

FACTORS ASSOCIATED WITH PROXY KNOWLEDGE OF EMPLOYMENT-RELATED INFORMATION

Paul A. Mullin, Bureau of Labor Statistics

Bruce E. Tonn, Oak Ridge National Laboratory

Paul A. Mullin, BLS, PSB Room 4915, 2 Mass. Ave., N.E., Washington, DC, 20212

KEY WORDS: proxy-self correspondence, information transmission, relationship

The Intrahousehold Communications Study was conducted to identify factors associated with proxy knowledge, with the ultimate goal of improving household survey respondent rules. In this part of the study, survey reports concerning labor force participation were examined. Household members serving as subjects in the study responded to Current Population Survey (CPS) questions for themselves and for other household members who were present. The correspondence between proxy and self reports was used as a measure of proxy knowledge. Use of this measure is justified because it is likely that correspondence is primarily determined by proxy, not self, report accuracy. Furthermore, in the analyses presented here, proxy-self correspondence was frequently examined as a function of factors whose variability should only affect proxy report accuracy. For example, if proxy education affects correspondence, this very likely represents changes in proxy rather than self report accuracy.

Conceptual Analysis

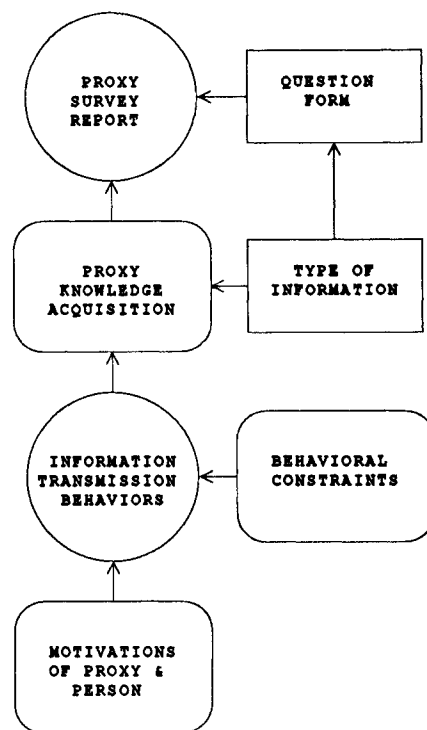
A simple account of factors that affect proxy report accuracy might stipulate that proxy reporting is determined by proxy knowledge, the form of the question and the type of information requested (see Figure 1). Proxy knowledge is itself determined by the quantity and quality of the information transmitted to the proxy about the person in question. This information transmission behavior is motivated by various human needs and is expressed in the context of a number of different constraints that affect the behavioral expression of these needs. These factors are discussed in more detail below.

Question and Information Factors

Proxy knowledge that is expressed in a survey report may vary as a function of the type of information requested from the proxy and the form of the question. The CPS questions were rated by the experimenters according to whether the

information requested was *dynamic* or *static* over time, a dimension that is relevant to the frequency at which information must be transmitted for the possession of valid knowledge. *Yes/no*, *multiple-choice* (including 2 open-ended), and *numerical* questions were distinguished from the CPS items.

Figure 1
Factors associated with proxy report accuracy.



Information Transmission Behaviors

Proxy knowledge is determined by the acquisition of information about other household members through information transmission behavior. Four kinds of proxy information transmission behaviors were distinguished: (a) *communication*, either verbal or written, with others, (b) *observation* of the person in question, (c) *participation* in relevant activities, and (d) *exposure to documentation* such as pay stubs or time cards.

Motivations and Behavioral Constraints

Information transmission behavior occurs as the result of the motivation to meet various human needs. The behavior is expressed in the context of a large number of constraints, including: (a) *pure constraints* such as temporal, physical, and economic conditions; (b) *social constraints* such as the general cultural, ideological, legal, and social climate; and (c) *personal constraints* such as individual socioeconomic, marital, educational, and employment statuses (Tonn, 1984).

By generally considering human needs and these behavioral constraints, it was possible to generate a number of expectations concerning the level of information transmission among proxy-self pairs. It was expected that spouses would know the most about each other, parents would know much about children, children would know something about the parents, and siblings would know relatively little about each other. Similarly, household members of the same sex may communicate more to one another than members of the opposite sex. And, in opposite sex communication women may know more about men than the reverse because men tend to dominate conversations with women (Deakins, Osterink, Hoey, 1987; Eakins & Eakins, 1978). Furthermore, household members of similar ages may communicate more than members of different ages, except that younger members may seek advice from older members.

