. Second, we needed a mechanism for recording relatively frequent behaviors. These frequencies could then be compared to both self and proxy reports. Finally, we wanted to be able to manipulate the amount of knowledge a person had about his/her partner's behavior. A laboratory experiment in which couples (married or living together as married) searched for information about a vacation they could win met these criteria.

There are three phases in the experiment. First, subjects individually examined vacation options via an interactive computer program (Search Phase), which was developed using the SEARCH MONITOR software package (Brucks 1988; Brucks and Neebe 1988). This program allows subjects to search for information about a particular topic. The software records a variety of information about this task, including the information screens examined, the amount of time spent looking at screens, and the order in which screens were examined. Menus are presented describing four beach locations and three cities. For each location or city, menus can be selected describing accommodations, restaurants, and activities. The format of these options is shown in Exhibit 1. Then, in the Choice Phase of the experiment, couples selected vacation accommodations, restaurants and activities. Finally, subjects individually completed a computer-administered survey about the decision process and about the options both they and their partner examined during the search task (Survey Phase). Exhibit 2 describes the content of these questions.

Sample

Thirty-two couples were recruited to participate in a study on "household decision-making" via newspaper ads and flyers. One couple's data was lost due to a computer error. Thus, the data from 31 couples are used in the analyses to be reported. To encourage participation, couples were entered in a lottery to win a vacation worth $600 that they would plan. Entry in the lottery also increased the relevance and the realism of the task (see Kleppe 1990 for a similar example). Subjects' ratings of the realism of the study and their interest in the task were quite high (on a 7 point scale, with 7 = very realistic/interesting, M's: realism = 5.27; interest = 5.90).

Because of the nature of the incentive, the recruitment procedures, and the demographics of the university community, subjects tended to be younger and had lived together for a shorter period of time than those in the population as a whole. Eighty-one percent of the subjects were younger than 36, and over one-third were younger than 26, while only 8.1% were over the age of 45. Likewise, almost two thirds of the couples (58%) had lived together 3 years or less, while only 23% had lived together for more than 10 years.

Design

Two variables were manipulated in the experiment: (a) the amount of time the couple spent discussing their vacation together during the Decision Phase, and (b) the format in which self and proxy reports were elicited during the Survey Phase. All subjects had a total of ten minutes to decide on their vacation. In the Low discussion condition, however, subjects spent five minutes planning the vacation individually, and five minutes planning with their partner. In the High discussion condition, the couple spent the entire ten minutes planning the vacation together.

The format of the questionnaire also varied. In the all self/all proxy question order condition, all of the questions about subjects' own search behavior were asked prior to questions about their partners'. This created a buffer between the self and proxy questions. In the alternate self/proxy condition, a set of behaviors were first asked about ones' self, followed immediately by similar questions about their partner's behavior. Thus, there was no buffer between the self and proxy items. This manipulation affected only self reports, and will not be discussed further in this paper.

The SEARCH MONITOR software records the number of times a menu was accessed, the order in which menus were accessed and the amount of time subjects spent looking at each menu. This information was used to create a variable indicating the number of options for which information was obtained for places, accommodations, restaurants, and activities, overall and for beaches and cities separately.

Subjects' estimates of both their own behavior and their partner's behavior were compared to actual behavior. In addition, the correlation between subjects' proxy reports and their partner's self report provides an indication of convergence.

RESULTS

The mean number of actual options observed for each item and the mean levels of reporting by self and proxy are shown in Table 1 as well as the relative accuracy of self and proxy reports, as indicated by the ratio of reported behavior to actual behavior. Numbers less than one indicate under-reporting, while numbers greater than one indicate over-reporting. Proportions close to one indicate more accurate reporting. Not surprisingly, the accuracy of self reports was greater than the accuracy of proxy reports for all items. In fact, self reports were highly accurate (over 90%) for reports about all options except activities. For self reports of activities, accuracy improved with item specificity. For example, self reports of total activities were 69% accurate, compared to reports of daytime and evening activities, with accuracies of 95% and 105% respectively.

