THE SURVEY OF INCOME AND
PROGRAM PARTICIPATION

A COGNITIVE APPROACH TO REDESIGNING MEASUREMENT IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

No. 150

K. H. Marquis, J. C. Moore and
K. E. Bogen
Bureau of the Census

U. S. Department of Commerce BUREAU OF THE CENSUS

A COGNITIVE APPROACH TO REDESIGNING MEASUREMENT IN THE SURVEY OF INCOME AND PROGRAM PARTICIPATION

Kent H. Marquis, Jeffrey C. Moore and Karen E. Bogen[1]
United States Census Bureau, Washington DC 20233

KEYWORDS: Measurement errors, Cognitive research

The Survey of Income and Program Participation (SIPP) is a major information resource for policy scientists and planners in the United States. By helping analysts learn about the principal determinants and correlates of income and program participation among U.S. residents, SIPP helps improve our understanding of the financial well being of families and leads to meaningful improvements in national welfare and entitlement programs, employer mandates, and national tax policy.

The ability of SIPP to meet its objectives rests heavily on the quality of the survey measurements that determine its estimates. In this paper we will briefly examine the quality of selected SIPP measurements, examine some explanations for the observed errors, describe some changes we are proposing to make in the measurement design and discuss the evaluation of those changes.

1. KINDS AND AMOUNTS OF MEASUREMENT ERRORS

1.1 SIPP Record Check

We used the SIPP record check study to evaluate SIPP measurement quality. It compared responses from sample persons residing in 4 states (Florida, New York, Pennsylvania, Wisconsin) to values in administrative records for participation (yes, no) for 8 months in 8 programs: Aid to Families with Dependent Children (AFDC), Food Stamps (FOOD), Unemployment Insurance (UNEM), Workers' Compensation (WORK), Civil Service Retirement (CSRetire), Social Security (OASDI), Supplemental Security Income (SSI), and Veterans' Pensions/compensation (VETS). The SIPP data are from the first 2 interviews of the 1984 panel. We matched using variables such as name, address, social security number and date of birth. For a more complete description of the record check study, see Moore and Marquis (1989) or Marquis and Moore (1990).

1.2 Estimates of Bias Due to Response Errors

To provide a summary picture of the effects of SIPP response errors for program participation measures, we have condensed the main descriptive findings from the record check study into Table 1. The eight programs are in the left column. The remaining columns show the percent biases in estimates of means and correlations due to response errors. The methods used to obtain these estimates are explained in Marquis and Moore (1990).

If one were estimating parameters for program status variables, such as the proportion of sample people participating in AFDC, or the correlation between AFDC status and income, the first two data columns are relevant. For the record check sample, the first column indicates that estimates of program participation rates will be too low, generally underestimated by 3 to 20 percent, due to the

Table 1. SIPP Response Errors Have Substantial Effects on Parameter Estimates

Program	Program Status – Effect on		Program Change – Effect on		
	Mean	Correlation	Mean Off Seam	Mean On Seam	Correlation
AFDC	-39%	-35%	-64%	+ 20%	-78%
FOOD	-13	-19	-32	+135	-65
OASDI	+ 1	- 7	- 6	+132	-89
SSI	-12	-18	-34	*	-95
UNEM	-20	-33	-32	+ 28	-58
WORK	-18	-51	+ 3	+ 80	-71
CSRetire	- 8	-11	*	*	*
VETS	- 3	-16	*	*	*

* Percent change bias undefined, true change = 0.

response errors. The 39 percent underestimate shown for AFDC is atypically high because AFDC is often called "General Welfare" in Pennsylvania (and Pennsylvania respondents form a big chunk of the sample for this analysis). We need to take account of these net underreporting biases in redesigning the measurement procedures.

Looking at the second column, correlations involving program participation variables will be underestimated also, by anywhere from 7 percent to 51 percent, depending on the program. To give an example, suppose we had a perfect measure of income for people in our sample and correlated the SIPP food stamp participation measure with it. Suppose further that the true correlation were .5. The observed correlation, using measured participation, would be 19% less or (1-.19) * .5 = .41. This kind of bias makes it harder for analysts to detect true relationships with program participation. The implications for redesign are that we should broaden our attention beyond the net response bias to include the variable response errors that affect estimates of association, such as the correlation or the regression coefficient.

