Cleo R. Jenkins, Bureau of the Census<br>CSMR, Washington, DC 20233

## KEY WOPDS: Self-administered Questionnaires, Institution Surveys

The Schools and Staffing Survey (SASS) is conducted by the Census Bureau for the National Center for Education Statistics. It is a relatively new set of integrated surveys first conducted in the 1987-88 and 1990-91 school years and scheduled to be conducted every three years hence. Selfadministered questionnaires, of which there are eight, are mailed to public school districts, and to both public and private school administrators, the schools themselves, and to teachers within the schools, asking questions about enrollment, teaching positions, as well as other school and teacher characteristics.
This survey has recently been the focus of questionnaire design research at the Census Bureau. One particular Schools and Staffing Survey, the Public School Questionnaire, was chosen for in-depth study. This survey had an especially high pre-edit failure rate in 1991 (Jenkins, 1992). This means that information from the 1991 school questionnaire disagreed with comparable data for the same school from a survey conducted by the National Center for Education Statistics, known as the Nonfiscal Surveys of the Common Core of Data (CCD).
This paper describes both the methods that were used to conduct this study and some results of the research.
I. METHODOLOGY

One objective of this research was to gain in-depth knowledge about questions that had high pre-edit failure rates in the 1991 surveys. Another was to test newly developed questions. A condensed version of the Public School 1991-92 Field Test Questionnaire (SASS-3A) served this purpose well.

Once the scope of the questionnaire was defined, the researchers used their intuition and experience in questionnaire design to uncover potential problems in the questionnaire and to develop questions to probe respondents' understanding of the items.
After establishing the protocol, twenty in-depth interviews were conducted, four in each of five mid-western states: Oklahoma, North Dakota, South Dakota, Nebraska, and lowa. These particular states were chosen because they exhibited the highest pre-edit failure rates in student and teacher counts in 1991. Together the National Center for Education Statistics (NCES), the agency that sponsors the SASS, and the Demographic Survey Division (DSD), the division within the Census Bureau responsible for conducting the SASS, provided the researchers with a list of approximately ten schools within each state. The reason for supplying the researchers with more than the final four schools was to allow for scheduling conflicts and refusals.
Not wanting to burden respondents, while at the same time wanting to study the reject phenomenon, a compromise was reached in which three of the schools selected in each state were not in any other SASS. The remaining school, however, was a 1991 pre-edit failure. A final constraint on sample selection was that the schools needed to be within a few hours' drive of the major city in which the researchers were based in each state.
The Public School Questionnaire is addressed to the
school principal. During the actual survey, a label is affixed to the right-hand side of the cover page. The researchers mimicked this (see Figure 1).
The researchers contacted the principals, explained the nature of the study, and established a date and time to conduct the interview. The final sample consisted of respondents who were willing to participate. The interviews tended to last an hour and a half. They were tape-recorded and summaries of the interviews were written (see Jenkins, 1992, for summaries of each interview).

## II. RESULTS AND DISCUSSION

The results of the cognitive interviews are discussed in the remainder of this paper. I have decided to focus on errors that resulted from the misunderstanding of concepts; the layout of the questions themselves, and finally, the use of records. I use two examples from the cognitive interviews to describe each of these errors in detail (see Jenkins et al., 1992, for a complete review of the results of the cognitive interviews).

## A. Misunderstanding of Concepts

The cognitive interviews revealed two concepts that were widely misunderstood: one was respondents' understanding of the school for which they were to report and the other was the classifying of employees in full-time or part-time positions. The many reasons for these misunderstandings are described below.

