Reliability of Proxy Response in the Current Population Survey

Lawrence E. Boehm
U.S. Bureau of Labor Statistics

Abstract

The Current Population Survey (CPS) uses one respondent to collect labor force data on all the members of a household. The acceptance of proxy responses is suspected of increasing the potential for response error in the survey. This paper reviews relevant findings and theory and presents a laboratory study which evaluated the reliability of proxy responses in the CPS. Two members of households were interviewed individually using the CPS questionnaire. Subjects answered questions for themselves (self response) and for the other family member (proxy response). Thus it was possible to compare the proxy response to the self response for each person. Each person also provided a confidence rating of their ability to report reliable answers and a rating of their knowledge of the other person's job or job search. Results address the reliability of proxy responses, the relationship between confidence, knowledge, and performance, the overall agreement of self and proxy respondents, and the impact of proxy responses on labor force classification. Alternative collection methods are discussed.

1. The use of Proxy Data in the CPS

Self-other differences in knowledge and cognitive processing are of practical importance to survey researchers because a number of national surveys allow "any responsible" adult member of a household to respond for all the members of that household (Dippo, 1989). Such proxy responses are permitted in the CPS and account for approximately 50% of the interviews conducted. The CPS is a monthly survey of approximately 59,000 households in the United States, from which monthly estimates of labor force status (employed, unemployed, and not in the labor force) and related characteristics are developed. Because of the volume of interviews conducted and the time constraints imposed by the monthly release of the labor force data, interviewers are permitted to accept proxy responses for members of the household that are not at home at the time of the interview. Accepting proxy responses for absent household members is convenient and economical, but the quality of data obtained from proxy respondent has been questioned (Hill, 1987; Moore, 1989; but see O'Muircheartaigh, 1986).

2. Psychological Theories of Information Processing

Because the CPS interview focuses on information about the respondent as well as other household members, social psychological theories of information processing may apply. According to attribution theory, people often explain their own behavior by attributing it to situational factors, whereas people explain the behavior of others by attributing it to the person's personality or disposition. A simplified scenario of attribution might go like this: I drive my automobile fast because I am late for work; you drive fast because you are reckless and irresponsible. Such differential explanations arise because people have access to different sources of information. Differential attributions affect the representation people hold of themselves and others, and, as a result, they can influence the organization of information about themselves and others in memory (Jones and Nisbett, 1972).

The way people organize and recall information about themselves and others affects the decision strategies they use when answering questions. Kuiper and Rogers (1979) have demonstrated that different cognitive processes are used depending on whether the incoming information is associated with the self or with another person. The basic experimental design required subjects to decide whether certain personality characteristics described themselves and another person. After the task was complete, difficulty and confidence ratings were obtained, and recall for the personality characteristics was tested. Over several experiments, Kuiper and Rogers found that decisions made in reference to oneself were easier to make, that subjects had higher confidence in self-referent decisions, made such decisions faster, and in some cases, recalled more personality characteristics. When recall levels were equivalent, subjects answered questions about themselves more quickly, indicating that they processed information about themselves more efficiently and recalled it more easily. Apparently the more organized self-knowledge provides not only a context in which to encode and retrieve information, but also leads to qualitatively different cognitive processing. Kuiper and Rogers also provided evidence that the level of familiarity with the other person can influence the cognitive processes applied to the situation. Incidental recall of a familiar person's characteristics was more than twice that for characteristics of an unfamiliar reference person.

