

MEASURING THE EFFECTS OF THE TRANSITION ON FIELD PERFORMANCE¹

Clyde Tucker and Brian Kojetin, Bureau of Labor Statistics

Richard Blass and Ronald Tucker, Bureau of the Census

Clyde Tucker, BLS, 2 Massachusetts Ave., N.E., Room 4915, Washington, D.C. 20212

KEY WORDS: Current Population Survey, Quality Control, Interviewer Effects

1. Introduction

Survey researchers have long been interested in the effects that interviewers have on the data they collect. In summarizing this research, Groves (1989) characterizes researchers who have studied interviewer effects as approaching the topic from either the standpoint of reducing measurement error or estimating the size of the errors (Bailey et al. 1978; Cannell et al. 1981; Fowler and Mangione 1985; Groves and Magilavy 1986; Hanson and Marks 1958; Hill 1991, 1993; O'Muircheartaigh 1986; Stokes and Yeh 1988; Tucker 1983). This paper is concerned with the estimation of errors or at least the potential for error. It carries research on interviewer effects a step further, however, by examining how changes in technology and questionnaire design can affect interviewer performance. Previous work has been done in this area (Blom et al. 1991; Bradburn and Frankel 1991; Sebestik et al. 1988), but the present research is the most comprehensive to date. In this case, the redesign of the Current Population Survey (CPS) is the subject of study. Comparisons are made of the behavior of individual interviewers from both before and after the implementation of the CPS redesign.

2. A Model of Interviewer Effects

Figure 1 outlines our model of interviewer effects. We examine two types of effects. One type is the direct measure of interviewer performance, such as response rate or proxy reporting. The other is more indirect--the possible distortion or bias in substantive survey results. This distortion can, to be sure, arise from nonresponse or proxy reporting (as indicated by the arrow from these variables to the substantive ones), but interviewers also can bias survey results independently of the other aspects of their performance.

Both types of effects are the products of a complex causal process. At the highest level, the national political and social context can place constraints on interviewers, although these constraints often can be affected by the survey organization (Kojetin, Tucker

and Cashman 1994). The influence of these particular variables, however, can only be estimated over a long period of time. Given the recent changes in the CPS, at least the effects of changing survey procedures at the national level can be considered. Below this level, administrative differences across the regional offices could cause interviewer performance to vary.

Of course, regional office effects must be disentangled from differences resulting simply from the natural variation of social and economic conditions among the regions, especially in the case of substantive variables. So the characteristics of an interviewer's assignment area as well as the distributional features of his or her respondents should be included as controls in any model. These factors are known to influence both the level of nonresponse and the prevalence of response error. (DeMaio 1980; Sudman and Bradburn 1973; Schuman and Presser 1981). Finally, although the evidence is somewhat weak (Groves 1989), differences among the interviewers, themselves, are likely to be related to interviewer performance. (Couper and Groves 1992). Respondents reactions may be related to the interviewer's race, gender or age, especially when the questions concern subjects related to these characteristics. The interviewer's own expectations may also affect his or her performance (Singer and Kohnke-Aguirre 1979).

3. Study Design

3.1 Overview

In recent years there has been a trend toward the use of computer technology for the collection of survey data from households. Besides CATI (Computer Assisted Telephone Interviewing) centers which have been in operation for a number of years, the availability of lightweight laptop computers has spurred more widespread use of CAPI (Computer Assisted Personal Interviewing). The use of computers in personal interviews changes the role of the interviewer in important ways (Bateson and Hunter 1991). Couper and Burt (1993) note that although "... the technical feasibility of CAPI (in terms of hardware and software) has been demonstrated in a number of studies in the U.S. and abroad...we still have little empirical

¹The views expressed in the paper are those of the authors and do not necessarily represent those of the Bureau of Labor Statistics or the Bureau of the Census.

information on the impact of this new technology on the people involved in the process of data collection, particularly the interviewers."