Policy planners in government agencies are often interested in measures of program change, either joining the program or leaving it. In the last three data columns, we show the biases involved in estimating parameters with change variables, omitting estimates for programs with little or no true change. First, note that the effect of response errors on estimates of the mean depends on where the change estimates are taken. If the change in program participation takes place between two months measured within the same interview, we say that it takes place "off the seam." If the change is measured for a pair of months covered in adjacent interviews, we say that the change took place "on the seam" between the two interviews. According to the estimates in the "off" seam column, SIPP response errors generally cause a downward bias in the estimated proportion of program changes when the pair of measures comes from the same interview. On the other hand, if the change measure is "on" the seam, the biases are positive and

often large. This indicates that an analyst will overestimate, usually by a considerable amount, the proportion of status changes for a program. Redesigned measurement procedures need to address the seam bias.

The final column indicates that correlation estimates involving measures of program change will be severely underestimated due to SIPP response errors. The modal bias estimate is around -75%; this indicates that an analyst might estimate a correlation of .12 when the true correlation was .50. Although not shown in the table, the size and sign of the bias for estimates of change correlations do not depend on whether the measure is taken on or off the seam. This is consistent with Young's (1989) empirical results. Because biases appear so severe for an important use of the data, it is important that the redesign focus on getting precisely accurate reports of the date, amount, kind, and recipient of each income payment. This kind of precision is necessary to meet the data quality requirements for modeling program status change.

To summarize, the record check results show that SIPP response errors can cause analysts to misestimate program status parameters such as means and correlations. Biases are especially large for estimates involving measures of program change where means are subject to a seam bias and correlations are severely attenuated. Goals of a measurement redesign should include reducing the severity of these response error effects.

To remedy these effects, we need explanations of, and solutions for, the full range of misclassification errors, not just the net underreporting bias. Next we consider some possible measurement error explanations that may provide guidance toward improved measurement procedures.

2. POSSIBLE REASONS FOR MEASUREMENT ERRORS

2.1 Forgetting and Construction Models for Memory

As a prelude to our discussion, let us mention, in greatly simplified form, two approaches to the explanation of human memory and recall: forgetting and construction.

The forgetting approach views memory as a repository of specific experiences (e.g., memory traces). Memory errors are failures to retrieve a stored experience (or forgetting). The availability or ease of retrieving a stored experience is partly a function of its age; memories are said to decay over time. Memories of your own experiences are more readily available and detailed than memories about someone else's experiences. Survey researchers (e.g., Neter and Waksberg, 1966) have added the concept of telescoping memory errors, the misremembering of when an experience happened (originally the concept referred to remembering something as closer to the present than it truly was, analogous to looking at it through a telescope).

Using administrative record data, Marquis and Moore (1990) asked whether the errors were all forgetting errors (omissions) and whether the probability of forgetting was greater for events occurring in the more distant past (decay). They examined whether errors were greater from proxy than self respondents and conducted some indirect tests of telescoping (was there more overreporting in recent time periods and more underreporting in past periods?). They concluded the following:

-Response errors are not all omissions (underreports) of participation. While there are usually more underreports of the programs, we observed many overreports in programs such as Social Security, Unemployment Compensation, and Food Stamps.

-Omissions (underreports) are generally not more likely to occur for past (four months ago) relative to recent (one month ago) program participation. We tested this prediction for each of the 8 programs. In only one of those 8 tests was there a statistically significant effect in the direction predicted by forgetting theory.

-Those results are relevant to half of the evaluation of telescoping also (and are not in the predicted direction). We also did the 8 tests for whether overreports were greater one month ago compared to 4 months ago. In none of those tests did we find a statistically significant effect in the predicted direction.

-In the 16 tests for self-proxy differences in misclassification errors (8 programs, 2 waves tested separately), three were statistically significant but the sign was opposite the predicted direction (we found that self reports contained more misclassification errors in these instances).

We do not conclude that forgetting theory is wrong but we suspect it is not a very useful model for understanding and fixing the majority of response errors that occur in SIPP. Marquis and Moore (1990) examined the evidence for some other traditional explanatory approaches and found them lacking also.

