

## DISCUSSION

Seymour Sudman, Survey Research Laboratory

Seymour Sudman, SRL, University of Illinois at Urbana-Champaign, 909 W. Oregon, Urbana IL 61801

One of the roles of a discussant is to find the common themes that unite these papers. I see two major themes emerging in this session. One that emerges from the first two papers deals with new uses of cognitive methods to improve the quality of survey data. The second theme discussed in the last two papers deals with the use of proxy respondents. I'll discuss the papers in the order they were presented.

The paper by Sirken and Hermann, unlike the other more applied papers in this session, is a broad conceptual look at the interaction between survey research and cognitive psychology. Their Figure 1 gives a persuasive description of how the process operates.

One point that seems to be implied in their discussion, but not made quite explicit is that transfers of information and technology do not occur automatically, and not even by having interdisciplinary teams. What are really needed are interdisciplinary researchers who publish in both survey and psychology journals. Bradburn, with his work on memory, and Schwarz, with his work on context effects, are examples, and there are, of course, many others. The current cohort of interdisciplinary researchers were initially trained mainly in psychology, although not always cognitive psychology, and developed their other skills on the job. The next cohort, already developing, are students of the current cohort and are getting their initial training and research experiences while still in school. As this cohort starts publishing, one would hope to see more impact of their work on basic cognitive research.

Let me comment on Sirken and Hermann's proposals to foster more survey oriented basic cognitive research. They suggest extending work that has been done in the health area to other subject matter areas such as labor, agriculture, energy, and justice. This is a sensible idea, but the question is "how to bell the cat?" Given NSF's limited resources, most of the research funds would need to come from the agencies, and by necessity, most of their resources go to more applied research. Fortunately, as illustrated by the published work to date, much of the research can be done with students in laboratory settings so that large amounts of research funds are not always needed.

Their second suggestion is to improve the infrastructures of cognitive research laboratories by undertaking projects to investigate cost and error effects of alternate cognitive methods. A useful first step before undertaking new experiments might be to do a careful analysis (possibly a meta analysis) of cost and error experiences to

date. There is no discussion of how cognitive laboratories should be staffed, but this important issue is raised in the Saffron paper, to which I now turn.

The Saffron paper described three new approaches in the development of population census forms for the 1996 New Zealand Census. I'm going to discuss only the first, the use by SNZ of cognitive evaluation. A significant organizational aspect of their work was that the design team did almost all of the cognitive work themselves.

When cognitive laboratories were first being set up at major data collection organizations they were generally staffed by cognitive psychologists who did all the interviews and interpreted the results to questionnaire designers. In recent years, as illustrated by this paper, the trend has been to democratize the process by spreading the cognitive work out to a broader range of professionals including not only questionnaire designers, but also senior field staff. I would predict that this trend will continue.

There are two reasons why this trend makes sense. First, flexibility is increased and communications problems are reduced by having questionnaire designers directly involved in cognitive interviews. Second, the methods used in cognitive laboratories have been sufficiently described in articles and in books such as Judy Tanur's book Questions About Questions and our books Answering Questions and Thinking About Answers so that survey designers who are not cognitive psychologists are comfortable, after some practice, with using these methods.

I couldn't agree more with Kim that trying to produce quantitative reports from cognitive interviews used to develop questionnaires is inefficient and pointless as a method of persuading subject matter specialists about the cognitive difficulties their proposed questions cause respondents. I would suggest that the next desirable change in this process would be to involve the subject matter specialists directly in doing cognitive interviewing themselves, or at the very least observing the interviews as they occur. This would significantly reduce the communications problems faced by survey designers interacting with subject matter specialists.

The paper by Kojetin and Tanur has some interesting things to say about how people in a household learn about what others in the household are doing and how this differs between adults and youths, but it has several problematic aspects. The most serious problem in the described study is the sample size which was too small to get enough reports that a household member had looked

for work in the past four weeks. In recognition of this, the analysis was extended to the last three years, but this raises several new concerns that need to be considered about retrieval of information from memory. In this study, when differences are seen between self and proxy reports, one does not know whether this is because the proxy never had information about the target or had information that was forgotten or misremembered.