The transition of the CPS from paper and pencil to CAPI provides a valuable opportunity to examine the effects of changing technology on the survey workforce. Furthermore, information about interviewer effects, in general, might be obtained, because these effects could be especially pronounced when interviewers are undergoing major changes in the way they perform their jobs. Complicating this analysis will be the fact that important changes were made in other aspects of the survey design. In particular, major modifications to question wording accompanied the transition to CAPI.

With the transition of CPS to CAPI in January, 1994, nearly 2000 supervisors and field representatives (interviewers) were trained to use laptop computers and the CPS CAPI instrument. The analyses reported here were restricted to a set of 1000 field representatives who worked on CPS continuously from October, 1993 (prior to the redesign) through May, 1994, five months after implementation of the redesign. Supervisors and new CPS interviewers were excluded. In some analyses, the number of interviewers is slightly less because of missing data.

3.2 Measures

As mentioned above, two types of interviewer performance measures, our dependent variables, were used. This information was available for the last three months of 1993 and the first five months of 1994. The direct measures were the overall nonresponse rate (including refusals and noncontacts) and the percentage of interviews conducted with proxy respondents. The substantive measures for each interviewer were the percentage of their respondents who were employed, the unemployment rate of their respondents, and the labor force participation rate of their respondents (the proportion classified as either employed or unemployed).

Information about the interviewer's length of tenure with the Census Bureau and their employment status (regular or sporadic) was obtained from records in the Census regional offices. Because the training program for the redesigned CPS spanned several months, data on how far in advance of the transition a particular interviewer was trained was collected. The interviewer's regional office along with some information about the previous CPS experience of that office's administrative personnel also were known.

A variety of other variables were included as controls for differences in the interviewers' assignment areas. Distributional characteristics of an interviewer's

respondents with respect to age, race, ethnicity, sex and marital status were used. Information about the entire assignment area included urbanicity and extent of poverty.

3.3 Analytical approach

The analyses involved hierarchical multiple regressions conducted in several stages to examine the independent effects of different factors. Separate models were created for each dependent variable. Some of these stages were used to introduce controls and others followed from the hypothesized sequence in our model of interviewer effects. The value of having data for the three months prior to the transition was that they could be used to construct baseline measures for more accurately assessing the effects of the changes in procedures, including interactions with prior conditions.

4. Results

4.1 Changes in Interviewer Performance during the Transition

As can be seen in Figure 2, there was a dramatic shift upward in the overall Type A rate for the CPS during the transition in January. Compared to the previous year, it has also been much more unstable during the first five months of 1994. Although the overall level has come down, it is still above the previous level. Furthermore, it is not yet clear at what level nonresponse rates will stabilize. There was also a shift upward in the percentage of reports given by proxies from the baseline of about 44% to the transition rate of about 47% that has remained quite stable. Both of these overall shifts raise concerns about interviewer performance and data quality during the first 5 months of the CPS CAPI transition.

To examine these changes in interviewer performance more carefully, we have focused our investigation, as previously mentioned, on interviewers who were doing the CPS with PAPI in October, November, and December of 1993 and made the transition to the CPS CAPI system in January through May of 1994. They not only represent the largest core of the CPS interviewing staff, but they also are the ones for whom the effect of the transition from PAPI to CAPI can be most thoroughly examined. Figure 3 shows the distributions of the average Type A rates for this core group of interviewers during the Baseline period from October to December, 1993 and the transition period from January through May, 1994. It can clearly be seen that the distribution has shifted and that it is more positively skewed after the transition than it was during the baseline period. Similarly, Figure 4 shows the distributions of proxy rates for

interviewers during the baseline and transition periods. The distribution after the transition has shifted markedly from the baseline and has become more concentrated at the higher percentages.

4.2 Modeling Interviewer Performance

Given these clear shifts in interviewer performance, the next phase of our analyses concerned the explanation of these changes. To examine what factors might be related to the changes in interviewers' Type A rates, we regressed their Type A rates during the first five months of the CAPI transition on their baseline Type A rates, creating a residual variable representing the changes not accounted for by knowing the baseline Type A rate. We then entered variables into this regression equation to attempt to explain the residual variance. We entered a set of variables representing assignment area characteristics, the characteristics of respondents who were interviewed and indicators for the regional office. After controlling for all of these changes from baseline on Type A rates, we entered interviewer characteristics including their length of experience at the Census Bureau, whether they were regular part-time or intermittent interviewers, their caseload, and when they were trained on CAPI.