An alternative conception of human memory, advocated first by Bartlett (1932), we term Constructivist. In this view, memory is an active, inferential, constructing process often guided more by a person's ideas of what should have been than guided by what truly was. Experiences such as the paycheck received three weeks ago are not encoded directly in memory storage, but enough general knowledge is retained and available to come up with an approximation of the characteristics of the paycheck such as its gross and net amounts and its date. The approximation is a product of applying rules of what should have been, such as the usual amount of a paycheck, or the amount of this week's paycheck and, because I didn't get a raise, the inference that the amount three weeks ago was the same.

Ross (1989) has shown that people have their own theories about themselves that they use to construct their personal histories. Personal histories include such things as the personal financial history that SIPP asks about. Ross also shows that peoples' theories about the self over time have a consistency bias which can easily lead to biases in the recall of change. In their classic study, Nisbett and Wilson (1977) show that people often do not perceive themselves as having changed when they truly have.

So, while the constructivist theory is an alternative explanation of reporting errors, it is subject to a different set of biases that affect the reporting of consistency and change.

2.2 Observational Research

The constructivist position receives support from some exploratory observational research that SIPP staff members undertook recently (Marquis, 1990). Because we had been unsuccessful in uncovering the major causes of SIPP response errors through data analysis, we carried out some exploratory observations guided by contemporary principles of cognitive science. Staff members learned to use "think aloud" and direct questioning techniques. They accompanied trained SIPP interviewers and interrupted at selected places to learn whatever the respondent could reveal about (1) the respondent's comprehension of the task, (2)

what the respondent actually recalled, and (3) how they decided what to answer. Observers kept notes, tape recorded the interviews, and prepared written summaries of their impressions and experiences.

Recall and Response Formulation. From our review of the summaries we concluded:

-No observer ever mentioned a respondent who directly recalled an entire 4-month payment history for any routine payment source (e.g., job, investment return, government program).

-Instead (and often encouraged by interviewers), people routinely used simple rules or heuristics, combined with a few recalled facts, to construct the 4 month payment streams. They created plausible stories about the past based on what they thought should have happened. They used the heuristic strategies as a substitute for detailed, direct recall and as a substitute for checking their personal records.

So, with respect to recall and response formulation, the observation research indicated that people use heuristic strategies, recalling only a few facts and using them with oversimplified rules to construct their answers.

Comprehension. The observation reports also addressed apparent misunderstandings of questions. However, we were quite puzzled by the dual result that the misunderstandings were frequent but not necessarily limited to a few "problem" questions. After looking over the detailed notes, we hypothesized that many respondents (and perhaps interviewers also) did not understand the general goals of the section of the questionnaire in which "their" problem question was embedded. If they had only realized that this section was about government programs or about health insurance or that it wasn't about welfare, for example, much of the misunderstanding might have been avoided. To support this notion we noted that the current SIPP questionnaire does not contain statements that introduce each new section or that provide a transition from one topic to another. Indeed, it is sometimes true that questions on various topics are mixed together (for technical or efficiency reasons). Frequent and complex branching rules (skip patterns) occur throughout that rely on the interviewer to remember something about the respondent such as age and marital status, for example. Some interviewers forget and, rather than look up the information, ask it again, further disrupting any topic "context" that might have been created by the group of preceding questions. We suspect that many of the misunderstandings resulted from the lack of a simple explanation of section goals, and a straightforward pursuit of those goals with an uninterrupted set of relevant questions.

So, if people aren't recalling their experiences directly, we should not be surprised that the errors they make aren't the ones predicted by the forgetting model, which assumes direct recall. But the "constructed" past, according to the record check data, isn't always a good approximation of reality either. The remedies for errors due to using simple heuristics are quite different than the remedies for forgetting, memory decay, and proxy biases. We discuss our approach to remedying the response error problem next.

3. CHANGES IN THE SIPP MEASUREMENT DESIGN

In this section we will describe some changes to SIPP measurement procedures that we propose to develop and test. Our emphasis will be on cognitive concepts such as recall, response formulation, contexts, comprehension and motivation.

