

# USING THE 1996 NATIONAL CONTENT SURVEY TO MEASURE THE IMPACT OF ALTERNATIVE SKIP PATTERNS ON THE COLLECTION OF EMPLOYMENT DATA

David A. Raglin, United States Bureau of the Census  
DSSD, Room 2406-2, Washington, DC 20233

**Key Words:** Questionnaire design, Decennial Census, self-administered questionnaire, response error, nonsampling error

## 1. INTRODUCTION

The long form of the United States Decennial Census collects information on employment, including whether or not the person has a job, and if not, whether the person is looking for a job. This information is used to help determine a person's status in the labor market--employed, unemployed, or not in the labor market.

However, the definitions of those terms are complicated. For example, if a person is laid off from their job, but expects to return (for example, during the model changeover in an automobile plant), and are not looking for a job, they are considered to be unemployed. If their layoff is permanent or there is no return date designated, and they are not looking for a job, they are considered to be out of the labor market; if they are looking for a job, they are considered unemployed.

The main source of employment data in the United States is the Current Population Survey (CPS), conducted by the Bureau of the Census and the Bureau of Labor Statistics. The CPS is a monthly survey, and the data are collected by either personal interview or by telephone with a household member. The CPS asks a series of questions, with skips, to determine employment status, which is easy to do in an interviewer-administered survey. However, the decennial census long form data are collected in a mail out/mail back format, making it more difficult to get high quality data on a complicated subject. In 1990, the Census showed a higher unemployment rate, 6.3 percent, than did the CPS for April 1990, 5.2 percent (not seasonally adjusted) (Palumbo, 1993). One possible reason for the discrepancy was confusion about layoff versus unemployment. The 1990 Census asked people whether they were "on layoff", but did not distinguish between temporary and permanent layoff, and did not ask whether the person expected to return to work.

As part of the preparations for Census 2000, the Census Bureau conducted the 1996 National Content Survey (NCS), which tested alternative versions of questions. The next section describes some details of the 1996 NCS design. One area of the form tested was the non-work status section. New versions of that section were developed which asked more detailed questions on

layoff. There were four versions of the non-work status questions tested against the 1990 Census wording. About 10,000 people in about 4,000 households responded to each version. All of the test versions asked more complicated questions with more complicated skip patterns than did the 1990 Census version.

One concern was the ability of respondents to follow the complicated skip patterns correctly to give data that would be more accurate. This paper will analyze how well the respondents followed the more complicated skip patterns. Gap errors, where the respondent skipped to the wrong question and therefore left out information they should have provided, are distinguished from loop errors, where the respondent answered the next question when they should have skipped it. Gap errors are considered more serious than loop errors because needed data was not provided. Loop errors can often be corrected in the editing process. The percentage of forms with gap and loop errors for each version will be compared, with the most prevalent errors detailed. The effects of these changes on the data will also be highlighted.

## 2. DESCRIPTION OF THE 1996 NATIONAL CONTENT SURVEY

The 1996 NCS was a survey designed to test content, coverage, and forms design issues in preparation for Census 2000. There were thirteen versions of forms tested in the 1996 NCS; seven of them short forms and six of them long forms. Data were collected for up to seven people on a form--first for the first person, then housing questions, and then for the other people in succession. Many different content issues were tested by having different versions of questions on different forms.

The 1996 NCS sample design reflects the goal of the NCS: to collect information related to coverage and content to help design the forms for the 2000 Census. The universe of the NCS was housing units in 1990 Census questionnaire mailback areas in the continental United States only that were not selected in any post-1990 Census evaluation or test survey. The tracts were divided into two strata--Low Coverage Areas (LCA) containing

---

This paper reports the results of research undertaken by Census Bureau staff. The views expressed are attributable to the author and do not necessarily reflect those of the Census Bureau. The author would like to thank Tom Palumbo, Henry Woltman, Debbie Griffin, Wendy Davis, Cleo Jenkins, David Chapman, Rita Petroni, and Michael Kornbau for their insightful suggestions.

tracts with large amount of minorities and renters based on the 1990 Census--groups for which there were more coverage problems in 1990; and High Coverage Areas (HCA) making up the residual. LCA areas were sampled more heavily than HCA areas. The data presented here have been weighted to reflect the differences in sampling rates. Housing units were sampled in clusters of up to thirteen neighbors, with each housing unit in a cluster receiving a different form, minimizing differences in units that might affect the comparisons between forms.