The results of these analyses can be seen in the first two columns of Table 1. Specifically, the first row shows the percentage of variance of the interviewer performance variables accounted for by the interviewer baseline levels. The second line shows the contribution of assignment area variables. The third line of Table 1 shows the results of entering respondent characteristics on the interviewer performance measures and the fourth line adds the regional office to the regression equation. In each of these steps the change in R^2 is significant. Finally, the interviewer characteristics are entered on the last line, but this change in R^2 is not significant for either type A rates or proxy rates.

Similar analyses were also allowing a stepwise inclusion of interviewer characteristics at the last step. For type A rates, whether or not the interviewer is consistently part-time compared to intermittent was significant, with interviewers who were part-time showing marginally less change in their Type A rates during the transition than intermittent interviewers. None of the interviewer characteristics significantly added to the prediction of proxy rates (not shown).

The changes from baseline to the transition are shown by regional office in Figure 5 for nonresponse rates. In particular, the changes in Type A nonresponse rates for this group of interviewers were not uniform across the country. There was less

variability across regional offices in the changes in percentage of proxy reports.

4.3 Modeling Interviewer Level Substantive Measures

Similar analyses were also conducted on the labor force survey estimates, including Percent Employed, Unemployment Rate, and Labor Force Participation. In addition, the interviewer performance measures were also included with the interviewer characteristics in the equations predicting labor force status. As can be seen in the last three columns of Table 1, inclusion of assignment area, respondent characteristics and regional office resulted in a significant change in R^2 . Finally, interviewer characteristics and performance measures were entered on the last line, but this change in R^2 is not significant for any of the substantive outcome measures.

Similar analyses were also conducted entering baseline, assignment area, regional office, and respondent characteristics and then allowing a stepwise inclusion of interviewer characteristics and performance measures. For percent employed and labor force participation, greater interviewer experience was associated with less change in these substantive outcomes. Analyses of unemployment rate revealed that higher interviewer caseloads were associated with greater changes in unemployment rate.

5. Conclusions

The present research was an initial attempt to enrich the research literature integrating the study of interviewer effects and performance during a major transition in data collection mode in an ongoing survey. We were able to identify and obtain baseline measures of interviewer performance and substantive survey variables prior to the transition for comparison to post-transition performance and results along with some information about the interviewers. The baseline data made each interviewer their own control, and additional control variables for the assignment area and respondent characteristics were also utilized to limit competing explanations for interviewer effects.

The transition of the CPS to a CAPI instrument in January of 1994 clearly had an impact on the field performance of the interviewing staff. The present investigation found only slight impact of interviewer characteristics on changes in interviewer performance during the transition after controlling for the interviewers' baseline performance, assignment area, respondent characteristics, and regional office characteristics. The significant differences among Census Bureau regional offices that were found point to variation in administrative procedures and training.

Also, there could be an interaction between changes to the survey and geographic characteristics.

The evidence that interviewer characteristics impacted upon changes in the substantive results of the survey, after controlling for the interviewers' baseline level of labor force characteristics, assignment area, respondent, and regional office characteristics was quite small. Only limited information on interviewer characteristics was available, and other variables, such as interviewer's expectations or their confidence with the computer and new CPS instrument, may have been useful predictors or their performance. It is also clear that during this transition period there were real changes in the labor market which are currently included in our models only indirectly though differences across regional offices, assignment area characteristics, or respondent demographic characteristics.

The amount of variance explained in unemployment and Type A nonresponse rates is a good deal less than that accounted for with the other three variables. This may be explained by the large relative variances in these two measures, given that the average proportions for both are less than .10. The substantial increase in variation explained in the proxy rate, percent unemployed, and labor force participation may be the result of skewed proportions in some demographic groups for certain interviewers, particularly with respect to age.

