THE REVISED CPS QUESTIONNAIRE: DIFFERENCES BETWEEN THE CURRENT AND THE PROPOSED QUESTIONNAIRES

Jennifer M. Rothgeb, Bureau of the Census and Sharon R. Cohany, Bureau of Labor Statistics
Jennifer M. Rothgeb, Bureau of the Census, CSMR, Washington, D.C. 20233-4700

Keywords: Labor Force, Questionnaire Design

I. INTRODUCTION

The Current Population Survey (CPS) is a monthly survey of approximately 60,000 households conducted by the Bureau of the Census for the Bureau of Labor Statistics (BLS). It is the primary source of information on the labor force and has remained essentially unchanged since the last major revisions in January 1967.

Beginning in 1986, BLS and the Bureau of the Census conducted a number of research projects to guide the development of a revised CPS questionnaire. Alternative questionnaires were developed based on the results of this research as well as recommendations made by the Levitan Commission in 1979. (See Campanelli, Martin, and Rothgeb 1989; Edwards, Levine, and Cohnay 1989; Fracasso 1989; Gaertner, Cantor, and Gay 1989; Martin 1987; National Commission on Employment and Unemployment Statistics 1979; and Palmisano 1989a, 1989b.)

The objectives of the CPS questionnaire redesign are five-fold: (1) to better operationalize existing definitions; (2) to reduce the potential for response error; (3) to implement minor definitional changes; (4) to expand the labor force data available and improve longitudinal measures; and (5) to utilize the capabilities of computer-assisted interviewing. (See Copeland and Rothgeb 1990 for a fuller discussion.)

The testing of alternative questionnaires was accomplished in a two-phase computer-assisted telephone interviewing/random digit dialing sample test (CATI/RDD). During Phase I, two alternative questionnaires ( Versions B and C) were tested against the current questionnaire (Version A). Phase I was conducted from July 1990 to January 1991 and data were obtained for over 72,000 persons. During Phase II, one alternative questionnaire (Version D) was tested against the current questionnaire. Phase II was conducted from July 1991 to October 1991 and data were obtained for over 32,000 persons. The results of both phases were used in the development of the final revised questionnaire which will become the official questionnaire in January 1994.

Question evaluation was based on item nonresponse measures, response distributions, respondent debriefings, interviewer debriefings, and interview monitoring (interviewer/respondent behavior coding.)

II. HIGHLIGHTS OF REVISIONS TO QUESTIONNAIRE

In this paper, we focus on some of the revisions that did not receive much attention in previous ASA papers (Copeland and Rothgeb 1990; Rothgeb et al. 1991). The revisions to the questionnaire are discussed in four sections, which are organized into general revisions, employment related revisions, unemployment related revisions, and revisions primarily affecting those not in the labor force.

A. General

1. Definition of Reference Week

In interviewer debriefings conducted during 1988, interviewers reported that the current major activity question "What were you doing most of LAST WEEK, working or something else?" was unwieldy and sometimes misunderstood by respondents. In addition, respondents were unsure what we meant by "last week" (BLS 1988). A 1988 respondent debriefing found that only 17 percent of respondents had definitions of "last week" that matched the CPS definition of Sunday - Saturday of the reference week. (Campanelli et al. 1991).

In the revised questionnaire, an introductory statement has been added with the reference period clearly stated, as follows: "I am going to ask a few questions about work-related activities LAST WEEK. By last week I mean the week beginning on Sunday, August 9 and ending Saturday, August 15." This statement also serves to define the survey content, so the major activity question was deleted.

2. Direct Question on Presence of Business

The definition of employed persons includes those who work without pay for at least 15 hours per week in a family business, and BLS produces monthly estimates of such persons. Currently, however, there is no direct question on the presence of a business in the household. The revised questionnaire includes such a question, which will be asked only once for the entire household prior to the labor force questions, and reads "Does anyone in this household have a business or a farm?" The primary purpose of this question is to screen for households that may have unpaid family workers, not to obtain an estimate of household businesses. (See Rothgeb et al. 1991; Copeland and Rothgeb 1990; and Martin 1987.)

