William Locander, University of Houston Seymour Sudman, University of Illinois, Urbana-Champaign Norman Bradburn, University of Chicago

Introduction

Response effects in surveys are a major concern to users and gatherers of survey data and have received wide attention in the literature. Sudman and Bradburn list more than 900 articles in their recent book <u>Response Effects in Surveys</u>. (35, 36) Two of the most important variables that have been studied are question threat (6, 12, 13, 14, 15, 17, 21, 26) and method of administration. (4, 5, 7, 8, 9, 10, 18, 20, 22, 23, 25, 28, 29, 31, 32, 33, 34, 38, 40, 42, 43, 44, 45) Usually these have not been studied jointly, nor has there been any validation of response.

The purpose of this study was to examine the joint effects of question threat and method of administration on response distortion. Another major objective was to study the randomized response model which has been described in several recent articles (2, 3, 11, 16, 19, 30, 41) as a technique to reduce or completely eliminate response distortion of threatening or personal questions. A brief discussion of this method is given below.

Random Response Model

The randomized response model was developed by Warner (41) as a technique to reduce response distortion of threatening or personal questions. By using a probability mechanism, the respondent answers one of two questions selected randomly, but the interviewer does not know what question was answered. Generally, the respondent would answer one of the following questions "Yes" or "No":

I am a member of Group A.

I am not a member of Group A.

By knowing the probability of answering each question, the sample size, and the total number of "yes" replies; the true proportion of the population that are members of Group A can be estimated. Warner feels that the potential advantages of the technique depend on the actual cooperation that is achieved by the model. Warner notes that the randomized response technique can be used to estimate distributions other than a simple dichotomous variable. For example, estimating the proportion of a population in particular income classes can be accomplished by asking the respondent to make five separate randomized responses of whether or not he is in each of the five classes.

Horvitz, Shah, and Simmons (19) suggest that the technique developed by Simmons of using unrelated questions in the randomized response model is a valuable modification. It is felt that this will help to overcome the respondent's suspicions and thus increase cooperation. With this modification, one question should be threatening and the other innocuous and unrelated. The questions should read as follows:

> I am a member of Group A. I am a member of Group B.

With this method, the unrelated question data are treated as a separate sample for estimating purposes. As long as the probabilities are not equal, the sample estimate can be obtained in the following manner:

Given:

- π_A = true proportion with attribute A
- P₁ = probability that the statement "I am a member of Group A" is answered by the respondent
- P₂ = probability that the statement "I am a member of Group B" is answered by the respondent
- π_v = true proportion with attribute y
- λ_1 = proportion of "yes" answers

The sample estimate can be determined by:

$$\pi_{y} = \frac{\lambda_{1} - \pi_{y} (1 - P_{1})}{P_{1}}$$

Design of the Study

The study looks at four interview techniques: face-to-face, telephone, self-administered, and the random response model. Figure 1 shows the design including the four levels of threatening questions. It was planned that there would be fifty respondents per cell and therefore a total sample size of eight hundred.

The threat dimension includes questions about the ownership of a Chicago Public Library card, voter registration and voting behavior, involvement in bankruptcy, and being charged with drunken driving. These four topics were chosen because a priori we believed that the level of threat increases as one goes from a question on having a library card to one on being charged with drunken driving. In addition, for these questions it was possible to obtain validation information from public records. Thus, in the results it is possible not only to see what differences there are by method of administration and threat, but also the actual response error.