The interviewers on which we focused the present analyses represented a large core of the interviewing staff who went through the CAPI training program and made the transition from PAPI to CAPI. We did not include new hires, or interviewers who had worked on the CPS CAPI instrument as part the research overlap sample from July, 1992 to December, 1993. It is also likely that these interviewers had an impact on the overall field performance during the transition and deserve some attention also, but we cannot perform analogous baseline comparisons for these interviewers. In further research we also plan to explore the impact of interviewers on household level analyses and to create multi-level models .

6. References

Bailey, L., Moore, T. F., & Bailar, B. (1978). An Interview Variance Study for the Eight Impact Cities of the National Crime Survey's Cities Sample. *Journal of the American Statistical Association*, 73, 16-23.

Blom et al. (1991). *Computer Assisted Data Collection Production Test in the Labour Force Surveys: Report of Technical Tests*. CADAC Project, Statistics Sweden.

Bradburn, N. M. & Frankel, M. (1991). A Comparison of Computer-Assisted Personal Interviews (CAPI) with Personal Interviews in the National Longitudinal Survey of Labor Behavior-Youth Cohort. *Proceedings of the Bureau of the Census Annual Research Conference*, 389-97.

Cannell, C. F., Miller, P. V., & Oskenberg, L. (1981). Research on Interviewing Techniques. In S. Leinhardt (Ed.) *Sociological Methodology*. San Francisco: Jossey-Bass.

Couper, M. P. & Burt, G. (1993). Interviewer Reactions to Computer-Assisted Personal Interviewing (CAPI). *Proceedings of the Bureau of the Census Annual Research Conference*, 429-450.

Couper, M. P. & Groves, R. M. (1992). The Role of the Interviewer in Survey Participation. *Survey Methodology*, 18, 263-277.

DeMaio, T. J. (1980). Refusals: Who, Where and Why. *Public Opinion Quarterly*, 223-233.

Fowler, F. J., Jr., & Mangione, T. W. (1985). *The Value of Interviewer Training and Supervision*. Final report to the National Center for Health Services Research, Grant #3-R18-HS04189.

Groves, R. M. (1989). *Survey Errors and Survey Costs*. New York: Wiley.

Groves, R. M. & Magilavy, L. J. (1986). Measuring and Explaining Interviewer Effects in Centralized Telephone Surveys. *Public Opinion Quarterly*, 50, 251-266.

Hanson, R. H. & Marks, E. S. (1958). Influence of the Interviewer on the Accuracy of Survey Results. *Journal of the American Statistical Association*, 53, 635-655.

Hill, D. (1991). Interviewer, Respondent, and Regional Office Effects on Response Variance: A Statistical Decomposition. In P. Biemer, R. M. Groves, L. E. Lyberg, & S. Sudman (Eds.) *Measurement Errors in Surveys*, New York: Wiley, pp. 439-459.

Hill, D. (1993). Response and Sequencing Errors in Surveys: A Discrete Contagious Regression Analysis. *Journal of the American Statistical Association*, 88, 775-781.

Kojetin, B., Tucker, C. & Cashman, E. (1994). *Response to a Government Survey as Political Participation: The Relation of Economic and Political Conditions to Refusal Rates in the Current Population Survey*. Paper presented at the Annual Meeting of the American Association for Public Opinion Research, Danvers, Massachusetts.

O'Muirheartaigh, C. (1986). Correlates of Reinterview Response Inconsistency in the Current Population Survey. *Proceedings of the Bureau of*

the Census Annual Research Conference, 208-234.

Schuman, H. & Presser, S. (1981). *Questions and Answers in Attitude Surveys*. New York: Academic Press.

Sebestik, J., Zelon, H., DeWitt, D., O'Reilly, J.M., & McGowan, K. (1988). Initial Experiences with CAPI. *Proceedings of the Bureau of the Census Annual Research Conference*, 389-97.

Singer, E. & Kohnke-Aguirre, L. (1979). Interviewer Expectation Effects: A Replication and Extension. *Public Opinion Quarterly*, 43, 245-260.

