DISCUSSION

Philip E. Converse, Survey Research Center, The University of Michigan

I have read and heard our papers today with a good deal of interest, since I have been involved myself in fairly large-scale panel operations. The Survey Research Center at the University of Michigan has done several of these studies, both in the area of consumer economics and in the area of political behavior.

My predominant reaction to most of the material in these papers is one of agreement. That is, over that portion of the findings or observations for which we have any analogous findings or experience, I think there is little that I would be inclined to challenge on any factual grounds.

Where the broad themes of the Waksberg-Pearl paper are concerned, for example, it is certainly true that data on gross change often has a rather different cast than knowledge of net change alone would lead one to deduce. It is also true that there are cost-efficiency arguments for panel studies, although my impression is that much of this gain can evaporate if the logic of the panel design is fully realized. Such a full panel design would certainly require that the investigator follow even those people who change their residences between waves of the panel, unlike the procedure used in the Current Population Survey. This is true simply because gross changes in other attributes among these movers may be quite different than gross changes among those who "never left home." But as soon as one begins to follow movers, of course, the expenses per interview begin to rise very rapidly at the same time so that the cost-efficiency argument for panel studies which is presented in the Waksberg-Pearl paper can be rapidly nullified. Finally, as at least two of our papers argue, people may respond somewhat differently to the same items on the second, third or fourth waves than they did in taking the first interview "cold."

It is on this latter truth--the presence of contamination, practice effect or what you will--that I am choosing to focus in my remarks, in part because it is a matter on which two of these papers make extended comment, taking a rather different view, and in part because we in our panel work have seen similar effects and have still a different impression of them.

Let us start, then, with the common-sense assumption that whatever these effects of repeated interviews may be, they are likely to vary in magnitude and perhaps in nature as a function of various conditions of the investigation. Under "conditions" we might list a number of mechanical design features such as the periodicity of the interviewing, for example. We might list as well features of a less mechanical nature such as the style and the content of the interview. And finally we might even want to include the intentions of the investigator as something which can vary, so that the same reinterviewing effect that delights one investigator may shatter another who has different ends in view.

Now I bring up some of these variable conditions because we know very little about their impact, although these papers are a good start on them. And I bring them up because the panel studies I have been associated with differ rather radically from those discussed here with respect to a number of these conditions. Hence it may be worth trying to put together things that we have noted in our studies with some of the observations in these papers and do a little triangulating to see what reasonable guesses about these effects we may go on to make.

In order for us to understand differences in conditions of investigation I must give at least a brief description of the panel study we have done which departs most notably from those described today. Our panel studies have been election studies, hence the content has been political. The fact that they are election studies also helps to determine the periods of interviewing. Between 1956 and 1960 we conducted a 5-wave national panel study. The waves were phased as follows. The first interview was held just before the 1956 presidential election and was followed by a brief interview after the election. The third wave did not take place until after the off-year congressional election of 1958. There was a similar two-year gap between the third and fourth waves. The fourth wave was taken just before the 1960 presidential election and the final wave involved interviews just after that election.

This was the basic timing of the panel. Now let me suggest some of what I would see to be the primary differences between this design and those which we have heard discussed. First, of course, we are dealing in these three designs with radically different sample sizes from the 300-odd involved in the Ferber study to the enormous case numbers of the Current Population Survey. Our panel involved roughly 1500 cases and hence lies somewhat between the two. Secondly, our design differed from each of the others described in terms of the fact that we attempted to follow people shifting their place of residence between waves.

I think that these two initial differences are largely irrelevant for the comments I wish to make, but there are a number of further differences which are extremely relevant. First, the lapse of time between waves--the periodicity of the interviewing--is much different in our study from any of those described here. Our design involved two lapses of virtually two years rather than the maximum of eight months between the fourth and fifth interviews in the Current Population Survey and the much shorter period described in the Harris and Ferber papers. Secondly, our design was more like the Ferber and Harris studies and quite unlike the current population survey in the lack of governmental suasion, moral if not legal, which is convenient in securing those second, third and fourth interviews from the respondents. Third, our study was like the Harris in subject matter; it was political rather than economic and hence there was a focus on relatively "soft" data--political attitudes--rather than some of what we may think of as relatively "hard" information about employment and personal finances. Finally, our study was not designed primarily as a methodolological investigation. This means that there was no attempt at external validation as in the Ferber study, nor were there the control samples available to the Current Population Survey. At the same time there was a fair amount of internal evidence on contamination and in 1960 for the fourth and fifth waves of the panel there was indeed an independent fresh sample with some overlapping content which permitted comparisons.