The second area of interest is calibration, the relationship between a person's numerical judgment of confidence or subjective probability and the relative frequency of the event being judged (Einhorn and Hogarth, 1978). In other words, calibration is the relationship between self-rated confidence and performance. In studies of calibration, the most common finding is that people are usually overconfident in the accuracy of their performance (Lichtenstein, Fischhoff, & Phillips, 1977). For example, people have been found to be overconfident (i.e., poorly calibrated) in their chances of answering questions correctly (Lichtenstein, 1977). Oskamp (1965) demonstrated that confidence in a decision increases as a function of the quantity of information available, yet there is not a corresponding increase in the accuracy of the decision. In some cases, those most confident in their decisions are the least accurate (Holsopple & Phelam, 1954). In a study of the relationship between comprehension and confidence, Glenberg and Epstein (1987) found that confidence was based on a self-classification strategy rather than on actual comprehension. That is, the person's perceived expertise, rather than actual expertise, was responsible for the level of confidence. Furthermore, feedback about the outcome of a decision can bias retrospective judgment rather than correct it--a hindsight bias (Arkes, Wortman, Saville, & Harkness, 1981; Fischhoff, 1975). That is, people knowledgeable of the outcome of a difficult decision are frequently overconfident in their ability to have correctly predicted the event. Thus, there is a well documented discrepancy between confidence and performance. The present study sought to extend the data base on confidence and performance to decisions about familiar people and their job or job search.

3. Goals of the Study

The present study had several goals: (a) to determine for each item the level of relative reliability between proxy respondents and self respondents, (b) to examine the relationship between self-rated knowledge, confidence, and performance, (c) to determine the overall percentage of disagreement between self and proxy respondents, (d) to evaluate the utility of confidence or knowledge ratings as screening procedures to identify qualified proxy respondents, and (e) to determine the effect of proxy responses on labor force classification.

Data were collected through laboratory interviews, rather than through a field survey. There are several advantages to testing self-proxy pairs in the laboratory: 1) it reduces the respondent self-selection problems of field studies by assigning respondents to different conditions in the laboratory; 2) it eliminates the problem of time lags associated with re-interview studies; and 3) it assures that the respondents are uninfluenced by the other household members.

METHOD

Subjects

Eighty-four individuals representing 42 households participated in the study. Volunteers were recruited using flyers and radio announcements and were reimbursed $15 for their time and travel expenses.

Procedure

Pairs of volunteers were brought into the laboratory and interviewed individually with the CPS questionnaire. While one member of the pair was being interviewed, the other member waited in the reception area. All respondents signed a consent form which informed them that the information they provided was confidential and used for research purposes only. All respondents were explicitly told that their participation in the study was voluntary and that they could decline to answer any question they chose. Basic demographic information was collected (e.g., age, education, years of acquaintance).

Each subject then completed two CPS interviews: one for themselves and one for their partner. Self-proxy pairs were randomly assigned to one of four counterbalanced interviewing sequences. Subjects rated their confidence in their ability to answer questions about themselves and their partner's job or job search. Confidence ratings were obtained from each respondent before and after each interview (1=very low confidence, 6=very high confidence). Before the interview began, proxy subjects indicated how much they knew about their partner's job or job search activities using a similar rating scale.

Results

Self and proxy respondents' answers to selected labor force questions were collected, persons were classified into labor force categories, and correlations were computed for each question (excluding the open-ended questions). The results reflect the reliability of the proxy respondent relative to the self respondent. Reliability does not imply validity; the self and proxy can agree on an incorrect answer. For example, both members of a pair answered "no" to question 25F (Is ... covered by a union or employee association contract?), when in fact the correct answer was "yes."

The CPS classifies respondents into one of three labor force categories (employed, unemployed, or not in the labor force) based on their responses to the questionnaire. The responses from both self and proxy respondents were reviewed to determine which labor force classification would result. The results are shown in Table 1. In the cases where self and proxy responses resulted in different labor force classifications, it was not possible to determine which classification was correct. For the purposes of this study it does not matter which is the "correct" classification; the point is that accepting responses from proxies can lead to a different labor force classification. That proxy responses make up a sizable proportion of CPS interviews only magnifies the issue.

Differences in labor force classification were pronounced among individuals whose self responses led them to be classified as unemployed or not in the labor force. That is, proxy respondents often gave answers leading to different labor force classifications. (See Table 1.) Among individuals whose responses led them to be classified as employed, self-proxy labor force agreement was very high (95%). However, self and proxy respondents were more likely to disagree when a choice was made between full-time and part-time employment (87% and 90% respectively). Overall, this study found that self and proxy responses result in the same labor force classification 83% of the time.

