RESPONSE ANALYSIS SURVEYS FOR ASSESSING RESPONSE ERRORS IN ESTABLISHMENT SURVEYS

Karen L. Goldenberg, Shail Butani, and Polly A. Phipps. Bureau of Labor Statistics
Karen L. Goldenberg, Bureau of Labor Statistics, 2 Massachusetts Ave. N.E., Room 4985, Washington, DC 20212

KEY WORDS: Response error, Response Analysis Surveys, establishment surveys

1.0 Introduction

Research on survey data quality and response errors has generally been based on data from surveys of households and individuals. Response error, the difference between a respondent's answer to a question and the true answer (Sudman and Bradburn, 1974), has become the subject of an extensive literature (e.g., Bradburn, 1983; Groves, 1989; Sudman and Bradburn, 1974), but relatively little work has addressed the problem in establishment surveys. Establishment surveys have human respondents, and thus are subject to many of the same types of response errors as surveys of households or individuals[1], but they differ from household surveys in two important respects: the type of data collected, and the respondent's role in providing those data. Both type of data and respondent's role are components of the survey task.

In this paper, we briefly examine the survey task in household and establishment surveys, and look at cognitive approaches to survey design and error reduction, especially Response Analysis Surveys (RAS). We present a framework for organizing establishment survey response errors according to their sources and component elements, and illustrate these errors with examples identified in the RAS studies.

2.0 The Task in Household and Establishment Surveys

Household and establishment surveys differ in the type of data requested and the nature of the respondent's role in providing those data. In a household survey, the task is "to obtain information from a sample of respondents about their (or someone else's) *behavior and/or attitudes.* The respondent's role is to provide that information; the interviewer's, to obtain the information in the manner prescribed by the researcher" (Bradburn, 1983:291, emphasis added). While the respondent and the interviewer can both be sources of response errors, Bradburn considers their contribution to overall survey response error much less important than contributions of the task.

In establishment surveys, the task is to obtain information from a sample of responding establishments about *some aspect of the organization.* Since establishment surveys frequently use self-administered questionnaires rather than interviewer-administered questionnaires, the respondent's role is to locate the source of the information and to provide it. The questionnaire functions as the researcher's means of collecting the data.

[1] For purposes of this discussion, we refer to a *household survey* when the unit of analysis is an individual or a household.

The cognitive information processing model (Tourangeau, 1984) is a useful way to view a survey respondent's role in an interview. The interviewer asks a question, and the respondent must first understand what that question means (comprehension). The respondent then searches his or her memory for information with which to answer the question (retrieval), decides what information to relate and how to relate it (judgment), and transmits the response to the interviewer (communication). The wording of the question shapes the respondent's comprehension and influences the retrieval process (e.g., Schwarz and Hippler, 1991).

The response process is similar in an establishment survey conducted with a mail questionnaire. Frequently, the respondent must seek answers to questions from records or some external source (generically, an information system). Instead of an interviewer demonstrating the task to the respondent, the questionnaire must convey the task on its own. In this sense, the respondent mediates between the questionnaire and the information system (Edwards and Cantor, 1991). Like the interviewed household respondent, the establishment respondent must understand the question and decide what information is required to answer it (comprehension). The respondent must then search his or her memory to determine whether or not the information is available in the information system, and if so, where. If the data are available, the respondent must access those data (retrieval), possibly by communicating the data requirement to a third party. The respondent must decide what information to report (judgment) and compile that information. Finally, the respondent must present the information in the format requested on the data collection form or questionnaire (communication).

3.0 Cognitive Approaches to Survey Design and Error Reduction

In the past few years, increasing numbers of survey designers have incorporated a cognitive approach into the design or review of questionnaires. The cognitive design approach allows the researcher to understand a questionnaire from the perspective of the respondent, and thereby to identify and correct potential response errors that result from discrepancies between the researcher's concept of the survey task and the respondent's understanding of it (Freedman, 1988; Gower and Nargundkar, 1991). Cognitive techniques relevant to designing establishment surveys conducted by mail include focus groups, think-aloud interviews with in-depth probing, and detailed retrospective interviews. These techniques are powerful tools for revealing possible response errors, but they have limitations. They typically involve only a small number of participants, are frequently conducted in a laboratory

setting, and are extremely labor-intensive and time-consuming to implement. Inferences made from the application of these techniques to a larger population must be made with caution.

A Response Analysis Survey (RAS) is another technique for evaluating questionnaires from the perspective of the respondent, an approach that complements the other procedures. It is essentially a respondent debriefing about a survey, a retrospective analysis conducted after a respondent has completed the main survey, and on a larger scale than that possible for pretesting with cognitive methods. It uses a structured questionnaire and generates quantitative data about how respondents answered questions, about the records (i.e., the information system) available for answering those questions, and about the real burden imposed by a survey. A RAS may be administered to a stratified subsample of respondents to the initial survey, so the results can be generalized to the larger survey sample or to a sample selected for a particular characteristic or problem. It is especially useful in the context of a recurring survey, where findings can be used to improve the next data collection cycle, or to evaluate changes made in an earlier version (DeMaio, 1983).

4.0 BLS RAS Studies

The Bureau of Labor Statistics (BLS) uses the RAS as a tool to assess response errors in recurring and one-time establishment surveys. This paper reviews the results of five different RAS studies: the 1989 and 1990 Occupational Employment Statistics Wage-Pilot Surveys (OESWP), the Employee Turnover and Job Openings (ETJO) Pilot Survey, the Current Employment Statistics Non-Wage Cash Supplement (NWCS) Survey, and the 1991 Hours at Work Survey (HWS). All of the RAS studies were conducted as telephone interviews with a sample of respondents who had previously completed a regular survey. Most RAS participants were mail respondents to the regular surveys, but some were originally interviewed by phone as part of the nonresponse follow-up procedure. We briefly describe each of these surveys.

1989 and 1990 Occupational Employment Statistics Survey Wage Pilots. The Occupational Employment Statistics (OES) Survey is an ongoing BLS survey which collects data on employment in specific occupations by industry. The 1989 OES Wage Pilot investigated the feasibility of collecting wage information in addition to employment data. The OESWP questionnaire asked establishments to report the total number of workers in detailed occupational categories, and instructed them to distribute all workers in an occupation into seven wage ranges. The questionnaire contained definitions of the types of workers and wages that were to be included.

The 1989 RAS looked at records availability and respondent comprehension of concepts and instructions for the OESWP. With a sample of about 150, we uncovered a number of errors in the way wages were reported. We determined that some of these errors resulted from the wording and layout of instructions, and used these findings to modify the questionnaire layout for the 1990 survey. The RAS following the 1990 survey was conducted in four states, with a total of 221 completed interviews. It tested the improvements made as a result of the 1989 RAS and addressed the issues that were a problem in 1989, as well as additional aspects of data availability and respondent comprehension.

Current Employment Statistics Non-Wage Cash Supplement. The Current Employment Statistics (CES) Survey is an ongoing BLS data collection program that produces monthly reports of jobs and payrolls. The Non-Wage Cash Supplement (NWCS) was an addendum to the CES in 1989 that obtained data from employers on payments other than wages made to employees. We asked respondents to indicate whether their establishment made certain types of non-wage cash payments (e.g., bonuses), and to report both the dollar amounts of those payments and the total 1988 payrolls. We tested two questionnaire designs that obtained basically the same information but with different layouts.

The NWCS RAS sample was designed to focus on potential problem areas for future survey improvement. A total of 350 respondents answered questions that evaluated overall data quality, memory errors or misunderstanding, and the implications of the different form layouts.

Employee Turnover and Job Openings Pilot Survey. BLS developed the ETJO survey in response to a Congressional mandate to the Department of Labor to develop a methodology for measuring national labor shortages. The focus of the survey was to study the feasibility of measuring the number of job separations, new hires, wages of new hires, and job openings by major occupational group within an industry. Developmental work on this survey made heavy use of several cognitive techniques, including focus group analysis and pretesting with retrospective think-aloud protocols. The spring, 1991 RAS resulted in 420 completed interviews. Questions emphasized the respondents' adherence to instructions and definitions, sources of data used to complete the form, and response burden.

Hours at Work Survey. HWS is an ongoing national survey that obtains measures of productivity data. The survey asks responding establishments for the number of hours its production or nonsupervisory employees were paid the previous year, and the number of hours those people were actually at work (i.e., hours paid minus paid leave). While intended as a mail survey, BLS collects a substantial proportion of the data as nonresponse follow-up telephone interviews.

The purpose of the HWS RAS was to evaluate a recurring survey and to identify areas for improvement. This RAS resulted in 458 completed interviews, and examined data availability, respondent understanding of and conformity to survey procedures, and overall response burden. In addition, it specifically addressed the question of differences between mail and telephone respondents and the data provided by each group. Results of this RAS will be used, in conjunction with other cognitive methods, to redesign the questionnaire and survey procedures.

