

TOWARD A QUALITY IMPROVEMENT SYSTEM FOR FIELD INTERVIEWING: PUTTING CONTENT REINTERVIEW INTO PERSPECTIVE

Patrick J. Cantwell, John M. Bushery, Bureau of the Census, Paul P. Biemer, Research Triangle Institute
Patrick J. Cantwell, Bureau of the Census, Washington, DC 20233*

KEY WORDS: Falsification, quick contact, interviewer evaluation, statistical process control, response bias.

of the data, how reliable it is, and when changes to the system may be affecting responses.

1. Introduction

In conducting demographic surveys for our sponsors, the Census Bureau contacts some of the original sample households a second time shortly after the initial interview. Depending on the survey, the respondent is asked again all or part of the questionnaire for one or more household members. This second contact, sometimes by personal visit but usually by telephone, is called *content reinterview*.

There are three reasons the Bureau conducts a program of reinterview and related activities, all of them involving the quality of the data produced by our surveys. These are:

- ▶ to deter and detect falsification, i.e., fabrication of interviews, or any major procedural violations, such as purposely misclassifying occupied households as vacant. Without some type of deterrent activity, even good interviewers may occasionally break the rules under the pressure of time or performance ratings.
- ▶ to control and improve interviewer performance. Currently, we select interviewers for the reinterview sample, recontact some or all of their workload, and either pass or fail the interviewer based on the results of the reinterviews.
- ▶ to collect information to improve the survey process. Several times, reinterview has identified problems of understanding with particular questions. It may indicate that respondents are not interpreting the questions as intended, or that interviewers are not following procedures properly.

Although these goals of reinterview are common to all the Bureau's demographic surveys, certain sponsors are interested in monitoring other aspects of survey quality. In particular, reinterview is also used

- ▶ to measure and reduce response error. By monitoring the levels and changes of response bias or variance, we can sometimes evaluate the quality

To assist in addressing these goals, the Census Bureau conducts a series of field activities, including training programs for interviewers, observing interviewers (especially during their first few months), editing responses turned in to the regional office, and checking that area listings are correct. However, content reinterview continues to play a large part in each of the goals mentioned above.

As an example, until this year, the reinterview program for the Current Population Survey (CPS) consisted of two parts.¹ At the centralized telephone facility in Hagerstown, Maryland, the reinterview attempted to replicate the original interview to provide a measure of response variance. Thus, a regular interviewer--rather than a supervisory staff member--made the call, the entire core questionnaire was repeated, and no reconciliation was done. Checking for falsification is not so important here since all centralized telephone interviewing is subject to call monitoring.

In the field a different set of reinterviewing procedures was implemented. Approximately 25% of CPS reinterviews were conducted in a manner similar to that of Hagerstown, i.e., unreconciled reinterviews. In the other 75%, a supervisory staff member recontacted the household, asked the entire core questionnaire, and reconciled any discrepancies between the initial and reinterview. When conducted in this manner, content reinterview tries to address all the purposes stated above--falsification, performance improvement, survey improvement, and measuring response bias (rather than variance).

We question whether, in trying to satisfy so many of the goals of a quality control system, content reinterview is addressing any of them efficiently and properly. Consider falsification. While content reinterview may well deter most and detect some instances of falsification, the respondents pay the price by repeating answers about some or all household members. In the CPS reinterview, for example, reinterview respondents generally answer, for the second time in a week, the entire core questionnaire for all eligible household members. We doubt it is necessary to place this burden on respondents. To measure response error, the reinterview could ask about

a subset of household members. To address falsification, it may not be necessary to repeat the questionnaire for *any* members.

Second, it is unclear whether content reinterview is the best tool to evaluate interviewer performance. Based on the results of the reinterview, the interviewer passes or fails the performance check. The philosophy of total quality management is to assist the worker to improve her performance by providing steady, constructive feedback. The statistical process control methods we will describe shortly may enable us to measure the performance of the regional office as a whole and its interviewers individually. By monitoring the processes, management can decide when sweeping action involving the entire regional office is necessary, as well as when individual attention and training is warranted.

Finally, we question whether the present setting for content reinterview leads to good measures of bias and variance. By trying to accomplish so much, the reinterview often yields inferior results obtaining the actual value of a characteristic (when estimating response bias) or replicating the original interview (when estimating response variance).