It was expected that a proxy's educational level might be related information seeking and acquisition. Household size was considered important because with increasing numbers of household members, the time spent in communication with a given household member will be reduced on average (Boussard, 1975). The total number of hours that the proxy and person in question usually work might indicate the amount of time they have available for communication. Conversely, it might reflect the need for increased communication to coordinate household activities because of dual-job holding. Finally, we reasoned that if the labor force participation (employed, unemployed, not-in-the-labor-force) of the proxy and person matched, they might be more likely to communicate than pairs that did not share similar situations. Total household income was included in the analysis in the attempt to account for some differences between households.

Proxy-Self Correspondence Measures

The correspondence of the proxy and self reports was scored in several ways. *Knowledge ratings* scored correspondence for each question on a 1 to 10 scale in order to give the proxy partial credit for having some knowledge of the relevant information. Scoring methods differed depending on the form of the question. For yes/no and nominal scale multiple-choice questions, tens and zeros were given for hits and misses, respectively. For multiple-choice questions that permitted multiple answers, scoring was the percentage of matches out of the total number of answers given in the self report. For ordinal scale multiple-choice questions, credit was assigned by dividing the number of response options separating the proxy and self answers by the total number of response choices and multiplying by 10. For numerical questions, ranges were designated and credit was assigned as the number of ranges separating the proxy and self answers divided by the total number of ranges and multiplied by 10. Open-ended questions were rated by the experimenters and were scored as a 0, 5, or 10.

Survey response matching scored correspondence in an all-or-none manner. This method was used for the analysis of information type and question form because it was applied in a consistent manner across different question types.

For both the knowledge rating and the survey response matching methods, only questions answered by both the proxy and self were used. The CPS contains different questionnaire paths for persons in different situations, and proxy and self may diverge in the paths they take. Although limiting the analysis to questions both persons answered somewhat curtailed the variance of the scores (i.e., increased the average scores of those who diverged), it was a better option than accepting the noise introduced for diverging pairs by the differing lengths of the questionnaire paths. Most divergences occurred after only a few questions, and thus, the average knowledge ratings and response match percentages reflected the difference between proxy-self pairs that diverged and those that did not (e.g., for the knowledge rating method: $F(1) = 76.06, p < .05$).

Labor force classification matching scored correspondence according to whether or not the proxy and self reports were categorized by the same labor force classifications. Fourteen percent of the respondent pairs were classified differently on the major classifications (i.e., employed, unemployed, or not-in-the-labor-force), and 27%

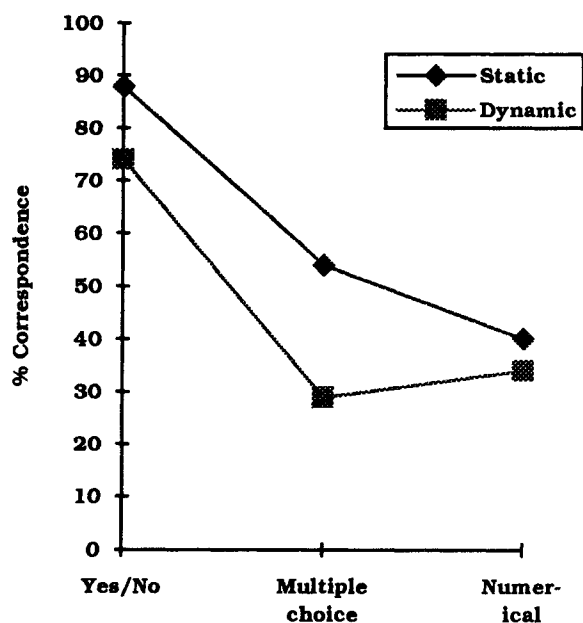
were classified differently when more detailed classifications were used (i.e., employed full-time and working, employed part-time and working, employed full-time and absent last week, employed part-time and absent last week, unemployed and on layoff, unemployed and looking for work, and not-in-the-labor-force).

Results

Effects of Question Form and Information Type

A categorical general linear model analysis (SAS CATMOD) was conducted on the survey response match scores at the individual question level for the 2894 questions that were asked of the 190 proxy-self pairs. The analysis revealed that the main effects of question form, $X^2(2) = 236.7$, $p < .05$, the main effect of the static/dynamic dimension, $X^2(1) = 17.02$, $p < .05$, and their interaction, $X^2(2) = 6.15$, $p < .05$, all significantly affected proxy-self correspondence. The responses of proxy and self matched on 84% of yes/no questions answered by both proxy and self, 54% of the multiple-choice questions, and 38% of the numerical questions. As expected, proxy and self matched more often for the static questions than for the dynamic questions (63% versus 59%).

Figure 2
Proxy-self correspondence as a function of question form and type of information requested.