5.0 Response Errors in Establishment Surveys

Response errors in establishment surveys conducted by mail can be associated with different aspects of the task, the respondent's role performance, or the information system. They may result from the way the data collection instrument measures the correct value, from the respondent reporting incorrect information, from a mismatch between data in an establishment's records and the task definition (Federal Committee on Statistical Methodology, 1988), or from other sources.

In an establishment survey, the unit of analysis is the establishment, but the respondent is a person with a specific role in that establishment's organization. Therefore, "respondent characteristics" include both attributes of the establishment, such as size and industry, and characteristics of the responding individual, such as position in the organization and familiarity with the information system.[2] Characteristics of the establishment help to shape the information system. Larger firms have more complex organizational structures, more complex information needs, and therefore more complex data systems than smaller firms. Size of firm may also affect the number of people with knowledge of and access to data required for the questionnaire. Industry also plays a role. There is some evidence that manufacturing firms' record-keeping systems are better able than those of other firms to differentiate production workers from other workers (Goldenberg, 1993; Scott, 1983).

Characteristics of the information system may affect the respondent's ability to provide the correct data. Information must be present and available in a format compatible with the questionnaire. If it is not there, or will require extensive work to compile, it influences the respondent's perception of the survey task. The expected and actual amount of time required to complete the data collection form—that is, both the perceived and the real response burden—contribute to overall response rates and the quality of the data provided in the survey.

Finally, the "respondent" in an establishment survey may be more than one person, although researchers rarely obtain that information. Response errors may result if each participant has different knowledge of the subject matter or differential access to the information system. If questionnaire completion requires multiple respondents, the interaction of the group of respondents may affect the values reported (Federal Committee on Statistical Methodology, 1988; Fecso and Pafford, 1988).

Table 1 summarizes these errors as they apply to establishment surveys conducted by mail, and lists some of their components. We group response errors into those originating in the survey task, in the information system, and in the respondent. While the table provides a convenient organizing structure, we note that the response errors are not independent of each other. For example, an error that originates in the task, such as a question that does not produce accurate recall, is also a memory error when its source is viewed as originating in the respondent.

Table 1. Sources and Components of Establishment Survey Response Error

Source of Error	Components
Task	• Overall presentation • Data collection instrument - Question/instructions wording - Question/instruction layout and formatting • Confidentiality and sensitivity • Mode of administration • Response burden
Information System	• Data content and structure • Access and data retrieval
Respondent	• Memory • Use (or non-use) of information system

6.0 Response Errors Identified in RAS Studies

In this section, we use the structure from Table 1 to guide a discussion of establishment survey response errors and their various components. We review the literature on each type of error, and then present findings from RAS studies that add to our understanding of that error.

6.1 Errors associated with the task

Task components include aspects of the data collection instrument, such as length and wording of questions, question order, and layout and placement. On a self-administered establishment survey questionnaire, which may look like a series or matrix of labeled spaces, "questions" include the placement and wording of instructions and definitions as well as the actual questions, since they are part of the information conveyed to the respondent. The topic of the survey may lead to response errors, if respondents consider establishment-related information sensitive or confidential. Response burden is also an issue because the task may involve a substantial amount of effort, and because it is imposed on top of other responsibilities. Mode of administration is presumed to be constant in a mail survey, but many establishment surveys use telephone follow-ups to mail nonrespondents. There may be mode effects that result, or totally separate sources of error, if the follow-up uses different data collection procedures.

6.1.1 Overall Presentation of the Task

The general appearance of a self-administered establishment survey may affect a person's initial reaction to it, and therefore the willingness to respond. If a person agrees to participate in the survey, his or her first look at the task may affect data quality by influencing the amount of effort the person is willing to

[2] Other demographic characteristics of the responding individual (e.g., age, education, gender) may be associated with various sources of error, either independently or through role characteristics such as organizational position. However, it appears that the demographic variables have not been studied in an establishment survey context.

invest. Dillman (1978) notes that a "professional" appearance to a questionnaire enhances the importance of the survey to the respondent. He and others stress that the questionnaire should look as easy to complete as possible, with attention to language and structure, form design, color, type face, and avoidance of crowding text (Dillman, 1978; Dippo and Herrmann, 1991; Sudman and Bradburn, 1982). Cognitive investigations of two Canadian surveys, the Survey of Employment, Payrolls and Hours (SEPH) and the Census of the Construction Industry, showed that respondents initially found the questionnaires formidable and expected to invest considerable time and effort in completing them. Afterwards, the respondents reported that the form took less time than they expected. Repeat respondents on the SEPH had a very different, and more favorable, reaction to the questionnaire (Gower and Nargundkar, 1991).

None of the BLS RAS studies directly addressed overall presentation of the task. However, we have evidence from the ETJO survey that initial perception may have an effect on response. During the pretests, several respondents indicated that their first impression of the form was that it looked time-consuming. The survey package also affected perception. It consisted of a solicitation letter, an information sheet, a booklet with a detailed listing of occupations for a specific industry, and a questionnaire. During follow-up telephone calls, the interviewers received many comments indicating that the volume of information in the package made the survey task appear far worse than it really was—especially for small firms with little job turnover, where the owner or manager could accurately complete the questionnaire in a few minutes (BLS, 1991; Phipps et al., 1993).

6.1.2 Data Collection Instrument

The questionnaire or data collection form defines the task, through the content of the questions and the structure and layout of the form. The wording of the questions and instructions to respondents, their presentation on the page, and their sequence can all contribute to response errors. We measure response errors associated with the instrument by looking at the extent to which respondents follow directions and report data according to our definitions.

Question and Instruction Wording. It is the responsibility of a survey designer to ensure that each concept in the questionnaire or data collection instrument has a clear and unambiguous definition. If the survey designers start with ambiguously-defined concepts, respondents may report different data than what the researchers intended (Freedman, 1988).

Commonly-used procedures for specifying clear definitions to the respondent include providing a list of covered criteria or categories, and specifying a time or reference period. However, these procedures can themselves introduce response errors if inclusion criteria or reference periods are ambiguous. Goss et al. (1989) illustrate respondent difficulties with unclear reference periods and multiple reference periods.

Underlying the whole issue of question and instruction wording is the problem of language. Concepts, questions, and instructions must be described in words, which can have more than one meaning, and can be interpreted in different ways by different respondents (Dippo and Herrmann, 1991; Groves, 1989). Questions must convey both concepts and a frame of reference (Kalton and Schuman, 1982). Establishment surveys often require the use of technical language, but respondents may not understand key terms (e.g., Cox et al., 1989). Instructions must convey to respondents the actions they are to take in providing answers to questions. Shortcomings in any of these areas may result in misunderstandings or misinterpretations and lead to response errors.

Insights from RAS Studies. All of the BLS RAS studies looked at data quality, which we define operationally as respondent compliance with concepts and definitions. Overall, the studies have found data quality to be high. However, each study has at least a few examples of response errors because respondents misunderstood or did not comply with instructions.

We identified four types of problems in the RAS studies that derive from question and instruction wording: understanding concepts and terminology, adherence to definitions, use of correct reference periods, and incomplete instructions. These categories are not totally independent of each other. For example, a respondent whose understanding of a concept differs from that of the questionnaire designer may not report according to the specified definition.

One of the more interesting examples of misunderstanding a term comes from NWCS. Some respondents thought that "cash payments" meant payments in currency, rather than checks. Another example, from HWS, is of respondents who treated paid leave as the amount an employee was entitled to use rather that the amount actually taken.

The OESWP and HWS RAS studies both offer examples of nonadherence to definitions. Some respondents erroneously omitted overtime from wage payments or reported work hours. A number of HWS respondents also incorrectly included managers and executives in data they were reporting for production workers. While these errors may result from inadequate records, we have some evidence of misunderstanding instructions on the subject.

Question and Instruction Layout and Format. There is a large and growing body of evidence, most obtained from cognitive studies, suggesting that the format of a self-administered questionnaire can affect the data reported on the form (e.g., Gower and Dibbs, 1989; Gower and Nargundkar, 1991; Dillman et al., 1993).[3] Apart from the general appearance of the form, the location of instructions relative to questions may contribute to their being noticed, read, or ignored; respondents tend not to refer to separate instructions unless they need them (Gower and Dibbs, 1989; Gower and Nargundkar, 1991). Pretest respondents to the U.S.

[3] A recent paper by Sanchez (1992) demonstrates that the layout of interviewer-administered questionnaires can also affect data quality.

Census of Construction Industries relied on section headings for instructions, which resulted in errors when details were not part of the question heading (DeMaio and Jenkins, 1991).

