Journal of Official Statistics, Vol. 20, No. 2, 2004, pp. 319–340

# Using Vignettes in Cognitive Research on Establishment Surveys

*Rebecca L. Morrison, Kristin Stettler, and Amy E. Anderson*

The use of vignettes in questionnaire development for household and demographics surveys is well documented and diverse. Vignettes are often presented as short narratives that describe a particular situation of interest. Respondents are asked to interpret the situation and then apply it to the survey instrument being studied.

Vignettes have not been as well documented in establishment survey cognitive testing. This article reviews the household and establishment literature on the use of vignettes. Next, we present some examples from the U.S. Census Bureau of vignettes that were modified from the traditional household design to accommodate the establishment survey setting. Finally, we discuss the characteristics of vignettes, and how they might be used effectively, particularly in establishment surveys.

*Key words:* Mock records; pretesting; questionnaire development; household vs. establishment.

## 1. Introduction

There are many different questionnaire development techniques available to the survey researcher. These can vary in scope and size from the expert review, to the cognitive interview, to the pilot test. Cognitive interviews employ a wide variety of methods, including retrospective and concurrent think-aloud, follow-up probes, and vignettes.

While vignettes have become a generally accepted cognitive technique for household and social surveys, their use in establishment surveys is not well documented. (This article uses "household surveys" and "social surveys" interchangeably.) However, their utility and their unique way of presenting situations to respondents make them a valuable addition to the establishment survey designer's toolbox.

Vignettes are used in establishment surveys for the same reasons they are used in household surveys. They are designed to provide a more realistic situation for the respondent to use as the basis for answering specific survey questions, and lend insight into

[1] Rebecca L. Morrison, U.S. Census Bureau, ESMPD, Room 3110-4, Washington, DC 20233-6200, U.S.A. Email: Rebecca.L.Morrison@census.gov
Kristin Stettler, U.S. Census Bureau, ESMPD, Room 3110-4, Washington, DC 20233-6200, U.S.A.
Amy E. Anderson U.S. Census Bureau, ESMPD, Room 3108-4, Washington, DC 20233-6200, U.S.A.
**Acknowledgments:** This article reports the results of research and analysis undertaken by the U.S. Census Bureau staff. It has undergone a U.S. Census Bureau review more limited in scope than that given to official U.S. Census Bureau publications. This report is released to inform interested parties of ongoing research and to encourage discussion of work in progress. The authors wish to express their gratitude for the helpful review comments provided by Jennifer A. Guarino, Tony Hak, Elizabeth M. Nichols, Diane K. Willimack, Dave Tuttle, Gustav Haraldsen, and the anonymous reviewers.

the cognitive response process. Furthermore, they allow analysis of respondents' judgments by varying the situations, and can be used to gauge potential sources of measurement error. However, vignettes used in establishment surveys sometimes take a different form than those used in household surveys.

The purpose of this article is to show that vignettes from the social survey setting can be used in the establishment survey setting, provided the differences between households and establishments are addressed and accommodated. These differences and their implications for the design of vignettes should be documented so that their use by establishment survey researchers can be expanded.

Section 2 provides background concerning household and establishment surveys, and the vignettes used in each setting. In Section 3, examples from U.S. Census Bureau establishment survey research are provided. The last section discusses the implications for vignettes in general, and for vignettes in the establishment survey setting in particular.

## 2. Background

### *2.1. The cognitive response model*

The cognitive response model of how people answer survey questions is often attributed to Tourangeau (1984). He describes four processes – comprehension, retrieval, judgment, and reporting. Comprehension involves understanding the question and instructions, interpreting the language/grammar and structure of the question, identifying the information sought, and linking key terms to relevant concepts. In the retrieval step, a retrieval strategy and cues are generated, data is retrieved from records or memory, and missing details are filled in. In the judgment step, respondents assess the completeness or relevance of the memories or data, integrate the material retrieved, and make estimates based on partial retrieval or incomplete data. Finally, reporting involves mapping the judgment onto a response category or editing the response. A prior step, encoding, is sometimes added to the model (Eisenhower et al. 1991). Encoding refers to how people initially record information, or where that information is stored, either in memory or records.