For households that have a family business, direct questions will be asked about unpaid work in the family business of all persons who were not reported as working last week. Estimates from the CATI/RDD Phase II test demonstrated that unpaid family workers constituted a significantly larger proportion of persons reported as employed (at work) in the revised questionnaire, which included the additional questions about unpaid family workers, than in the current questionnaire, which does not have direct questions about such workers (1.0% vs .2%; Ns=9,036 and 9,327 respectively; df = 1; p < .000.)

B. Employed Related Revisions

1. Revised "At Work" Question

Having a direct question on the presence of a family business not only improved the estimates of unpaid family workers, but also permitted a revision of the "at work" question. In the current questionnaire, this question reads: "LAST WEEK, did you do any work at all, not counting work around the house?" In the revised questionnaire, the wording reads, "LAST WEEK did you do ANY work for (either) pay (or profit)?" (The parentheticals are read when there is a business or farm in the household.) The phrase "work for pay (or profit)" better operationalizes the concept of work. (Martin 1987)
In Phase II, 25 percent of persons (in households with family businesses) not reported to be working for pay had done unpaid work. A direct question on unpaid workers improves the measure of unpaid family workers by reducing reliance on volunteered responses and reducing response error.

2. Direct Question on Multiple Job Holders

Currently, the actual hours question reads: "How many hours did you work last week at all jobs?" During the interviewer debriefings in 1988, it was reported that respondents do not always hear the phrase "at all jobs." Some respondents who work at two jobs may only report hours for one job (BLS 1988.)

The revised questionnaire includes a question at the beginning of the hours series to determine whether the person is a multiple jobholder. Multiple jobholders are asked about their hours on their main job and other job(s) separately. These new questions permit monthly estimates of multiple jobholders to be produced.

3. Hours Series

The current question on "hours worked" reads "How many hours did you work last week at all jobs?" The hours data are important because they are used to determine the sizes of the full-time and part-time labor forces. It is unknown whether respondents report exact actual hours, usual hours, or some approximation of actual hours.

In the revised questionnaire, a new hours series was adopted utilizing a recall estimation strategy to obtain a better measure of actual hours and to address the issue of work schedules more completely. The revised questionnaire first asks about the number of hours a person usually works. Then, separate questions are asked to determine if a person worked any extra hours or took hours off. Then, a question on the number of actual hours worked last week is asked. For multiple jobholders, separate data are obtained on hours worked at each job.

Data from Phase II indicate that the recall strategy used in the revised hours series was slightly more sensitive in obtaining estimates of persons working part time (less than 35 hours per week.) The proportion of employed persons who worked full time was 69 percent for Version A and 67 percent for Version D. Mean hours reported were 39.0 for Version A and 37.9 for Version D. (See Rothgeb et al. 1991.) It should also be noted that the new hours series will provide monthly estimates of usual hours worked for all employed persons.

4. Industry/Occupation - Reordered Series

In the current CPS questionnaire, after the industry and occupation data are collected, the interviewer obtains the "class of worker" information, that is, whether a person works for a private organization, a government unit, is self-employed, or is an unpaid family worker.

Monthly estimates of class of worker status are published by BLS. During monitoring of interviews in Phase I of the CATI/RDD test, researchers often heard the interviewers incorrectly "verifying" the person's class of worker status instead of asking the question as worded. The most frequent "verification" heard was "And you're employed by a private company?" Respondents frequently confirmed this without ever hearing the other response options. It is suspected that this occurs since the interviewer receives a description of the person's industry and occupation prior to asking the class of worker question.

In the revised questionnaire, the I/O series has been reordered so the class of worker questions are at the beginning of the series, prior to any questions on industry and occupation. This resequencing of the series allows the industry and occupation questions to be tailored to the specific person's class of worker status based on the responses to the class of worker items.