Correlations between self and proxy responses for selected items are provided in Table 2. For continuous data (e.g., question 20A-hours worked), the correlations were computed using Pearson's product moment r. For dichotomous responses, the phi coefficient was used which, in this case, is equivalent to Pearson's r. For the dichotomous variables the p values were obtained from a Chi-squared test of the correlations. The means for the earnings and hours questions are presented in Table 3.

When self and proxy respondents were compared across all questions (open and closed), it was found that proxy respondents disagreed with the self respondents on 30% of the questions. For questions providing continuous data (e.g., earnings per hour), the experimenter used a predetermined criterion to score self-proxy disagreement. Although the correlations between self and proxy are generally high and statistically significant, the finding that self and proxy respondents disagreed on 30% of the questions suggests that proxy respondents have considerable difficulty with some questions (this includes both open- and closed-ended questions).

Approximately 20% of the sample respondents were unemployed. Self-proxy disagreements occurred primarily for questions 22C and 22F. For questions 22C ("How many weeks has ... been looking for work?") and 22F ("When did ... last work at a full-time job or business...?"), self respondents gave lower estimates (means = 7 and 5 weeks, respectively) than proxy respondents (means = 11.5 and 10 weeks, respectively). These data also suggest that the respondents started seeking work before they were separated from their last job.

There was no significant change over time in confidence for either self, $t\,(83) = -0.13$, $p > .05$, or proxy respondents, $t\,(83) = -1.19$, $p > .05$. Put another way, confidence was unchanged by the administration of the questionnaire. As in previous studies of subjective confidence ratings, the respondents in this study were very confident in their ability to answer questions for themselves and their partners. Almost all of the subjects tested rated themselves as slightly confident or above. Similarly, more than 90% of the proxy respondents considered themselves to be average or above in their knowledge of the other's job or job search.

Although proxy respondents' self-rated knowledge of the other person's job was significantly correlated with their pre- and post-confidence ratings (r=.60, $p < .001$, and $r = .54$, $p < .001$, respectively), other results suggested that both proxy knowledge and confidence ratings are unrelated to the reliability of their answers. The correlations between proxy pre-test and post-test confidence ratings and the proportion of self-proxy disagreements were r= -.04, $p = .36$ and r = -.15, $p = .11$, respectively. The correlation between self-rated knowledge and the number of proxy disagreements was r=.13, p =.15.

Intuitively, one might suspect that the longer someone has known another person the better proxy he or she would be. However, the number of years the self and proxy have been acquainted does not appear to be related to self-proxy agreement. The correlation between years acquainted and the proportion of self-proxy disagreements was r = -.035, p =.39.

Males and females did not differ significantly in their confidence ratings for themselves or others, or their knowledge of the other's job.

DISCUSSION

Although significantly correlated with self respondents, proxy respondents reported hourly and weekly earnings below the self respondents' figures. It is not uncommon for proxy respondents to provide data that result in different labor force classifications.

The finding that the length of time that the self and proxy respondents have known each other was unrelated to the reliability of their answers is counter-intuitive but not unusual. The finding that confidence ratings were unrelated to performance is consistent with the literature on the relationship between confidence ratings and performance criteria. For example, Lichtenstein, Fischhoff, and Phillips (1977) have shown that people are often grossly overconfident in the accuracy of their performance. In some cases, confidence is related to performance, but this finding is relatively less common (e.g., Glenberg and Epstein, 1987). The results of this study suggest that the use of self-rated knowledge and confidence would not be useful for screening proxy respondents.

The restricted range of relationships may be one reason this study revealed no relationship between confidence, knowledge, and accuracy. The relationship between confidence and performance may depend on the breadth of the knowledge domain. Glenberg and Epstein (1987) found no relationship between confidence and performance *within* a knowledge domain (e.g., physics), but a relationship between confidence and performance when comparisons were made *across* knowledge domains (e.g., physics and music). Thus, the lack of a confidence-performance relationship in this study may be due to the comparative closeness/overlap of knowledge for household members (i.e., within domain knowledge). Future research could address the relationship between confidence, accuracy, and type of relationship.