Any concern over bias or contamination introduced in the later stages of repeated interviewing naturally divides itself into two parts: (1) the concern over panel mortality or panel attrition, and biases which may thereby result; and (2) contamination of those respondents who stay with the study through all or most of the waves.

With respect to panel mortality, apparently the problem was not a pressing one in the Current Population Survey. However, it was a problem in the Ferber study and it was a problem in ours. In absolute terms, the attrition was very large over this period of time. In our second wave, for example, we were able to reinterview 91 percent of those respondents whom we had interviewed in the first wave. This figure dropped to 70 percent in the third wave, 63 percent in the fourth wave and 61 percent in the fifth wave. Now, of course, death and senility, ignored in these raw figures, take on significant proportions over a four-year period. Hence a better statement of our panel mortality in the sense which it interests us would undoubtedly be closer to two-thirds of the original respondents being successfully reinterviewed in the final two waves.

One way of asking what kind of bias this rather large attrition may have created for us is to compare the 1956 characteristics--soft ones as well as hard--of 1960 survivors with the original pool of 1956 respondents from which they were drawn. You will note here that there is no question of real change in characteristics over the time period since we are referring to the original 1956 characteristics of the 1960 survivors.

With such an examination across a very large number of variables involved in the study, the bias turned out to be remarkably slight in at least one sense. For virtually all variables the survivors were at most one or two percent different with respect to variables (partitioned three, four and five ways) from the original pool from which they had been drawn. There was one major exception. On measures of political involvement, this discrepancy increased to three or four percent; that is, the 1960 survivors included a three to four percent overrepresentation of people who in 1956 had registered themselves as being highly involved and, conversely, was depleted by three to four percent among those least politically involved. Further inspection showed that the other smaller discrepancies which did occur tended to turn up in characteristics correlated with political involvement. That is, there was a slight increase of a percent or two in the proportion of the survivors who had had college education, and involvement and education have some fair positive correlation with one another. This configuration of results suggested a very clear pattern: a systematic loss of those less interested in the subject matter of the study without any noteworthy loss along any other systematic lines save those correlated with the first. Such results, of course, would lead one to suspect that if the subject matter had been different, the nature of selection of respondents who persisted with the study would have looked slightly different as well.

Without further comment on this at the moment, let us turn to the problem of contamination of those who were successfully reinterviewed, for this information sheds interesting light on the dynamics of those whom we did not succeed in reinterviewing. Where such evidence of contamination is concerned there were some interesting differences between the general flavor of our results and those which you have heard reported today. That is, in the Ferber study as well as the Waksberg-Pearl study one of the major changes in response occurs between the first and ensuing interviews. This is true of our panel in terms of sheer proportion of mortality, but where contamination is concerned our findings were just the reverse: the signs of contamination started very slowly if at all through the third interview and only began to be noticeable in 1960 at the time of the fourth and fifth interviews.

I suspect that these are not contradictory results but that they do point up some of the crucial variables involved. First, one might mention the expectation of being reinterviewed as a fairly crucial matter. It is my understanding that in both of the other studies the respondents were apprised that they would be reinterviewed. In our study they were not so forewarned. We did not say that we were not coming back, of course, but we made no point of suggesting that we would. The second variable, which interacts with the first, is the sheer time lapse between interviews (in our case two years), a matter which would tend to reduce expectation of being reinterviewed even further. As I have suggested, we found almost no evidence of contamination in the 1958 interviewing. Perhaps the sole potential instance which came to light was the respondent who remarked to the interviewer in effect, "If I'd a known you were coming I would have studied up." However, by 1960 more than a handful were beginning to say to the interviewer such things as "I bet my husband you'd be back this fall."