The results in Table 1 indicate that proxies tended to under-report their partner's behavior. There are several possible reasons for this. First, because of the limited amount of time available for discussion of options, knowledge about the partner's specific behavior may have been limited. Debriefing interviews suggested that most couples began the decision phase by briefly describing the cities or beach areas about which they collected information. Once the couple agreed on a vacation place, however, they often did not discuss specific options (e.g., restaurants or accommodations) about places not considered. Thus, if respondents were using a counting strategy, they would not have complete information available in memory about their partner's behavior and reports would be low. Further, if respondents were using a counting strategy, they may have had difficulty retrieving specific information, even when it was available in memory. Cuing specific places in the question (e.g., beaches or cities) resulted in somewhat higher accuracy for proxy reports. This suggests that the information may have been available, but was difficult to retrieve. Finally, respondents may have used an estimation strategy, but may have used a conservative strategy in making proxy reports. For example, respondents may have used their own behavior as an anchor and then adjusted down to be conservative (Bickart et al. 1991).

Proxy reports about activities were least accurate. Subjects' ratings of the importance of activities in planning a vacation were lower than those for all other vacation attributes except restaurants. The mean importance rating for activities was 4.92 on a 7 point scale (7 = Very Important), compared to 5.68 for location, 5.92 for accommodations, and 4.65 for restaurants. Thus, subjects may not have spent as much time discussing activities they did other options. Further, there was only one screen for each activity, compared to three screens for accommodations and restaurants. Hence, subjects spent less time looking at information about any one activity, relative to accommodations and restaurants. This reduced elaboration may have resulted in poorer recall of activities information.

Table 2 shows the correlations between reports (self and proxy) with actual behavior and the correlation between self and proxy reports, or convergence. Again, self reports were always more highly correlated with behavior than were proxy reports. More importantly, in all cases convergence was higher than the accuracy of proxy reports, although some of these differences are not large. This indicates that there is shared error between the self report and the proxy report. Couples are more similar in their reports about their partner than they are accurate. Thus, convergence appears to be an inadequate measure of accuracy of proxy reports about behaviors.

Finally, Table 3 shows the mean absolute difference between proxy reports and actual behavior as a function of the discussion condition. We expected that proxy reports would be more accurate when discussion was high, and thus this difference would be smaller. This hypothesis was supported, as the mean absolute difference was consistently larger in the low (versus high) discussion condition. Thus, people's ability to report about their partner improved as their knowledge about their partner's behavior increased.

IMPLICATIONS AND FUTURE RESEARCH

In this paper, we present a new method for investigating the accuracy of proxy reports. In a pilot study, we found that the convergence between self and proxy reports tends to overstate accuracy. In other words, convergence tends to be higher than accuracy when a validation source is used. The convergence score appears to contain shared error between the self and proxy report. Thus, similar others tend to be similarly biased in their judgments about another person as that person is him/herself. This finding suggests that we should be cautious about using convergence as an indication of the quality of proxy reports, at least for reports of behavior.

Second, the accuracy of proxy reporting appears to vary with the respondents' knowledge about their partner's behavior. This suggests that we may want to screen potential proxies on discussion about the behavior of interest. Third, the accuracy of proxy reporting appeared to increase with the specificity of the question. Specific questions may provide cues that aid the respondent's recall.

The method used in this experiment is limited in several ways. First, in the pilot study, the sample was not representative of the population as a whole. This limitation could be remedied by recruiting subjects using a probability sample design. Further, when the goal of the research is to make theoretical statements about respondent thought processes, it may be better to have a homogeneous sample (Calder, Phillips, Tybout 1981). A more significant limitation with this method may be differences in response strategies due to the artificial nature of the task. For example, in this pilot study, respondents appeared to use a counting strategy (versus estimation). Other research suggests, however, that behaviors about others are more likely to be estimated than counted (Bickart et al. 1990). Thus, if we are interested in learning how people use general knowledge structures about others to make specific judgments, it may be more realistic to query behaviors after a delay.

Even with these limitations, we feel that this experimental approach can be useful in helping us to understand the cognitive aspects of proxy reporting. As our research program suggests, we feel it is important to use multiple methods in investigating these issues. The experimental approach is especially important for studying cause and effect relationships between memory structure, questionnaire design, and the accuracy of proxy reports. In the future, we plan to use this method to examine how the relationship between the respondent and the target person affects the accuracy of proxy reporting (e.g., roommates versus couples). In addition, we plan to manipulate various aspects of the questionnaire design (e.g., question wording and order) to help us gain insight into these relationships.