3.1 Recall and Response Formulation

We plan to give more emphasis to the constructivist model of memory and less emphasis to traditional notions from the forgetting approach. This means we want to preempt the respondent's use of simple heuristic strategies and substitute more accurate ones. We also need to stimulate recall of all sources of income payments, including those that do not come to mind immediately.

Our principle tool involves substituting personal records of individual income payments for the current practice of recalling monthly totals by source. Instead of asking how much you got from your job in November, December, January and February, we will ask how often you are paid and then ask you to get your pay stubs and report each gross paycheck amount starting November 1st and continuing through today. This is not what many practitioners would call a "cognitive approach." We agree. We have gone to this approach because it is clear to us that most human beings cannot provide the information required by SIPP accurately enough using just recall and naturally occurring heuristic rules. Alternative approaches relying more upon recall are possible (see discussion below) but we judge them both more difficult for the respondent and less likely to produce error-free information.

But records can be omitted or lost, just as human experiences can be forgotten. What do we plan to do if a respondent doesn't get records or gets them but throws them away? There are two basic approaches here: Training and Guided Complex Heuristics.

During the first interview in the panel, we expect that many households will not have a complete set of records available. We will ask them to report using what is available and we will offer to come back later if there is a chance that the missing records can be recovered. If the respondent does receive records but doesn't ordinarily save them, we will give the person a special folder in which to save future records (such as pay stubs, bank interest statements, welfare statements, etc.) for future SIPP interviews. If the respondent does not receive records with income payments (a characteristic of some government transfer programs, for example) we will give the respondent a preprinted form and provide training in what to enter after each check arrives in the mail. We will ask the respondent to keep the form with the special record keeping folder for use in later interviews. Our plans are to implement a telephone reminder system between the early interviews in the panel to prompt households, who ordinarily don't keep income records, to maintain their records.

But, especially in the first interview, recall will be necessary for income that cannot be substantiated with records. Our plan is to teach interviewers how to help respondents to construct realistic models of their payment stream from each source. Specifically, we will ask the respondent to think of what determines the size of each (gross) payment, in effect, to construct a model of what causes the amount to increase or decrease (e.g., hours worked, leave taken, overtime rates, varying numbers of days in a pay period, etc), and then to try to recall and reconstruct the variations in the causal variables, using whatever records are available as an aid. We will alert interviewers to the idea that real world payment streams are not just a repetition of the same amount week after week. After going through this experience, we suspect respondents will view the personal record keeping strategy as a welcome alternative to having the interviewer apply the Guided Complex Heuristics approach.

3.2 Cue Giving to Minimize Underreporting

Net underreporting, according to error results cited earlier, is a measurement problem in SIPP. And for the personal records approach to work, we must elicit all sources of income and program participation from each household. Such underreporting could be due to memory failures or deliberate decisions not to report. For the memory possibility, we have adopted the classical memory cueing approach of the recognition list. We derived a list of over 50 of SIPP's highest priority sources of income and created a short (5 min. or less) section that asks whether anyone receives those kinds of income. The section is designed to use as many memory-jogging cues as possible in hopes of eliciting reports of all relevant sources of income. In case net underreporting is due to deliberate withholding, we end the section with a question about sources that "you thought of but decided not to mention" and sources that "you don't want the tax people to know about."

3.3 Procedures to Minimize the Seam Bias

To reduce any seam bias that remains after generally reducing response errors, we will use dependent interviewing, an overlapping reference period and reconciliation.

We don't know exactly what causes seam biases in SIPP and other surveys (e.g., Hill, 1987). Speculation usually implicates a recall decay process whereby past participation and other characteristics are remembered as more like the present than they truly were. To correct such distortions, we will use a dependent procedure that calls the household's attention to any sources of income reported in only one of the two (most recent) interviews. We will ask whether the income was received in both reference periods and we will try to pin down (with personal records) the exact start and stop dates. To further pin down start and stop dates near the seam, we will follow the lead of Statistics Canada (Michaud, et al., forthcoming). We will use partly overlapping reference periods and reconcile any differences in income sources, dates and amounts.

