An Analysis of the Accuracy of Spousal Reports

GSS Methodological Report No. 35

November, 1985

Tom Smith
NORC
University of Chicago

Like the proverbial horsetrader, survey researchers believe it is best to "go straight to the horse's mouth." Self-reports are accepted as the most accurate, while proxy reports by informants are considered less reliable (Sudman and Bradburn, 1982). Self-reports are judged superior because the individual either has unique, personal knowledge of the subject under inquiry, such as about subjective feeling states or attitudes towards various issues, or because the individual's recall of events and personal experiences is likely to be more complete and detailed.

By far the most work has been done in the area of health care. Numerous studies report that incidents of illness, disability, physician care, and even hospitalization are underreported by informants (Cartwright, 1957; Cannell and Fowler, 1963; Haase and Wilson, 1972; Kovar and Wilson, 1976; Kovar and Wright, 1973; Nisselson and Woolsey, 1959; Marquis and Cannell, 1971; and Cannell, Marquis, and Laurent, 1977-but see the partial challenge to this in Andersen, Karper, Frankel and Associates, 1979). Various other studies report notable variation between proxy and self reports, but without the presence of validation criteria that makes it possible to know which are actually more accurate (Bailor and Rothwell, 1984; Scanzoni; 1965; Ferber, 1955; Ballweg, 1969; Clark and Wallin, 1964; Brooks and Bailor, 1978; Haberman and Elinson, 1967; and Levinger). Studies that find no difference between informant and self-reports on some or all variables are comparatively rare (Ballweg, 1969; Singer, 1972-73; and Ferber, 1955).

Yet despite the acknowledged superiority of self-reports, survey research frequently collects proxy information from informants. In two situations proxy reports are actually considered preferable to self-reports. They are favored when individuals are unable or unlikely to give as accurate information as an informed person close to the individual. Most frequently

this involves having parents report on the experiences (e.g. health care) of their minor children (Nisselson and Woolsey, 1959). Similarly, caretakers are often interviewed about patients and charges who are physically or mentally incapacitated (Singer, 1972-73). Second, it is believed (but not as firmly established) that informant reports about behaviors that are threatening to a respondent will be more complete than self-reports (Bradburn, Sudman, and Associates, 1979; Sudman and Bradburn, 1982; Marquis and Cannell, 1971). It is thought that social desirability effects will cause respondents to underreport the negative behavior (e.g. drug use, temper outbursts, non-voting), while informants will be less likely to suppress the deviant behavior.

These two circumstances however account for only a small proportion of the instances where surveys resort to proxy reports in lieu of self-reports. Overwhelmingly the motivation for using proxy reports is their great convenience and low cost. Interviewing each adult in a household is usually considered unnecessary when the information being sought is either 1) household level information accessible to any adult family member (e.g. family income, home tenancy, automobile ownership), 2) a joint behavior shared with the informant (e.g. number of hours spent together, date of a couple's marriage), or 3) a basic demographic or observed behavior that is commonly known by household members (e.g. spouse's religion, employment status, tobacco use, age).

PROXY REPORTS FOR SPOUSES

The General Social Surveys (GSS), which are national full probability samples of households, have followed these standard practices. The GSS select one random respondent per household (using a Kish table). The selected respondents report mostly on personal demographics, behaviors, and

attitudes.   Respondents also report on a set of household variables (e.g.

number of earners, family income) and certain characteristics of their

spouses.[1]

While the GSS does not collect any information to validate the

individual accuracy of proxy reports about spouses, it is possible to judge

their aggregate accuracy by comparing them with self-reports. For married

couples, the self-reports of the randomly selected respondents and of

respondents' spouses represent the same universe (married people) and

therefore their distributions should be the same. For example, for married

couples the educational distribution reported by respondents should be the

same as the educational distribution that respondents report for their

spouses. Using this as our standard of judgment, we compared self and spouse

reports on the GSS.

