

THE CORRELATES OF ITEM NONRESPONSE IN A LARGE SCALE MAIL SURVEY

Donald A. Hughes, Sears, Roebuck and Co.

Item nonresponse can be a particularly serious problem in mail surveys where there is no interviewer involved who can urge the respondent to complete the entire questionnaire and to write in "Don't Know" or "Refused" in those cases where the respondent completing a mail questionnaire might leave a blank. Also, as the author has shown in a previous paper (1) there is a propensity on the part of respondents in mail surveys to indicate a "zero" by leaving the answer space blank for those questions requiring a precise numerical answer.

PAST RESEARCH

Previous work in this area by Ferber (2) and Craig and McCann (3) showed age, sex, occupation, and education to be important correlates of item nonresponse. Of these, age was the most consistent predictor of item nonresponse with the correlation being positive.

METHOD

A mail survey was conducted simultaneously in 27 separate geographic markets (roughly equivalent to SMSA's) in 1977. The questionnaires were identical, except that one question concerning frequency of shopping at selected outlets was customized for each market and the number of stores listed varied from a low of ten in one market to a high of 22 in several others. While Craig and McCann could find no indication that length of questionnaire influenced item nonresponse, the variation in the number of answers required of a respondent in this survey to a single question gave us the opportunity to investigate the impact of increasing a list within a single question on the questionnaire.

Another series of questions in this survey listed 20 outlet attributes and asked that the respondent indicate which outlet they associated most closely with each attribute. The list of attributes was constant across all 27 markets.

Two step-wise linear regressions were conducted. In the first, the total item nonresponse rate for the shopping frequency question in each of the 27 markets was used as the dependent variable. This item nonresponse rate varied from 18.4% to 41.4%. In the second step-wise linear regression, the total item nonresponse rate for the store attribute series of questions was used as the dependent variable. This rate varied from 20.6% to 30.1%.

The independent variables used were:

<u>Variable</u>	<u>Range Over 27 Markets</u>
1. Number of Outlets Listed on Shopping Frequency Portion of the Questionnaire	10-22
2. Return Rate for the Survey	32.2%-54.6%

<u>Variable</u>	<u>Range Over 27 Markets</u>
3. Dwelling is a Single Family House	63.5%-90.3%
4. Own Home	62.8%-86.4%
5. Married	69.8%-86.9%
6. Children Under 21 in Home	39.9%-62.8%
7. Respondent Not Employed	44.4%-65.4%
8. Chief Wage Earner in White Collar Occupation	47.7%-70.1%
9. Head of House Retired	7.0%-31.2%
10. Head of House 55 or Over	22.5%-42.1%
11. Median Household Income	\$12,950-\$19,980
12. Have a Sears, Penney or Wards Catalog	41.4%-79.3%
13. Have a Major Charge Account	75.3%-85.7%
14. Respondent is Male	35.3%-43.6%

The first step-wise linear regression, using the omission rate for the shopping trip question as the dependent variable, produced the following results:

<u>Step</u>	<u>Independent Variable</u>	<u>Correlations (R)</u>	<u>Multiple R<sup>2</sup></u>
1.	Number of Outlets Listed on Shopping Frequency Section of Questionnaire	+0.8364	0.700
2.	Median Household Income	-0.0639	0.864
3.	Return Rate for the Survey	-0.5033	0.893
4.	Percent of Respondents Having Any Major Charge Account	-0.3291	0.908

The F Ratio to enter was set at 2 and no other variable met that requirement.

The second step-wise linear regression, using the item nonresponse rate for the store attribute series of questions as the dependent variable, produced these results:

<u>Step</u>	<u>Independent Variable</u>	<u>Correlations (R)</u>	<u>Multiple R<sup>2</sup></u>
1.	Percent of Respondents Age 55 and Over	+0.7289	0.531
2.	Return Rate for the Survey	-0.3445	0.738
3.	Median Household Income	-0.5320	0.772

Step	Independent Variable	Correlations (R)	Multiple R <sup>2</sup>
4.	Number of Outlets Listed on Questionnaire	+0.3758	0.848
5.	Percent of Respondents Having a Sears, Penney or Wards Catalog	-0.4852	0.863
6.	Percent of Respondents Who are Married	-0.0058	0.877

No other variable met the requirement of an F ratio of 2 to enter. (See Exhibit No. 1 for the complete correlation matrix covering both regressions.)

#### DISCUSSION

The rate of item nonresponse for the shopping frequency question was heavily influenced by the number of outlets the respondents had to consider in searching their memories for recall of shopping activity. As the task increased, so did the item omission rate. Previous work cited by Craig and McCann indicated that Activity, Interest and Opinion (AIO) type questions generally had low item omission rates, even though these questions normally number far more than the upper limit of 22 for the number of outlets for which shopping frequency was requested. It would appear that a series of questions, such as AIO statements, which are varied, require little thought in answering and can be answered by checking a box or circling a number, can be presented in great numbers in a questionnaire without fear of extensive item nonresponse. However, questions requiring memory search and the writing in of a precise number will produce ever higher rates of item omission as the number of questions in the series increases. This would reinforce Ferber's finding that:

"Evidently, the greater thought required to answer the ---- questions did contribute to higher item nonresponse rates." (2 - page 408)

Median Household Income surfaced as a significant correlate in both regressions. Since educational levels were not asked in this survey, income is probably a rough surrogate for educational attainment.