Our earlier work has suggested that we be especially cautious in using dependent interviewing procedures since they have the potential to do harm. According to some exploratory modeling with the record check data (Marquis and Moore, 1989), the seam phenomenon is just a byproduct of the response errors made every month. These errors are slightly less correlated across the seam than between months in the same interview. The lower correlation causes more apparent changes to occur. Thus, anything we do to lower the within-interview response errors may help to lower the seam bias. But the danger in dependent interviewing is that we may only increase the error correlation across interviews, preventing improvements in data quality from one interview to the next. An increased correlation could make the seam bias disappear without curing the underlying response error problem.

To minimize increases in the response error correlation, we will introduce two refinements into the dependent interviewing approach: (1) it will be bi-directional, allowing correction of errors in both directions (vs. transmission of errors in one direction) and (2) the procedures will be introduced late in the interview, after obtaining income reports that are not conditioned by the dependent procedure. In these ways, we hope to gain the advantages of dependent interviewing without experiencing the disadvantages.

3.4 A Context Approach to Comprehension

We have observed many instances of misunderstanding questions and suspect that we could do a lot, besides simple rewording, to help respondents understand what we really want them to do. First is to continually inform the respondent of the purposes and goals of each section of the questionnaire, giving the respondent some bearing on where we are and where we are going. Related to this, we have attempted to simplify the structure of the entire interview, eliminating complex branching (skip) patterns which research suggests may be responsible for some response error variance (Hill, 1989). A simpler structure should help interviewers understand the goals of the section and, consequently, be of greater help to confused respondents who need clarification. Complex skip patterns often depend on the interviewer's remembering esoteric details about each person in the household. Sometimes interviewers, confronted with a branching decision, will reask the questions to get decision information. This interrupts the flow of the questionnaire and may destroy whatever hypothesis the respondent has formed about the goal of this section. On the basis of prior observation research findings, we have reworded some of the SIPP core questions but we suspect that the more global efforts to establish and maintain the comprehension contexts will be of major value.

3.5 Quality Context

Perhaps the most controversial change is the way we propose to deal with conflicts between seeking quality and achieving either efficiency or high response rates. The Census Bureau's interviewers are recognized as extraordinarily successful in obtaining public cooperation for government surveys and with relatively low costs per interview. To do this, interviewers may sometimes conduct themselves in ways that might compromise quality. While there is no research evidence either way on this issue, we will change some of these procedures in ways that we think will reduce measurement errors.

Setting. We emphasize to interviewers that SIPP is not a doorstep interview. It should be conducted inside the household with access to records and relatively free from distractions. We ask interviewers to call back, if necessary, to secure an appropriate setting.

Time. Time is money and current practice is often to get through the interview as quickly and efficiently as possible. This may result in a rushed pace, a reluctance to give explanations, skipped questions, failing to call back for missing information and/or avoiding the requirement that all present adults respond for themselves. Interviewers may discourage the respondent from consulting personal records because it is especially time consuming.

The new approach will recognize that it takes a considerable amount of time, at first, to do a new job well. Through training and feedback from tape recorded observations, we will encourage interviewers to moderate the pace of the interview, take the time needed to get full and complete answers (waiting while the respondent assembles personal records and providing the necessary training in how to accumulate and interpret personal records), give full explanations when needed, and schedule additional appointments to assure self-response from all adults. (The self-response requirement is relaxed in later waves if the proxy respondent has the sample person's records.)

Answer Quality. The way interviewers handle obvious instances of poor answer quality will change. This includes the interviewer's reaction to inadequate answers, "don't know" answers, and outright refusals to answer. Currently, interviewers try to avoid refusals by doing nothing that might irritate the respondent, such as pushing hard for quality responses by probing inadequate replies, resolving "don't know" answers (e.g., by callbacks), and trying to convert item refusals. The new procedures will not penalize interviewers for refusals in later interviews and will encourage them to actively seek adequate answers, even at the risk of a later refusal.

Process Control. In the past, even when it wanted to, the Census Bureau has not had an efficient way of evaluating how well interviewers do inside the household. Supervisors can observe interviewers but we all suspect that such interviews are conducted "by the book" and may not be typical. For the current research, we are planning to tape record each interview as a matter of routine practice, draw a sample of tapes to evaluate, and provide quick feedback from those evaluations to the interviewer. Our evaluation dimensions emphasize the quality-oriented behaviors mentioned here.