1. Respondents' Understanding of the School For Which They Are to Repont

A very important concept and one which affects the way respondents answer every item on the questionnaire is their understanding of the unit for which they are supposed to report. The cover page contains a very important instruction in the paragraphs on the lett-hand side of the page that tells respondents to "Please complete this questionnaire with information about the SCHOOL name on the label." (See Figure 1.) This is the unit for which respondents are supposed to report. I have categorized the result of this understanding into three types: general agreement between their understanding and the intent of the questionnaire, ambiguity between the two, and finally, disagreement between the two.
The first group is made up of nine respondents whose understanding of the school for which they were to report generally agreed with the questionnaire's intent. Basically, respondents in this first group were inclined to report in terms of the school named on the questionnaire label, whether they read the school named there or not. Six of the nine respondents actually read the label.
Respondents who were principals over schools that clearly stood apart (i.e., functioned separately and/or were not in close proximity to any other school) seemed to fall into this first group. These respondents were not confused as to their school's identity. For the most part there was a clear demarcation such that the principals did not consider reporting for any other school(s).

Group II was made up of eight respondents whose understanding of the school for which they were to report was ambiguous. Generally, this was the case in relatively small school systems in which two or three schools comprised the entire school district. Often the schools were
housed in one building or they were housed in separate buildings that were clustered around one another. There was a principal for each of the two or three schools within the district, but the principal of the school named on the label saw himself as capable of reporting for the other school(s), if he thought, however begrudgingly, that's what was being asked of him. Because of the schools' close association with one another, the line of demarcation wasn't as clear for these respondents.

These respondents had alternative definitions of the unit for which they could report and they relied on the questionnaire to inform them which one to use. On one hand, they could define their school as the grades over which they had jurisdiction. Because of their organization, however, it was conceivable to them that the questionnaire might be asking about the entire school system, kindergarten $(K)$ through 12th grades. As a result, they were quite receptive to cues from the questionnaire. Unfortunately, these cues were conflicting. Half of them began to complete the questionnaire by reading the cover page. Generally, they read through the title information and then the first two paragraphs on the lett-hand side of the page. Because the paragraphs refer to the label, they turned the questionnaire sideways to look at the label. The other half of the respondents, however, never noticed the school named on the label. Neither the instruction referring to the label nor the school's name itself is prominent. Both are buried among a lot of information on the cover page. In fact, the school's name is not only buried, but it is turned sideways (see Figure 1).
After providing their name and address in item a, which is the first question on the form, these respondents turned to item $\mathbf{b}$. Hem $\mathbf{b}$ asks if the school serves students in any of grades 1 through 12. If the school doesn't, they are instructed to return the questionnaire to the Census Bureau. If they do, they are to continue. The reference to "grades 1 through $12^{\prime \prime}$ in this item seemed to trigger these respondents into thinking that the questionnaire might be asking about the entire school system rather than just their school. They weren't sure, but they now had a reason to believe this was the case.
After answering item c , which asks about their School State Identification Number, they turned to item 1. The question of item 1a asks for what grade levels the school offers instruction, and the question of item 1b asks how many students were enrolled in each grade on October 1 of the school year. In both cases, prekindergarten through 12th grade answer categories are provided (see Figure 2). The answer categories seemed to provide these respondents with more evidence that the questionnaire might be asking about the whole school system. As a result, some began interpreting item 1 as asking about the entire school, but most didn't.
Most waited until they reached item 2, which asks how many students were enrolled in the school in grades K-12 on or about October 1 of this school year. For the most part, these respondents voiced their ambiguity at item 1 , but still they answered item 1 in terms of their school. In some cases, they may have done this simply because it was easier, but in other cases, it seemed that these respondents needed more evidence before they could be swayed into reporting for the entire school system. And the fact that item 2 seemingly asks for the number of students in grades "K-12" became the evidence they needed.
Once made, this interpretation was continuously
reinforced by the many "K-12" references in the questions that follow item 2 , until eventually it became solidified in the minds of some respondents. They stopped questioning the unit for which they should report and began to report for the entire school system. This is not to say, however, that this was painless. The fact was they needed to go through a great deal of work to obtain information to answer for the entire school.
H is not surprising, therefore, that their interpretations weren't always the final determinant of how they reported. Sometimes the records they had on hand became the limiting factor. This meant that although their interpretations might be relatively consistent from item 2 onward, their answers were not necessarily consistent. Sometimes they answered in terms of the school system. This was often the case with item 2. This item requests a summary statistic they often had on hand. It asks how many students were enrolled in the school on or about October 1 of this school year. At other times, they answered in terms of their school only, as was often the case with item 3 . This item requests information they couldn't conveniently obtain. It asks for a breakdown of the student population into ethnic categories.
Also, it should be noted that some respondents continued to express ambiguity. These respondents didn't settle on one definition, but instead interpreted questions in which they noticed the "K-12" reference as asking about the entire school system and questions in which they didn't notice this reference as asking about their school.