For example, Biemer and Forsman (1992) have determined that at least 50% of the errors in the original interview are not detected by the reconciled reinterview process. Sudman, Blair, and Menon (1992) has found that a large portion of the detected differences are the result of different respondents in the initial and reinterview. O'Muirheartaigh (1991), using CPS data, has determined that the reinterview survey without reconciliation may underestimate the response variance in the original survey by 30% or more and that this bias is not uniformly distributed across the subpopulations often compared in response variance studies.

Part of the problem might be that, in designing the reinterview sample to check for falsification and to evaluate the interviewer's performance, the typical number of reinterview cases is quite large. Supervisory staff have indicated that they often lack the time to conduct the reinterview as intended. Sometimes they settle for a respondent different from the preferred one. Often they cannot determine satisfactorily the source of the discrepancy or which response is closest to the true value.

The approach we are advocating is to separate the goals of the quality improvement system and to address each through different activities. Although one activity might address other goals indirectly, it would be designed intently for its primary goal. This system would

- 1) check for falsification *quickly and cheaply* and with a minimum of respondent burden,
- 2) evaluate the performance of interviewers individually and as part of the entire regional office staff, and provide constructive feedback as necessary, and
- 3) *using a smaller sample*, conduct a content reinterview designed to create the most favorable conditions (a) to elicit the "true value" of the target characteristic (when measuring response bias), or (b) to replicate as nearly as possible the initial interview (when measuring response variance).

The sections which follow address in order the three components just listed. However, our discussion of the third activity is limited to part (a).

2. Detecting and Deterring Falsification

A Field Representative (FR) (interviewer) falsifies when he or she knowingly deviates from the specified procedures to avoid interviewing, classifying, or listing units. Falsification also might include accepting proxy information when self-response is required, interviewing ineligible proxies, using the telephone when a personal visit is required, or fabricating answers to individual questions in an interview. Falsification also occurs when an FR knowingly deviates from the specified procedure and tries to conceal it.

Much of our information about falsification in Census Bureau surveys comes from the Interviewer Falsification Study (Schreiner, Pennie, and Newbrough, 1988). The database, recently updated to include information on all confirmed cases of falsification from 1982 to 1992, records characteristics of the interviewer caught and relevant details of the case.

How extensive is the problem? In the ten years of the study, we confirmed 305 cases of falsification; however, the rate varies by survey. In the Current Population Survey (CPS), with a work force of about 1500 FRs, only 113 cases of falsification were confirmed in ten years, about 0.8 percent of the current FRs per year. A rough calculation shows that this level of detection is consistent with about 1.3% of the FRs falsifying 5% of their assignments, or about 5.1% of the FRs falsifying 1% of their assignment. On the other hand, the New York City Housing and Vacancy Survey is conducted in an area which has a disproportionately high share of obstacles in completing cases. In the NYCHVS, 5.5% of the FRs were found to have

falsified data.

2.1 Current Methods to Control Falsification

The Census Bureau currently relies primarily on its reinterview program to detect and deter falsification. The extent of the reinterview ranges widely among surveys. The CPS recontacts one household member, but then asks all questions about each eligible person in the housing unit. In the National Crime and Victimization Survey, one household respondent is selected randomly for a complete reinterview. In one-time surveys, like the National Survey of Fishing, Hunting, and Wildlife-Associated Recreation, a "quick check" is often conducted to make sure an interview actually occurred.

The typical reinterview design selects a sample of FRs each month and a subsample of their workloads for the month they fall in reinterview. Survey supervisors or supervisory field representatives (SFRs) conduct the reinterview.

Because cost is a major factor, we usually conduct reinterviews by telephone. In the Bureau's five major surveys, if the sample unit has no telephone (or no one answers) the interviewer often accepts a noninterview, i.e., the case remains unchecked. To contain costs, units requiring personal visit reinterviews that are more than 50 miles from the regional office or the reinterviewer's home are considered ineligible for reinterview. This weakens the Bureau's falsification protection, especially if FRs are aware of the policy.

In addition to detecting falsification, the Bureau is trying to remove demoralizing factors, such as excessively long questionnaires and complex, difficult, or personal questions. Supervisors and SFRs try to provide their staffs with contact and support. Some regional offices have a "pre-observation program" in which newly hired FRs visit households with an experienced FR, before receiving training or assignments, to see exactly what the job entails. This gives FRs not suited to the job a face-saving chance to quit before they or the Bureau makes a big investment in their training.