The universe that the results of this survey reflect is not the entire United States--it is only the universe of mail respondents in the 1990 Census questionnaire mailback area in the continental United States. The data have not been adjusted to account for those who did not respond by mail nor for those areas not in the mailback universe. There is no indication that there is a real difference in either the number or characteristics of the respondents to the different forms. The mail response rates for the five forms in question were similar, ranging from 62.2 percent for control form DS-2A to 65.4 for one of the experimental forms (DS-2C) (Leslie, 1996). That result is not surprising since the forms and mailing packages looked alike and were of about the same length, and mail response rate was not a key variable being tested with the long forms. However, because of the universe differences, this data cannot be directly compared to the results of the 1990 Census or any current survey.

### 3. CONTENTS OF FORM VERSIONS

There were five versions of the non-work status questions. In all of the versions, the respondent was asked earlier in the questionnaire if they were fifteen years or older, and if so, they were to continue. Later on, they were asked if they worked last week. If they did not, they were to answer questions regarding their non-work status. The non-work status questions are analyzed in this paper. Below is a description of the features of each form followed by Table 1, which summarizes the key points:

Control (Form DS-2A): This form contained the same questions as the 1990 Census. The person was asked whether or not they worked last week. If they did, they answered a series of questions on the job location and their commute. If they did not, they skipped to the non-work status questions. The first question asked if they were temporarily absent from a job on layoff; on vacation, temporary illness, labor dispute, etc.; or not temporarily absent from a job. They were then asked if they had looked for a job in the last four weeks, and, if so, if they could have taken one if offered.

Respondent-Defined Layoff (Form DS-2C): This form attempts to have the respondent distinguish between temporary versus permanent layoff. After the work last

week question, if they worked, they gave the number of hours they worked, and skipped out. If they did not work last week, they were asked if they were on layoff, with the choices being temporary layoff, permanent layoff, or not on layoff. If they were not on layoff, they were asked if they were temporarily absent from a job on layoff, on vacation, temporary illness, labor dispute, etc. If they were on permanent layoff, they were asked if they looked for work in the last four weeks, and if they had, if they could have gone to work. If they said they were on temporary layoff, they were just asked if they could go to work last week.

Government-Defined Layoff, Layoff First (Form DS-2E): This form asks laid off workers a question on when they will be returning to work. This is used to identify people who meet the official definition of layoff.

Persons who answered yes to the work last week question answered job location and commuting questions similar to the control version, otherwise skipping to the non-work status questions. They were first asked if they were on layoff. If they were, they skipped to the third question in the section asking if they would be returning to work soon, the same question as in the other government-defined versions. If they answered no to that question, they skipped to the fourth question where they were asked if they had looked for work in the last four weeks, and, if so, if they could have started a job if offered. Persons who said they were not on layoff skipped to the second question which asked if they were temporarily absent from work because of vacation, temporary illness, or labor dispute. If they were, they skip to the fourth question, where they were asked if they had looked for work in the last four weeks, and, if so, if they could have taken one if offered. This version of the non-work status section is most like the version used in the CPS.

Government-Defined Layoff (Form DS-2B): After the work last week question, persons who worked gave the number of hours they worked and skip out. Persons who did not work last week were asked if they were temporarily absent from work because of vacation, temporary illness, or labor dispute. This is the key difference between this form and the previous one--on this form, temporary absence is asked first and on the other one, layoff was asked first.