Group III was made up of three respondents whose definitions simply didn't agree with the questionnaire's. Two of these respondents had jurisdiction over both the elementary and secondary portions of a relatively small school, with both portions housed in one building. In another case, the respondent was principal of both the middle and high school portions of the school system, which again were housed in one building. In these cases, the respondent's definition of his school was clearly different from the school named on the label, and the problem was that the questionnaire tended to reinforce this wrong definition. In fact, two of these respondents never looked at the school named on the label.
2. Respondents' Understanding of Full-time Versus Part-time Status.

Another concept respondents had a very difficult time with was that of full-time versus part-time employment, as asked for in item 30 (see Figure 3). To understand why respondents misreported, it may be best to begin with a situation in which respondents were likely to report correctly. They were likely to correctly report an employee as part-time if that employee was exclusively part-time and the job itself could be considered full-time. For example, respondents were likely to report an instructional aide as part-time if that aide only worked for part of the day, meaning he/she didn't work the rest of the day, and there were others who did work all day as an aide. In this case, the part-time aide could be compared to a full-time aide and there wasn't any confounding information with which to be confused (i.e, any other assignment or job). As a result, the situation was clear to them, but this was also one of the less frequent situations.
The more frequent situations were less clear. For example, employees who worked at jobs that by definition could never be considered full-time jobs were difficult for respondents to categorize. This was the case with bus drivers. Respondents could agree that bus drivers always work less than a full day, but they couldn't agree if that meant they should be
categorized as part-time or full-time. The reason they couldn't agree on this is that the bus driver's job is not fulltime relative to other full-time jobs, such as the principal's job; however, it is full-time if the unit of comparison is limited to a bus driver's job. Looked at from this perspective, it is as "full" a job as a bus driver's job can get.

Also, problems arose when an employee worked part-time in more than one assignment, but full-time at the school. One reason respondents misunderstood this concept was that they were used to thinking in terms of an employee's employment status at the school overall and not by assignment. Take, for instance, an aide at the school, who works full-time, but whose assignment is divided between being an Instructional aide and librarian aide. More often than not, the respondent would report this employee as a full-time instructional aide and full-time librarian aide. The same happened with a teaching principal. He reported himself as a full-time teacher and then again as a full-time administrator. In these instances, respondents thought of the employees as full-time and had difficulty thinking of them as part-time.

Respondents also had difficulty if an employee worked part-time at this school, but full-time for the school district, meaning the employee was shared among the schools. In the smaller schools, many of the staff were shared, including librarians, guidance counselors, clerical staff, the student support services staff, and the other support staff. Here again, respondents had a tendency to report these employees as full-time.

## B. Format Considerations

Errors occur when an item is laid out such that respondents don't see, and consequently don't read, information that is necessary to correctly answer the item. Respondents commonly overlooked information that was placed beyond what they considered to be the answer space, including "none" boxes and skip instructions. As a result, they were likely to spend a great deal of time and energy trying to answer questions that didn't apply to them, as demonstrated below with item 15. There were also instances in which an entire item was laid out poorly, as demonstrated below with item 29.