Stokes, L. & Yeh, M. (1988). Searching for Causes of Interviewer Effects in Telephone Surveys. In R. M Groves, P. P. Biemer, L. E. Lyberg, J. T. Massey W. L. Nicholls, & J. Waksberg (Eds.) *Telephone Survey Methodology*, pp. 357-373. New York: Wiley.

Sudman, S. & Bradburn, N. M. (1974). *Response Effects in Surveys: A Review and Synthesis*. Chicago: Aldine.

Tucker, C. (1983). Interviewer Effects in Telephone Surveys. *Public Opinion Quarterly*, 47, 84-95.

Figure 1. Model of Interviewer Effects

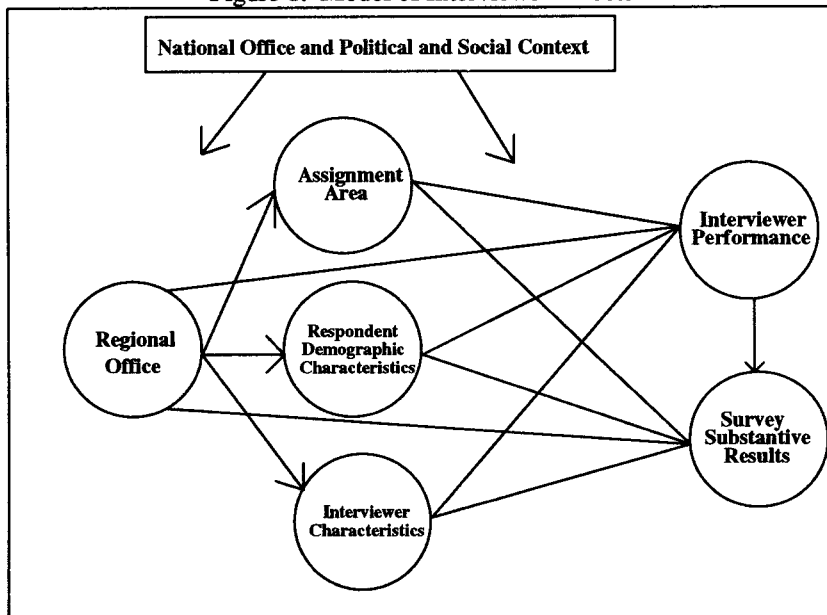


Figure 2. CPS Type A Rates and Refusal Rates

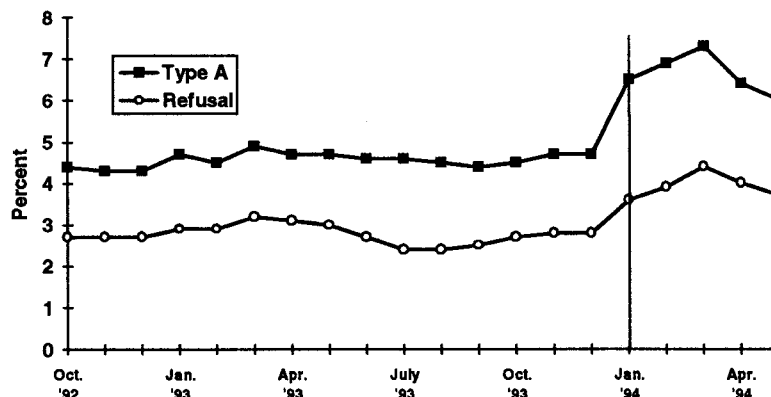


Figure 3. Distribution of Interviewer Type A Rates

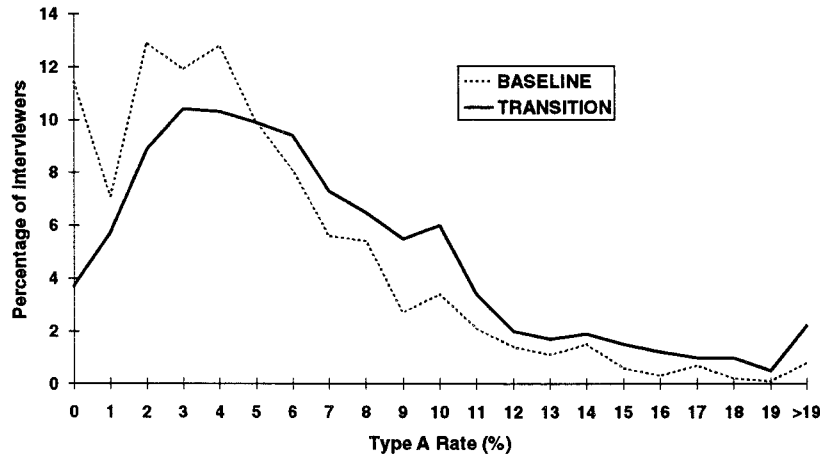