Questionnaire. There are a couple of things we will do to the questionnaire to create and maintain the quality context. First, we will explain that quality answers are very important and that is why we are doing certain things, such as requiring the use of personal records, or going through a lengthy series of recognition questions. Second, we feel that the current questionnaire begins with a task that is so difficult that respondents may feel that we cannot really want accurate, complete answers (week-by-week details about employment over a 4 month period). In addition, we have simplified all sections of the questionnaire and we have moved the difficult section to a later position in the interview.

3.6 Motivation

We have taken several steps to turn the SIPP core interview into a genuinely collaborative effort between the respondents and the interviewer. Suchman and Jordan (1990) advocate this approach to assure that all interacting parties negotiate and share the meanings intended by the questionnaire designer. While we agree with this social anthropological interpretation, as social psychologists we recognize the power of the respondent's high degree of involvement (without tangible reward) to create interest, commitment and compliance with the legitimate requests of this and future interviews.

To create more involvement we will begin the interview with a "free recall" section in which the interviewer merely explains the goal to the respondents (a complete list of your income for the last 4 months based on personal records), lets the respondents structure and control the task, helps respondents find the relevant information from records, and leaves each household member with a real sense of accomplishment at the end of the process. The interviewer shows the form to the household members and encourages them to watch as it gets filled out in response to the information they give. While the remaining sections of the interview are not as free-wheeling as this, they do have a similar structure (explaining the goals and assisting the respondent to solve whatever problems arise in providing the information requested).

4. EVALUATION ISSUES

These measurement procedures are novel and they should be evaluated thoroughly before being seriously considered for adoption in SIPP production interviews. Below we briefly raise some central evaluation questions and describe how we plan to address them.

Our general approach is dictated by a time schedule which calls for implementing a redesigned questionnaire in January 1995. The short time frame does not allow a series of planned experimental tests of key aspects of the new procedures. Instead, we have about a year to work out the full set of operational details in a series of pilot studies. Then we will conduct a single evaluation study using experimental and control interviewing procedures and administrative records both to create the sample and to validate selected responses. We have adapted this general evaluation procedure to address what we consider are the major risks of the new cognitive procedures. We discuss these risks next.

4.1 Evaluation of Risks

The new procedures may not reduce underreporting and may increase overreporting. Therefore, we will use "partial-design" record checks (Marquis, 1978) to assess separately both underreporting and overreporting. Although this approach will not allow us to get a direct estimate of net bias, it will allow us to get relative estimates of the two components of net bias for the standard and new SIPP interviewing procedures. This should provide important guidance for making decisions about whether to adopt the new procedures.

We have designed the evaluation study to include one seam (between waves one and two) to measure the seam bias and we plan to use the record check to detect whether seam errors have truly lessened or merely shifted somewhere else.

It is very likely that interviews using the new procedures could require more time, at least in the first interview. We will measure the time spent on major activities, relying on interviewers to make accurate time entries and time estimates. We will use tape recordings if necessary for in-depth analyses of the timing of selected interview processes.

We plan to use a relatively large number of interviewers in the evaluation study and we are looking into costs of interpenetrating the assignments to estimate interviewer variance. If those costs are prohibitive, we may be able to model the differences and make corrected estimates.

We are also looking into using the administrative record feature to evaluate the trade-off between increases in noninterview biases and decreases in measurement errors that may accompany the new procedures.

4.2 Preliminary Field Experience

We want to end on a less formal note. While we have raised and addressed the major evaluation issues above, there are more practical issues of immediate concern: Will households be willing to use their records? If they are willing, can they make any sense out of them? Do people resent tape recordings and refuse? Is efficiency really suffering? Is there an increased involvement and commitment?

For various reasons, these questions were compelling enough to cause us to hire a private contractor to conduct

several dozen early interviews using key elements of the described technique so we would know where the "real" problems lay. There was no systematic sampling and adults were paid $15 for participating.

To our surprise, the pilot interviews suggested we need be as concerned about persuading experienced interviewers about our procedures as we are about persuading households. Experienced interviewers cautioned that we cannot successfully make this kind of demand in the ordinary household, let alone in low income households. We have learned that it is important to pay attention to these expectations in training and perhaps spend a considerable amount of time putting them in the proper perspective.