Tourangeau's cognitive response model was expanded to fit the establishment survey setting by Edwards and Cantor (1991). They broke the encoding process into two steps – record formation and respondent selection. Record formation acknowledges that establishment surveys often ask for data that is maintained in company records, rather than in a respondent's memory. However, a respondent's knowledge of records is still an important factor in their ability to respond accurately to survey questions; a respondent has to know what information is kept, and where it can be found. Respondent selection is determined not only by who is the most knowledgeable about the survey subject matter, but also by functional roles in the establishment, level of authority, and knowledge of and access to the relevant information systems. In cases where information is not centralized, multiple respondents are often required.

### *2.2. Cognitive testing methods*

Forsyth and Lessler documented the primary cognitive laboratory research methods used to "identify those aspects of survey questions that promote accurate response" (1991).

They divided the methods into four types – expert evaluation, expanded interviews, targeted methods, and group methods. Expert evaluations include interactional behavior coding, cognitive forms appraisal, and analysis by questionnaire experts. Expanded interviews include concurrent think-aloud interviews, follow-up probes, memory cue tasks, and retrospective think-alouds and probe questions. Targeted methods include paraphrasing, free-sort classification tasks, dimensional card-sort classification tasks, vignette classifications, rating tasks, response latency, and qualitative timing. Finally, group methods include focus groups, group interviews, and group experiments. It is the experience of the three authors that the most commonly used cognitive research methods for establishment surveys are concurrent think-aloud interviews, retrospective think-aloud interviews, and probe questions.

Vignettes are brief descriptions of hypothetical situations that are presented to respondents, who often then answer a series of survey questions based on the information given (Forsyth and Lessler 1991; Gerber et al. 1996). In essence, they attempt to create a reality for respondents. Most times, vignettes are constructed to cover the "gray area" of the concepts being tested (Alexander and Becker 1978; Gerber et al. 1996). These are situations usually not common in the population of interest or in the population from which the cognitive interviews are recruited. The relative rarity of some situations makes them a potential source of error for survey instruments that fail to take them into account.

### *2.3. Why vignettes are used*

The various cognitive research methods described in Section 2.2 are generally not used on their own. Rather than rely solely on one method, researchers often choose a combination of methods that address their research questions in the most effective way. Vignettes are used in conjunction with other cognitive research methods for four primary reasons having to do with the identification of sources of error arising from the cognitive response process.

(1) They provide a concrete situation for respondents and "more closely approximate a real-life decision-making or judgment-making situation" (Alexander and Becker 1978). Rather than having to think in the abstract, respondents are provided vignettes with a context within which they can make decisions regarding how to answer specific survey questions.
(2) Vignettes lend insight into the cognitive response process, especially with respect to the judgment and communication steps (Alexander and Becker 1978; Gower and Nargundkar 1991; Martin and Polivka 1995). Respondents are asked to make decisions based on the information in the vignette, record their answer on a self-administered form or report it to a researcher, and then answer the researcher's probing questions based on that vignette. In addition, vignettes can be helpful in pointing out problems with wording and comprehension (Martin and Polivka 1995).
(3) Vignettes are used to aid in the determination of what information is used by respondents when they answer questions (Gerber 1994; Gerber et al. 1996). When the vignette's situations are systematically altered, it becomes easier for researchers to determine what bits of information are critical for respondents' decisions (Alexander and Becker 1978).

(4) Vignettes can gauge potential measurement error sources, since researchers know the right and wrong answers for each vignette (Gower and Nargundkar 1991). In contrast, when respondents answer for themselves, it can be difficult for researchers to know the "true" answer.

The value of vignettes, however, is only as good as the assumptions researchers must make when using them as part of their cognitive testing. First, researchers using vignettes assume that respondents will give the same answers for the vignette as if the situation actually applied to them (Martin and Polivka 1995). Second, vignette usage assumes that respondents will understand circumstances that may be outside their experience or knowledge. Indeed, Goldenberg (1996) found that "respondents do not need a clear understanding of (the topic of interest) in order to answer questions about it."