Another aspect of question placement is context. The position of a question in a questionnaire relative to other questions can influence the responses to questions, and therefore introduce response errors, because of the cognitive activity aroused by prior questions. In an interview, prior questions can lead a respondent to interpret a subsequent question differently, so that the response is affected by the earlier question. If the questionnaire is self-administered, respondents can examine the entire form prior to answering any questions, so the presence of *later* questions may also create a context influencing responses to earlier questions (Sudman and Bradburn, 1982).

Since establishment surveys generally obtain factual data, is it reasonable to look for effects from the sequence of the questions? While the answer appears to be yes, the order effects may be a function of other question characteristics such as sensitivity. Cox et al. (1989) noted that some respondents were unwilling to report aggregate debts or cash holdings in a balance sheet, although these respondents had already provided dollar amounts of individual debts.

Insights from RAS studies. The OESWP surveys also show that respondents do not use separate sets of instructions unless they have to. An integral part of the OES is a booklet listing detailed occupations for the industry being surveyed, which participants are to consult to place their employees into occupations. Seventy percent of RAS respondents in 1989 and 67 percent in 1990 said they used the separate instructions. Of this group, 55 and 57 percent used them only when needed (Phipps 1990, 1991).

The Nonwage Cash Supplement RAS tested the effect of question placement on respondent comprehension. This survey used two different questionnaire designs, a "short form" that requested payroll totals near the top, followed by questions on types of payments, and a "long form" that began with questions on types of payments, and requested totals below the types of payment questions. We found that 82 percent of respondents who received the short form, but only 72 percent of respondents who received the long form, provided the requested payroll totals. Many long form RAS respondents indicated that they had overlooked the items (Phipps, 1989).

The ETJO RAS also provides evidence of the importance of questionnaire layout for data content. Respondents were to include laid off employees in the category of job separations, unless those layoffs were temporary. They were also to include transfers as separations and new hires. The RAS shows that respondents were reasonably successful with the former activity, which we attribute to the presence of a column heading asking for "Number of Job Separations (exclude temporary layoffs)." They did less well with internal transfers, for which there was no reminder in the column heading (BLS, 1991).

The OESWP surveys gave us an opportunity to modify questionnaire formats and evaluate the effect of those changes through RAS studies. In the 1989 questionnaire, only 56 percent of respondents correctly excluded premium payments such as overtime from reported wage data. Instructions for what to include and exclude appeared in paragraph form and were relatively easy to overlook. We changed the instructions in the 1990 questionnaire to lists of what to include and what to exclude, and the percentage reporting correctly increased to 78. Further inquiry in the 1990 RAS suggests that records were more of a problem than respondents overlooking or misunderstanding the instructions (Phipps, 1991).

6.1.3 Confidentiality and Sensitivity

An establishment may have explicit or implicit policies concerning the information it reveals about itself (Edwards and Cantor, 1991), particularly if it considers that information confidential or of value to competitors. If respondents feel that answering questions about their establishment is somehow threatening, they may withhold important information. However, none of the RAS studies directly addressed the issue of confidentiality or sensitive information.

6.1.4 Mode of administration

Most of the research looking at mode effects in surveys has been conducted for household surveys, and has focused on personal versus telephone interviews. Mail surveys have only recently been included. Dillman (1991) reviewed mail and telephone surveys conducted in the 1980s and concluded that some differences exist between responses to questions asked using both modes. Bradburn (1983) suggests that we should expect mode effects for sensitive topics, and that anonymous methods such as mail questionnaires should obtain more accurate reports of behavior. However, these conclusions are based on household surveys of attitudes and behaviors and may not apply to establishment research that obtains factual data.

Insights from RAS studies. Mode effects are a concern in establishment surveys because standard survey practice often includes a telephone follow-up to mail survey nonrespondents. Only one of the RAS studies reported here conducted a systematic analysis of mode effects. ETJO contained an experiment in which we attempted to collect data from half the units in the smallest size group by mail and half by means of CATI. All establishments received a survey package, but the CATI group got a letter informing the firm that an interviewer would call in about a week to collect data over the phone. The intended experiment was confounded somewhat because 19 percent of the CATI-designated sample mailed in their questionnaires, and 44 percent of the mail-designated sample responded by telephone, mainly during follow-up (BLS, 1991).

The HWS RAS uncovered substantial differences between mail and CATI respondents in the information they reported and in the sources of data they used. We also identified industry and employment structure

differences between mail and CATI establishments (Goldenberg, 1993).

6.1.5 Response Burden

Response burden is a component of any survey task, but it takes on additional magnitude for establishment surveys. The chore of completing a questionnaire is imposed upon other work responsibilities, and is especially heavy for owners or managers of small firms. Survey designers must recognize that there is a real cost for establishment personnel to participate in surveys, especially if the participants are executives or senior managers (Gower and Nargundkar, 1991). The burden includes both number of questions, the effort required to research or manipulate data sources, and that involved in preparing information in the needed format. While it is often measured as the amount of time needed to complete a questionnaire, it can contribute to response errors through the information requested and the ease with which that information can be extracted from company records. A high response burden also means more opportunities for errors.

Insights from RAS studies. The NWCS RAS looked at response burden by asking respondents if they spent more time reading and understanding the directions for the survey or preparing the data. Sixty-two percent said the directions took more time, 21 percent said preparing data took longer, and 17 percent said they were about equal. When we probed to see whether particular items on this form took longer than others, we learned that compiling annual payroll totals and breaking out payrolls for production workers were something of a problem for respondents to report. While the non-wage cash payments caused some difficulty, it was less than that created by the need to separate production workers from all workers.

Another way to look at response burden is to consider the number of types of records a respondent must consult to complete a questionnaire. The 1990 OESWP RAS, the ETJO RAS, and HWS all asked respondents whether they had used any of these sources of information: their memory, personnel records, payroll records, or something else. Respondents could say yes or no to each source of data.

In the 1989 OESWP, three-fourths of respondents indicated that they used one source to prepare the questionnaire. This finding suggests that the addition of wage information to the OES may not have created much additional burden. The comparable figure for 1990 respondents was a much-lower 39 percent who used one source for the data. Only half of that group consulted payroll records as their source, while 46 percent (generally respondents from small establishments) cited memory as their sole information source. The difference between the two years may reflect different record-keeping practices in the service industries that were the subject of the 1990 OES as compared with the manufacturing industries in the 1989 survey.

Perhaps, the number of sources consulted is not an adequate measure of response burden. In the HWS, approximately two-thirds of respondents indicated that they consulted a single data source for hours paid or hours at work information. Of that group, just under three-fourths looked at payroll records, and the remainder cited memory as their only source of information. However, time to complete the survey was extremely high, suggesting that the content and structure of the information source plays a major role as well.

We might note that if a respondent needs to examine both payroll and personnel records, it increases the likelihood of involving more than one person in the response process. If the firm is a large one, payroll and personnel may be in two separate departments guided by different policies. Additional respondents usually bring additional sources of response error through different interpretations of questions and different knowledge of the subject matter.

6.2 Response Errors Associated with the Information System

The external records or information source from which the respondent obtains data for the questionnaire comprises another source of response errors. We call this source the *information system*, and it consists of the collection, processing, and quality control procedures an establishment invokes for the information it maintains (Edwards and Cantor, 1991). An information system can be any set of records, manual or automated. While use of external record sources is not truly unique to establishment surveys (e.g., Kominski, 1991), the need for accessing records is far more widespread in establishment surveys than in household surveys (Edwards and Cantor, 1991). Variables associated with the information system that may influence response errors include the content of data in the system, the way those data are structured or aggregated, and a user's ability to access the system and retrieve data from it.

6.2.1 Data Content and Structure

An establishment questionnaire solicits specific types of data. The most fundamental type of response error arises if the data requested on the questionnaire are not available at the establishment, because then the respondent cannot answer the question.

Time lag is another aspect of data content. While payroll records are as current as the last completed payroll, other information may not be updated quite as often. The more current the information in the system, the more accurate the retrieved data. Compiling recent information that has not yet entered the information system adds to the response burden and increases the potential for transcription errors.

At the same time, the way data are structured or organized has significant implications for an establishment survey. The questionnaire seeks data aggregated or broken out according to very specific rules, and respondents who are asked to report familiar information but in an unfamiliar way may make errors in doing so (Freedman, 1988).

Imposition of survey definitions on existing information systems creates numerous difficulties for data collection. When Cox et al. (1989) surveyed small businesses about their finances, they found that for

proprietorships, the proprietor's personal and business finances were not normally separated—that is, the business data were not readily available. They usually could be separated, but not easily. The same study pointed to inconsistent accounting practices across firms, which conflicted with the need to collect consistent financial data across sample units.

Another characteristic of the information system that relates to both data content and data structure is the extent to which the information system has been automated. In an era marked by the widespread use of personal computers, and at a time when businesses increasingly turn to service bureaus to handle their data requirements, it is hard to imagine a firm without access to automated records. However, the content and structure of an automated information system may be less amenable to manipulation by individual establishments than the contents of an equivalent manual system, especially if an establishment works within the confines of purchased record-keeping software or contracts for a fixed set of services from an outside organization.