During the contracted interviews, no respondent refused to be interviewed just because of the tape recording requirement (and nobody granted an interview but refused to be tape recorded). One roommate, who did not hear the introductory explanation, did refuse to participate when he learned that we wanted him to report his income using personal records and to be tape recorded. He felt his income was none of our business. So the tremendous anxieties we experienced by introducing radical procedures such as requiring personal records and tape recordings were unjustified. These procedures were readily accepted by households in this early test.

On the other hand, efficiency may suffer a great deal. We didn't do any formal timing but, when we listened to tapes, we noticed very long pauses while respondents disappeared somewhere to find records. There could be half-a-dozen or more pauses like that in a single interview. If the first interview does become longer and more costly, we might view it as an investment, especially if respondents become well organized and can go through subsequent interviews without excessive time-outs to search for records.

From listening to the tape recordings, the level of respondent involvement has surprised us. We are used to observing rather passive respondents who mostly listen to questions from the interviewer and seldom have an answer to share. When income is reported we are used to hearing a respondent do some simple figuring and then assert that the amounts for all the months were about the same, with maybe a small adjustment here and there. The new tape recorded interviews, however, are more collaborative and active. Record based reporting is something almost any adult can do. It is clear what steps need to be taken and there are usually some successful problem solving efforts required to meet the goals. As each problem is solved, the participants tend to feel good and welcome the opportunity to confront the next challenge. A lot of information gets reported and everybody seems enthusiastic in their adoption of the quality ethic.

We do not know how these experiences will generalize. Households were paid $15 for participating and this may give our new procedures a better reception than they will get in unpaid SIPP interviews. We will conduct further research in the next few years.

NOTE

[1] This paper reports the general results of research undertaken by Census Bureau staff. The views expressed are attributable to the authors and do not necessarily reflect those of the Census Bureau.

REFERENCES

Bartlett, F. (1932). Remembering. Cambridge University Press, Cambridge.

Hill, D. (1989). "Analysis of SIPP Reinterview Data: A Poisson Model of Response Error." 1989 Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 375-380.

Hill, D. (1987). "Response Errors Around the Seam: Analysis of Change in a Panel with Overlapping Reference Periods." Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 210-215.

Marquis, K. (1990). "Report of the SIPP Cognitive Interviewing Project." Unpublished Report, Census Bureau. Washington DC.

Marquis, K. (1978). "Inferring Health Interview Response Bias from Imperfect Record Checks." Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 265-270.

Marquis, K. and Moore, J. (1989). "Response Errors in SIPP: Preliminary Results." Proceedings of the Fifth Annual Research Conference, Census Bureau, Washington DC, pp. 515-536.

Marquis, K. and Moore, J. (1990). "Measurement Errors in SIPP Program Reports." Proceedings of the Sixth Annual Research Conference, Census Bureau, Washington DC. Also available as Report No. 9008 in the SIPP Working Paper Series (June, 1990).

Michaud, S., Egan, M., Lemaitre, G. and Murray, T.S. (Forthcoming). "Invisible Seams? The Experience with the Canadian Labour Market Activity Survey." Proceedings of the 1991 Annual Research Conference, Census Bureau, Washington DC

Moore, J. and Marquis, K. (1989). "Using Administrative Record Data to Evaluate the Quality of Survey Estimates." Survey Methodology, Vol. 15, pp. 129-143.

Neter, J. and Waksberg, J. (1966). "A Study of Response Errors in Expenditures Data from Household Interviews." J. American Statistical Assn., Vol. 59, pp. 18-55.

Nisbett, R. E. and Wilson, T. D. (1977). "Telling more than we know: Verbal reports on mental processes." Psychological Review, Vol. 84, pp. 231-259.

Ross, M. (1989). "Relation of Implicit Theories to the Construction of Personal Histories." Psychological Review, Vol. 96, pp. 341-357.

Suchman, L. and Jordan, B. (1990) "Interactional Troubles in Face-to-Face Survey Interviews." J. American Statistical Assn., Vol. 85, pp. 232-241.

Young, N. (1989). "Wave Seam Effects in SIPP." Proceedings of the Section on Survey Research Methods, American Statistical Association, pp. 393-398.