### *2.4. Vignettes in the household survey setting*

Vignettes in household survey pretesting have generally taken two forms: traditional and mock records. The first type – traditional – presents respondents with a brief narrative describing a situation or scenario that is ambiguous relative to the concept being measured in the survey question(s) of interest. Respondents might then be asked to choose a category label that best describes the circumstances within the vignette (Forsyth and Lessler 1991). Gerber et al. (1996) used traditional vignettes in their study of household rosters. Their vignettes frequently took the form of one- or two-sentence descriptions of living situations, such as the following examples used to test the household roster for the U.S. Population Census, which has a reference date of April 1:

- Sandy's husband, Peter, left on a business trip on March 15 and won't return until April 30th. Should Sandy list Peter on her Census form?
- Craig and his wife have a house in Pennsylvania. Craig's job is in Washington, D.C. so he stays with his mom in D.C., Monday through Thursday of the week. Where should Craig be listed on a census form?

The second type – mock records – work in a similar fashion. Here, respondents are given a fictitious document and are asked to answer survey questions with respect to the information contained within it. For instance, Schober and Conrad (1997) presented respondents with what they called "fictional scenarios." In order to measure response accuracy using standardized and flexible interviewing approaches for obtaining consumer purchases in the Consumer Expenditures Survey (CES), these fictional scenarios took the forms of apartment floor plans and purchase receipts.

### *2.5. Differences between household/social and establishment surveys*

Establishment surveys differ from household surveys in a number of ways (Cox and Chinnappa 1995; Rivière 2002; Willimack et al. 2004), including the following:

- Many establishment surveys are required by law.
- The data collected in establishment surveys tends to be factual in nature – financial information, employment figures, etc., – rather than opinions, attitudes, or data on personal behaviors.

- Establishment surveys often measure technical concepts with precise definitions, and records do not always easily match with response categories.
- Data requested in establishment surveys may require multiple sources.
- In some cases, it may be necessary to get authorization to release data outside the establishment.
- Respondents in establishment surveys may be unwilling or unable to retrieve records during cognitive interviews.
- Most establishment surveys are self-administered, in order to encourage the use of records.

Willimack et al. contend that these differences impact establishment survey testing in several ways. First, establishment survey response is labor-intensive, requiring record look-up and multiple respondents. Surveys are usually not completed in a single sitting. As a result, it is difficult for researchers to observe the respondent completing the questionnaire in a traditional think-aloud cognitive interview. Second, response burden in establishment surveys is high and limits cognitive testing. Respondents are not motivated to complete draft survey forms prior to the researcher's visit, thus limiting the effectiveness of retrospective debriefings. Third, the data requested and found in records varies across respondents. Researchers do not know what components are included or excluded from the record-keeping systems, which makes it difficult, if not impossible, to measure response errors.

### *2.6. Vignettes in the establishment survey setting*

Because of these limitations to testing in the establishment survey setting, vignettes are used in establishment survey questionnaire development for three primary reasons. First, researchers often cannot directly observe the retrieval step of the cognitive response process because the record look-up process may be labor-intensive, time-consuming, or inconvenient (Stettler et al. 2001; Willimack et al. 2004). Second, asking respondents to answer survey questions based on the information in a vignette gives the researcher the chance to observe the judgment and communication steps of the response process, especially with respect to identifying problems with the format or layout of the questionnaire that could not be seen unless respondents actually entered data onto the form (Stettler et al. 2001). Third, respondents at establishments may hesitate to report answers because of their concern for accuracy – if a number is not reasonably accurate, they may not write it on the survey form, even in the setting of a pretest.

The two types of vignettes used in household survey pretesting can be applied to the establishment survey setting. A search and thorough review of papers in the published and conference literature indicating the use of traditional vignettes in establishment surveys yielded little (Willimack et al. 2004). Goldenberg (2002) used paragraph-length business information, similar to the vignettes used by Gerber et al. (1996), to test questionnaires that would be used to convert employers from the Standard Industrial Classification system to the North American Industrial Classification System.

While the use of traditional vignettes in the establishment setting was not documented in the available literature, several research studies were identified that used mock records containing simulated data. These mock records generally looked like fictitious business

records for an establishment or organization, and were usually some sort of financial records analogous to the respondent's own records. Respondents were presented with the mock records and were then asked to answer survey questions for that organization, based on the information provided. For instance, at one stage of the redesign of the Occupational Safety and Health Survey (Kydoniefs 1993), respondents were given mock documents with detailed information that should be reported on the survey. Researchers chose this strategy so that record keeping across respondents could be standardized and problems could be attributed to the questionnaire rather than to variations in business records.