The respondents in the face-to-face bankruptcy cell had all declared bankruptcy in the recent past. The respondents in the drunken driving cells had all been charged with drunken driving in a time period not less than 6 months or more than 12 months from the date of the study. The respondents in the library card and voting behavior cells were drawn from a random sample and validated from Chicago Public Library and city voting records. This was done after the questionnaire was administered. We ignore in this experiment possible errors in lists of library card holders and voters such as mis-filing or misspelling of names. There is no reason to believe that such errors would be related to method of

It was recognized that our a priori judgement of threat might not be the same as the respondent's. Thus, while we thought admission of bankruptcy to be highly threatening because it is viewed as a personal failure, some people could see it as a shrewd business tactic to alleviate debt. After the main part of the interview was completed, respondents were asked about how threatening they found the questions. The responses to these questions were combined to form an acute anxiety scale. A measure of chronic anxiety was also obtained so that response effect could be related to chronic, acute and total anxiety. Chronic anxiety was measured by the Bendig Short Form (1) of the Taylor Manifest Anxiety Scale (37).

The interviewing was done in Chicago by interviewers trained and supervised by the National Opinion Research Center of the University of Chicago. Interviewer assignments were randomized over the different methods of administration, but interviewers were matched on race.

Interview Completion Rates

The rate of completed interviews varied by method and group. Table 1 shows the percent completed by method and sample. The overall completion rate was 72.2% of 941 interviews. This is about average for a sample in a large city. Use of the telephone achieved the highest interview completion rate. It did better than the other three methods across all sample types except in the bankruptcy sample when comparing it to personal interviewing.

Table 1

Percent of Completed Interviews by Method of Administration and Sample

	Sample				
Methods	General		Drunken		
	Samp1e	Bankrupts	Drivers	Total	
Personal	76.0	70.3	57.1	67.8	
Interview	n=125	n=54	n=63	n=242	
Telephone	89.9	68.3	77.8	76.6	
-	n=109	n=60	n=63	n=232	
Self	75.4	59.3	47.5	60.7	
Administered	n=114	n=59	n=61	n=234	
Random Response	77.6	67.2	58.1	67.6	
-	n=116	n=55	n=62	n=233	
TOTAL	79.7 n=464	66.2 n=228	60.1 n=249		

It is interesting to note that the telephone was relatively more successful in getting completed interviews with drunken drivers. The selfadministered technique where the interviewer left a questionnaire and picked it up later did about as well as personal interviewing and the random response model in the general sample. However, self administration did not do nearly as well for the bankruptcy and drunken driving samples where a large fraction of the respondents had not finished high school. The random response method achieved completion rates similar to personal interviewing.

It was much more difficult to locate the bankrupts and drunken drivers and this factor was the major source of not completing interviews with these groups. About 90 per cent of the non-interviews with bankrupts and 80 per cent of noninterviews with drunken drivers were due to the interviewers inability to locate the respondent.

RESULTS

The main findings of the study are shown in Table 2. The data are presented as proportions of distortion for each of the twenty cells representing different conditions. The results represent the proportion of respondents in each condition who gave distorted answers. The proportion of distortion in each cell is defined as the absolute value of the difference divided by the total sample size:

$$Distortion = \frac{Response - Validated}{Total N}$$

Table 2 is organized so that the threat dimension increases for most method conditions from low to high distortion.

In all cases, except the random responsebankruptcy cell (.00), the proportions of distortion increase as threat increases. The threat dimension has been reordered because the voter registration question was least distorting across all four methods. The library card question appeared to be more threatening than voter registration. Voting in the primary election had a higher rate of distortion than the bankruptcy question in all method conditions. (Some of this may have been due to errors in the list of primary voters.) The drunken driving question, with the exception of the random response cell, had the highest distortion rate.

Looking down the columns of Table 2, the method treatment findings are not as clear cut. At the low end of the threat dimension (voter registration and library card), the self-administered method does have a lower distortion rate than the two more personal methods (face-to-face and telephone interviewing). As question threat increases, self-administration does not continue to reduce distortion relative to the other two methods. In the drunken driving cell, self-administered forms had a higher bias than the more personal methods. The fact that the self-administered technique did no better in reducing bias for threatening questions than the other methods is in keeping with Cannell and Fowler's (5, p.254) findings. They found that comparisons of self-administered and personal interviewing data showed no method effect in the reporting of threatening material.