Data from Phase II of the CATI/RDD test demonstrate that the reordered series resulted in a smaller proportion of persons being classified as "private company" in Version D than in Version A (68.8% vs 73.2%; X2 = 14.640; Ns = 3738 and 3393; df = 1; p = .001.) Also, a larger proportion of persons were classified as "self employed/unincorporated" in Version D than in Version A (10.5% vs 8.3%; X2 = 8.878; df = 1; p = .01.) It is suspected that these differences may be due to the household business/farm question at the beginning of the questionnaire, the use of direct questions on class of worker status, and the resequencing of the series. The reordered series also resulted in lower referral rates for both the industry and occupation items. That is, fewer industry and occupation descriptions were unable to be coded.

5. Industry/Occupation Data for Second Job

In the revised questionnaire, for multiple jobholders, industry, occupation and class of worker data are collected for their second job. (These data will be collected only in the outgoing rotations.) For multiple jobholders identified as "self-employed unincorporated" on their main job, it will be necessary to collect class of worker data on the second job in all months to enable production of monthly estimates of multiple jobholders. This is necessary in order that persons who are self-employed (unincorporated) at both jobs can be eliminated from estimates of multiple jobholders. According to the official definition, such persons are not considered to be multiple jobholders. This will allow a more accurate measure of multiple jobholders and will provide information necessary for comparing estimates between the CPS and BLS's establishment data.

Currently, industry and occupation (I/O) data are not always consistent from month to month for the same person in the same job. This is due, in part, to the fact that over half the data are collected through proxy interviews (persons responding for other household members) and the respondent often varies from one month to the next. Further, it is sometimes difficult for respondents to describe their own occupation consistently from month to month.

Moreover, changes in occupation and industry (gross flows) are measured at the 3-digit level, that is, the most detailed classifications, which can imply very subtle distinctions, particularly among occupation groups. Currently, about 32 percent of the continuing sample report a change in their 3-digit occupation classification from one month to the next, and about 16 percent report a change in their 3-digit industry classification.

In order to obtain more consistent data and make full use of an automated interviewing environment, the revised questionnaire uses dependent interviewing for months-in-sample 2-4 and 6-8 households. Dependent interviewing uses information collected during the previous month's interview in the current month's interview.

Specifically, respondents are given the name of their employer in the previous month and asked if they still work for that employer. If they answer "no," they are asked the independent I/O questions. If they answer "yes," they are asked if the usual activities and duties of their job have
changed since last month. If there was a change, they are asked the independent questions on occupation, main activities or duties, and class of worker. If their duties did not change, they are asked to verify the previous month's description of their occupation and activities. If this information is still accurate, no further coding is required. If the information is not accurate, they are asked the independent questions on activities and class of worker.

Data from Phase II indicate that only 5 percent (n = 5386) of persons eligible for dependent interviewing reported changing employers between months. An additional 2 percent (n = 5081) reported that their duties had changed between months and another 2 percent (n = 4894) said that the previous month's description of their duties was inaccurate.

Independent measures of I/O obtained from respondent debriefings showed that 21.7 percent (n = 1064) of persons reporting no change in employers changed 3-digit industries between consecutive months, and 39.0 percent (n = 1014) changed occupations, even though they reported no change in activities (CPS Questionnaire Development Workgroup 1991). These high rates of change reflect response and coding errors rather than true change, since, by their own reports, respondents' employers and job duties had not changed.

A problem in evaluating I/O data was the lack of current information on true levels of monthly gross flows. A 1975 study estimated that 4.6 percent of persons changed industries and 9.9 percent changed occupations (both at the 3-digit level) over a 2-month period (Collins 1975).

A more up-to-date measure of true change was accomplished by the use of "expert coders," who analyzed job descriptions (from Version A) collected for three pairs of consecutive months from the same respondent to determine if a true change at the 3-digit level had in fact occurred (Cantor 1991). Under contract to BLS, Westat Inc. designed the expert coding test and analyzed the results. As displayed in Table 1, the study found that estimates of true change ranged from 3.8 - 4.2 percent for industry codes and 5.9 - 7.4 percent for occupation codes.

Thus, according to the expert coding study, the current questionnaire greatly overestimates gross flow rates, while the revised questionnaire obtains rates that are close to the "true" values. Note that the estimates for Version D include both dependent and independent I/O measures.