In another study documented by Moy and Stinson (1999) and Schechter et al. (1999), participants were given one of two sets of fictitious job applications containing information about Hispanic and Latino origin – the researchers referred to these as "dummy records." Respondents were asked to aggregate information from the applications and report it on the test questionnaires. They then answered a series of questions concerning their comprehension and interpretation of various sections of the survey instrument.

As discussed in Section 2.5, establishment survey researchers face issues different from those faced by social survey researchers. Consequently, they have modified vignettes from the social survey setting to address these differences.

## 3. Methodology

Vignettes in the household survey setting have generally been used to test survey concepts. In the establishment survey setting, too, vignettes have been used for this purpose, as well as to test form layout and electronic instrument design. This section describes some of the vignettes that have been used by the authors in questionnaire development and testing of establishment surveys at the U.S. Census Bureau. Examples of both traditional and mock records types of vignettes are given, along with an example of another variation on vignettes – a mock questionnaire. These examples are intended to highlight the forms that vignettes can take and what can be learned from vignettes when they are used with other cognitive research methods, not to offer a critique of the vignettes themselves. An evaluation of alternative vignettes and other comparative research on pretest results described below would be useful, but is beyond the scope of this article.

### *3.1. Traditional vignette testing survey concepts*

Prior to the 2002 Economic Census, questions about the complex concept of co-employment were tested to determine how respondents in establishments would interpret them. Co-employment (also known as "employee leasing") is an employment arrangement where an employee-leasing firm contractually assumes responsibility for managing key human resource and employer services for a client firm. Functions handled by the leasing firm include payroll, employee benefits, unemployment, and workers' compensation. For employees, the transfer to an employee-leasing firm is almost transparent because their original employer retains the supervisory role, even though their paychecks are issued under the (tax) Employer Identification Number of the leasing firm.

There are several complications regarding co-employment that affected the researchers' ability to pretest questions about it. First, these types of arrangements are rare. Second,

while "co-employment" is the terminology advocated by the employee leasing industry, it was unclear how pervasive this terminology was and whether co-employers would recognize themselves in the questions. Third, there are a variety of alternative employment arrangements by which people can work at businesses that may be confused with co-employment, including temporary employees, contract employees, and consultants. Finally, many firms use payroll services that do many of the same functions as leasing firms, but without reporting the client company's employees under their Employer Identification Number. Both co-employers and non-co-employers could confuse these alternative arrangements with co-employment, resulting in measurement error in employee and co-employee counts.

Because of the concern about the method and wording used to ask about co-employees, cognitive interviewing was conducted to test the new questions (Stettler et al. 2001). The goal for the cognitive interviewing was two-fold: 1) to understand how respondents understand and interpret a very difficult technical concept and 2) to make any necessary changes in order to ease reporting for respondents with and without co-employees.

In 2001, 40 cognitive interviews were conducted with establishments in seven industries. Respondents represented co-employers, non-co-employers, and Professional Employer Organizations (PEOs, or leasing firms).

Respondents were asked to use data from their own firms to complete the newly drafted employment and payroll questions, which incorporated co-employment data. The core cognitive research methods of concurrent think-aloud with follow-up probes and a retrospective debriefing were used during this portion of the cognitive interview. From these methods, the researchers uncovered two major findings: 1) it was necessary to ask about employment and co-employment in two separate questions, and 2) the terms "leased employees" and "leasing company" should be used, instead of the cognitively more difficult "co-employment" and "PEO."

After respondents completed the questions using data from their own firms, they were asked to complete the same questions for hypothetical companies described in five vignettes, which outlined various employment arrangements. After each vignette, researchers asked probing questions about the respondents' definitions and distinctions among co-employees, "leased employees," temporary employees, contract employees, and consultants.

Like traditional vignettes used in social surveys (Gerber et al. 1996), each vignette was only a few sentences long. For the establishment survey setting, these vignettes used language familiar to respondents who were knowledgeable about payroll and employment arrangements within their companies. The vignettes were intentionally vague, to allow respondents wide latitude in their interpretations, and so the researchers could find out what sort of information is crucial to the judgment process. The following are two examples:

- JKL is a large telecommunications company with 300 workers. JKL has outsourced its data processing department to MNO Computing Inc., which provides all 40 computer specialists in JKL's data processing department. Where should the 40 computer specialists and 260 other employees be reported on the economic census form?