The relation between personal interviewing and telephoning is also interesting. At low threat levels, the distortion on the face-to-face interview is slightly lower than on phone interviews. As the questions become more threatening, face-to-face interviews have larger errors than phone interviews. This supports, although weakly, Hyman's theory that the degree of social



Impersonal

n = 800

Figure 1. Study design--question threat by method of administration

	Threat				
	Voter	Library	Bank-	Vote	Drunken
Methods	Registration	Card rupto	ruptcy	Primary	Driving
Face-to-face	.15	.19	.32	.39	.47
	n=92	n=93	n=38	n=80	n=30
Telephone	.17	.21	.29	.31	.46
	n=89	n=97	n=41	n=77	n=46
Self-Administered	.12	.18	.32	.36	.54
	n=80	n=82	n=31	n=74	n=28
Random Response	.11	.26	.00	.48	.35
	n=61	n=61	n=26	n=50	n=23

Proportion of Distorted Responses

involvement or "physical presence" of the interviewer can contribute to response distortion. As the interviewer is removed to a telephoning situation and the questions become more personal, the physical absence of the interviewer tends to reduce social involvement and telephoning becomes less distorting than personal interviewing.

The random response model tended to produce higher variances across threat treatments. The range of results ran from (.00) distortion in the bankruptcy cell to .48 in the voting cell. In the highest threat condition, random response yielded the lowest rate of distortion. However, in the March primary voting question the model yielded the greatest distortion. In a somewhat threatening bankruptcy question, the distortion was zero.

The bankruptcy variable is unique in that part of the sample might not have perceived the question as threatening but rather an opportunity to tell of a shrewd business maneuver. This point causes difficulty in evaluating the raw proportions. However, in the drunken driving condition, the model was the lowest in response bias. Initially as one looks across the treatments, it is difficult to evaluate the model's performance in total. The random response technique did not, however, remove <u>all</u> error from the responses.

Another way of looking at the data in Table 2 is to note that response errors may be due either to over-reporting a socially desirable act such as owning a library card, being registered to vote or voting in the primary election, or under-reporting a socially undesirable act such as being involved in a bankruptcy proceeding or being charged with drunken driving. Generally, admitting a socially undesirable act is more threatening, but Table 1 indicates that not voting in the primary election was considered more threatening than having declared bankruptcy.

It is evident that random response procedures are least effective in reducing overreporting of socially desirable acts. If anything, they are even worse than face-to-face interviews. (The randomized response questions were, of course, asked in the context of a faceto-face interview.) Self-administered and telephone interviews work best on reducing overstatements. The same results were found by Weiss (43).

On the other hand, randomized response procedures are most effective in reducing underreporting of socially undesirable acts while selfadministered forms are least effective, with faceto-face and telephone methods in the middle.

In order to test the statistical significance of the results in Table 2, an analysis of variance was run on the raw data. Table 3 presents the results of these computations. It can be seen that the threat treatment was significant at the .01 level. The method effect, however, was not significant.

Table 3

Analysis of Variance of Proportional Data

Source	Sum of Squares	df	Mean Square	F Ratio
Treatment A				
(method)	.75	3	.25	.516
Treatment B				
(threat)	13.40	4	3.35	5.41*
Residual	5.808	12	.484	

Significant at .01 level

With only one observation per cell it is impossible to determine within cell variance. Kirk (24 p.227) outlines a procedure for handling ANOVA with n=1. Tukey's (39) test procedure for one degree of freedom was modified by Kirk for a factoial design experiment. If Tukey's F-test for nonadditivity is insignificant, the interaction term may be used to test the treatment effects. The interaction in the data for this study was not significant.

The data were also transformed using an arcsin transformation. (Winer (46), pp. 399-400) An analysis of variance was run on the transformed data to see if the results differed from those reported in Table 3. Again the method effect was not significant and the threat effect was significant at the .01 level.