2.2 New Methods--Quick Contacts

A content reinterview may not be the most efficient means to detect falsification. We offer two alternatives. One is a telephone quick contact--a very short telephone call. The second method is a short questionnaire mailed to a subset of original respondents with a return envelope. In either case, only a few questions would be asked, such as

- ▶ Did someone from the Census Bureau contact you or someone in your household? When?
- ▶ Was this contact by telephone or in person?

The telephone quick contact to deter and detect falsification requires less of the reinterviewer's time than a content reinterview, is less repetitious (it does not repeat any of the original interview), and places a smaller burden on the respondent. (On the other hand, some have argued that, with the time required to recontact households--including unsuccessful attempts--we may just as well keep the respondent a little longer and complete a content reinterview, asking some or all of the questions.)

Mailings sent to vacant households could have one of three resolutions:

- ▶ returned by the U.S. Postal Service, indicating that the housing unit was vacant or did not exist,
- ▶ returned by a resident, indicating occupancy (including usual residence elsewhere or member of the armed forces), or
- ▶ no return (undetermined status).

The mailout check would elicit responses from most respondents much more cheaply than a content reinterview or a telephone quick contact. Further, it would reach households without a telephone which are currently reinterviewed in person. Again, it does not repeat any of the original interview. Because respondents are not required to answer the telephone or doorbell, it could prove less annoying than a reinterview call. The chief drawback would probably be a lower response rate.

How important is a high response rate to the method's feasibility? Recall that our target is to deter falsification. Because FRs will not know which respondents will fall into the mail check nor which ones will return their cards, the deterrent effect may still be there. One question for the research is whether the response rate for the difficult-to-enumerate cases will be high enough to make a mail falsification check effective.

Research may indicate that a mixed-mode reinterview will work most efficiently. We might use telephone quick contact for households with telephones and the mailout check for households requiring a personal visit. Another option is a maximum mail-out check **followed** by a telephone quick contact for the **nonrespondents**.

2.3 Quick Contact Research Plans

To compare quick contact approaches for detecting falsification, we plan to measure response rates, costs, and falsification detection rate. The Bureau has decided to implement such a study on the CATI/CAPI Overlap (CCO) Panel, which fields the new CPS questionnaire, for two reasons. Only the CCO survey instrument will be used in CPS beginning January, 1994. In addition, the CCO had already planned a sample of telephone quick contacts to check for falsification.

During the CCO, one-sixth of all interviewers are scheduled for a falsification check each month. There are two main components of the falsification check and research. First, for a randomly selected portion of their completed interviews and type B/C noninterviews (vacant or demolished households), a telephone quick contact will be attempted by the assigned reinterviewer if a phone number is available. In all, approximately 400 telephone quick contacts are planned each month across all regional offices to check for falsification during the CCO. Our intent is to measure for this mode of reinterview the completion rate, the distribution of the number of calls per case, the cost and time spent per case (as compared to content reinterview), the population coverage attained, and the detection of falsification.

Second, mailout questionnaires will be sent to a sample of households. We will select from completed interviews as well from type B/C noninterviews, where a telephone number is often not available. At the outset of the experiment, about 750 mail questionnaires per month are planned throughout the country. As indicated in the last section, mailings sent to households which are actually vacant should be returned by the U.S. Postal Service. Here we hope to gauge the mail response rate, the distribution of return times, the cost per case, what happens to questionnaires mailed to vacant or demolished homes, and the detection of falsification.

In addition to the activities just described, others are planned on a smaller scale in one or more months of the study. From those households which did not send back mail questionnaires, we will follow up a sample with a **telephone** quick contact. This will help us resolve mail nonresponses--whether they simply did not reply to our reinterview questionnaire, or whether no one can be reached at the address.

From those households who did not send back mail questionnaires and could not be reached by telephone (no number available or no response), we will follow up a sample with a short **personal visit**. As well as resolving cases of reinterview nonresponse, we can measure how often the U.S. Postal Service returns mail

from vacant or demolished addresses.