First, we examined the completeness of information from respondents

and spouses (Table 1). Except for one tie, more missing information occurs

for spouse reports than self-reports. The higher level of missing information

comes about equally from more "no answers" - resulting from failure to ask

follow-up questions about the spouse, and higher levels of "don't knows" -

resulting from respondent's lack of knowledge about the spouse. On the other

hand, even among spouse reports, the level of information is so nearly

complete that the relative loss of information is trivial.

Second, among those giving responses we compared distributions.

Overall, the distributions were very close. Only five of sixteen comparisons

were significantly different and given that most comparisons involved six to

eleven thousand cases, significant differences of unimpressive magnitude can

_____

[1]Additional information is collected on parents and respondent's
family of origin, but we will not be dealing with this information.

· be detected. No differences were detected on any of the SES variables and
religion showed no notable differences. Differences did occur, however, among
some labor force and ethnicity variables. More full-time employment was
reported for spouses than for respondents (55.2% vs. 48.2%). Similarly, more
people working over 34 hours in the previous week are reported for spouses
than respondents (84.1% vs. 80.6%). In contrast to the higher level of
current labor force involvement reported for spouses than for respondents,
among those not currently in the labor force respondents were more likely to
report having worked for a year or more than spouses (82.1% vs. 76.9%). The
higher labor force participation of spouses results from two forms of non-
response bias: an undersampling of full-time employees in the early block
quota surveys and an underrepresentation of males on the full-probability
surveys. Full-time employed people are underrepresented on the block quota
samples that were used in 1972-1974 and on half samples in 1975 and 1976
(Stephenson, 1979). Conversely, males are underrepresented on full-
probability surveys that were used on half-samples in 1975 and 1976 and in
1977-1985 (Smith, 1979). Correcting for either the block quota undersampling
of full-time employees or the full probability underrepresentation of males
reduces the gap between spouse reports and self-reports and when both are
adjusted for simultaneously the gap falls from 4.0% to 1.6% and becomes
insignificant.

For ethnicity respondents were asked "From what countries or part of
the world did your ancestors come?" Up to three mentions were recorded. A
parallel question was asked for spouses. Considerably more ethnicities were
mentioned by respondents than they mentioned for their spouses. Respondents
mention an average of 1.54 countries while for spouses only 1.24 mentions were
given. When we used mentions as the unit of analysis, respondents mentioned

more countries from the first wave of immigrants (English, Scots, French,
etc.) and fewer from later waves. (Although the difference of 4.1% is
significant at only the .11 level, we believe that the 1986 GSS will show this
difference to be reliable). However, when we used preferred ethnicity, the
difference disappeared (to a difference of only 0.6%). It appears that
respondents lack complete knowledge of the ethnic complexity and diversity of
their partners (and therefore mention fewer ethnic backgrounds for their
spouses), but that they do know their partner's main ethnicity (the omitted
ethnicities as a group appear to be minor or secondary identifications). (For
similar findings see Smith, 1983). We suspect that the lower level of ever
working reported for spouses is also the result of the proxy having less
complete information about the spouse. Respondents may not know or not
remember their spouse was employed either for a short duration or long ago
(possibly even before they met).

Overall, it appears that spouse reports on basic demographics produce
distributions undistinguishable from self-reports.[2] The differences on labor
force status do not result from the lower accuracy of proxy reports, but
rather from two forms of non-response bias. Only on ethnicity (and possibly
ever worked) are spouse reports clearly inferior, producing less complex
representations of ethnic diversity and probably undercounting the proportion
of people with ancestors from the first immigrant wave. Yet even in this
situation no difference appear on the main ethnicity of respondents and
spouses.

---

[2]In addition a comparison of bivarate associations between respondent
variables and spousal variables showed no significant differences, suggesting
that random error is not greater among spouse reports than self-reports.

REPORT BY HUSBANDS AND WIVES OF JOINT AND HOUSEHOLD CONDITIONS

Next, we examined how the reports of husband and wives differed on either joint or household variables. In this comparison we are not evaluating whether self-reports differ from proxy reports, but whether reports by husbands differ from reports by wives on the same attribute. As above, this comparison is not done on the micro level, since we do not have separate reports for the same couples, but by aggregate level comparison of the reports of husbands and wives as a group.