6. Earnings

Currently, persons are asked if they are paid by the hour, asked for an hourly wage and then asked for weekly earnings. In the revised questionnaire, the earnings series was designed to first request the periodicity for which the respondent finds it easiest to report earnings and then request an earnings amount in that periodicity. (Because data on weekly earnings are published quarterly by BLS, earnings data provided by respondents in periodicities other than weekly are converted to a weekly estimate during later processing of the data.)

The current version forces everyone to report earnings on a weekly basis, even though that may not be the easiest way for the respondent to recall and report earnings. Data from Phase I of the CATI/RDD test indicate that only 14 percent (n = 853) of non-hourly wage workers in Version B were paid at a weekly rate and less than 25 percent (n = 1623) of non-hourly wage workers in Version C found it easiest to report earnings as a weekly amount. Results from Phase II indicate that the revision in the earnings series led to significant differences in "don't know" responses. In Version A, "don't know" responses were received for nearly 19 percent of sample persons. This contrasts sharply with 13 percent in Version D (X² = 39.10; df = 1; p = <.000; n= 3467; np = 2985). (Refusal rates were 8 percent for both versions.) These results indicate how much more information can be obtained from respondents if they are provided with reporting options as opposed to a forced periodicity. Reporting a weekly earnings amount apparently was a much more difficult task for respondents than reporting in their selected periodicity.

The data do not indicate any significant differences in earnings estimates. Average weekly earnings for Version A did not differ significantly from Version D ($465 vs $456; F = 0.89; df = 1; p = .3447). Other revisions include a question to determine if a person usually receives overtime pay, tips or commissions. If so, then a lead-in is included in the earnings questions that reminds respondents to include overtime pay, tips and commissions when reporting earnings. If a respondent reports that it is easiest to report earnings on an hourly basis, then a separate question is asked regarding the amount of overtime pay, tips and commissions usually received.

Since BLS needs to obtain estimates of hourly wage workers in order to study minimum-wage workers, an additional question is asked of persons who don't report that it is easiest to report their earnings hourly: "Even though you told me it is easier to report your earnings (annually), are you PAID AT AN HOURLY RATE on this job?"

Without this question, Phase II data indicated that estimates of hourly wage workers were drastically lower for Version D than Version A (38.28% vs 61.20%; Ns = 2696 and 3402; X² = 298.141; df = 1; p = .000.) With this question, the estimates of hourly wage workers were significantly higher in Version D than in Version A, but the difference was not as extreme (66.93% vs 61.70%; Ns = 2800 and 3402; X² = 15.5366; df = 1; p = .000.)

Other enhancements to the earnings series have been made since the CATI/RDD test. Several refusal conversion statements have been added, as have probes for estimates whenever "don't know" responses are provided.

C. Unemployed Related Revisions

1. Job Search Methods

The concept of unemployment requires, among other criteria, an active job search during the past 4 weeks. In the current questionnaire, the following question is asked to determine if a person conducted an active job search.

"What has ... been doing in the last 4 weeks to find work?"

checked with--

| Public employment agency | 0 |
| Private employment agency | 0 |
| Employer directly | 0 |
| Friends and relatives | 0 |
| Placed or answered ads | 0 |
| Nothing | 0 |
| Other | 0 |

Interviewers are instructed that all passive job search methods should be coded in the "nothing" category. Only active job search methods for which no appropriate response category exists are to be coded as "other."
In the revised questionnaire, several additional response categories have been added and the response options have been re-ordered and reformatted to more clearly represent the distinction between active methods and passive methods, as displayed below. "What are all of the things you have done to find work during the last 4 weeks?"

**ACTIVE**
- Contacted: employer directly/Interview
- public employ agency
- private employ agency
- friends or relatives
- school/university employment center
- Sent out resumes/
- filled out applications
- Placed or answered ads
- Checked union/professional registers
- Other active

**PASSIVE**
- Looked at ads
- Attended job training programs/courses
- Other Passive
- Nothing

The revisions to the job search methods question grew out of concern that interviewer errors were being made because of confusion in the use of precoded response categories. This was evident even before the analysis of the CATI/RDD test. In 1986 Martin conducted an examination of verbatim entries for the "other" category and found 22 percent (n=88) of the "other" responses should not have been included in the "other" category, but rather should have been marked "nothing", and another 14 percent were too vague to determine whether an active job search method was undertaken (Martin 1987). Research was also conducted during 1989 which determined that the current set of response categories was not adequate for accurate classification of active and passive job search methods (Fracasso 1989).