Although data will be collected on the two quick contact methods used to check falsification from October 1992 through September 1993, budget or analysis of response rates may compel us to stop earlier. When this part is completed, one or more **focus groups** of FRs will be conducted to elicit their reactions to the methods. In particular, we wish to learn if either method appears to be a stronger or weaker deterrent to falsification than content reinterview.

3 Using Statistical Process Control (SPC) to Monitor FR and Regional Office Performance

Statistical process control (SPC) is an essential tool for managing the performance and data quality generated in any survey quality improvement system. The Census Bureau is researching how to develop and use SPC tools effectively. Other organizations have already started.

Italy's National Institute of Statistics has begun using SPC techniques in its survey quality control system for the Italian household surveys. This system aims "to prevent, correct and evaluate non-sampling errors" (Masselli and Signore, 1991). The National Opinion Research Center (NORC) at the University of Chicago uses a computer-based Survey Quality Assurance System to provide "a single point of entry for survey managers and NORC clients to monitor progress on all key activities and engage in a joint management process" (from NORC's quality profile statement). Since the application of SPC to survey data collection is a recent development, the Census Bureau needs to gain experience using SPC techniques on our survey performance data.

The purpose of SPC is to identify the causes of poor quality and to maintain an acceptable level. Alone, it cannot improve quality. But using SPC to monitor survey data and performance will help distinguish between common causes and special causes of error and take appropriate corrective action.

Common causes are unassignable to one person and are, therefore, a problem with "the process." For example, inadequate maps may contribute to listing errors in area segments. We can use appropriate problem-solving strategies to improve this aspect of the process, rather than assigning fault to FRs or supervisors.

SPC can also signal when **special causes** of error are operating. Special causes are aberrations from the normal survey process. They can include FRs or regional office supervisors who need additional training or instruction. For example, an FR who does not know

how to read segment maps may make numerous errors listing area segments. The corrective action for special causes could be feedback about errors committed, retraining, or extra coaching. SPC can indicate when corrective actions are appropriate.

The Bureau is taking a number of specific steps to achieve successful SPC systems for its surveys. Some of these are:

- Identify where SPC can be applied and conduct the necessary research to apply the methods.
- Train and empower supervisors, SFRs, FRs, and CATI interviewers to use SPC and other statistical tools to analyze the data for their surveys. Currently, regional office supervisors use only simple averages and rates to improve performance.
- Provide timely and user-friendly access to the data at the regional offices as well as at headquarters.
- Change the "corporate culture" in the regional offices, at the telephone centers, and at headquarters to emphasize quality--as measured by carefully chosen performance indicators--as well as response rates, productivity and cost containment. And
- Provide and use effective means to take corrective actions, and determine how effective these actions are.

3.1 SPC in the National Health Interview Survey (NHIS)

The NHIS research will first develop and test the feasibility of SPC as a primary tool for survey management. It will concentrate on developing SPC for the NHIS, but would also identify new measures and improve feedback systems. A "process action team" at the Census Bureau has begun the preliminary steps of this feasibility study and has documented their progress in McCall and Rogers (1992).

This study will address traditional performance measures, such as response rates, item nonresponse, reinterview results, and questionnaire edit results. It will also measure aspects of performance not currently monitored, such as supplement response and data quality, timeliness of completing assignments, productivity and itinerary planning, and the cost implications of improved response rates. Specific measures will depend on information collected in the

first phase of the study.

In the first phase, using focus groups, questionnaires, and supervisor and SFR conferences, NHIS managers and supervisors will be asked what survey and administrative data they use or would like to have for managing FR performance. Further, they will be asked about the feedback they give and receive from FRs. What information is the basis of the feedback? Does the feedback achieve its purpose? What system does management have to obtain feedback from the FRs?

The second phase will test the use of SPC in one or more regional offices. The Charlotte regional office has been recommended as the first choice, as we are already working with them to investigate SPC in the NHIS. To start, we must produce the data the first phase suggests would be useful in an SPC system, provide the software and other tools needed to analyze these data, and train the NHIS managers, supervisors, and SFRs in SPC and other techniques. The system will then be operated and modified as necessary over a period of months.

The next step is to evaluate the success of the system. By monitoring performance data, we can observe whether FRs who need feedback or corrective action get it, and whether the action helps the FR improve her performance and maintain that level. One element of the system is to monitor the variability of the performance measures among assignments. Can SPC distinguish (i) between individual FR performance problems and the variability expected among FR assignments; (ii) between problems and the ordinary month-to-month variation within an individual FR's assignments; and (iii) between common causes (issues involving the entire survey process) and special causes (individual problems)? Finally, the managers and supervisors must indicate how much burden the analysis is, and whether they find it helpful and worth continuing.