We first examined whether item non-response differed between husbands and wives. Only on one variable out of the 18 household/joint variables did the amount of missing information differ significantly. Wives were less likely to report family income than husbands (91.7% vs. 94.7%). This difference results entirely from higher don't knows among wives. This probably reflects the lower level of female labor force participation and perhaps a reluctance among a portion of male breadwinners to discuss their earnings with their wives. (Women who did not give income information are older, less involved in the labor force, and more from the South than women who gave family income.)

Looking at the substantive response to the household/joint variables we find significant differences on nine of 20 comparisons. Wives tend to report slightly large households than husbands. However since the actual information on household size is reported by an informant on the household enumeration form who may not be the respondent, we can not clearly credit these differences to variant reporting patterns of husbands and wives. Wives also report more earners per household than husbands (1.69 vs. 1.64). We suspect that this results from husbands being more likely to forget part-time or discontinued employment by wives during the previous year. Conversely men

report more households having guns and hunters than wives do. Since hunting and gun ownership are heavily concentrated among men, we believe that wives are probably forgetting some instances (% with hunter: husbands-35.2, wives-32.5; % with guns: husbands-57.2, wives-54.5). On income we get mixed results. For a 12 category income classification employed from 1973 to 1985 and a 17 category classification used since 1982 we find no notable differences. But for a 16 category scale used from 1977 to 1981 men report significantly higher incomes than wives (mean categories: husbands-11.25, wives-10.84). Some two partner interview studies have also found a net tendency for wives to report lower incomes than husbands, so we believe that there is probably a real difference. One study (Haberman and Elinson, 1967) credits such differences to housewives reporting take-home pay rather than gross income and from wives underestimating the income of husbands working on commission.

Finally, comparatively large differences are observed on two of the three subjective measures we included in the analysis. Wives are more likely to rate the families financial situation as average (60.9% vs. 51.5%) and less likely to consider their marriages as very happy (64.1% vs. 67.2%). In each of these cases respondents give their subjective evaluation of a family financial or marital matter. Thus, we would not necessarily expect to find similar levels. While in the aggregate the marriages and incomes being evaluated represent the same universe, the subjective evaluations of the condition do not have to match. (Similarly subjective questions about sexual intercourse tend to show more disagreement than objective facts about intercourse, although these too seem to diverge because of selective perception-Levinger, 1966 and Clark and Wallin, 1964). This finding confirms previous indications that spousal reports tend to agree more on hard questions and vary more on soft items (Ballweg, 1969 and Scanzo, 1965).

No differences are found on several of the age categories of household

composition, home ownership, presence of telephone, racial composition of

neighborhood, or satisfaction with family income.[3]

CONCLUSION

Overall proxy reports for spouses were as accurate as self-reports.

Significantly higher levels of item non-response were found for proxy reports,

but the level of missing data was nevertheless negligible. Proxy reports were

probably sufficient because the attributes measured (religion, education,

occupation, etc.) were major, basic demographics. On the one attribute

(ethnicity) that involved complex and less salient information, proxy reports

proved to be less complete. Similarly, reports of husbands and wives tended

to diverge in two circumstances. On household attributes that the husband was

probably more knowledgable about and or more interested in (guns, hunting, and

family income) reports differed significantly (although in each case the

discrepancies were not large). And on subjective evaluations differences

appeared.

When dealing with major, concrete attributes, proxy reports are the

preferable method of data collection because they are as accurate and less

costly. But when proxy reports are extended into complex and detailed areas

about which the spouse has either superior knowledge or an unique perspective,

proxy reports are less accurate than self-reports.

---

[3]Additionally, a comparison of nine bivariate associations showed no
notable difference between husbands and wives.