Random Response Model

In addition to the results of Table 2, there are other observations that one might make about this relatively new technique. Generally the model was well received by both interviewer and respondent. During the course of the interview, only 5% of the sample who used the random response model said it was confusing, silly, or unnecessary.

The interviewers were asked to evaluate each respondent's reaction to the random response box. Table 4 shows the results of each question by sample type. The general sample and bankrupts appear to follow the same distribution. However, the percentages of "yes" responses for the drunken drivers drops off somewhat.

Chronic and Acute Anxiety

Chronic anxiety was significantly related to response distortion for personal interviewing and self-administration but not for telephoning. The telephone method appeared to be more stable than the other methods across anxiety groups.

One of the principal reasons for taking acute measures was to validate the threat dimension. Generally, the dimension was validated with library card and voting behavior at the low end and court and traffic questions at the high end. Response distortions were significantly different between the high and low acute groups. The acute effect was significant at the .05 level but method of administration proved to be insignificant. There were significant differences between high and medium acute groups with personal interviewing and between high and low acutes using self-administered forms. Telephoning did not show any significant differences between groups that differed in acute anxiety. The distortion rate for the telephone method was also the lowest across acute groups. The random response model produced acute anxiety scores that were generally higher than or close to personal interviewing. Further discussion of these relations is found in Locander (27).

DISCUSSION AND IMPLICATIONS

It is clear from the findings of this experiment that no data collection method is superior to all other methods for all types of threatening questions. If one looks for significant differences, then none of the methods differ among themselves and one is free to use whatever procedure is most convenient.

If one accepts the results at face value, then each of the data-gathering methods is best under certain conditions. The randomized response procedure gives the lowest distortion on threatening questions asking about the performance of socially undesirable acts. It is obvious, however, that one does not always obtain unbiased answers using random response models. The 35 per cent understatement of drunken driving is still a major response bias, although lower than for other methods. The use of randomized response procedures does not permit any multivariate analysis of the relation between the threatening question and independent variables, unless very large samples are screened. For many uses, the loss of information from using randomized response would not be compensated by a modest reduction in response bias.

Self-administered procedures are slightly better than other methods for reducing the overstatements on questions asking about performance of socially desirable acts, but are worst on questions that ask about undesirable acts. In addition, the cooperation rate is lowest for self-administered questionnaires.

There do not appear to be any meaningful differences in response bias between telephone and face-to-face interviews in this experiment except that, for this large city sample, the

Table 4

Interviewer's	Answers	to Qu	estions	About	the	
Random Response Model						

		Response			
Question	General Sample		Bankrupts	Drunken Drivers	
 Do you think the respon- dent understood the use of the random response box? 	Yes	90 %	89.2%	78.4%	
	No	8.9%	5.4%	18.9%	
	DK	1.1%	5.4%	2.7%	
2) Do you think the respon-					
dent accepted the ex- planation of the box	Yes	92.2%	89.2%	78.4%	
and believed that his/ her answers really were private?	No	6.7%	5.4%	5.4%	
	DK	1.1%	5.4%	16.2%	
		n=90	n=37	n=37	

sample cooperation was highest by telephone. This study again indicates the usefulness of telephone procedures especially in metropolitan areas.

Ultimately the conclusion is that highly threatening questions have high response biases that are not greatly affected by the way in which the question is asked, even if privacy is preserved.