3.2 SPC in the CATI/CAPI Overlap Survey

To develop a successful SPC program in future survey systems, we must update the traditional measures of FR performance and quality to reflect operations in a Computer Assisted Survey Information Collection (CASIC) environment.

Fundamental differences between the NHIS and the CPS make this area of development and testing important. Unlike the NHIS, the CPS (i) will employ CATI, CAPI, and decentralized telephone interviewing, (ii) is a panel survey, (iii) continues to place primary emphasis on its "core" questionnaire, rather than supplements, and (iv) has its data processed by the

Census Bureau. As the CPS better represents the broader range of continuing surveys than the NHIS, the CPS CATI/CAPI Overlap Panel would be the most appropriate testing vehicle for this track. However, the proposal is still under consideration.

Although we have not yet developed the details of this study, it is essential to producing an SPC program which can address performance and quality issues in a CASIC environment. Some of these issues follow.

- FR assignments will change greatly in composition due to the broadened use of centralized telephone interviewing. Further, assignment composition and size will become more variable among FRs, and from one month to another for individual FRs. The higher variability in assignments will increase the variability of commonly used performance measures, such as response rates and production rates, making it more difficult for supervisors to decide when corrective action is appropriate. We need to provide more powerful statistical tools to deal with this added variability.
- CASIC will make some of the current quality issues irrelevant, such as interviewer branching errors (skip-pattern errors).
- CASIC will give us the opportunity to measure new aspects of the survey process: length of interview, and distribution of completed interviews and noninterviews by time of day, day of week, number of contacts and attempted contacts, etc.
- CASIC can provide the opportunity for real-time access of content data--if we design the system to provide it--permitting better management of data quality. Current plans for CASIC do not include this possibility.

4 "True Value" Reinterview

When reinterview responses can be regarded as very nearly the truth, they can be compared with responses from the original interview and help indicate which items in the survey are most prone to error. However, the Bureau's reinterview surveys, as currently designed, suffer from a number of operational problems which diminish the accuracy of their responses, introducing bias into the measurements. In the introduction to this paper, several sources are cited to support this statement.

Current content reinterview procedures for most cases in the Bureau's demographic surveys specify that

- the reinterviewer be the supervisor of the field representative (FR) who conducted the original interview,
- the reinterview be conducted by telephone, when possible,
- proxy responses may be taken in the reinterview,
- the wording of the reinterview question be more or less identical to the original wording,
- discrepancies between the interview and reinterview are shown to the reinterview respondent and an explanation is requested (reconciliation),
- the reinterviewer (the supervisor) determines which response is correct after reconciliation, and
- the reinterviews are conducted on a very tight schedule with little time available to follow-up nonrespondents.

All of these features of the design can introduce bias into the true-value reinterview process. The purpose of our research is to investigate the extent to which they affect reinterview accuracy and to recommend cost effective alternatives.

4.1 Exploring Alternative Approaches For Obtaining the Truth

For the initial interview in a survey, the goal is to produce the most accurate data within specified cost and time constraints. For a true-value reinterview, the goal is the same, but the constraints are different. Because the data are not tabulated and published as in the original interview, the variance requirements are not as strict, and thus, the sample size can be much smaller. Reinterview does not have the same data release deadlines, so that procedures which require more time to execute may be considered. Together, these differences can allow greater time and attention allotted to the reinterview.

Still, there is a point of diminishing returns beyond which additional expenditures will not increase the accuracy of survey characteristics appreciably. One purpose of the proposed research is to help the designers of reinterview surveys determine this point.

The features of the Bureau's current design for reinterview give rise to eight areas of reinterview

research. The research being planned or under way attempts to address the first five of these areas.

a. **The reinterview introduction.** The current procedure is to introduce the reinterview as a check on the FR's work. However, the respondent and reinterviewer can be influenced by this introduction to the point where response quality may suffer. The present research explores alternative introductions which enhance respondent cooperation while encouraging truthful responses and their accurate recording.