Table 1

	% Missing		Probability
Labor Force Status			
Self	0.0	(11620)	.088
Spouse	0.1	(11620)	
Time of Day Worked			
Self	0.2	(587)	.090
Spouse	0.9	(569)	
Days Worked			
Self	0.7	(587)	.924
Spouse	0.7	(569)	
Total Hours Worked			
Self	0.2	(5950)	.0001
Spouse	0.7	(6324	
Occupation			
Self	0.4	(10818)	.092
Spouse	0.5	(10663)	
Industry			
Self	0.7	(10818)	.008
Spouse	1.0	(10663)	
Ever Worked			
Self	0.1	(4475)	.009
Spouse	0.4	(4164)	
Years of Schooling			
Self	0.2	(11620)	.0001
Spouse	1.2	(11620)	
Highest Degree Earned			
Self	0.4	(11620)	.0001
Spouse	1.2	(11620)	
Current Religion			
Self	0.3	(10460)	.001
Spouse	0.6	(10460)	
Protestant Denonmination (Current)			
Self	0.5	(6848)	.005
Spouse	0.9	(6823)	
Religon Raised			
Self	0.8	(10460)	.0001
Spouse	1.3	(10460)	
Protestant Denomination (Raised)			
Self	0.7	(6968)	.0002
Spouse	1.3	(6919)	

Table 2

A Comparison of Response Distributions from Self and Spouse Reports			
Variables	χ^2 Probability	Means	T-Test Probability
Labor Force Status			
Self	.0001	--	
Spouse		--	--
Time of Day Worked			
Self	.104	--	--
Spouse		--	--
Days Worked			
Self	.754	--	--
Spouse		--	--
Total Hours Worked			
Self	.0001	40.72	S
Spouse		41.35	
Occupation			
Self	.855	--	--
Spouse		--	
Prestige			
Self	--	39.96	NS
Spouse		40.05	
Industry			
Self	.380	--	--
Spouse		--	
Ever Worked			
Self	.0001	--	--
Spouse		--	
Years of Schooling			
Self	--	11.99	NS
Spouse		11.97	
Highest Degree Earned			
Self	.613	--	--
Spouse		--	
Current Religion			
Self	.879	--	--
Spouse		--	
Protestant Denonmination (Current)			
Self	.692	--	--
Spouse		--	
Religon Raised			
Self	.741	--	--
Spouse		--	
Protestant Denomination (Raised)			
Self	.051	--	--
Spouse		--	
#Mentions of Ethnicity (0-3)			
Self	.0001	1.54	S
Spouse		1.24	

Table 2 (continued)

A Comparison of Response Distributions from Self and Spouse Reports			
Variables	χ^2 Probability	Means	T-Test Probability
Ethnicity (Country)			
Self	.114	--	--
Spouse		--	

S = Significant
NS = Not Significant

Table 3

A Comparison of Reports on Household and Joint Variables
by Husbands and Wives

Variables	Probability
Household size	.003
# under 6	.177
# 6 to 12	.000
# 13 to 17	.259
# Adults	.489
# of Earners	.006
Income (12 Categories, 1973-1985)	.155
Income (16 Categories, 1977-1981)	.033
Income (17 Categories, 1982-1985)	.080
Homeownership	.619
Phone in home	.233
Gun Ownership	.028
Union member (R or Spouse)	.463
Hunter (R or Spouse)	.038
Blacks live in neighborhood	.065
Blacks living close	.192
How close are Blacks	.151
Marital Happiness	.000
Satisfaction with family income	.298
Family relative income rank	.000

References

Andersen, Ronald; Kasper, Judith; Frankel, Martin R.; and Associates, Total Survey Error: Applications to Improve Health Surveys. San Francisco: Jossey-Bass, 1979.

Bailar, Barbara A. and Rothwell, Naomi D., "Measuring Employment and Unemployment," in Surveying Subjective Phenomena, editing by Charles F. Turner and Elizabeth Martin. 2nd Vol., New York: Russell Sage, 1984.

Ballweg, John A., "Husband-Wife Response Similarities on Evaluative and Non-Evaluative Survey Questions," Public Opinion Quarterly, 33 (Summer, 1969), 249-254

Bradburn, Norman M. and Sudman, Seymour, and Associates, Improving Interview Method and Questionnaire Design. San Francisco: Jossey-Bass, 1979.