If they were not temporarily absent, they were asked if they were on layoff. If they were, they were asked if "this person (had) been informed that he or she will be recalled to work within the next 6 months or been given a specific date to return to work". This question is key to determine if a person is on layoff according to the official definition. If they answered no, they were asked if they had looked for work in the last four weeks, and, if so, if they could have started a job if offered.

**Table 1: Summary of the Features of Each Form Regarding Non-Work Status**

Form Name Form Description	DS-2A Control	DS-2C RDL	DS-2E GDLLF	DS-2B GDL	DS-2D GDLWD
Several commuting questions after “work last week” question	Yes	No	Yes	No	Yes
Skip reminder provided before non-work status section	No	No	Yes	No	No
Layoff or temporarily absent question first in non-work status section, or together	Together	Layoff	Layoff	Temp Abs	Temp Abs
Type of layoff not defined, defined by the respondent, or defined by the government def.	Not Def	Resp’dent	Gov’mnt	Gov’mnt	Gov’mnt

**Table 2: Work Last Week Responses Versus Whether They Started the Non-Work Status Section**

Form Name Form Description	DS-2A Control	DS-2C RDL	DS-2E GDLLF	DS-2B GDL	DS-2D GDLWD
Number responding to WLW question	7,294	7,247	7,435	7,665	7,490
Percent who said yes to WLW question	58%	58%	57%	54%	55%
Of those, percent who started NWS- <b>incorrect</b>	55%	53%	18%	54%	56%
Of those, percent didn’t start NWS- <b>correct</b>	45%	47%	82%	46%	44%
Percent who said no to WLW question	42%	42%	43%	46%	45%
Of those, percent didn’t start NWS- <b>incorrect</b>	24%	11%	19%	12%	21%
Of those, percent who started NWS- <b>correct</b>	76%	89%	81%	88%	79%

**Government-Defined Layoff With Distance (Form DS-2D):** This form is similar to the government-defined layoff version. The difference is that if they answered yes to the work last week question, instead of just giving the hours they worked, they answered job location and commuting questions similar to the control version. This creates a large distance between the work last week question and the non-work status questions.

**4. DEFINING THE PEOPLE IN THE ANALYSIS**

As has been mentioned, the non-work status questions are a small part of the long form, and response to those questions is affected by earlier questions. Analysis was done to examine any problems with the correct respondents getting to the “work last week” question, and no serious problems were discovered.

However, there was a potential problem with inconsistencies in the answers to the work last week question and later parts of the form. After the work last week question, there were question(s) for people who worked last week followed with the non-work status questions for people who did not work last week. If they said they worked last week, they were supposed to answer the question(s) for people who worked and skip over the non-work status questions. If they did not work last week, they were supposed to skip the question(s) for those who worked, fill out the non-work status section and continue with the questionnaire.

There were two problems respondents the respondents could have had: (1) they responded that they worked last week but they answered the non-work status questions intended for people who did not work last week, and (2) they responded that they did not work last week but they did not answer the non-work status questions. The second problem is the more serious, since items on the form that should be completed are not. If someone who worked completed the non-work section in error, their responses can be edited out.

How often did the respondents have such problems? Table 2, based on people fifteen years of age or older who answered the work last week question, shows the percentage who answered that they worked and did not work, and, of those people, how many started the non-work status section (defined as answering the first non-work status section question). All data are unweighted.

Note that while the percentages of people who said they worked and did not work last week were similar across forms, the percentages who correctly started or did not start the non-work status section were widely different. Maybe the most startling result in this table is that for four of the forms, more than half of the people who said they worked started the non-work section. Research showed that a vast majority of those people also completed the question(s) for working people as they should have--the indication is that they just kept on going into the non-work section. There was a skip instruction after the the last question to be filled out by people who

worked last week, but it was not highlighted and was off to the right of the answer box. Jenkins and Ciochetto (1993), reported by Jenkins and Dillman (1983) suggest that skip instructions not in the direct navigational flow (as an instruction off to the right of the answer box would be) can be missed. They also reported work by Turner, et al. (1992) that information to the right of an answer box, will be seen only if it made to look important. The skip instruction in this case was not highlighted.