## 1. "None" Boxes and/or Skip Instructions

Item 15 asks a series of questions about limited-English proficient students (see Figure 4). Part a of this item asks "How many students attending this school as of October 1, 1991, were identified as limited-English proficient (LEP)." In response to this question, quite a few respondents made the mistake of reporting " 0 " on the answer line because they didn't notice the "none" box that was placed about half an inch beneath the answer line. The cognitive interviews revealed quite a bit about how respondents interpret questions that don't apply to them from this.
Respondents who had previously had LEP students but who didn't have any now used their past experience to answer part b, which asks what methods were used to identify LEP students. They reported the methods they had previously used to identify LEP students. Another respondent whose school had never had any LEP students answered the best he could by marking the "other" answer category and writing in "never been a problem." It became evident as a result of this research that respondents commonly marked the "other" box and wrote something in when they thought they were supposed to answer a question, but they couldn't understand it. Either it was
ambiguously worded or it wasn't applicable to them, as was the case here.

In general, respondents who had previously had LEP students came to realize that part c , which asks about the number of LEP students in specified programs, didn't apply to them and correctly skipped to the next item at this point. These respondents seemed to be familiar with the notion of limited-English proficiency and its acronym. This helped them realize that this question didn't apply to them.
Unfortunately, respondents who never had LEP students just plowed away, trying to answer questions they shouldn't have. This was probably due to the fact that only the acronym LEP is used in this question and although it was defined previously, they weren't really familiar with the notion of limited-English proficiency in the first place, let alone its acronym.
It became obvious as respondents tried to answer this part of the item that they didn't really know what programs (1) through (6) were, since they didn't have and never had any LEP students. Consequently, they transiormed these answer categories into something that had meaning to them. All sorts of misinterpretations arose as a result. One of the more reasonable interpretations was to think it was asking for the number of regular students in the listed programs. In this case the respondent reported "none" in all but the fourth category, where he reported all of his special education students. In other words, he didn't change the meanings of the individual programs per se, just the population to which they applied.
Matters really broke down, however, when respondents not only thought the question applied to regular students, but they changed the meanings of the programs as well. This happened most for the first two programs. These were written such that respondents couldn't comprehend the entire sentence, but they could find meaning in individual words. For instance, one respondent thought the first category (subject matter in home language) was asking for the number of classes in grades 7 through 12. This respondent seemed to key in on the words "subject matter." To her these words were associated with the number of classes in grades 7 through 12. To understand this, one must realize that usually subject matter is taught in subject matter classes in grades 7 through 12, and not at the elementary level. Another respondent interpreted the second category (maintaining fluency in home language) as asking if the school offered foreign language instruction in Spanish. Obviously, this respondent noticed the word "Spanish" in the example and extrapolated from that a program that had meaning to him. The point is these respondents were not answering the questions asked of them.
2. Hem Layout

Hem 29 asks a series of questions abóut teaching vacancies in the school. There is a problem with the layout of part d in this item (see Figure 5), which asks how difficult or easy it was to fill the vacancies in the listed fields. The first problem was that not all subparts of the question applied to all schools. Elementary schools weren't sure how to answer parts (3) through (9), since these are subject matter courses not offered at the elementary level. It didn't seem quite right to mark "no vacancy in that field" when the truth was they didn't even have that field. Conversely, high schools weren't sure how to mark the first two categories.

Also, there was a special problem with the layout of the ninth category (9), vocational-technical education. Very few respondents read the follow-up question on the left-hand side
beneath this category. Consequently, they didn't understand this part of the item. They were supposed to mark how difficult or easy it was to fill the vacancies they had in vocational-technical education in the boxes on the right-hand side. Then they were supposed to identify the subfield(s) of vocational-technical education to which that mark applied in the follow-up list of subfields. However, respondents were misled by the fact that the boxes fall under the "no vacancy in that field" column. Because they tended to see the list of subfields on the right-hand side as just a continuation of the fields specified on the left-hand side, they often continued right on down the column, marking these boxes (as they had the others) to indicate they didn't have these fields.

## C. Use of Records

The most striking aspect of respondents' use of records was how varied their recordkeeping systems were. They ranged from slips of handwritten papers that were found in the top drawers of their desks or hanging near them to more formal systems. Some used report forms that came from files in either their office or the secretary's office; some even used state-of-the-art computer databases.