During development of the alternative questionnaire for Phase II, two additional passive categories were included--"looked at ads" and "attended job training programs/courses"--and two additional active categories were included--"contacted school/university employment center" and "checked union/professional registers". Based on results from Phase I, it was decided to expand the response categories even more for the alternative questionnaire in Phase II. (Analysis of Phase I verbatim responses indicated that interviewers used many of the precoded response categories inappropriately. Also, interviewers had difficulty coding relatively common responses such as "sent out resumes" and "went on interviews"; therefore, the response categories were further expanded in Phase II to reflect these more common job search methods.)

A comparison of the Version D categories to their comparable categories in Version A yielded a different distribution of responses, as shown in Table 2. Note that the passive job search methods in Version D have been recoded to "nothing" for comparability with Version A.

The differences in response distributions in the "other" and "nothing" categories most clearly exemplify the impact of the additional response categories. The additional passive response categories in Version D resulted in 24 percent of the responses being coded (recoded) to "nothing" for Version D compared to only 5 percent for Version A. Over 25 percent of the responses in Version A were coded as "other" compared to only 10 percent in Version D. Analysis of the verbatim responses for the "other" categories indicates that almost half of the difference in estimates between the questionnaires occurs because 30 percent of "other" entries in Version A were incorrectly coded since they were actually passive responses. Eliminating incorrectly coded passive job search methods in the "other" category explains in part the decrease from 26 percent to 10 percent in this category. The additional active response categories are also responsible for some of the decrease because responses that were put in the "other" category in Version A are coded to a more appropriate category in Version D.

Analysis of verbatim responses collected during Phase II indicated that expanded response categories reduced interviewer coding errors. (See Table 3). Thirty-two percent (n=133) of the Version A "other" responses should have been classified as passive and coded as "nothing," in contrast to 8 percent (n=48) in Version D that were coded as "other active" but should have been coded as passive. It should also be noted that 34 percent (n=38) of Version D "other passive" entries should have been coded as "active". Multiple responses are permitted in Version D it is possible for both an active and passive response to be given. This is not possible in Version A. Therefore, although an incorrect entry of "other," "other active," or "other passive" may lead to misclassification, it is unknown if other responses provided would "correct" the erroneous entry.

Table 3 also shows that a larger proportion of responses are coded as "other" in the current questionnaire (Version A) than are coded as "other active" or "other passive" in the revised questionnaire (25.9%-other vs 7.2%-other active vs 5.4%-other passive). Therefore, although both questionnaires have the potential to produce misclassification of unemployment status due to erroneous entries in the job search item, the current questionnaire, with a large percentage of "false actives," has the greater likelihood of misclassification.

2. Duration of Job Search and Layoff

The duration of unemployment is published monthly by BLS. Currently, this information is collected by the question: "How many weeks have you been looking for work?" This wording forces people to report in a periodicity that may not be meaningful to them, especially for the longer-term unemployed. Also, asking for the number of weeks (rather than months) may lead respondents to underestimate the duration. In Version D, the question was revised to read "As of the end of LAST WEEK, how long had you been looking for work?" Respondents selected the periodicity themselves and interviewers could record the duration in weeks, months, or years.

A higher mean duration was estimated from Version D than from Version A. 17.9 weeks (n=607) versus 14.2 weeks (n=710). Since these results could reflect differences in collection methods (Version D automatically updates duration for persons unemployed in consecutive months, whereas Version A independently obtains duration estimates each month) as well as question wording, this analysis was repeated just for persons in their first month in the sample. The mean duration was still significantly higher in Version D than in A, 19.1 weeks (n=226)
An analysis of Phase I (Version B) results revealed that a very high proportion (75 percent, n=79) of people reporting durations of one or two months gave a duration in weeks that was exactly four times their monthly report. (Version B respondents job searches of one or two months were also asked to provide an estimate in terms of weeks.) To avoid clustering of answers around whole months, the final revised questionnaire asks persons who report duration in whole months (between one and four months) a follow-up question to obtain an estimated duration in weeks.