Possible approaches to introducing the reinterview include (a) to check the FR's performance, (b) to verify the respondent's previous answers, (c) to ensure the quality of the data, or (d) to improve the survey.

b. **The prior notification of reinterview.** Currently we give respondents little or no prior notification that they may be recontacted. Several options are available: (a) no prior notification, (b) the FR informs the respondent or provides a written statement *at the time of the initial interview*, (c) the FR informs the respondent or provides a written statement *after each interview*, or (d) the regional office sends out written notification to households to be reinterviewed just prior to the date of the reinterview (e.g., 3 days prior).

Respondents for panel studies are generally told that they will be contacted again in a specified number of months. They are often upset when another interviewer contacts them for a reinterview several days or weeks later. On the other hand, early suggestion of a reinterview may provide an excuse for respondents to refuse the entire process. The fourth option notifies only those households targeted for reinterview. However, it is also more costly, involving a separate mailing from the regional office.

c. **Reinterview respondent rules.** One of the basic tenets of survey methodology is that self response is almost always preferred to proxy response, the exception being highly sensitive or embarrassing information. In most Bureau reinterview surveys, the usual rule is to maximize response by the initial respondent. This reflects the current goal of checking the original FR's work.

To obtain more accurate information in the reinterview, we must explore alternative respondent rules. For example, for family, household, and housing units items, we can develop criteria and screening questions which will identify the individual in the household who can provide the most accurate information on these items, the "most knowledgeable

respondent." For person-level data, we can explore a self-response rule for reinterview.

d. **Choice of reinterviewer.** The SFR currently is the reinterviewer for most Bureau reinterview surveys. But if we can obtain accurate responses to improve the survey, alternatives to the SFR, such as experienced FRs, can be reasonably considered.

e. **Resolution of discrepancies.** When there is a difference between the original and the reinterview response, the reinterviewer attempts to discreetly identify the source of the discrepancy, with the help of the reinterview respondent. This process of "dependent reconciliation" may well put the respondent on the defensive and could elicit inaccurate responses. Sensing this, reinterviewers might not press the respondent to revise his responses to obtain the most accurate information. An alternative approach which has been used in Sweden (Bergman et al., 1991) is to conduct an in-depth reinterview without reconciliation.

These five research topics can largely be explored generically for all Bureau surveys since they are somewhat independent of the survey content. The next three design features are more survey specific and the optimal reinterview design for these may differ substantially among surveys.

f. **Timing of reinterview.** Currently, the reinterview is conducted as soon as possible after the close-out of the survey period in order to provide timely feedback to FRs regarding their performances. If we reinterview too soon after the initial contact, some respondents simply remember previous answers and repeat them. If we wait much longer, they might forget the required information or events. In determining the appropriate interval between interview and reinterview, the particular survey content should be considered. In the current research, the timing of the reinterview will not be manipulated. However, thorough coverage of this topic for Bureau surveys may be found in Bailar (1968).

g. **The reinterview content and questionnaire design.** Perhaps the most essential components of the reinterview are its questionnaire and procedures. The questionnaire must be designed to obtain the most accurate responses. Because this typically requires longer, more burdensome questioning, only a subset of items from the original survey can often be covered in the reinterview. Our research, however, will not examine different sets of questions, but will use a single reinterview instrument.

h. **Mode of interview.** The usual mode of interview for Bureau reinterview surveys is the telephone (when available) and face-to-face, otherwise. However, for most survey items, convention says that the preferred mode of interview is face-to-face. Given the emphasis in this research toward developing reinterview methods which obtain highly accurate data for small-scale survey improvement studies, we will use face-to-face interviewing exclusively in this research.

4.2 Overview of the Research Plans

Currently, there is no plan to investigate research issues (f) through (h). In this section, we briefly describe the plans for areas (a) through (e). Our immediate interests in improving the accuracy of the data collected in "true-value" reinterviews are

- a) To test alternative methods for introducing the reinterview to the respondent which result in higher cooperation rates and increase respondent interest in participating in the reinterview,
- b) To test the effect of a prenotification letter on respondent and interviewer acceptance of the reinterview,
- c) To investigate the effect on the accuracy of the reinterview data of using a most knowledgeable respondent (MKR) rule for obtaining reinterview responses,
- d) To investigate the effect on data accuracy and interviewer and respondent acceptance of a "no reconciliation" approach to obtain the true survey values, and
- e) To determine the effect on costs and data accuracy of using regular FRs--rather than supervisory personnel--to conduct the reinterviews.