Insights from RAS studies. The Hours at Work Survey RAS specifically addressed the issue of records maintained by establishments. The questionnaire asks for two sets of data, hours paid and hours at work (i.e., hours paid minus paid leave) for all production or nonsupervisory workers in the establishment, by quarter and for the entire year. In the RAS, we asked whether the establishment keeps hours data for individual employees, and found that just over four-fifths do. Roughly two-fifths of that group also produce quarterly and/or annual summaries for each worker, with manufacturing establishments (production workers) somewhat more likely to summarize the information than nonmanufacturing (nonsupervisory) firms.

Since one component of hours at work information is hours of paid leave, we asked if respondents maintained records of hours at work, paid leave, or both. While 57 percent of the units have both types of records, 21 percent only have paid leave, 16 percent only have hours at work, and 6 percent have neither (Goldenberg, 1993). It is unclear how the respondents whose firms maintain neither hours paid nor paid leave information completed the HWS forms.

The issue of data availability was also a subject in the NWCS. The survey asked for payroll totals and nonwage cash payment totals for all employees and for production workers. Respondents who did not provide the information frequently indicated that they could not obtain the data from their records. Of the 159 respondents who did not enter production worker payroll totals on the form, 13 percent said it was because they had no separate records for production workers. An additional 18 percent gave reasons related to nonavailability of the data from their records. Similarly, 41 percent of respondents who did not report nonwage cash payments omitted them because the data were not present in their records (Phipps, 1989).

6.2.2 Access and Data Retrieval

While the establishment's information system must contain the data required for a survey, the person responsible for completing the survey must have knowledge of and access to that information. Equally important, the person must understand the relationship between the questions in the questionnaire and the content of the information system (Gower and Nargundkar, 1991). In a small business, the person compiling the information may be an owner or manager, but in a larger organization the responsibility may be assigned to an accountant or to a clerk. The individual must have the authority to use the system, or lacking that, enough knowledge to explain to a third party exactly what information is needed and how to get it.

In addition, an establishment needs a means of retrieving those data, especially if the information system contents do not conform to the definitions specified by the questionnaire. Respondents from small or independent establishments, or firms that maintain manual records, should be able to complete the questionnaire, if the data exist. However, an establishment may be a component of a larger corporate organization that handles all data processing. It may use a service bureau to maintain employee records, or maintain records in-house with a software package that has limited retrieval capabilities. In any of these situations, the respondent may provide data that are easy to obtain, regardless of definition.

Insights from RAS studies. The RAS results demonstrate that an establishment may have "computerized" data, but not the data requested in a survey. For example, in the 1990 OESWP RAS 39 percent of respondents said they used a computer listing to fill in the questionnaire, but nearly four-fifths of them (about a third of all respondents) still had to manually calculate the number of employees for each wage group (Phipps, 1991).

The HWS RAS looked more directly at data retrieval. We asked if the firm produced summary reports of hours data for its own purposes. We found that 72 percent of responding establishments did so, but just over two-fifths of them had a report limited to production or nonsupervisory workers. We also learned that the availability of this summary report has a direct effect on data quality. Establishments with summaries limited to production or nonsupervisory workers were more likely to conform to survey definitions. Manufacturing establishments were more likely to have a report limited to production workers (49 percent) than nonmanufacturing industries were to have a report limited to nonsupervisory workers (34 percent).

These studies tell us that having information in a company's records, and having information available in the format needed for a survey, are two entirely different matters. If the unit already compiles information as needed for a survey, overall data quality will be higher and response burden will be lower.

6.3 Response Errors Associated with the Respondent

The respondent is a third source of errors in survey data. In the discussion that follows, we present several components of response error that we attribute to the respondent and that are relevant to establishment surveys. These sources of error include memory and the respondent's use (or non-use) of the information system. These error components are not necessarily independent of error components from the task or from the information system.[4]

6.3.1 Memory

Memory plays a large role in determining the accuracy of a respondent's reporting. Most of the literature on memory errors in survey responses is based on household surveys, but refers to the recall of factual information and is therefore relevant to establishment surveys. Bradburn (1983) identifies two types of memory errors: forgetting, or omission of events (results in underreporting), and telescoping, or recalling events as occurring later in time (results in overreporting).

Memory is relevant to establishment survey respondents in several ways. First, most establishment survey questionnaires have at least a few questions that can be answered without consulting records. (For example, the Hours at Work questionnaire asks for the types of paid leave the firm offers.) Second, some questions require that a respondent remember events or types of activities. Third, a respondent must use memory to determine whether an information element is available in establishment records, and if so, how to access it. Finally, some respondents will use memory instead of the information system as their primary source of data (Section 6.3.2).

Insights from RAS studies. The NWCS RAS specifically addresses memory errors. Most were "errors of omission," i.e., payments that should have been reported on the survey form but were not. We asked all RAS respondents if their establishments gave Christmas or executive bonuses, or merit, incentive or employee recognition awards during the preceding calendar year. If the establishment gave any of the payments, we asked whether the respondent had included them in the figures they reported on the survey form. If not, we asked the reasons for any omissions. The questioning revealed that from one-third to one-half of establishments making specific types of payments failed to report them. There were a total of 106 unreported payments, half of them attributable to 22 respondents. While the largest share of underreporting was due to the unavailability or limitations of records, 15 of the 106 omissions were due to reporters simply forgetting the payment. When we offered cues to aid recall, we found that respondents remembered not only the specific types of payments in the examples, but also others in the category we asked about (Phipps, 1989).

[4] In fact, Bradburn (1983) considers memory a task error variable rather than an respondent variable.

6.3.2 Use (or Non-Use) of the Information System

A respondent can choose to obtain data for an establishment survey from an information system—or the respondent can report from memory. In most cases, this is the difference between providing accurate data (within the constraints of the respondent's comprehension of the question and the limits of the data available) and providing estimates. Bradburn (1983) notes that there is evidence to support the "commonsense belief" that using records will increase the accuracy of information reported on a survey. However, he adds that records are not a panacea, and they do not totally eliminate response errors. This should be clear from the discussion of other sources of error in Sections 6.1 and 6.2.

Even if respondents extract information from establishment records, they may differ in their attention to the detail they enter on a questionnaire. Cox et al. (1989) observed that some respondents looked at records with exact figures and reported rounded data. When they surveyed small businesses about their finances, these researchers found that respondents were willing to indicate the presence of certain financial products (e.g., types of loans), but that the reported dollar amounts were often estimates (Cox et al., 1989). Gower and Nargundkar (1991) observed that higher level officials in a business were less likely to be concerned with the details that lead to quality data (e.g., differences in the questionnaire relative to data in their information system), and to make estimates instead. They found that office managers, accountants, and comptrollers were more concerned with detail and with the accuracy of the information they were compiling, turning to multiple sources as needed.

Insights from RAS studies. The RAS studies all included questions about the sources of data respondents used for information about specific topics, and several also asked about the use of estimated data. We will focus here on memory as an information source, and the extent to which it was *the* information source for certain data elements. Respondents who provide data from memory (or based on their knowledge of an establishment) are more prone to make response errors as a result of forgetting, telescoping, or making incorrect estimates.

Three of the five RAS studies specifically included memory as a possible source of information, along with payroll and personnel records. In most cases, memory was one of the sources used--not unreasonable for certain types of data. Forty-nine percent of respondents to the 1990 OESWP cited memory as a data source, as did 30 percent of respondents to the HWS. In the ETJO survey, 22 to 24 percent of respondents cited memory as one source of information on employee separations, new hires, and job openings, while the vast majority of respondents in establishments with no employment turnover (79 percent) indicated that memory was one of their data sources.

In all of these studies, the number of respondents whose *only* source of information was memory is relatively small, and they are concentrated in small estab-

lishments. Almost two-thirds of the memory-only respondents in the 1990 OESWP were in firms with fewer than 10 employees, as were 77 percent of memory-only ETJO respondents whose firms had no employment turnover activity. Both of these surveys ask for data covering a relatively short and current time period, which would help to minimize both telescoping and forgetting errors. The small size of the firms also means there are fewer events for a respondent to try to recall, which should contribute to greater accuracy.

A striking finding from the studies with both mail and telephone components is that memory-only respondents are overwhelmingly respondents to telephone nonresponse follow-ups, while respondents who use records are more likely to respond by mail. The telephone interview may, by its very nature, discourage respondents from checking records. For example, the ETJO survey had 16 respondents (of 144) who indicated that memory was their only source of data. Thirteen of the 16 were telephone respondents.