The one form that was an exception to this skip problem was the government-defined layoff, layoff first form, where only 18 percent of the people who worked started the section, due to an additional instruction before the non-work section to remind respondents to fill out this section only if they worked last week.

The more serious problem is people who said they did not work last week but did not complete the non-work section. Those percentages are much lower than for the error discussed previously, but they are still high. Note that they seem to be lower (no significance testing was performed) for the two forms where there was only one question for people who worked between the work last week question and the non-work status section.

The question at this point is to decide which people should be included in the analysis. Since this is a study of the skip pattern of the non-work status section, only people who started the non-work status section should be included. Furthermore, the indication is that people who filled out the non-work status section but previously said they worked filled out the section in error. Therefore, people were included who: (1) were fifteen or more years older, (2) did not work last week, and (3) answered the first question in the non-work status section.

There was an incorrect skip instruction in the DS-2E form for the second through the seventh persons. Therefore, it was decided to use only data from the first person for all forms for the analysis in this paper. Generally one respondent fills out the whole form and they may "learn" how to complete the form; so using only the first person should not cause a problem. The number of cases per form range from 1,021 to 1,385.

## 5. COMPARISONS OF FORMS

Comparison between forms will be made in three ways: percentage of forms with gap errors, percentage of forms with loop errors, and the percentages of people who said they (1) were on layoff, (2) were temporarily absent from work, (3) were unemployed but not on layoff, and (4) were not in the labor force. Since the analysis universe is people who did not work last week, those are all of the possibilities for these people. There were not enough data to classify some people into one of those four categories--that percentage is also reported.

All data in these tables are weighted, variances were calculated using the complex sample design, and a significance level of  $\alpha = .10$  was used in comparisons.

### 5.1. ONE VERSUS TWO LAYOFF CATEGORIES

In the control form, the respondent just has a choice of "layoff", while in the respondent-defined layoff form, they have a choice of temporary and permanent layoff. If they mark that they are permanently laid off, they are not considered on layoff--they are either unemployed or out of the labor force, depending on whether or not they looked for a job. Also, in the control form, layoff and temporary absence are choices in the same question, while in the respondent-defined layoff form, layoff is the first question and temporary absence is the second question. The suspicion is that many people who responded that they were laid off on the control form were never going to return to work and therefore are like any person without work. The expectation is that the percentage of people laid off should be lower for the respondent-defined layoff form, with those people more likely to be classified as out of the labor force.

**Table 3: One Vs Two Layoff Categories**

Form	DS-2A	DS-2C	
Form Description	Cont.	RDL	
# layoff categories	One	Two	Sig. Diff.
# people in analysis	1,021	1,102	
Any gap errors (%)	1.2	10.9	Yes
Any loop errors	14.4	24.8	Yes
Official layoff (%)	5.1	2.1	Yes
Temporarily absent	2.6	4.8	Yes
Unemp--no lay off	5.1	9.1	Yes
Not in labor force	87.2	84.0	Yes
Unable to code	0.7	10.5	Yes

The results received from respondent-defined layoff form differed significantly from the control form. Since the control version had just one simple skip, the opportunity for gap errors was small, so it is no surprise that the percentage of gap errors is lower. The goal of differentiating the type of layoff was to lower the number of people mistakenly thought to be on layoff. The evidence is that the goal was accomplished--the percentage of people in the study on layoff was less than half for the respondent-defined layoff form as it was for the control form, with that difference being found significant. However, the hypothesis was further that these people who were mistakenly classified as on layoff were really not in the labor force. That was not true--the percentage of people out of the labor force was actually significantly lower for the respondent-defined layoff

form, not higher. Instead, the percentages of people temporarily absent and unemployed but not on layoff, were higher for the respondent-defined layoff form.