The Census Bureau's Survey of Income and Program Participation (SIPP) provides an ideal opportunity for investigating true-value reinterview at this time. Bureau scientists will be testing a cognitively designed version of the SIPP questionnaire this fall (see Moore, Bogen, and Marquis, 1992). Their plan involves conducting the current SIPP interview in 350 households and giving the cognitively designed version to another 350. Administrative records will be used on all 700 households to assess the accuracy of the SIPP responses. We will use a subsample of the first set,

those households receiving the usual SIPP questionnaire.

The reinterview research will be pursued in three phases. The first involves conducting focus groups to determine promising methods for resolving issues (a) through (e) above. The focus groups will be conducted with SIPP data analysts, SIPP respondents or potential respondents, and SIPP interviewers and supervisors. This phase will develop feasible combinations of these factors for testing in two field experiments.

To identify the MKR, the focus groups will try to identify various family structures, who keeps the income information, and how to identify knowledgeable respondents, such as the wage earner, the bill payer, and the record keeper. To our knowledge, this information has only been used once in a large national survey, the National Medicare Expenditure Survey, to identify the preferred respondent to report on health care utilization and expenditures for family members. However, such an approach has not been used in identifying reinterview respondents. The product of these focus groups would be a set of criteria to be tested in the next phase.

The second phase of the research consists of a small field experiment to test procedures to introduce the reinterview, to notify respondents, and to identify the MKR. The most promising procedures developed in the first phase will be compared with the current SIPP procedures. The field experiment, involving 100 interviews, will assess interviewer and respondent reactions to the methods.

The results of the experiment will be analyzed in several ways. First, the reinterviewer will complete several post-interview questions to give a subjective rating of how well the MKR rule appeared to work. The reinterviewers will also be asked to take part in a debriefing conference at the end of the first phase's field period to discuss the strengths and weaknesses of the experimental procedures. Respondent reactions will be obtained at the end of the reinterview on such issues as how difficult it was to provide the requested information. A ratings task may also be incorporated to assess how confident the reinterview respondent was that the information he gave was accurate. Finally, data on the cost of conducting the reinterview will be analyzed, including the number of callbacks required to reach the MKR and the total length of the reinterview.

The objective of the third phase of the research is to test and evaluate the "no reconciliation" approach to resolving discrepancies and the efficiency of conducting reinterviews using regular FRs rather than supervisory field staff. The plan is to compare the accuracy, costs, response rates, and operational complexity under the several alternatives.

Approximately 200 reinterviews will be conducted on a subsample of the 700 households participating in the SIPP cognitive interview research described above. These 200 households will be selected from the 350 households administered the regular SIPP interview. The reinterviews will be conducted after the second wave of the cognitive interview experiment.

The "no reconciliation" reinterview questionnaire will embody the latest developments in SIPP cognitive interview design. In brief, these interview procedures rely much less on respondent recall and much more on the use of household records to obtain income and program participation data. Using records whenever possible, respondents are asked to help the interviewer fill out a worksheet for each income source that provides the amount and the date of payment. The approach incorporates a number of additional improvements to ensure accurate data, such as the use of flashcards to jog the respondent's memory about potential sources of income, and obtaining responses from all respondents in the household as a group.

Research to date on this approach has been promising. This research into the "no reconciliation" approach will borrow heavily from these developments. Our hypothesis is that the procedure can provide accurate SIPP data in the reinterview without the need to reconcile discrepancies. The sample of size 200 will be split between this approach and the usual SIPP reinterview procedure. Administrative records will be accessible for all 200 cases to assess the relative accuracy of the two approaches.

5 Summary and Other Research

While the content reinterview program as currently implemented succeeds in at least partially meeting its intentions, we suggest that changes be investigated. Instead of requiring one format to address each of the important goals, it may be more productive and cost effective to address the goals through separate activities. This paper describes research under way to explore how feasible and effective these methods can be.

- 1) *A system of telephone quick contacts, mailout questionnaires, or a combination of the two might replace part of the current reinterview sample, thereby addressing falsification more cheaply and with less burden on the respondent. The CATI/CAPI Overlap Panel provides the setting to test these alternatives.*
- 2) *Statistical process control applied to FR performance data at and below the regional*

office level can become a contributing part of the feedback cycle to FRs. What information to use is being solicited from field staff across the country. How to develop and implement these procedures will be examined in one of our regional offices on an experimental basis.