As mentioned earlier, the CATI/CAPI technology makes it possible to automatically update duration of job search and layoff for persons who are unemployed in consecutive months. For these persons, the previous month’s duration will be updated without re-asking the duration questions. This should reduce respondent burden and enhance the longitudinal capability of the CPS. This revision should also produce more consistency in the month to month estimates of duration. Respondent debriefing data from Phase I indicated that estimates of duration of unemployment were consistent (as indicated by increases of 3, 4, or 5 weeks) from month to month only 25 percent of the time when duration was obtained independently (CPS Questionnaire Development Work Group 1991). Therefore, a short series of questions was added to Version D to identify persons who, although they had looked for work in the past year, had not looked for work since their last job, in order to exclude them from the estimate of discouraged workers.

Currently, the question asking about reasons for not looking for work allows multiple responses. (A complex priority scheme determines whether the sample person should be classified as discouraged if both discouraged and nondiscouraged reasons are reported.) In the revised questionnaire, the question asks for the primary reason. This lets the respondent determine whether a discouraged or nondiscouraged reason is the primary reason, as opposed to having that decision reached during a priority scheme during processing.

With the criteria of job search in the last 12 months (job search since last worked if worked in last 12 months), availability, desire for a job, and providing a discouraged reason for not looking, only .3 percent of persons not in the labor force in Version D were classified as discouraged workers, compared with 1.1 percent in Version A.

To summarize, the new criteria for discouraged workers will include a job search during the last 12 months (and since last worked if worked within the past 12 months), desire for a job, and availability to have taken a job during the reference week. The new questionnaire asks the discouraged worker series in all months, not just the outgoing months as is currently done. This change will allow publication of a discouraged worker series monthly rather than quarterly.

III. QUESTIONNAIRE IMPLEMENTATION AND FUTURE ACTIVITIES

Beginning in July of 1992, the revised questionnaire is being administered in a fully automated environment of CATI/CAPI to a national overlap sample survey of 13,000 eligible households per month for 18 months. The overlap data will also be processed using the revised labor force edit and processing system which will be used in "live" CPS starting in January 1994. The results from the overlap sample will be compared with data from the current CPS to benchmark differences in the labor force estimates due to the new questionnaire, new modes of interviewing, and other design changes. The revised questionnaire will be fully implemented in January 1994.

REFERENCES


| Table 1. Average Month-to-Month Change for Industry and Occupation Codes |
|-----------------------------|-----------------------------|-----------------------------|
|                             | Version A | Version D | Expert Coders |
| IND | 23% (1426) | 5% (1361) | 3.6% - 4.2% (258) |
| OCC | 39% (1392) | 7% (1392) | 5.9% - 7.4% (406) |

| Table 2. Percentage of all jobseekers reporting each method |
|-----------------------------|-----------------------------|-----------------------------|
|                             | Version A | Version D |
| Public employment agency    | 13.2       | 16.0       |
| Private employment agency   | 5.2        | 7.1        |
| Employer directly           | 60.1       | 72.2       |
| Friends or relatives        | 11.6       | 13.0       |
| Placed/answered ads         | 27.0       | 21.0       |
| Nothing                     | 4.7        | 23.6       |
| Other                       | 25.9       | 10.3       |
| TOTAL                       | 773        | 706        |

| Table 3. JOB SEARCH METHODS (In Percent) |
|------------------------------------------|-----------------------------|-----------------------------|
| Other                                    | Other Active | Other Passive |
| in Version A                             | in Version A | in Version D |
| TOTAL                                    | 26           | 7            | 5            |
| False Active in Version A                | 32           | -            | -            |
| False Active in Version D                | -            | 8            | -            |
| False passive in Version D               | -            | -            | 34           |

654