Why was that so? The theory was that many of these permanently laid off people were not looking for work and therefore not in the labor force, but research showed about 70 percent of the permanently laid off people met the official definition of being unemployed. Therefore, the percentage of people who were unemployed was not significantly different between the two forms.

## 5.2 RESPONDENT- VERSUS GOVERNMENT-DEFINED LAYOFF

In the respondent-defined layoff forms, the respondent determined the layoff status by checking temporary or permanent layoff, while in the government-defined layoff, layoff first form, the respondent answers a question (returning to work in the next six months or have a specific date) which determines their layoff status. With both forms, layoff is the first question in the non-work status section and temporary absence is the second question. The expectation is that there will be a higher proportion of people classified as on layoff for the respondent-defined layoff forms than for the government-defined layoff, layoff first form, because people will consider themselves on temporary layoff when by government definition they are on permanent layoff.

**Table 4: Respondent- Vs Government Layoff**

Form	DS-2C	DS-2E	Sig. Diff.
Form Description	RDL	GDLLF	
Layoff determined by	Resp	Gov't	
# people in analysis	1,102	1,087	
Any gap errors (%)	10.9	14.5	Yes
Any loop errors	24.8	32.9	Yes
Official layoff (%)	2.1	1.6	No
Temporarily absent	4.8	5.6	No
Unemp--no lay off	9.1	8.8	No
Not in labor force	84.0	84.0	No
Unable to code	10.5	13.2	No

The respondent-defined layoff form had less gap and loop errors than did the government-defined layoff, layoff first form. The government-defined layoff, layoff first form has the most complicated skip pattern of any of the forms, so that result is not a big surprise. However, those were the only significant differences in the results of the forms, which was not the hypothesis. The percentages of people on layoff, temporarily absent, unemployed but not laid off, and not in the labor force were not significantly different between the forms. Even the percentage of people that could not be coded directly was not

significantly different between the forms, which indicates that the gap and loop errors are at least not making it hard to classify the person's non-work status.

## 5.3 LAYOFF VS TEMPORARY ABSENCE FIRST

This comparison was designed to examine the effects of question ordering effect. No significant differences were found between the government-defined layoff and government-defined layoff with distance forms, so their data was combined to increase the power. They hereafter are referred to as the government-defined layoff, absence first forms, and they begin with a question asking if the person was temporarily absent from their job for any reason but layoff last week. If not, they were asked if they were on layoff.

The government-defined layoff, layoff first form starts, as the name implies, with a question asking if the person was on layoff, and if not, asks them if they were temporarily absent from their job (without mentioning the layoff exception). The latter form asks the questions like the CPS does, but the skip pattern is more complicated. For example, no matter how they answer the temporary absence question, which is the second question, they are supposed to skip over the third question and go either to the fourth or the fifth questions.

The hypothesis is that the government-defined layoff, layoff first form will have more gap and loop errors due to the more complicated skip pattern. There also might be more people temporarily absent from work and less on layoff for the government defined layoff, absence first forms since temporarily absent is first.

**Table 5: Layoff vs Temporary Absence First**

Form	DS-2E	2B&D	Sig. Diff.
Form Description	GDLLF	GDLAF	
First NWS question	Layoff	Absence	
# people in analysis	1,087	2,527	
Any gap errors (%)	14.5	15.3	No
Any loop errors	32.9	36.4	Yes
Official layoff (%)	1.6	1.2	No
Temporarily absent	5.6	10.6	Yes
Unemp--no lay off	8.8	6.8	No
Not in labor force	84.0	81.3	No
Unable to code	13.2	14.7	No

The hypothesis that there would be more gap and loop errors on the more complicated government defined layoff, layoff first forms turned out to not be correct. There was no statistically significant difference in gap errors, and there were actually less loop errors on the government defined layoff, layoff first form.

There was no difference in the percentage of people

laid off between the two forms, but as hypothesized, there were a higher percentage of people on temporary absence for the government defined, absence first. There was no difference in the percentage of people in the total unemployment or unable to code directly lines. What explanation is there for the results that were seen?