Several procedural changes could be considered to improve the quality of CPS data. One strategy would be to modify the current respondent rules and not accept proxy responses. Such a modification in the respondent rules could increase the quality of the data obtained and would lessen the respondent burden felt by any one respondent. The operational concern would be impact on response rates. CPS studies form the 1960's suggest that the number of interviews lost because of a self-respondent procedure may be small, that such a change may have little effect on labor force classification, but would cause a moderate increase in enumeration costs (Deighton, 1967). A similar but less radical change might be to exclude proxy responses for items that are found to be difficult for proxy respondents to answer accurately or to restrict proxy respondents to certain relationships. Another strategy might be to allow proxy respondents to refuse to answer a question rather than guess.

The use of dependent interviewing may be an alternative to alleviate some of the proxy errors reported here. Dependent interviewing is a procedure in which the previous month's responses are used in the current month's interview. If the first interview was conducted only with self respondents, then in subsequent months proxy responses could be accepted in which the proxy could accept or update the previous month's data. This procedure has several advantages. First, the data obtained from a self respondent would probably be accepted as valid by an uncertain proxy respondent. The result would be a decrease in well-intentioned but unreliable estimation by proxy respondents. Second, the continued use of proxy respondents for interviews other than the first and fifth month would preserve the cost saving benefits of accepting proxy responses. A third benefit of using proxy respondents in connection with dependent interviewing is that proxies are less motivated to provide a socially desirable answer. That is, a proxy respondent may be more likely to report undesirable circumstances accurately (e.g., the loss of a job, low income, or a less than desirable occupation). Hence, the proxy may correct faulty self responses motivated by social desirability. Fourth, dependent interviewing would reduce the burden on both proxy and self respondents. Finally, knowledge gained by the proxy respondent about the other member's answers may stimulate discussion between the members of a household as they attempt to resolve the discrepancies between their understanding of the persons job or job search and the person's self-reported activities. Such discussion could increase the accuracy of subsequent proxy interviews. There are number of data items which could make use of dependent interviewing techniques (usual hours worked, industry/occupation, duration unemployed), although the main labor force classification is based on activity in the prior week and therefore could not use dependent techniques.

SUMMARY

The results suggest that proxy respondents feel confident providing data for other household members, but that the data may be biased, relative to self responses, in some areas. Confidence, self-rated knowledge, and the length of the self-proxy relationship would not be useful predictors of proxy performance. Future research should examine the use of dependent interviewing in connection with proxy respondents and perform a more in-depth study of the confidence/performance relationship.

Acknowledgments

Thanks are due to the following people at the U.S. Bureau of Labor Statistics who reviewed this paper and provided helpful comments: Kennon Copeland, Maria Fracasso, Alan Tupek, and Jack Bregger. I also wish to thank Charles McNeill for his help in conducting the interviews and organizing the data.

References

Arkes, H. R., Wortmann, R. L. Saville, P. D., & Harkness, A. R. (1981), "Hindsight Bias Among Physicians Weighing the Likelihood of Diagnoses," *Journal of Applied Psychology, 66,* 252-254.

Deighton, R. (1967), "Some Results on Experimentation with Self-Respondent Interviewing Procedures," Bureau of the Census memorandum.

Dippo, C. S. (1989). "The Use of Cognitive Laboratory Techniques for Investigating Memory Retrieval Errors in Retrospective Surveys," *Proceedings of the International Statistical Institute.* Paris, France.

Einhorn, H. J. & Hogarth, R. M. (1978), "Confidence in Judgment: Persistence of the Illusion of Validity," *Psychological Review, 85,* 395-416.

Fischhoff, B. (1975), "Hindsight = Foresight; The Effect of Outcome Knowledge on Judgment Under Uncertainty," *Journal of Experimental Psychology: Human Perception and Performance, 1,* 288-299.