ACKNOWLEDGEMENTS

This research has been funded by the National Science Foundation Grant No. SES8821362. We thank Michael Guiry and Stephen Holden for their help in conducting the experiments, and Stanley Presser for helpful comments on an earlier draft of the paper.

REFERENCES

Bickart, Barbara A., Johnny Blair, Geeta Menon, and Seymour Sudman (1990), "Cognitive Aspects of Proxy Reporting of Behavior," in *Advances in Consumer Research*, Vol. 17, eds., M. Goldberg, G. Gorn, and R. W. Pollay, Provo, UT: Association for Consumer Research, 198-206.

Bickart, Barbara A., Geeta Menon, Norbert Schwarz, Seymour Sudman, and Johnny Blair (1991), "Gender Differences in Proxy Reporting," paper presented at the American Association for Public Opinion Research Conference, Phoenix, AZ.

Bickart, Barbara, Geeta Menon, Seymour Sudman and Johnny Blair (in press), "Context Effects in Proxy Reporting," in J. Sherry and B. Sternthal (eds.), *Advances in Consumer Research*, Vol. 19, Provo, UT: Association for Consumer Research.

Brucks, Merry (1988), "Search Monitor: An Approach for Computer-Controlled Experiments Involving Consumer Information Search," *Journal of Consumer Research*, 15 (June), 117-121.

Brucks, Merry and Eloise Neebe (1988), *Search Monitor User's Guide* (Version 4.0).

Calder, Bobby J., Lynn W. Phillips, and Alice M. Tybout (1981), "Designing Research for Application," *Journal of Consumer Research*, Vol. 8 (September), 197-207.

Corfman, Kim P. (1991), "Perceptions of Relative Influence: Formation and Measurement," *Journal of Marketing Research*, Vol. 28 (May), 125-136.

Kleppe, Ingeborg (1990), "Teoretisk versus Subjektiv Beslutningsinnfl Ytelse: En Eksperimentell Studie," unpublished doctoral dissertation, Department of Marketing Economics, University of Bergin.

Mathiowetz, Nancy and Robert M. Groves (1985), "The Effects of Respondent Rules on Health Survey Reports," *American Journal of Public Health*, 75, 639-644.

Menon, Geeta, Seymour Sudman, Barbara A. Bickart, Johnny Blair, and Norbert Schwarz (1990), "The Use of Anchoring Strategies by Proxy Respondents in Answering Attitude Questions," paper presented at the American Association for Public Opinion Research Conference, Lancaster, PA.

Moore, Jeffrey C. (1988), "Self/Proxy Response Status and Survey Response Quality: A Review of the Literature," *Journal of Official Statistics*, Vol. 4, 2, 155-172.

Sudman, Seymour, Barbara Bickart, Johnny Blair, and Geeta Menon (in press), "A Comparison of Self and Proxy Reporting," in Norbert Schwarz and Seymour Sudman (eds.), *Autobiographical Memory and the Validity of Retrospective Reports*, New York: Springer-Verlag.

EXHIBIT 1

Description of Vacation Information

I. Beach Vacations (4 areas)
   A. Accommodations
      1. Room Description & Rates
      2. Restaurants
      3. Facilities & Activities
   B. Restaurants
      1. Menu
      2. Atmosphere
      3. Prices
   C. Activities - Description

II. City Vacations (3 cities)
   A. Accommodations
      1. Room Description & Rates
      2. Restaurants
      3. Facilities & Activities
   B. Restaurants
      1. Menu
      2. Atmosphere
      3. Prices
   C. Daytime Activities - Description
   D. Evening Activities - Description

EXHIBIT 2

Survey Questions

1. Locations (Beach Areas/Cities):
   - Total (#)
   - Beach Areas (# & ID)
   - Cities (# & ID)

2. Accommodations:
   - Total (#)
   - Beach Areas (# & ID)
   - Cities (# & ID)

3. Restaurants:
   - Total (#)
   - Beach Areas (# & ID)
   - Cities (# & ID)

4. Activities:
   - Total (#)
   - Beach Areas (# & ID)
   - Cities (#)
     a) Daytime (# & ID)
     b) Evening (# & ID)