Brooks, Camilla A. and Bailer, Barbara A., "An Error Profile: Employment as Measured by the Current Population Survey," Statisical Policy Working Paper 3. Washington, D.C.: GPO, 1978.

Cannell, Charles F., and Fowler, Floyd J., "Comparison of A Self-enumerative Procedure and a Personal Interview: A Validity Study," Public Opinion Quarterly, 27 (Summer, 1963), 250-264.

Cannell, Charles F.; Marquis, Kent H.; and Laurent, Andre, "A Summary of Studies of Interviewing Methodology," Vital and Health Statistics, Series 2, No. 69, March, 1977.

Cartwright, Ann, "The Effect of Obtaining Information From Different Informants on a Family Morbidity Inquiry," Applied Statistics 6 (March, 1957), 18-25.

Clark, Alexander L. and Wallin, Paul, "The Accuracy of Husband's and Wives' Reports of the Frequency of Marital Coitus," Population Studies, 18 (November, 1964), 165-173.

Ferber, Robert, "On the Reliability of Responses Secured in Sample Surveys." Journal of the American Statistical Association, 50 (September, 1955), 788-810.

Haase, Kenneth W. and Wilson, Ronald W., "The Study Design of an Experiment to Measure the Effects of Using Proxy Response in the National Health Interview Survey," American Statistical Association Proceedings, Social Statistics Section, 1972, pp. 289-293.

Haberman, Paul W. and Elinson, Jack, "Family Income Reported in Surveys: Husbands versus Wives," Journal of Marketing, 4 (May, 1967), 191-194.

Levinger, George, "Systematic Distortion in Spouses' Reports of Preferred and Actual Sexual Behavior," Sociometry, 29 (September, 1966), 291-299.

Kovar, Mary Grace and Wilson, Ronald W., "Perceived Health Status - How Good is Proxy Reporting?" American Statistical Association Proceedings, Social Statistics Section, 1976, Part 2, pp. 495-500.

Kovar, Mary Grace and Wright, Robert A., "An Experiment with Alternative Respondent Rules in the National Health Interview Survey," American Statistical Association Proceedings, Social Statistics Section, 1973, pp. 311-316.

Marquis, Kent H. and Cannell, Charles F., "Effect of Some Experimental Interviewing Techniques on Reporting in the Health Interview Surveys," Vital and Health Statistics, Series 2, No. 41. May, 1971.

Morgan, James N., "Some Pilot Studies of Communication and Consensus in the Family," Public Opinion Quarterly, 32 (Spring, 1968), 113-121.

Nisselson, Harold and Woolsey, Theodore D., "Some Problems of the Household Interview Design for the National Health Survey," Journal of the American Statistical Association, 54 (March, 1959), 69-87.

Scanzoni, John, "A Note on the Sufficiency of Wife Responses in Family Research," Pacific Sociological Review, 8 (Fall, 1965), 109-115.

Singer, Eleanor, "Agreement between "Inaccessible" Respondents and Informants," Public Opinion Quarterly, 36 (Winter, 1972-73), 603-611.

Smith, Tom W., "Sex and the GSS: Nonresponses Differences," GSS Technical Report No. 17. Chicago: NORC, 1979.

Smith, Tom W., "Problems in Ethnic Measurement: Over-, Under- and Misidentification," Proceedings of the American Statistical Association, Social Statistics Section, 1983, pp. 107-116.

Stephenson, C. Bruce, "Probability Sampling with Quotas: An Experiment," Public Opinion Quarterly, 43 (Winter, 1979), 479-496.

Sudman, Seymour and Bradburn, Norman M., Asking Questions: A Practical Guide to Questionnaire Design. San Francisco: Jossey-Bass, 1982.

Thompson, Donovan J. And Tauber, Joseph, "Household Survey, Individual Interview, and Clinical Examination to DeterminePrevalence of Heart Disease," American Journal of Public Health, 47 (September 1957), 1131-1140.