Triangulating across these several studies it seems clear that the expectation of being reinterviewed as well as the time lapse between interviews can drastically influence the degree to which the respondent is sensitized to the area of investigation in a manner which prepares him for any reinterviews. How much difference this sensitization makes in behavior patterns or response patterns is almost certain to be a function of the subject matter and, quite reasonably, how sensitive the individual was to the subject matter in the first place. In other words, we would propose that the effects of sensitization are greatest (although they may come on more slowly) in areas where the respondent is normally least sensitized. A standard cross section of the public tends to be less sensitized to politics than to the details of personal saving habits and probably less sensitive to the details of personal saving habits than to whether or not he happens to be employed. A number of studies suggest that inaccurate reports, even on sensitive matters like income, are very often nonmalicious: they represent honest ignorance or at least honest failure to retain details. In studies of personal time use, for example, the person who is to serve as a respondent is often subjected to the forewarning that he will be so asked because it improves the accuracy of his recall. It seems clear that sensitization of this sort is one of the prime effects in the Ferber study.

The question which remains, of course, is whether such sensitization is desirable or undesirable. Ferber is pleased at the phenomenon and we are horrified by exactly the same phenomenon, but we do not have to look far to see why this difference in reactions occurs. For here is where the whole intention of a study becomes crucial. Ferber has a right to be pleased with the effects of sensitization for it pulls his reports closer to the parameters which he hopes to estimate. We, in our work, had cause to be alarmed by exactly the same effects because there was reason to believe that this sensitization was promoting changes in actual behavior in the phenomenon which we wished to investigate. In political studies there are few variables more powerful in mediating crucial differences of behavior than political involvement and political information. In 1958 there was little reason to believe that the political involvement or the political information of our respondents had been affected by the repeated interviewing, but there were quite noteworthy indications that this was true by 1960. In point of fact, after the 1960 election study was completed there was a 5 percent surplus of people reporting that they had voted in the presidential election, a mysterious increment which remained by comparison with the fresh control sample after we had taken account of the differential panel mortality selecting toward those who were more involved. In other words, one effect of our repeated interviewing, and the sensitization which accompanied it, may actually have been to stimulate some people to vote who would not otherwise have voted!

It is clear in this instance, of course, that while a sensitization to politics may have given us better information in those cases where we asked the respondents to recall details of past behavior, analogous to the Ferber results, we were at the same time affecting our sample, and inadvertently drawing its sample characteristics out of line with many of the population parameters which we were trying to estimate. You will note that this contamination of those whom we successfully reinterviewed produces a bias which goes in the same direction as the panel mortality bias, so that the effects of the two put together are joint or cumulative.

I suppose that even in the Ferber case, we must keep in mind the possibility that the study, with its sensitization toward personal bank deposits, might actually have affected respondents' saving behavior, as well as their accuracy of recall of that behavior. Such a change, if it existed, would not of course show up as a discrepancy in the "validation" or "criterion" data used in the study. However, I think we would agree that such effects, if plausible at all, should be slight simply because the behavior in question is less "elastic" than the decision to vote in an election or to pay more rather than less attention to politics in the newspapers. Employment behavior would seem less elastic still: we might reasonably doubt that the Current Population Survey interviewing leads anybody to pick up jobs, drop them, or in other ways change their job-seeking behavior. Nevertheless, Mr. Waksberg assures me that the items employed to determine labor force status inevitably have their "soft" edges involving perceptions as to what a "job" is and the like. Furthermore, we note that response changes as a consequence of repeated interviewing seem most marked in the Current Population Survey case among teenagers for whom indeed labor force behavior is a little more "soft" and definitions more optional.

In closing we might note that it is somewhat simpler to attempt a diagnosis of the effects of repeated interviewing and the contextual conditions which make them greater or lesser than to suggest how we might go about controlling them in those cases, unlike Ferber's, where we find them damaging. If the period between waves is at all short, as it often must be, it is impossible to remove the respondent's expectation of being interrogated again in whatever area of behavior is involved in the investigation. Conceivably, in the long run if we wish to preserve the values of the panel design we may be driven to highly multiple-purpose studies of a sort which will permit sufficiently rapid rotation of content within and across waves that the respondent is kept, so to speak, "off balance," without any clear expectation as to the content areas to be investigated in the next interview.

XI

CONTRIBUTED PAPERS - II