The interaction between question form and information type is unexplained (see Figure 2). One might expect that numerical questions requesting dynamic information would be the most difficult, but these questions yielded better performance than multiple-choice/dynamic questions. A separate analysis using a 5-point rating scale for the dynamic/static dimension suggested that this result was not due to differences in the degree to which the two sets of questions are dynamic.

The results correspond to what might be expected concerning the different question forms and types of information, but must be interpreted cautiously as these are factors that may affect self report accuracy as well as proxy report accuracy. Still, it is likely that much of the effects are due to differences in the quality of proxy reports.

Effects of Information Transmission Behavior

Subjects in the study were asked to report the ways that they learned information about the other household members for whom they completed a proxy survey report. This measure was designed to identify different information transmission modes by which proxies acquire knowledge. It was hoped that the quantity of information transmission might be represented in the number of different modes selected. It is also possible that the quantity and quality of information transmission might vary as a function of the mode. These questions appeared to be relatively easy for proxies to answer. In contrast, proxies might have difficulty estimating the amount of time they spent acquiring information because information transmission is likely to be dispersed over time and because temporal measures may not be appropriate for some modes (e.g., viewing pay stubs, commuting together).

The number of information transmission modes significantly accounted for only about 3 to 4 percent of the variance in proxy-self correspondence on the knowledge ratings and survey response match measures. Weighting the data to reflect experimenter ratings of the relative quality of each form of information transmission for proxy knowledge acquisition did not change the results. Nor did including only questions that requested information that is highly dynamic over time.

With respect to labor force classification matching, however, the number of information transmission modes did distinguish ($p < .05$) between proxy-self reports that matched and those

that did not on the major classifications (see Table 1). This trend was evident for all four forms of information transmission, but simple comparisons revealed that only the difference for the participation mode was significant ($p < .05$).

Table 1
Mean number of information transmission behaviors as a function of the proxy-self labor force classification match correspondence.

<u>Mode</u>	<u>Match</u>	<u>No Match</u>
Overall	8.6	6.8 * $p < .05$
Communication	2.3	1.7
Observation	2.9	2.5
Participation	2.2	1.5 * $p < .05$
Exposure to Docs.	1.3	1.0

Table 2 shows the percentages of proxies reporting information transmission by the various modes. Conversations with the person in question were prevalent (reported by 97% of the proxies), as were conversations with others and conversations that were overheard. Communication by notes was fairly infrequent.

On average, each of the various kinds of observations were reported by 50% of the proxies. And, a moderate percentage of proxies reported participating in the labor force activities of the person in question, including 53% who reported visiting the person's workplace, and 33% who reported commuting with the person. Finally, substantial percentages of proxies reported viewing different media forms such as paychecks (42%), and accounting records (31%).

Because the information transmission measure was not significantly associated with proxy-self correspondence in the knowledge ratings and survey response match scores, it was dropped from further analysis using these measures. In future research, the measure of information transmission might be improved by collecting more direct information on the quantity and quality of transmission behavior. For example, virtually all proxies reported engaging in conversations, and this ceiling effect might be avoided by asking more detailed questions about the conversations, such as their frequency. The measures might also be improved by asking subjects to report the information transmission for each item of information, instead of asking for a single summary of information transmission for

the entire survey. Self respondents might additionally be asked to report details about the transmission of information to other household members who will serve as proxies.

Table 2
Percentages of proxies reporting different information transmission behaviors.

<u>Mode</u>	<u>Percent of Proxies</u>
<u>Communication</u>	
Conversations with <i>PERSON</i>	97
Overheard between <i>PERSON</i> & others	46
Conversations with others	34
Overheard among others	11
Notes by <i>PERSON</i> to you	14
Notes by <i>PERSON</i> to others	5
Notes by others about <i>PERSON</i>	4
Other conversations or notes	8
<u>Observation</u>	
Noticing comings and goings	65
Noticing clothing	51
Noticing preparations for work	48
Noticing work done at home	39
Other things noticed	17
<u>Participation</u>	
Visiting the workplace	53
Meeting people from work	53
Commuting together	33
Helping with applications & resumes	21
Working together	19
Helping search help wanted ads	0
Doing other things	17
<u>Exposure to Documentation</u>	
Pay checks or records	42
Account records	31
Daily planners or calenders	24
Paperwork	23
Reading other materials	6

Effects of Behavioral Constraints

A general linear model analysis (SAS GLM) was conducted to examine the relationship between various behavioral constraints and proxy-self correspondence. Because the group of households was not large and there was considerable confounding between membership in a particular household and the behavioral

constraints, the results must be interpreted cautiously. The effects of proxy-self relationship, in particular, are likely to be a result, in part, of differences among households.