The key is the temporary absence question on the government defined layoff, absence first forms. In a sense, layoff is a subset of temporary absence--people on layoff might easily consider themselves "temporarily absent" from their jobs. To put the temporary absence question first in order to simplify the skip pattern, therefore, the question was worded "...TEMPORARILY absent from a from a job or business for any reason except layoff?"--"temporarily" is highlighted in all capital letters, while the layoff exception is buried at the end of the line. The person could have missed the layoff exception, marked that they were temporarily absent and correctly skipped out of the section (missing the layoff question). They also could have read the question, thought, "yes, I'm on layoff", and answered yes and skipped out. The conclusion, then, is that the government-defined layoff, absence first form overestimates people temporarily absent from their jobs.

Why, then was there no significant difference in the number of people laid off between the two forms? The first question in the government defined layoff, layoff first form was about layoff in general, without defining it. A tabulation of the people who marked that they were on layoff on the government defined layoff, layoff first form indicated that some were on official layoff, but more were officially unemployed but not unemployed or out of the labor force. That caused a nominal increase in the officially laid off, unemployed but not laid off, and not in the labor force categories for the government defined layoff, layoff first form as compared to the government defined layoff, absence first forms, but not enough to be statistically significant.

## 6. CONCLUSIONS

The research reported in this paper shows that collecting complicated information in a self-administered paper-and-pencil format can be difficult. It is worth noting the large number of gap and loop errors in the test versions of the forms. The user-friendly design of the forms the Census Bureau is doing for the forms to be used in Census 2000 will hopefully minimize that problem.

It is clear that there was a problem with the control form overestimating the number of people on layoff. All of the test forms had significantly less people on layoff than the control form. However, those people who said that they were on layoff when they were not by definition did not move to the "not in labor force" category as first

hypothesized. The evidence is that many of those people are looking for work and therefore considered unemployed. When research is done on this issue in the future, a version that had a combined layoff and temporary absence question, as on the control form, followed by a government-defined layoff set of questions, as was done on the government-defined layoff forms, would be worth testing.

The Census Bureau will use the government-defined layoff, layoff first version of the non-work status section in Census 2000. A big factor in that decision was the compatibility of the definition of unemployment with the CPS. The government-defined layoff, layoff first form performed better than the government-defined layoff, absence first forms. The respondent-defined layoff forms had less gap and loop errors, but improved design of the forms should make that less of a problem.

---

## REFERENCES

- Jenkins, C., and Ciochetto, S. 1993. "Results of Cognitive Research on the Multiplicity Question from the 1991 Schools and Staffing Survey Student Records Questionnaire." A report submitted to the National Center for Educational Statistics, U.S. Bureau of the Census Memorandum, February 10, 1993.
- Jenkins, C., and Dillman, D. 1993. "Combining Cognitive and Motivational Research Perspectives for the Design of Respondent-Friendly Self-Administered Questionnaires." Prepared for the American Association for Public Opinion Research Conference, 1993. U.S. Bureau of the Census, Washington, DC.
- Leslie, T. 1996. "U.S. 2000 Census Test - Mail Response Analysis." U.S. Bureau of the Census 1996 National Content Survey DSSD Memorandum No. 2, November 13, 1996.
- Palumbo, T. 1993. "Sources of the Unemployment Differences Between the 1990 Census and the Current Population Survey (CPS)." *Proceedings of the American Statistical Association Conference, 1993*, American Statistical Association, August, 1993.
- Turner, C.F., Lessler, J.T., George, B.J., Hubbard, M.L., and Witt, M.B. 1992. "Effects of Mode of Administration and Wording on Data Quality" in C.F. Turner, J.T. Lessler, and J.C. Gfroerer, (eds), *Survey Measurement of Drug Use Methodological Studies*, Washington, DC: National Institute of Drug Abuse, U.S. Department of Health and Human Services, pp. 221-243.