In sum, our experience with respondents who use memory rather than the information system is that they have only a limited impact on overall data quality. Most come from small establishments, and their numbers overall are relatively small. The influence of memory—i.e., the non-use of records—seems to be stronger for telephone surveys, although we must be extremely cautious about drawing conclusions. These data suggest that telephone surveys are less than ideal for collecting complex data, at least on a one-time basis. The situation is somewhat different in the case of surveys that return to the same respondents regularly.

7.0 Discussion

Response Analysis Surveys are a useful tool for identifying and evaluating data quality. They provide a source of quantitative information about respondent behavior vis-a-vis a questionnaire, and they obtain reasons for that behavior. A RAS is particularly valuable when used in conjunction with cognitively-based questionnaire-design techniques, such as focus group discussions and think-aloud pretest interviews. It offers a way to examine issues identified as problem areas during form development, and to see if the resolutions to those issues were effective. They also provide a means of identifying other, previously unknown, response problems. If the survey is an ongoing one, a RAS provides a feedback mechanism and a framework for testing study modifications and future improvements.

The largest group of errors identified in the BLS RAS studies are errors associated with the data collection instrument, and specifically with the wording and layout of questions and instructions. We found problems with understanding concepts or terminology, adherence to definitions, use of reference periods, and missing or unclear instructions. Most of these errors are, at least in theory, *controllable*—and not unique to establishment surveys. They result from the way we ask for the data. Eliminating comprehension problems is no small matter, and may require extensive research to ensure that questions addressed to a specific audience can be understood by that audience as the researcher intended them to be understood. Part of this process may involve recognizing multiple audiences, and designing data collection procedures accordingly.

To a certain extent, we can include memory errors among the controllable errors. Questionnaire designers can incorporate cues to improve recall; the cues have the added benefit of improving comprehension as well as memory (Dippo and Herrmann, 1991). The memory errors identified in the NWCS RAS might have been avoided had we provided examples of the bonuses and awards of interest on the questionnaire.

Errors that survey designers can *not* totally control result from the structure, content, and retrieval capabilities of respondents' information systems. We found numerous examples in the RAS studies of reported information that does not meet our definitions because of the way the records are maintained, and of information that requires significant additional compilation to conform to the questionnaire. Even here, however, there are steps researchers can take to minimize response errors. Survey designers can consult with potential respondents while designing the questionnaire, to find out how respondents might maintain their data. We can request data in a form that respondents are able to provide, and we can ask respondents to identify data that may not meet our definitions.

As researchers, we may have to compromise as we define survey concepts. While definitions must be precise, the RAS studies indicate that the research community may not always (or even often) get exactly what it asks for. What we can do is to accept high-quality data that is *almost* what we want. The alternative—the current situation?—may be to obtain exactly what we want, but with low response rates, or to ask for exactly what we want and accept what we get, with or without an understanding of the data we are *really* being given. In either case, response analysis surveys can help us to understand the data.

References

Bradburn, N. M. 1983. "Response Effects." In P. H. Rossi et al. (eds)., *Handbook of Survey Research.* New York: Academic Press: 289-328.

Bureau of Labor Statistics (BLS). 1991. *Employee Turnover and Job Openings Survey: Results of a Pilot Study on the Feasibility of Collecting Measures of Imbalances of Supply and Demand for labor in an Establishment Survey.* Washington, D.C.

Cox, B., G. Elliehausen, and J. Wolken. 1989. *The National Survey of Small Business Finances: Final Methodology Report.* RTI/4131-00F. Research Triangle Park, NC.: Research Triangle Institute.

DeMaio, T., ed. 1983. *Approaches to Developing Questionnaires.* Statistical Policy Working Paper 10. Federal Committee on Statistical Methodology, U.S. Office of Management and Budget.

DeMaio, T. and C. Jenkins, 1991. "Questionnaire Research in the Census of Construction Industries." *Proceedings of the Section on Survey Research Methods.* Amer. Statistical Assn.: 496-501.

Dillman, D. A. 1978. *Mail and Telephone Surveys: The Total Design Method.* New York: Wiley.

Dillman, D. A. 1991. "The Design and Administration of Mail Surveys." *Annual Review of Sociology* 17: 225-249.

Dillman, D., M. Sinclair, and J. Clark. 1993. "The 1992 Simplified Questionnaire Test: Effects of Questionnaire Length, Respondent-Friendly Design, and Request for Social Security Numbers on Completion Rates." Presented at the 1993 Annual Research Conference, U.S. Bureau of the Census.

Dippo, C. and D. Herrmann. 1991. "The Bureau of Labor Statistics' Collection Procedures Research Laboratory: Accomplishments and Future Directions." In Federal Committee on Statistical Methodology, *Seminar on Quality of Federal Data. Part 2.* Statistical Policy Working Paper 20: 253-267.

Edwards, W. S. and D. Cantor. 1991. "Toward a Response Model in Establishment Surveys." In P. Biemer et al. (eds.), *Measurement Errors in Surveys.* New York: Wiley: 211-233.

Federal Committee on Statistical Methodology. 1988. *Quality in Establishment Surveys.* Statistical Policy Working Paper 15. Statistical Policy Office, U.S. Office of Management and Budget.

Fecso, R. and B. Pafford. 1988. "Response Errors in Establishment Surveys with an Example from an Agribusiness Survey." *Proceedings of the Section on Survey Research Methods.* Amer. Statistical Assn.: 315-319.

Freedman, S., "Control and Measurement of Specification Error in Establishment Surveys." *Proceedings of the Section on Survey Research Methods.* Amer. Statistical Assn.: 304-308.

Goldenberg, K. L. 1993. "Results of the Response Analysis Survey (RAS) for the Hours at Work Survey (HWS)." Internal report, Bureau of Labor Statistics.

Goss, Gilroy, & Associates. 1989. *Qualitative Research to Evaluate the Questionnaire for the Survey of Employment, Payrolls and Hours (SEPH).* Final Report prepared for Statistics Canada.

Gower, A. and R. Dibbs. 1989. "Cognitive Research: Designing a 'Respondent Friendly' Questionnaire for the 1991 Census." *Proceedings of the Fifth Annual Research Conference.* U.S. Bureau of the Census: 257-266.

Gower, A. and M. Nargundkar. 1991. "Cognitive Aspects of Questionnaire Design: Business Surveys Versus Household Surveys." *Proceedings of the Annual Research Conference.* U.S. Bureau of the Census: 299-312.

Groves, R. M. 1989. *Survey Errors and Survey Costs.* New York: Wiley.

Kalton, G. and H. Schuman. 1982. "The Effect of the Question on Survey Responses: A Review." *Journal of the Royal Statistical Society, Series A* (145, Part 1): 42-73.

Kominski, R. 1991. "Record Use by Respondents." *Proceedings of the Section on Survey Research Methods.* Amer. Statistical Assn.: 402-406.

Phipps, P. 1989. "Non-Wage Cash Supplement Response Analysis Survey." Internal Memorandum, Bureau of Labor Statistics, November 1989.

Phipps, P. 1990. "Occupational Employment-Wage Pilot Response Analysis Survey." Internal Memorandum, Bureau of Labor Statistics, January 1990.

Phipps, P. 1991. "Response Analysis Survey for 1990 OES Wage Pilot Survey." Internal Memorandum, Bureau of Labor Statistics, January 1991.

Phipps, P., S. Butani, and Y. Chun. 1993. "Designing Establishment Survey Questionnaires." *BLS Statistical Note 35.* Washington, D.C.: Bureau of Labor Statistics.

Sanchez, M. E. 1992. "Effects of Questionnaire Design on the Quality of Survey Data." *Public Opinion Quarterly* (56): 206-217.

Schwarz, N. and H.-J. Hippler. 1991. "Response Alternatives: The Impact of their Choice and Presentation Order." In P. Biemer et al.(eds.), *Measurement Errors in Surveys.* New York: Wiley: 41-56.

Scott, Stuart. 1983. "Availability of Employee Hours Data." *Proceedings of the Section on Survey Research Methods.* Amer. Statistical Assn.: 461-465.

Sudman, S. and N. Bradburn. 1974. *Response Effects in Surveys.* Chicago: Aldine.

Sudman, S. and N. Bradburn. 1982. *Asking Questions.* San Francisco: Jossey-Bass.

Tourangeau, R. 1984. "Cognitive Sciences and Survey Methods." In T.B. Jabine et at. (eds.), *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines.* Washington, D.C.: National Academy Press: 73-101.

USING FOCUS GROUPS TO IDENTIFY USER NEEDS AND DATA AVAILABILITY

Lynda T. Carlson, John L. Preston, and Dwight K. French
Energy Information Administration
Lynda T. Carlson, Energy Information Administration, EI-63, Washington, DC 20585

KEY WORDS: focus groups, data users, data providers, questionnaire development, issue identification

Introduction

Any organization with primary data collection responsibilities must balance the interests of two groups, the data users and the survey respondents. Data users frequently need more information, different types of data, or additional detail in order to address current issues. On the other hand, survey respondents often resist, citing increased burden or the unavailability of the requested data.