The cognitive interviews revealed that using records did not necessarily guarantee the data would be accurate. Errors arose when respondents didn't use appropriate records. Sometimes this was because they didn't have the appropriate records. Other times it was because they didn't think they had them. Still, other times it was because they didn't recognize the record was inappropriate. One very common error resulted from respondents thinking they didn't have records for the time period specified in a question when in fact they did, as demonstrated below with item 1b. Another more complicated error occurred when respondents applied misguided heuristics to the use of their records, as described below with items 2 and 3.

## 1. Use of Inappropriate Records

Item 1b asks how many students were enrolled in each of the listed grades on October 1 of this school year (refer to Figure 2). By law, schools are required to submit reports with student enrollment by grade for around October 1st to either the school district or the state. Respondents should have used this report to answer this question. It would have saved them from having to reproduce numbers and probably would have yielded more accurate data. However, a third of them didn't.

One reason for this may be that the question fails to tell them to use it, and in fact, it may even hinder them from considering it. According to the framers of this question, they had the official fall reporting date in mind when they used the date "October 1." They expected respondents to associate this date with the official fall reporting date. Among respondents who focused on the "October 1" date, however, this was either not enough to trigger them to think about their fall report, or if it did, it caused them to dismiss it. For instance, one respondent dismissed using the state report because it was dated September 10th rather than October 1st. He had the business office go through the trouble of producing October 1st numbers from their computer database when the report dated September 10th was already available. Although he reported for the right time period, the office spent more time than necessary answering this question.
Another reason respondents didn't use the fall official report was because they weren't aware of its existence. These respondents, who were the principals of the schools,
either weren't as familiar with the school records as their secretaries or they were new to the job. In these cases, the respondents just didn't realize they could comply with the reference period, so they did what they thought best: they ignored it and reported data for the time period they had.
Relying on other records forced this group to report data for a time period different from the one requested. They reported numbers for the beginning of the year, end of the first quarter (November 3rd), end of second semester (January 13th), as well as current figures.
2. Heuristics Applied to the Use of Records

Hem 3 asks for a break-down of students into ethnic categories (refer to Figure 6). The majority of respondents used a heuristic to answer this question. First, they relied either on their knowledge of the student population or on some kind of record to report the number of students in the ethnic categories in parts a through d. After this, they calculated the number of white students in part e by subtracting the total number of minority students from the total they had reported in item 2 a (refer to Figure 7). As a result, the total number of students reported in item 3 was consistent with the number reported in item 2a. However, the number of white students was not always accurate.
This approach was fine, as long as the record they used to answer item 3 was for the same time period as the record they had used to answer item $\mathbf{2 a}$. Then the data were not only consistent, but they accurately reflected the ethnic counts at a given point in time. However, since item 3 doesn't specify a time period, a few respondents answered item 3 using current data, whereas they had used records as of October 1st to answer item 2a. It wasn't obvious to these respondents that they might be introducing an error into the data by deriving the number of white students as they did.
Also, their method of calculating white students was flawed if the minority counts themselves were off, which was the case a number of times. For instance, one respondent reported the number of American Indians as of last year. He initially interpreted this question to be asking for last year's numbers because of the reference period given in item 2 b . In addition, he reported the wrong number of black students because he made a mistake when he manually counted up these students from a student list. When he was done reporting these wrong counts, he proceeded to calculate the number of white students by the method mentioned above. As a result, the white count was off as well.

Another respondent double counted the number of minorities she reported in parts a through d because of the way she answered here. According to the secretary, the school actually reported all minorities as American Indian on a report they submit to the Office of Indian Education. Since the respondent used this report to answer part a, she inadvertently reported all minorities as American Indian. Following this, she went on to report the minorities again in parts $\mathbf{b}, \mathrm{c}$, and d . As a result, the number of white students was also erroneous.
In these cases, the numbers didn't accurately reflect the ethnic counts, but the values reported in items 2 and 3 were consistent. In some cases, these mistakes seemed to be the result of respondents not paying close attention to what they were doing. In other cases, it seemed to be because the questionnaire asks for data the respondents didn't have in the requested format. And in still other cases the questionnaire asks for data with which the respondents weren't wholly familiar.