One example of this comes from Table 2 where there is about 80 percent agreement that there was a job search in the three year period, but lower agreement as to when the job search occurred. This difference might well reflect telescoping, the problem respondents have remembering dates about their own activities, as well as those of others in the household. This topic would seem useful to look at more closely as the analysis of the data continues.

The same memory objection can be raised about Table 3. How reasonable is it to expect respondents to remember how proxies learned of their job search or proxies to remember how they learned about an event in the distant past? Nevertheless, the data here look plausible, and differences between adult and youth targets make sense; for example, adult proxies are told less often and so guess more about youth targets than adult targets. I would speculate that many adult respondents might have used a schema in answering this question, such as, "my kid never tells me what s/he is up to, so I have to learn it from others or simply guess." Some cognitive testing of this question might have been informative.

Table 5 attempts to relate mode of transmission to accuracy. Rather than comparing the accuracy and completeness of those who did and did not use a given mode, I compared the accuracy and completeness of those who used different modes. There really isn't very much meaningful difference given the sample sizes and memory problems discussed above.

Table 4 is asking about time spent together and number of interactions in the past week and does not ask about the distant past, but still creates memory problems. One suspects that probably for weekday and possibly even for weekend time spent together many respondents retrieved or computed an average daily rate and multiplied by 5 or 2. Here also, however, the data look plausible and confirm our expectations. It might be interesting to compare the results of Table 4 with work by Robinson and others on time use. It appears that what is being tapped here is not what really happened last week, but a more generalized response on closeness of the relationship. This explains why there is a relationship between accuracy of reports of job search that occurred in the past and closeness of the relationship.

To link this paper with the first two, it is possible that the use of cognitive methods before the main study might have clarified some of the issues that I've raised.

The final paper by Goldman and his colleagues reminds us of a principal reason for using proxies-to reduce sample biases by increased sample cooperation. In this study, by using proxies, the cooperation rate went from 34 percent to 93 percent. The question arises, however, as in the previous paper about the quality of the data, here on a potentially sensitive question-does the medical care provider treat HIV+ patients. Ultimately, as I understand it, the aim here is to get a good sample of patients.

The data clearly indicate that a higher percentage of MDs (45%) report treating HIV patients than do proxies (about 30%). The question that arises is whether this is caused by sample selection or by proxy misclassifications.

As indicated in the previous paper, the most satisfactory way of determining this is to obtain information simultaneously from a sample of pairs of physicians and proxies. That wasn't done here since the primary purpose of the study was substantive, not methodological. Thus, the authors are left with the difficult task of attempting to parcel out sample selection effects using mathematical models.

As I understand it, interviewers were told to interview whoever was most readily available. This would tend to reduce the likelihood of bias. If proxies were only chosen if physicians refused, and if physicians who did not treat HIV+ patients were more likely to refuse, then one would see a lower percentage of proxies reporting treatment.

Alternatively, what could be happening is that some proxies are unaware of cases, especially when the physician treats few cases. That this is a plausible explanation is seen in the differences between the clinical and administrative proxies. Administrative and clerical staff are less likely to know what the doctor is doing than are nurses and other clinical staff. Of course, a combination of effects could be occurring. It should be noted that the sample bias in using proxies is not simply the difference in percentages of physicians treating any patients, but is really a difference in the number of patients identified in this way. That is, if most of the proxy errors are in settings where the physician treats very few cases, then the net effect on a sample of patients would be small. Still another explanation is that respondents lied. I find this much less likely since there is no really good reason why they should lie, and if they did lie why there should be differences between clinical and administrative staffs.

What are the practical conclusions from this for use of proxies in physician surveys? The results would certainly suggest the use of nurses and other clinical staff as proxies for physicians who are unwilling to cooperate. There is some modest reduction in completeness, but overall the reduction in sample biases outweighs the loss

to response errors. The use of administrative staff is more problematic. Although these are the easiest to reach, their knowledge of clinical issues of the type given in this study is inferior to that of the clinical staff.

I found all of these papers to be stimulating and congratulate the authors.