The Energy Information Administration (EIA) recently undertook a major expansion of a national survey of energy consumption by the manufacturing sector. This paper describes how the EIA solved dilemma of conflicting interests by using focus groups to obtain input from data users and data providers and to balance their conflicting interests.

Survey Responsibilities of the Energy Information Administration

The EIA is the independent statistical and analytical agency of the U.S. Department of Energy (DOE). The EIA engages in the whole spectrum of energy-related data collection, compilation, and publication. It produces data, analyses, and forecasts that are used by the program offices of the DOE, other Federal agencies, the Congress, university researchers, private-sector analysts, and others. Various organizations within the EIA conduct data surveys. Generally, these surveys are one of two types, supplier surveys and end user surveys.

Supplier Surveys

Supplier surveys are taken from a population of producers and marketers of the energy sources such as petroleum, natural gas, coal, liquefied petroleum gases (LPG), and electricity. These surveys collect data concerning production, sales, stock changes, and price of specific energy sources. The results of supplier surveys are tabulated and published, at aggregate levels, in EIA publications including, *Natural Gas Monthly, Petroleum Supply Monthly, Electric Power Monthly, Quarterly Coal Report,* and others.

The supplier surveys also are aggregated into broad groupings of types of customers—residential, commercial, industrial, transportation, electric utilities—based on rate classifications. These estimates are published by the Energy End Use and Integrated Statistics Division (EEUISD) in the *Monthly Energy Review* and the *State Energy Data Report* and provide end-use estimates for those customer classifications. The estimates for the industrial customers include manufacturing, mining, construction, agriculture, and forestry and fisheries.

End User Surveys

End user surveys gather data directly from consumers of the energy products. These surveys query a representative sample of the four major end-use sectors: residential, commercial buildings, manufacturing, and residential transportation. The information collected by the end user surveys includes the amounts of the various fuels consumed, the types of appliances or machines in use on the site, and characteristics of the consuming unit.

The EEUISD conducts four end user surveys: the Manufacturing Energy Consumption Survey (MECS), the Commercial Building Energy Consumption Survey (CBECS), the Residential Energy Consumption Survey (RECS), and the Residential Transportation Energy Consumption Survey (RTECS).

The EIA Evaluates Its Survey Operations

The EIA has turned a great deal of attention to the issues of improving the quality, timeliness, and effectiveness of its data-gathering operations. The

DOE's recent development of the National Energy Strategy (NES) required an in-depth evaluation of each aspect of the EIA's data collection efforts.

The EIA formed working groups of experts in energy modeling, projection, and estimation to conduct the evaluation and recommend improvements in existing systems. Working groups were responsible for all data analysis and forecasting activities needed to support the DOE's NES teams. The EIA formed one working group for each major end-use sector, defined according to the customer classifications (industrial, residential, etc.).

The working groups, after an exhaustive, iterative process of agency-wide cooperation, completed the development of a set of reference case forecasts for each major end-use sector. They also developed excursions from these baseline estimates, reflecting different penetration of various new technologies, and answering a series of "what if" questions to assess the effects of policy initiatives that differ from those represented in the reference case.

The analysts from the working groups pointed to several weaknesses in the surveys that the EIA was using then. The data collected from the industrial sector, in particular, were deficient from the data users' standpoint. The MECS is the main vehicle that the EIA uses to obtain information on energy consumption patterns in manufacturing, which accounts for 70 to 75 percent of industrial energy consumption. Excursions developed for the industrial sector were difficult to construct due to a lack of quantifiable information concerning the penetration of technology innovations and the differing uses of energy in various manufacturing industries. This prevented the analysts from developing technology-specific excursions. Instead, they employed a simple, highly aggregated econometric approach. That approach used price as the principal driver for choice of a specific fuel and production levels and energy intensities as the drivers for consumption. The NES models would be stronger if they contained comprehensive industry-level data. The EEUISD saw this as an opportunity to improve the quality and effectiveness of one of its major surveys—the MECS.

The Use of Energy in the Manufacturing Sector

Energy use in the manufacturing sector takes many forms and is highly complex. Manufacturers consume large quantities of energy sources such as natural gas, fuel oil, liquefied petroleum gases, coal, and electricity. The main use of energy by manufacturers is to provide heat and power for their manufacturing operations. Energy sources are also consumed as a raw material input to the manufacturing process. Relatively small quantities of energy are used to provide heating, cooling, and lighting to their physical plants.

Manufacturers buy much of their energy from outside vendors, but also can use energy that is produced as a byproduct from their manufacturing processes. Many manufacturers produce some of their own electricity onsite through cogeneration. Some of these cogenerating facilities sell or transfer surplus electricity directly to other manufacturers or to the local power grid.

Another factor that complicates tracking manufacturing energy consumption is that many manufacturers can substitute one energy source for another through fuel switching. If the relative prices of fuel sources change, the plant manager may increase the percentage use of the least expensive type of fuel. The MECS is specifically designed to produce energy consumption estimates for this complex sector.

The Background of MECS

In 1986, the Omnibus Budget Reconciliation Act, P.L. 99-509, required EIA to conduct triennial surveys of energy consumption in the manufacturing sector. The MECS, as required by Congress, must gather data on energy expenditures, fuel-switching capability, onsite generation of electricity, and byproduct energy use. The Energy Policy Act of 1992, P.L. 102-486, required that the MECS be conducted in at least a biennial basis. The MECS will become a biennial survey beginning in 1994.

The first two MECS provided data for the years 1985 and 1988 and had the same basic design, frame, and content. Those surveys also included questions on price of energy inputs and storage capacity for petroleum products.

In addition to providing input to EIA's forecasting models and DOE's policy analysis operations, the results of the MECS are used to produce several standard publications and numerous articles covering manufacturing energy consumption and related issues. The standard MECS publications are, *Manu-*

facturing Energy Consumption Survey: Consumption of Energy, 1988; Changes in Energy Intensity in the Manufacturing Sector, 1980-1988; and *Manufacturing Fuel-Switching Capability, 1988.* These publications, and companion publications for 1985, are available from the Superintendent of Documents. Similar publications for 1991 will be available beginning in 1994. The 1991 publications will include the information on the new issues identified in this paper.

The Industry Division of the Bureau of the Census serves as the data collection and compiling agent for the MECS. There are four major benefits resulting from this arrangement:

- The Industry Division is responsible for conducting the Annual Survey of Manufactures (ASM). The ASM sample serves as the sampling frame for the MECS and greatly facilitates the MECS sampling procedure.

- Under the provisions of Title 13, *U.S. Code of Federal Regulations*, respondents' confidentiality is guaranteed; only sworn Census Bureau employees may have access to the reports; and copies of the questionnaires retained by the respondent are immune from legal process. Confidentiality is a major concern of respondents.

- The EIA uses certain economic measures—value added, value of shipments, number of employees, and so on—in its analyses of manufacturing energy consumption. The ASM sample collects those economic data, and appends them directly to the MECS respondent records.

- Between 1974 and 1981, the Industry Division collected information on the consumption of purchased fuels and electric energy in a supplement to the ASM. Analysts can use these historical data, with the 1985 and 1988 MECS, to construct longitudinal analyses.

The Census Bureau updated its list of manufacturing establishments after its 1987 Census of Manufactures. That update coincided with the internal review of the MECS and the development of the NES. The EEUISD decided to take this opportunity to reexamine thoroughly the survey and set dual goals of maximizing the quality of data collected and minimizing unnecessary burden on the respondent. The EEUISD also felt that all the concerned parties—the data users, the data providers, and the survey staff—needed to provide input. To facilitate this dialogue, the EEUISD decided to employ a tool common in marketing research, the focus group.

The Use of Focus Groups in the MECS Redesign

During the original development of the 1985 MECS, the EEUISD staff requested and received extensive input from representatives of the manufacturing sector. Working closely with manufacturing trade associations, EEUISD staff visited over 20 major manufacturing facilities and corporate headquarters throughout the United States. The purpose of those visits was to learn first hand how manufacturers consume energy and to discuss energy-reporting concerns and problems. Those contacts were essential to the successful development of the 1985 MECS because they highlighted the major concerns of potential respondents and allowed EIA to address those concerns before conducting the survey. The use of focus groups to obtain input from data users and providers for the 1991 MECS was, essentially, a formalized extension of that original field work.

Focus groups are increasingly considered a cost-effective method of gauging public opinion regarding a new product or service. Organizations making use of focus groups range from advertising agencies to the Internal Revenue Service. Researchers often use focus groups to develop preliminary hypotheses that can be tested using formal, quantitative research methods.