## CONCLUSION

In this paper, I have described questionnaire research with the Public School Questionnaire from a cognitive perspective, meaning how and why respondents interpreted information as they did. Examples of respondent errors from the cognitive interviews were presented, including errors that resulted from the misunderstanding of concepts, the layout of the questions themselves, and from the use of records.

The cognitive interviews revealed that errors occur because information presented on the questionnaire is not always perceived as intended. Many respondents did not understand the school for which they should report. In large part, this was due to the fact that the school's name is hidden from view on the cover page and suggestive references to the entire school system are used throughout the questionnaire. In general, this error should be relatively easy to correct. Most respondents were inclined to report their school correctly, but were just confused by the questionnaire. On the other hand, many respondents didn't understand the concept of full-time versus part-time employment as intended by the questionnaire; however, this may be more difficult to correct because asking respondents to think as the questionnaire does is asking them to think in a relatively complex and foreign way.

The "none" boxes and skip instructions present respondents with problems, and this seems to be due to the method respondents use to answer questions. Once respondents answer a question, they seem to think the response task is over. As a result, they do not take in new information until they begin what they perceive to be the next "question-answer" cycle. Also, the layout of the questions themselves sometimes give respondents difficulty. However, mistakes such as these may be relatively easy to correct.

Respondents' use of records is one of the most complex areas of questionnaire research to study, since it requires indepth knowledge about respondents' records as well as how they use those records, and very little is known about this process to date. Certainly this is an area in need of further research. As demonstrated earlier, problems can occur when respondents use records. Some of the errors that were witnessed during the cognitive interviews may be correctable, some need further research, and some seem to be intractable. Errors that arise from questionnaire miscues, such as the use of inconsistent time periods and not providing clear references to particular records may be relatively easy to correct. However, mistakes that occur for other reasons may be difficult to correct. An error needing further research is one that arises because respondents do not have information in the requested format. In-depth studies are needed to design questions that ask for information in appropriate formats. An example of a mistake that may be intractable, however, is one in which respondents do not pay close attention to what they are doing.

The next step in this process will be to redesign the questionnaire using guidelines resulting from this research. The first and probably most important guideline is that the school's name and grade levels should be prominently displayed. The final step will be to conduct a test of alternative questionnaires. Discussions are underway on how best to conduct this test.

REFERENCES
Jenkins, C. J. 1992. "A Review of the Post-edit Item Response Rates, Pre-edit Reject Rates, and Edit Change Tallies from the 1991 Schools and Staffing Surveys." Internal Census Bureau memorandum. March 23, 1992.
Jenkins, C. J. 1992. "Cognitive Interviews on the Public School Questionnaire." Internal Census Bureau memorandum. April 30, 1992.
Jenkins, C. J., S. Ciochetto, and W. Davis. 1992. "Results of Cognitive Research on the Public School 1991-92 Field Test Questionnaire for the Schools and Staffing Survey (SASS-3A). Internal Census Bureau memorandum. June 15, 1992.

Jenkins, C. J., S. Ciochetto, and W. Davis. 1992. "Recommendation for Changes to the Public School 1991-92 Field Test Questionnaire for the Schools and Staffing Survey (SASS-3A) Based on Cognitive Interviews." Internal Census Bureau memorandum. Sept. 24, 1992.



Figure 2. Student Enroliment by Grade Item


Figure 3. Classification of Employees Item


Figure 6. Student E nrol 1 ment by Ethnicity Iten $^{\text {I }}$


Figure 4. Number of Limited-English Proficient Students Item


Figure 7. Total Student Enroliment ite