Glenberg, A. M., & Epstein, W. (1987), "Inexpert Calibration of Comprehension," *Memory and Cognition, 15,* 84-93.

Hill, D. H. (1987), "Response Errors in Labor Surveys: Comparisons of Self and Proxy Reports in the Survey of Income and Program Participation (SIPP)," *Proceedings of the Bureau of Census Third Annual Research Conference.*

Holsopple, J. G., & Phelam, J. G. (1954), "The Skills of Clinicians in Analysis of Projective Tests," *Journal of Clinical Psychology, 10,* 307-320.

Jones, E., & Nisbett, R. (1972), "The Actor and the Observer: Divergent Perceptions for the Causes of Behavior," In E. Jones *et al.* (Eds), *Attribution: Perceiving the Causes of Behavior.* Morristown, NJ: General Learning Press.

Kuiper, N. A., & Rogers, T. B. (1979), "Encoding of Personal Information: Self-Other Differences," *Journal of Personality and Social Psychology, 37,* 499-514.

Lichtenstein, S., Fischhoff, B., & Phillips, L. S. (1977), "Calibration of Probabilities: The State of the Art to 1980," In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases.* Cambridge: University Press.

Moore, J. C. (1989), "Self/Proxy Response Status and Survey Response Quality: A Review of the Literature," *Journal of Official Statistics* (in press).

O'Muircheartaigh, C. (1986), "Correlates of Reinterview Response Inconsistency in the Current Population Survey (CPS)," Proceedings of the Bureau of Census Second Annual Research Conference, U.S. Bureau of the Census, Washington, D.C.

Oskamp, S. (1965), "Overconfidence in Case-Study Judgments," *Journal of Consulting Psychology, 29,* 261-265.

Table 1. Breakdown of labor force classification by self and proxy respondent

	Self respondent		
Proxy Respondent	Not in the labor force	Unemployed	Employed
Not in the labor force	8	4	1
Unemployed	2	10	2
Employed	2	3	52

Table 3. Means for hours worked last week, hours usually worked, and earnings (hourly and weekly)

Hours worked last week

	Self	Proxy	Self-Proxy
mean	32.1	30.9	1.24
SD	19.3	20.1	12.4
n	42	42	

Hours usually worked

	Self	Proxy	Self-Proxy
mean	30.4	29.7	.68
SD	18.5	16.1	7.14
n	38	38	

Earnings per hour

	Self	Proxy	Self-Proxy
mean	$8.4	$6.99	1.47
SD	5.90	2.97	4.98
n	19	19	

Earnings per week

	Self	Proxy	Self-Proxy
mean	$769	$632	137
SD	731	495	293
n	14	14	

Note: Different sample sizes are due to the skip pattern in the CPS. Not all respondents are asked all questions and some proxy respondents follow a different path than the self respondents.

Table 2. The correlation (r) of self and proxy responses for CPS labor-force questions: employed people

ITEM	n pairs	r	p	%agreement
Questions eliciting continuous date				
20A Hours worked last week.	42	.75	.001	
25A Hours usually work per week	38	.92	.001	
25C How much does...earn per hour?	19	.53	.009	
25D How much does...earn per week?	14	.96	.001	
Questions eliciting dichotomous data				
20C Does ...usually work < 35 hrs	19	.20	*	79
20D Time off last week?	13	.64	.05	69
20E Work any overtime?	12	.43	n.s.	67
25B Is...paid by the hour?	38	.70	.005	92
25E Union member?	38	.44	.005	89
25F Covered by union contract?	38	-.08	*	87

Note: p values reflect the probability that a sample correlation this large or larger could occur under the null hypothesis that the population correlation is zero. p values are approximate because a nonprobability sample was used. Unequal n's are due to the fact that not all respondents are asked all questions and some respondents refuse to answer questions.

* For question 20C and 25F Self/proxy agreement was in fact high. The correlation coefficient does not reflect this relationship because of a lack of variability in the responses (i.e., almost all responses are in one cell).