Focus group research is strictly qualitative. A typical focus group contains 6 to 12 individuals who have an understanding of or opinion about the topic of interest. The group spends two to three hours in discussion of the issue. This entire session is generally audiotaped and frequently videotaped, allowing the researchers to examine the group's insights and observations in detail later. A moderator leads the focus group's activities and conversations. The moderator is often a trained psychologist or social scientist but can be a subject specialist. A successful moderator is one who keeps the discussion focussed on the subject area but does not guide the conversation or reveal personal opinions. The nature of the topic determines the specific type of moderator required; no single approach is always effective.

The EEUISD held two series of focus groups in the MECS redesign process. One series was conducted with data users from DOE, and the other, with data providers from the manufacturing sector. These focus groups were conducted between March 1990 through March 1991.

The focus groups were conducted in the traditional manner except that they were iterative in nature—the results of one group were used to refine the agenda for a subsequent group. That iterative approach was especially useful for sharing information between the data users and data providers[1].

Focus Groups Identify Data Users' Needs

The EEUISD staff held a series of 10 focus group discussions with analysts throughout the DOE and the EIA to crystallize the general impressions gained during the NES process into specific data requirements. These sessions allowed the full spectrum of MECS data users to share their opinions and communicate their individual data needs. As the series progressed, results from earlier sessions helped guide the participants. Inputs from the manufacturers' sessions, concerning feasibility and limitations, filtered into the later discussions.

These meetings were organized around the following general topics:

- **Level of Industry Detail**—The 1985 and 1988 MECS produced energy consumption estimates and related data for the 20 industry groups defined by 2-digit Standard Industrial Classification (SIC) codes. Data were also provided for the 10 industries (4-digit SIC) that generally consume the most energy. Did the data users want the MECS to expand its coverage of specific industries?

- **Specific Uses of Energy**—Both previous MECS collected information only on total energy consumption and consumption as a feedstock or raw material input. Should EEUISD revise the survey to collect more detailed end-use data?

- **Level of Geographic Breakdown**—The 1985 and 1988 MECS produced estimates for the four census regions—Northeast, Midwest, South, and West. Should geographic coverage be more detailed?

- **Data Quality**—Should the MECS collect only verifiable, quantifiable data, or are estimates acceptable in some areas?

- **Level of Energy Source Breakdown**—Previous MECS covered many energy sources, including byproduct energy sources such as coke oven gas, blast furnace gas, and pulping liquor. Is the coverage of specific types of energy sources sufficiently detailed? Should future MECS collect more information on how manufacturers use byproduct energy sources?

- **Frequency of Data Collection**—Should the MECS be conducted more or less frequently?

- **Need for Additional Economic Data**—Should the MECS include more economic data?

- **Need for Fuel-Switching Data**—The previous MECS reflected practical rather than technical switching capability. Should the MECS redefine fuel-switching capability?

- **Need for Environmental Data**—Should the MECS include this area of inquiry? Should the survey be linked to environmental surveys conducted by other Federal agencies such as the Environmental Protection Agency?

The data users felt that the geographic coverage and frequency of MECS were sufficient. Also, they felt that the definition of fuel-switching capability was appropriate. In addition, most data users felt that other Federal surveys provided sufficient coverage of economic and environmental data, so the MECS did not need to expand in these areas.

The data users' focus groups did, however, agree on some areas where the EEUISD could augment MECS. The users agreed that the MECS needed more detail in the following areas:

- **Level of Industry Detail**—Most of the groups expressed the desire to see the coverage of specific industries increased from the previous MECS.

[1]For a comprehensive discussion of the MECS focus group methodology and the findings, see Energy Information Administration, *Development of the 1991 Manufacturing Energy Consumption Survey*, DOE/EIA-0555(92)/2 (Washington, DC, May 18, 1992).

- **Specific Uses of Energy**—The users needed more information on how energy was used within the manufacturing plant.

- **Characteristics of Manufacturing Establishments**—The data users required more specific detail on the penetration of various advanced technologies into the manufacturing processes. What capital investments had the respondents made for energy conservation equipment or programs?

The EEUISD staff realized that they could achieve the required additional level of industry-specific detail relatively easily, by expanding the sample size. The other two areas of interest could require significant modifications to the survey and increase respondent burden. As the groups' progressed, it became clear that they would need feedback from the other forum, the data providers' roundtables.

The MECS Respondents Provide Input

While the data users' focus group series was continuing, EEUISD staff began the other series of discussions, those with the data providers. These focus groups, called "roundtables," would assess the possibilities for measuring energy use at the subestablishment or process level and address the future availability of data concerning energy conservation measures.

The roundtables featured representatives of seven energy-intensive industries: fertilizer, petroleum refining, steel, motor vehicles, pulp and paper, chlor-alkali, and olefins. Each roundtable included the industry experts such as plant managers, energy managers, and engineers. The roundtables were also attended by EEUISD staff responsible for the MECS, and were conducted by a professional moderator. The sessions lasted approximately four hours; the participants structured their discussion around the following specific objectives:

- To examine the current availability of manufacturing energy use data;

- To refine the EIA's understanding of energy use and energy-related decision making in manufacturing;

- To explain prior gains in energy efficiency in manufacturing; and

- To foster an appreciation of the factors that will drive future manufacturing energy consumption.

Members of the roundtable discussions received a packet of background documents prior to the meeting. The purpose of those documents was to ensure that all participants were familiar with the terminology, to provide the participants with materials to analyze the impact of new technologies and sources of efficiency gains, and materials to allocate energy consumption to end-uses. Each roundtable was audiotaped in its entirety and a written transcript and summary of the proceedings was prepared.

The use of the focus group format proved to be a uniquely valuable tool to tap into the knowledge and assess the opinions of the representatives of manufacturing. The EEUISD grouped its findings into three main areas: data availability, energy efficiency, and general comments concerning MECS. The manufacturers helped the EEUISD staff determine which desires of the data users they could meet easily, and which would place a greatly increased burden upon them.

The manufacturers collect data concerning energy usage for four main reasons: allocating energy costs among departments, advancing corporate goals such as cost minimization and quality control, determining return-on-investment and thus ranking capital expenditure projects, and meeting guidelines of government policy. The industry representatives generally view energy data collection from an economic perspective. The dramatic increases in energy costs during the 1970's made managers more aware of the ways in which they could control and limit energy use.

It also became clear that some uses of energy were easier to track than others. Meters that measure fuels purchased from outside vendors are the easiest and most accurate means to collect energy data. Larger, energy-intensive manufacturers used internal cost-accounting systems to partition costs to specific products. The roundtable participants discussed the manner in which they estimate the use of self-generated fuels. Roundtable members pointed out that government policy has played a major role in energy data collection. Specifically, the control and monitoring of emissions, as well as surveys like the MECS, have helped drive manufacturers to a heightened awareness of energy consumption.

While manufacturers routinely collected information on consumption of specific fuels and total energy consumption per unit output, few collected the data that the MECS users wanted most, energy consumption by end use. Still, most roundtable participants agreed that the plant or energy managers at their facilities could make reliable, intuitive estimates for these categories.

The manufacturers expressed some concerns about reporting energy use at the subestablishment level. These concerns were largely proprietary; most companies resist having to describe their production processes to that level of detail. The other concerns were technical. Many processes reuse energy from an earlier stage, or pass some on to the next. Tracking the use of steam in many plants is nearly impossible. Cogenerated power and fuels produced as byproducts are not straight-forward measurements either.

The 1991 MECS: Integrating Users' Needs and Respondents' Concerns

During the 12 months of focus group discussions, the EEUISD was able to continually update and redefine its ideas for improving the MECS. Insights provided by a data users' meeting could become part of the next roundtable's agenda. The respondents could discuss the merits and difficulties of the request and pass it back to the users through EEUISD. As a result, the 1991 MECS addresses several areas of particular concern to the data users, without overburdening the respondents. The EIA expanded both the methodology and the substance of the MECS questionnaire.

Methodological Changes

The methodological changes include changes to the size and timing of the MECS. The 1991 survey size was increased to 16,000 from 12,000 to allow for more specific detail by industry, while keeping relative standard errors at acceptable levels. The EIA mailed the 1991 survey to respondents earlier than the two previous editions. This would give them more time to respond, and to get the survey in their hands while the data were fresher. Also, the EEUISD redefined the measure of size for the MECS sample selection so that smaller manufacturers that used energy intensively would not be excluded. The MECS staff also made decisions in two cases to keep the survey unchanged. The survey would continue to be a subset of the Census Bureau's ASM and would continue with the same level of geographic coverage.

The EEUISD didn't need the focus groups to help make these methodological decisions. The MECS has several major new substantive changes that are the direct results of the focus group process. These changes deal with end-use breakdowns of energy consumption, building square footage, participation in energy management programs, and presence of new efficiency technologies.