- 3) *Research on new techniques for conducting a content reinterview to better measure bias is planned. We will try to determine improved methods to introduce the reinterview, to notify prospective respondents, to select the preferred household member, and to resolve discrepancies. Further, we hope to resolve whether FRs, as well as SFRs, may be used to conduct reinterview. This research is planned on a sample of SIPP reinterview respondents.*

Other topics relating to reinterview or a quality improvement system have been considered for research. Some of them follow.

- 4) *The use of SPC on survey data at the FR level to signal possible falsification. These data might include edit error rates and FR productivity rates. FRs would still receive supplemental reinterviews to check for falsification. Currently, only 8% percent of all confirmed falsifiers are caught by administrative analysis in the regional offices. Research is under way to determine whether analytic methods can effectively detect falsifiers.*
- 5) *Time and date stamping. This technique attaches times and dates to specific events recorded on the FR's laptop computer, e.g., when a specific case was retrieved or returned to the assignment database, when the interview began or ended, or when a new household member started an interview. This stamping tells us how much time specific activities take, and when and how frequently they occur. These time data would provide estimates on average time to complete interviews, time spent on noninterviews, and time required to travel--all information which is currently unreliable for many surveys. Further, unusual time sequences may alert us to problems with a questionnaire, with the survey process, or with individual FR performance.*
- 6) *The effect of reinterview on subsequent*

response rates. We have not seen an analysis, in any panel or longitudinal survey, of the response rates for the wave or interview *after* a household has been reinterviewed. Survey supervisors in several regional offices have expressed a fear that reinterview might hurt response in later waves among households whose participation is already precarious. However, this belief is not universally held. A simple task would be to compute these rates for one or several surveys, and compare them statistically with households who have not been placed in reinterview. Of particular interest is the effect on subsequent response rates of a reinterview in the first wave or month in sample.

ACKNOWLEDGMENTS

We would like to thank Irv Schreiner for his assistance in several technical matters, Paul Siegel for his helpful comments, and members of the Field Division at the Census Bureau headquarters and the Regional Offices for many insightful discussions and suggestions.

REFERENCES

- Bailar, B.A. (1968). "Recent Research in Reinterview Procedures," *Journal of the American Statistical Association*, 63, 1, p. 41-63.
- Bergman, L., Kristiansson, K., and Säfström, M. (1991). "Decentralized CATI Versus Paper-and-Pencil Interviewing: Effects on the Results in the Swedish Labor Force Surveys." Unpublished manuscript. Stockholm: Statistics Sweden.
- Biemer, P.P. and Forsman, G. (1992). "On the Quality of Reinterview Data with Application to the Current Population Survey," *Journal of the American Statistical Association*, in press.
- Masselli, M. and Signore, M. (1990). "Quality Control System and Field Work Control in the Italian

Surveys," paper presented at the International Conference on Measurement Errors in Surveys, Tucson, AZ, November, 1990.

- McCall, L. and Rogers, P. (1992). "Quality Management Approach to Field Procedures for the National Health Interview Survey," Proceedings of the 1992 Annual Research Conference, in press.
- Moore, J.C., Bogen, K.E. and Marquis, K.H. (1992). "A Cognitive Redesign of the Survey of Income and Program Participation: Development of Procedures and Initial Test Results," presented to the Census Advisory Committee of the ASA, April 9, 1991.
- O'Muircheartaigh, C.A. (1991). "Simple Response Variance: Estimation and Determinants," in Biemer, et al. (eds.), *Measurement Errors in Surveys*, Wiley, N.Y.
- Schreiner, I., Pennie, K, and Newbrough, J. (1988). "Interview Falsification in Census Bureau Surveys," *Proceedings of the Survey Research Section*, American Statistical Association, p. 491-496.
- Sudman, S., Blair, J., and Menon, G. (1992). "Respondent Reactions to Reinterviews," *Proceedings of the Survey Research Section*, American Statistical Association, in press.
- ¹ Recently, the unreconciled reinterview program has been discontinued to create funds for the research described in this paper.
- * This paper reports the general results of research undertaken by staff of the Census Bureau and the Research Triangle Institute. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau or RTI.