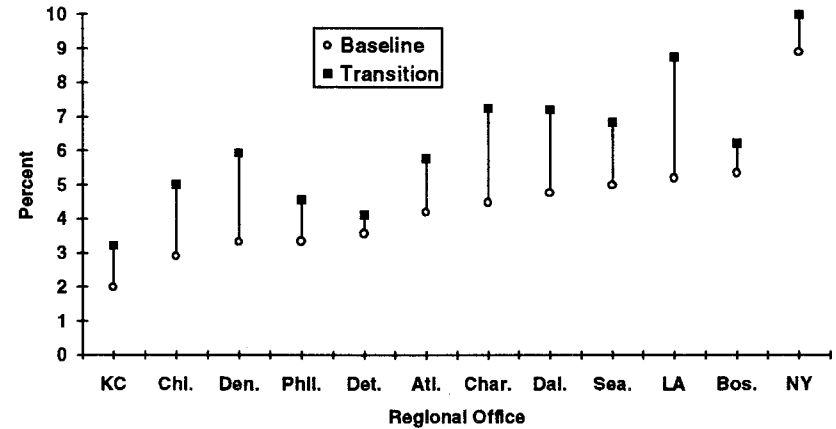


Figure 5. Type A Rates by Regional Office for Interviewers doing both Old and New CPS



353

Figure 4. Distribution of Interviewer Proxy Rates

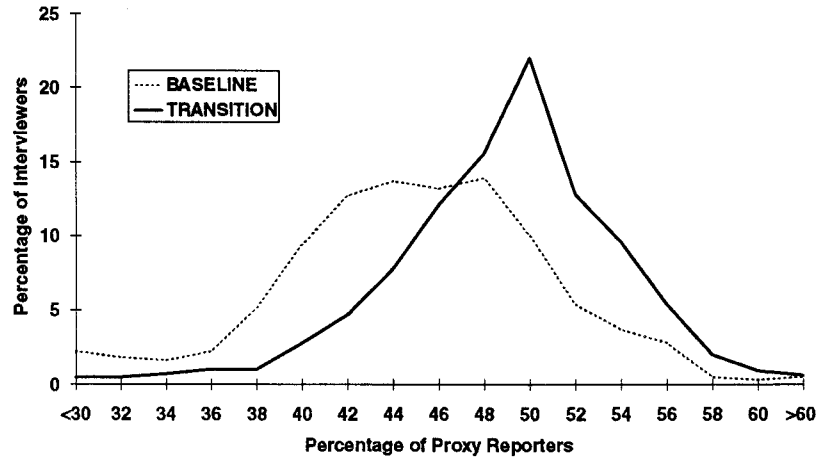


Table 1. Results of Hierarchical Multiple Regressions
(Values shown are Cumulative R²)

	Type A rate	Proxy Rate	% Employed	Unempl. Rate	L.F. Part.
Baseline	.284*	.277*	.520*	.289*	.471*
Assignment Area	.299*	.322*	.560*	.332*	.509*
Respondent Characteristics	.328*	.505*	.718*	.368*	.724*
Regional Office	.379*	.539*	.744*	.405*	.747*
Interviewer Characteristics	.383	.540	---	---	---
Interviewer Characteristics and Performance	---	---	.746	.411	.749

* change in R² significant at $p < .05$