Age was a powerful factor in item nonresponse for the series of outlet attribute questions, which is no surprise in view of findings in previous research cited. However, the impact of the overall survey return rate on item nonresponse was a new development. Other studies cited have dealt with a single universe where the return rate was a constant. In the surveys on which this paper is based, we have 27 universes with varying return rates. It is apparent that geographic areas which produce relatively low return rates also produce respondents who do a lower than average job of completing the questionnaire.

It is also of interest that the number of outlets

listed in the first part of the questionnaire had an influence on the level of item nonresponse in the subsequent series of questions concerning store attributes. Since the question was answered by writing in the identifying number of the store from a list in the front of the questionnaire, cities with longer lists of stores presented a more difficult task to the respondent.

#### CONCLUSIONS

The findings of this study reinforce those of Ferber concerning the tendency of item nonresponse to increase with age. Surveys of publics largely made up of senior citizens are, therefore, best conducted through telephone or in-home interviewing.

The fact that geographic areas which yield low total response to mail surveys also yield substandard item completion rates suggests caution in the use of mail surveys where the publics have a history of low cooperation in surveys. Myatt (4), in an unpublished paper, gives good evidence that response rates to mail surveys vary inversely from one city to another in proportion to the presence of minority groups in the city and to the size of the city.

Item nonresponse will vary within a questionnaire depending on several factors.

- a. A series of questions requiring memory search will have an increasing proportion of item nonresponse as the length of the series increases.
- b. A series of questions calling for a precise numerical response (and where a blank may be considered by the respondent as an adequate substitute for a zero) will have a high nonresponse rate.
- c. A series of questions asking for an expression of opinion by the respondent will suffer high item nonresponse rates where many respondents are indifferent to the item subject (e.g. credit terms).
- d. A long series of questions that can be answered by checking off alternative answers from a supplied list will suffer little item nonresponse provided the questionnaire does not induce boredom.

Finally, the growing popularity of mail surveys suggests that greater effort to design for minimal item nonresponse and to better analyze the meaning of those item nonresponses with which we must inevitably live, should be a major effort in the survey research community.

References

1. Hughes, Donald A., "A Statistical Approach to the Analysis of Item Omission in a Mail Survey," American Statistical Association 1978 Proceedings of the Section of Survey Research Methods, pp. 594-98.
2. Ferber, Robert, "Item Nonresponse in a Consumer Survey," Public Opinion Quarterly 30, (Fall 1966), pp. 399-415.
3. Craig, C. Samuel and John N. McCann, "Item Nonresponse in Mail Surveys: Extent and Correlates," Journal of Marketing Research XV, (May 1978), pp. 285-9.
4. Myatt, Larry H., An unpublished paper given at the American Statistical Association, Chicago Chapter Spring Conference, March 21, 1978.

Exhibit No. 1  
CORRELATION MATRIX

	Shopping Trip Nonresponse	Store Attri- bute Non- response	No. of Out- lets on Quas.	Survey Return Rate	% Live in House	% Own Home	% Married	% Children <21 at Home	% Respondent Not Employed	% White Collar	% Retired	% 55 + Over	Median Income	% Have Catalog	% Have Charge Account	% Male
Shopping Trip Nonresponse	1.00	-														
Store Attri- bute Non- response	-	1.00														
No. of Out- lets on Quest.	0.84	0.38	1.00													
Survey Return Rate	-0.50	-0.34	-0.49	1.00												
% Live in House	-0.12	-0.18	-0.07	0.24	1.00											
% Own Home	-0.03	0.10	0.12	0.25	0.56	1.00										
% Married	0.25	-0.01	0.38	-0.03	0.46	0.52	1.00									
% Children <21 at Home	-0.03	-0.53	0.21	-0.16	0.47	0.06	0.51	1.00								
% Respondent Not Employed	0.15	0.34	-0.07	0.19	0.38	0.36	0.19	-0.10	1.00							
% White Collar	0.01	-0.28	0.23	-0.43	-0.16	-0.37	-0.15	0.30	-0.53	1.00						
% Retired	0.25	0.65	-0.06	0.25	-0.11	0.41	-0.03	-0.66	0.53	-0.70	1.00					
% 55 + Over	0.33	0.73	0.01	0.15	-0.02	0.40	0.05	-0.64	0.60	-0.69	0.90	1.00				
Median Income	-0.06	-0.53	0.37	-0.30	0.03	-0.12	0.20	0.60	-0.43	0.56	-0.76	-0.68	1.00			
% Have Catalog	-0.41	-0.49	-0.38	0.10	0.20	-0.11	-0.31	0.24	-0.09	0.24	-0.36	-0.40	0.05	1.00		
% Have Charge Account	-0.33	-0.19	-0.16	0.00	0.25	0.16	-0.12	-0.08	-0.11	0.38	-0.17	-0.07	0.14	0.39	1.00	
% Male	0.36	0.15	0.54	-0.10	-0.31	0.22	0.40	0.04	-0.21	0.10	0.10	0.10	0.30	-0.32	-0.08	1.00