- Bendig, A.W. "The Development of a Short Form of the Manifest Anxiety Scale." Journal of Consulting Psychology, 20, 1956, 384.
- Boruch, Robert F. "Assuring Confidentiality of Responses in Social Research: A Note on Strategies." <u>The American Sociologist</u>, 6, November, 1971, 308-31.
- Boruch, Robert F. "Relations Among Statistical Methods for Assuring Confidentiality of Social Research Data." <u>Social Science Re-</u> <u>search</u>, 1972, Vol. 1, 4, 403-413.
- Cahalan, D. "Measuring Newspaper Readership by Telephone: Two Comparisons with Face-to-Face Interviews." Journal of Advertising Research 1, December 1960, 1-6.
- Cannell, C.F. and Fowler, F.J. "A Comparison of a Self-Enumerative Procedure and a Personal Interview: A Validity Study." <u>Public Opinion Quarterly</u>, 27, 1963, 250-64.
- Clark, J., and Tifft, L. "Polygraph and Interview Validation of Self-Reported Deviant Behavior," <u>American Sociological Review</u>, 31, 1966, 516-23.
- Colombotos, J. "Personal versus Telephone Interviews: Effect on Responses." <u>Public</u> <u>Health Reports</u>, 84, 1969, 773-782.
- Dohrenwend, B.S. "An Experimental Study of Directing Interviewing," <u>Public Opinion</u> <u>Quarterly</u>, 34, 1970, 117-125.
- Ellis, A. "Questionnaire versus Interview Methods in the Study of Human Love Relationships." <u>American Sociological Review</u>, 12, 1947, 541-553.
- Enterline, P.E., and Capt, K.G. "A Validation of Information Provided by Household Respondents in Health Surveys." <u>American</u> <u>Journal of Public Health</u>, 49, <u>1959</u>, 205-212.
- Eriksson, Sven A. "A New Model for Randomized Response," <u>International Statistics Review</u>, 1973, Vol. 41, 1, 101-113.
- 12. Ferber, R. "On the Reliability of Responses Secured in Sample Surveys." Journal of the <u>American Statistical Association</u>, 50, 1955, 788-810.

- Ferber, R. The Reliability of Consumer Reports of Financial Assets and Debts. Studies in Consumer Savings, No. 6, Urbana, Ill.: Bureau of Economic and Business Research, University of Illinois, 1966.
- 14. Ferber, R., Forsythe, J., Guthrie, H.W., and Maynes, E.S. "Validation of a National Survey of Consumer Financial Characteristics: Saving Accounts." <u>Review of Economics and</u> Statistics, 51, 1969, 436-444.
- Gray, P.G. "The Memory Factor in Social Surveys." Journal of the American Statistical Association, 50, 1955, 344-363.
- 16. Greenberg, B., Abul-Ela, A., Simmons, W., and Horwitz, D. "The Unrelated Question Randomized Response Model: Theoretical Framework." Journal of the American Statistical Association, 64, 1969, 520-539.
- Hagburg, E.D. "Validity Questionnaire Data: Reported and Observed Attendance in an Adult Education Program." <u>Public Opinion Quarterly</u>, 32, 1968, 453-456.
- Hochstim, J.R. "A Critical Comparison of Three Strategies of Collecting Data from Households." Journal of the American Statistical Association, 62, 1967, 976-989.
- 19. Horwitz, D.G., Shah, B.V. and Simmons, Walt R. "The Unrelated Question Randomized Response Model." 1967 Social Statistics Section Proceeding of the American Statistical Association, 67-72.
- 20. Hyman, H. <u>Interviewing in Social Research</u>. Chicago: <u>University of Chicago Press</u>, 1954.
- 21. Ito, R. "An Analysis of Response Error: A Case Study." <u>Journal of Business</u>, 36, 1963, 440-447.
- 22. Jackson, R.M., and Rothney, J.W.M. "A Comparative Study of the Mailed Questionnaire and the Interview in Follow-up Studies <u>Personnel and Guidance Journal</u>, 39, 1961, 569-571.
- 23. Kahn, R.L. "A Comparison of Two Methods of Collecting Data for Social Research: The Fixed-Alternative Questionnaire and the Open-ended Interview." Unpublished Ph.D. dissertation, University of Michigan, 1952.
- 24. Kirk, Roger E. Experimental Design: Procedures for the Behavioral Sciences, Belmont, California, Brooks/Cole Publishing Company, 1968, 227.
- 25. Knudsen, D.D., Pope, H., and Irish, D.P. "Response Differences to Questions on Sexual Standards: An Interview-Questionnaire Comparison." <u>Public Opinion Quarterly</u>, 31, 1967, 290-297.