The Collection of Energy End-Use Data

The data users' groups were unanimous in their call for more detail on the specific end uses of the energy consumed at manufacturers. They wanted a better picture of how the companies use different energy sources in producing their goods. However, the users split on how to best categorize the components of energy use. Some analysts suggested breaking consumption down by specific end-use. Others wanted to see the data by industrial process. Another group thought of allocating energy consumption by individual pieces or classes of equipment.

The roundtable participants initially rejected the idea of providing end-use energy consumption data because they did not normally collect data that way. However, most of the manufacturers agreed to the concept of estimating these data. They felt that most plant or energy managers would have a good intuitive sense of their consumption of different energy sources by various end uses. If they could express these values as a percent of total consumption of specific energy sources, which was already required by the MECS, the respondents felt they could provide numbers that would be sufficient for most analyses.

The data users agreed that such estimates would be an acceptable alternative to exact, measured values. The EIA proposed an energy end-use matrix to include in the 1991 MECS. This original matrix requested an estimated percentage allocation, by end-use of six energy sources: steam, electricity, coal, natural gas, distillate fuel oil including diesel fuel, LPG and natural gas liquids, and residual fuel oil. The EEUISD took this design back to the manufacturers at the next roundtable, seeking comments. The manufacturers felt steam would be too difficult

to break out due to the nature of steam use in plants. They also wanted it clearly stated that these were not to be complex engineering estimates, but only reasonable approximations. The EEUISD made these modifications, and developed the matrix shown in Figure 1.

Energy Efficiency

One of the primary goals of the NES was to examine potential improvements in energy efficiency that can be made by the various sectors of the U.S. economy. After developing the various forecasts that helped support the NES, many EIA analysts realized they required more information on the specific energy efficiency programs that the manufacturing sector employed. The following are examples of the specific types of data that the EIA analysts needed:

- Amount of capital invested in energy-efficiency projects, including estimated energy savings;
- Descriptions of energy-consuming equipment in manufacturing facilities, including remaining life expectancy and turnover rate;
- Participation in and the results of demand-side management (DSM) programs conducted by electric utilities;
- Characteristics of buildings used in the manufacturing sector;
- Types of equipment used to produce renewable energy and the specific uses of such on-site produced energy;
- Potential for converting processes using fossil fuels to electricity;
- Potential for using disposable waste products as fuel; and
- Reasons for switching fuels and the quantity of each fuel actually replaced.

The roundtable participants did not specifically address these requests, but many of the topics covered by the manufacturers had relevance to the general area of energy efficiency. First, the manufacturers made it clear that energy use was viewed as one of many production inputs, and that energy efficiency was justified chiefly by demonstrating economic benefits. The increases in energy costs of the 1970's and 1980's, along with expanded government emission monitoring requirements, have heightened awareness among managers to the complexity of energy use in their establishments. The introduction of computers for industrial process control has also allowed for improved tracking. Still, many manufacturers have not taken advantage of all the programs or equipment available to increase energy efficiency. The manufacturers cited shortage of capital as one major barrier to the installation of energy efficiency programs or projects.

Projects designed to improve energy efficiency must compete with all other capital expenditures for a company's fixed capital resources. Firms rank the desirability of capital investments. The highest priority is assigned to obligatory projects such as fulfilling environmental regulations. The second level typically includes non-deferrable items such as market opportunities or plant maintenance. Energy efficiency projects, which are generally not obligatory but are deferrable, must then compete with other projects, on a cost-benefit basis, for the remaining capital. Since it is often difficult to document the savings involved with energy efficiency projects, they are at a disadvantage to projects with a quantifiable positive cash flow. Thus, energy efficiency decisions are often tabled in favor of those projects with an apparent greater return.

Manufacturers have made substantial progress toward improving energy use in their facilities. Many projects that increase productivity or product quality, such as replacing old equipment, have a corollary benefit of improving energy efficiency. Others have instituted energy management plans or have participated in utility-sponsored DSM plans. The EIA data analysts showed an intense interest in the level of penetration of these activities into manufacturing.

The analysts and the MECS staff, after listening to what the manufacturers had to say, decided to include three sets of new questions related to energy efficiency improvements. These were designed to add as little as possible to the respondent burden.

The first set of new questions asks the total square footage of buildings on the manufacturing site, and the percentage that had controlled heating or cooling. Reasonable approximations were considered sufficient here. The second set lists energy management activities. Respondents simply mark yes or no,

Figure 1. End-Use Matrix for 1991 Manufacturing Energy Consumption Survey

Estimated Percent Consumption by End Use

REASONABLE APPROXIMATIONS ARE ACCEPTABLE - SEE INSTRUCTIONS

Energy sources can be consumed either directly, in equipment such as motors, furnaces, kilns, etc., or indirectly through conversion to steam and hot water in a boiler. All indirect use of the energy sources listed below are represented by the single percent entry for boiler fuel. The remaining categories are to split out direct use of energy. Thus, the percentages entered for boiler fuel and all direct uses should sum to 100 percent.

	Electricity	Coal	Natural Gas	Distillate Fuel Oil	LPG	Residual Fuel Oil
1. Copy Section I line 12 column 2 value for electricity to column 2. Copy Section II column 9 values for the other energy sources in columns 3 through 7.	Kilowatt-hours	Short Tons	1000 cubic feet	Barrels	Gallons	Barrels
2. Of the amounts listed on line 1, what percent of each energy source was used for the following purposes:						
A. BOILERS	%	%	%	%	%	%
B. DIRECT PROCESS USES						
1. Process heating (e.g., kilns, furnaces, ovens)	%	%	%	%	%	%
2. Process cooling and refrigeration	%	%	%	%	%	%
3. Machine drive (e.g., motors, pumps, etc., associated with manufacturing process equipment)	%	%	%	%	%	%
4. Electro-chemical processes	%					
5. Other Process (Please specify any other uses of energy in the "Remarks" portion of the questionnaire.)	%	%	%	%	%	%
C. DIRECT NON-PROCESS USES						
1. Facility heating, ventilation, and air conditioning	%	%	%	%	%	%
2. Facility lighting	%					
3. Facility support other than C1 or C2 above (e.g., nonprocess cooking, water heating, office equipment)	%	%	%	%	%	%
4. Onsite transportation	%		%	%	%	
5. Conventional electricity generation		%	%	%	%	%
6. Other (Please specify any other uses of energy in the "Remarks" portion of the questionnaire.)	%	%	%	%	%	%
TOTAL for all purposes	100 %	100 %	100 %	100 %	100 %	100 %

indicating their participation in such programs. Finally, the forms include checklists of energy efficiency technologies, and, again, asks for only a yes or no response.

Conclusions

Even though focus group research is purely qualitative in nature and cannot be generalized to a population of interest, it has proved to be an extremely useful tool for assessing user needs and data availability. The EEUISD believes that the simple changes to the questionnaire resulting from the focus groups make the 1991 MECS a stronger, more useful survey than its predecessors. There is, in fact, evidence of the success of the focus groups in redesigning the MECS questionnaire. All Federal surveys of 10 or more participants must be submitted to the Office of Management and Budget (OMB) for clearance. A part of this clearance process allows data users and providers and other interested parties to provide comments on the proposed survey. OMB received numerous comments on the 1985 and 1988 surveys. These comments covered a wide area including respondent burden, confidentiality, and usefulness of the results. The proposed 1991 survey was subjected to the same process. This time, however, no public comments were received.

Other evidence of the success of the focus groups lies in the responses to the 1991 MECS. Although the final results have not yet been tabulated, the questionnaires have been edited and there was no evidence that the new sections of the questionnaire had higher levels of respondent errors than the established sections. Finally, a response analysis survey was conducted to evaluate the effectiveness of the new sections of the 1991 MECS questionnaires. The response analysis survey was conducted by telephone with a purposive sample of about 200 respondents to the 1991 MECS.

The major purpose of this follow-up survey was to determine, for the new sections, whether the respondents understood the instructions, were able to provide reasonable estimates, and whether more accurate information could have been provided. The preliminary results of this survey indicate that the respondents had little difficulty in responding and that they could have provided more detailed and precise information.

The EEUISD has always valued the input of our data users and providers. Data providers especially are highly appreciative of being involved in the design phase of the survey and their input is invaluable in the questionnaire development process.

The usefulness of the roundtables conducted with representatives of the manufacturing sector extended beyond the input obtained from the direct meetings. Many of the individuals freely gave of their time and technical expertise to provide additional input on the wording of new questions and instructions for the 1991 MECS.

In addition, the summaries and transcripts of the roundtables have proved to be especially useful to modelers and policy analysts by providing them with a "reality check" of the characteristics and complications of energy consumption in the manufacturing sector.

The focus group approach used for the MECS was so successful that EIA plans to expand it to cover many of the other surveys that it conducts. The EIA has been granted a generic clearance from OMB to conduct these focus groups on an ongoing basis to aid in the future development of surveys.