In the analysis, relationship was coded into 5 classes: spouse, parent-child, child-parent, sibling, and other; proxy and self sex pairings into 4 classes; the proxy-self age difference into 3 classes: 0 - 10 years, 11 - 20 years, and 20+ years; and proxy education into 3 classes: less than a high school diploma, high school diploma to some college, and 2 year college degree and above. The total number of usual hours that the self and proxy worked was coded into 3 classes: 0 - 20, 21 - 55, and 56+; and household income into 3 classes: \$0 - \$20,000, \$20,001 - \$60,000, and \$60,001+. Because of serious a confounding between household size and relationship, household size had to be dropped from the analysis.

In the model, relationship was nested within the variable representing the labor force classification of the person in question by their self report (PERSLFS). This was done to control for differences between groups of subjects who followed different question paths. All other variables, except household income, were nested in the interaction between the PERSLFS and relationship. The nesting within relationship was necessary because the occurrence of variables such as the age difference and the proxy self sex pairings are severely confounded with relationship (e.g., all spouses are of different sexes). The match between the labor force classification of the proxy and person proved to be largely redundant with the total number of usual hours worked by the self and proxy, and it was dropped. The N was 186 because there were missing ages in 4 of the 190 pairs.

The general linear model was first applied to the knowledge ratings. The main effect of the PERSLFS (i.e., the question path) was not significant. Relationship, $F(14) = 3.04$, sex pairing, $F(16) = 3.41$, age difference, $F(5) = 5.37$, proxy education, $F(13) = 2.41$, total usual hours, $F(10) = 3.56$, and household income, $F(2) = 3.50$, each accounted for a significant part of the variance at the $p < .05$ level. Overall, the model accounted for 77% of the variance in proxy-self correspondence using 72 of 185 degrees of freedom, $F(72) = 5.2$, $p < .05$. This is an undoubtedly an inflated figure because differences among households could not be factored from the effects of the behavioral constraints. In addition, a considerable number of degrees of freedom were

consumed by the model due to the nesting of variables to control for grouping by question path and for the confounding with relationship and most other variables.

Similar results were obtained when the analysis was conducted using the percentage of survey response matches as the correspondence measure. However, total usual hours ($p = .08$) and household income ($p = .27$) only approached significance in this analysis. The remaining variables accounted for 65% of the total variance ($p < .05$).

The means of the knowledge ratings for the individual categories of each variable suggested that most of our expectations were supported. For relationship, spouses were best, followed by parent-child and child-parent, with siblings performing relatively poorly. An advantage of the "other" relationship class over the siblings could be ascribed to the presence of unmarried couples in that class. The age difference means indicated that proxy-self pairs who are close in age or who are distant in age perform better than those in the middle range of age differences, as expected. Proxy education, household income, and the match of the proxy and self labor force classifications all had positive effects on correspondence, as expected. Interestingly, total usual hours and correspondence were related positively, possibly suggesting that when there are two job holders in the household, more communication is necessary for effective coordination of household activities.

The sex pairing results did not conform to expectations, as all pairings except "female-female" were equal. The female-female mean correspondence was lower than the others because it contained a group of 4 sibling pairs. Removing spouses (different sexes) from the sex pairing data did not affect the sex pairing results.

Conclusions

The high rate of divergence in the labor force classifications of the proxy and self reports demonstrates the difficulty of obtaining valid information from proxies. One possible remedy is to not use proxies, but this is a dramatically expensive and impractical course to take. A second possibility is to collect only information that proxies can provide reliably. The present analysis suggests that proxy accuracy to an extent depends on the form of the question and the type of information requested. Further research is necessary to ascertain how survey designers might

incorporate the consideration of such factors in the questionnaire construction.

A third option is to identify variables that reliably indicate the person who is the most knowledgeable about other household members, and ask a few screening questions prior to the survey to select a household respondent. On the whole, our analysis suggests that simple sociodemographic variables may be useful for predicting proxy report accuracy. In future research, we also plan to investigate the suitability of measures of household responsibilities, as well as direct questions about who might be the best proxy reporter for the household.

References

- Boussard, J. (1975). *The large family system*. Westport, CT: Greenwood Press.
- Deakins, A. H., Osterink, C., Hoey, T. (1987). Topics in same and mixed sex conversations. In L. B. Nadler, M. K. Nadler, & W. R. Todd-Mancillas (Eds.), *Advances in gender and communication research*. New York: University Press of America.
- Eakins, B., & Eakins, G. (1978). *Sex differences in human communication*. Boston: Houghton-Mifflin.
- Tonn, B. E. (1984). A sociopsychological contribution to the theory of individual time-allocation. *Environment and Planning*, 16A, 201-223.