- 26. Lansing, J.B., Ginsburg, G.P., and Braaten, K. <u>An Investigation of Response Error</u>. Studies in Consumer Savings, No. 2. Urbana, Ill.: Bureau of Economic and Business Research, University of Illinois, 1961.
- Locander, William, <u>An Investigation of</u> <u>Interview Method</u>, <u>Threat</u>, <u>and Response</u> <u>Distortion</u>, <u>Unpublished</u> Doctoral Dissertation, Urbana, Illinois, University of Illinois, 1974.
- McCord J., and McCord W. "Cultural Stereotypes and the Validity of Interviews for Research in Child Development." <u>Child</u> <u>Development</u>, 32, 1961, 171-185.
- Mooney, H.W., Pollack, B.R., and Corsa, L. "The Use of Telephone Interviewing to Study Human Reproduction." <u>Public Health Reports</u>, 83, 1968, 1049-1060.
- 30. Moors, J.J.A. "Optimization of the Unrelated Question Randomized Response Model." Journal of American Statistical Association, September 1971, 66, 627-29.
- 31. Price, D.O., and Searles, R. "Some Effects of Interviewer-Respondent Interaction on Responses in a survey Situation." Proceedings of the American Statistical Association, Social Statistics Section, 1961, 211-21.
- Schmiedeskamp, J.W., "Reinterviews by Telephone." Journal of Marketing, 26, January 1962, 28-34.
- Suchman, E.A., and Phillips, B.S. "An Analysis of the Validity of Health Questionnaires." <u>Social Forces</u>, 36, 1958, 223-232.
- 34. Sudman, S., Greeley, A.M., and Pinto, L. "The Effectiveness of Self-Administered Questionnaires." Journal of Marketing Research, 2, 1965, 293-297.
- 35. Sudman, Seymour. "Toward a Theory of Response Effects in Survey Research." Proceedings of the American Statistical Association, Social Statistics Section, 1970, 14-16.

- 36. Sudman, Seymour and Bradburn, Norman, <u>Response Effects in Surveys</u>, Chicago: <u>Aldine Press</u>, 1974.
- Taylor, Janet A. "A Personality Scale of Manifest Anxiety." Journal of Abnormal and Social Psychology, 48, 1953, 285-290.
- Tittle, C.R., and Hill, R.J. "The Accuracy of Self-Reported Data and Prediction of Political Activity." <u>Public Opinion</u> Quarterly, 31, 1967, 103-106.
- Tukey, John W. "One Degree of Freedom for Non-Additivity." <u>Biometrics</u>, 1949, 5, 232-242.
- 40. Walsh, W.B. "Validity of Self-Report." Journal of Counseling Psychology, 14, 1967, 18-23.
- 41. Warner, Stanley, L. "Randomized Response: A Survey Technique for Eliminating Error Answer Bias." Journal of the American Statistical Association, 60, 1965, 63-69.
- Wedell, C., and Smith, K.V. "Consistency of Interview Methods in Appraisal of Attitudes." Journal of Applied Psychology, 35, 1951, 392-396.
- Weiss, C. "Validity of Welfare Mothers' Interview Responses." <u>Public Opinion Quar-</u> terly, 32, 1968-69, 622-633.
- 44. Williams; J.A., Jr. "Interviewer-Respondent Interaction." Sociometry 27, 1964, 338-352.
- Williams, J.A., Jr. "Interviewer Role Performance: A Further Note on Bias in the Information Interview." <u>Public Opinion</u> Quarterly, 32, 1968, 287-294.
- Winer, B.J. <u>Statistical Principles in Experimental Design</u>, McGraw-Hill, New York, 1971.