Note: # = Frequency Estimate
ID = Identification

TABLE 1
RELATIVE ACCURACY OF SELF AND PROXY REPORTS

Option	Actual[1]	Self Report[1]	Self Accuracy[2]	Proxy Report[1]	Proxy Accuracy[2]
Total Places	3.90	3.64	.93	3.42	.88
Total Accommodations	8.55	7.59	.89	3.30	.39
Total Restaurants	7.16	7.35	1.03	2.71	.38
Total Activities	9.45	6.51	.69	2.12	.22
# Beach Places	2.35	2.33	.99	1.96	.83
# Beach Accommodations	5.50	6.00	1.09	6.36	1.16
# Beach Restaurants	4.31	4.43	1.03	1.93	.45
# Beach Activities	4.89	4.31	.88	1.67	.34
# City Places	1.55	1.55	1.00	1.23	.79
# City Accommodations	3.05	3.21	1.05	1.81	.59
# City Restaurants	2.85	2.70	.95	1.47	.52
# City Activities	4.56	3.07	.67	1.10	.24
# Daytime Activities	2.58	2.44	.95	1.03	.40
# Evening Activities	1.98	2.08	1.05	1.02	.52

[1]Mean Level of Behavior: Based on n = 62 respondents
[2]Ratio of the reported behavior to the actual behavior

TABLE 2
ACCURACY AND CONVERGENCE OF SELF AND PROXY REPORTS

Option	Accuracy[1] Self Report r	(n)	Proxy Report r	(n)	Convergence[2] r	(n)
Total Places	.73*	(61)	.36*	(50)	.37*	(50)
Total Accommodations	.59*	(61)	.05	(60)	.12	(60)
Total Restaurants	.58*	(60)	-.08	(59)	-.06	(58)
Total Activities	.05	(59)	.08	(59)	.21**	(57)
# Beach Places	.86*	(62)	.46*	(50)	.59*	(50)
# Beach Accommodations	.79*	(61)	.40*	(61)	.44*	(61)
# Beach Restaurants	.83*	(61)	.17**	(61)	.23*	(61)
# Beach Activities	.60*	(61)	.21**	(60)	.31*	(59)
# City Places	.84*	(62)	.28*	(56)	.33*	(56)
# City Accommodations	.76*	(61)	.12	(62)	.22*	(61)
# City Restaurants	.59*	(61)	.37*	(62)	.47*	(61)
# City Activities	.61*	(61)	.34*	(60)	.48*	(59)
# Daytime Activities	.68*	(61)	.11	(61)	.22*	(60)
# Evening Activities	.63*	(62)	.20	(62)	.49*	(62)

[1]Correlation between report and actual behavior.
[2]Correlation between self and proxy report.
*r > 0, p < .05.
**r > 0, p < .10.

TABLE 3
MEAN ABSOLUTE DIFFERENCE BETWEEN PROXY REPORT AND ACTUAL BEHAVIOR BY DISCUSSION

Option	Low Discussion $\bar{x}$	(Std)	High Discussion $\bar{x}$	(Std)
Total Places	1.70	(1.29)	1.74	(1.45)
Total Accommodations[1]	7.72	(3.81)	5.06	(4.46)
Total Restaurants	6.68	(4.99)	5.52	(4.55)
Total Activities	9.11	(7.92)	6.97	(5.52)
# Beach Places	.88	(1.00)	1.00	(1.18)
# Beach Accommodations	4.67	(2.71)	4.00	(2.78)
# Beach Restaurants[1]	4.57	(3.41)	3.19	(2.68)
# Beach Activities	4.41	(4.43)	4.10	(3.82)
# City Places	1.14	(1.25)	.78	(.97)
# City Accommodations[1]	3.67	(4.06)	1.97	(2.79)
# City Restaurants	2.40	(2.88)	2.31	(3.05)
# City Activities[2]	4.54	(4.70)	3.06	(4.08)
# Daytime Activities[2]	3.03	(3.34)	1.84	(2.69)
# Evening Activities[1]	2.53	(2.65)	1.38	(1.86)

[1]Low < High, p ≤ .05
[2]Low < High, p ≤ .10