- Widgets.com manufactures custom widgets, and sells them online. Widgets.com has 20 part-time employees, who are leased from Leasing Services International. Where should the 20 employees be reported on the economic census form?

The use of vignettes to supplement the think-aloud and debriefing process allowed for more natural respondent interpretations of the various types of employment arrangements. Although the researchers probed on these topics during the part of the interview where respondents used their own records, respondents' answers were vague and hypothetical. Respondents were willing to discuss their knowledge of these employment arrangements in a more specific manner after being presented with the vignette. Sometimes the various arrangements had never been used by the respondent's business or had not been used during the reference period, yet respondents had knowledge or opinions about what each was. Vignettes, in other words, enabled respondents to apply their knowledge in novel and engaging ways. Often, the respondents reinterpreted the information provided in the vignettes in the context of their own company or industry, allowing a better understanding of the particularities of each employment arrangement. Using vignettes enabled researchers not only to confirm findings from the core cognitive research methods, but also to explore in-depth criteria that distinguished employee leasing from other employment arrangements. Researchers learned that leasing needed to be explained in terms of what it was not (not temporary employees, not contractors/consultants), rather than what it was, which enabled them to clarify the questions and instructions.

Additionally, because co-employment is rare, the researchers were concerned about "confusing the many for the few." It was helpful to see how respondents reacted to the co-employment questions in the context of these other arrangements. This is somewhat similar to how vignettes were used by Gerber et al., and not overly unique to the establishment setting. However, the questions involving employment arrangements that encompass a number of technical terms (jargon) with precise definitions, prescribed accounting practices, and tax consequences are unique to the establishment survey setting.

### *3.2. Mock records vignette testing survey concepts*

Goldenberg et al. (2002) used mock records when they compared similar data items collected on two surveys from two different government agencies: the U.S. Bureau of Labor Statistics' Current Employment Statistics (CES) Survey, and the U.S. Census Bureau's Annual Survey of Manufactures (ASM). Although both surveys request similar data about wages, production worker employment, and hours worked, they collect data in different ways. Researchers wanted to assess how the different presentations of these similar items and their instructions influenced respondents' strategies for response. Using cognitive interviews, researchers were able to evaluate respondents' behavior when they were answering both the CES and ASM questionnaires.

Four items asked on both the ASM and CES are total employment, number of production workers, production worker payroll, and production worker hours. Although it may initially appear that these questions ask for identical pieces of data, there are the following significant differences:

- The CES asks respondents to exclude irregular payments (e.g., bonuses) from reported figures; the ASM asks respondents to include these types of payments.
- The CES asks respondents to provide information about the production worker hours worked, while the ASM asks respondents about the production worker hours paid.
- The CES asks for total employment first and then asks the respondent to report subsets (i.e., production workers). The ASM collects employment by subgroup, then sums them, resulting in a total.
- CES instructions are printed on the back of the form in a bulleted format, while the instructions for the ASM are printed in a separate booklet and written in paragraphs.

At the time of the testing in 2002, both statistical agencies were legally prohibited from sharing information from their business registers, making cognitive interviews with actual business respondents impossible. As a result, cognitive testing was held in a laboratory, and participants were internal agency payroll or human resource staff members who were familiar with employment concepts and records. Vignettes in the form of mock records were used so researchers could assess how different presentations of similar questionnaire items and instructions influenced response strategies. Since the correct answers for each mock record were known, researchers could easily identify response errors.

Like Stettler et al. (2001), the researchers created mock records that closely resembled what real businesses use. In this case, the mock records looked like payroll records. As with real records, some pieces of information could be easily transferred from the records to the form, while other pieces of information required additional calculations.

Prior to the interview, subjects were given mock payroll records (see Figure 1 for an example) for a fictitious manufacturing establishment and asked to complete either the CES or ASM.

During the cognitive interview, after completing the form, participants were asked retrospective probing questions about it. Participants were then asked to complete the other form using the mock records in the presence of the researcher and were asked to "think aloud" as they went through it. While the records contained all the necessary information to complete the forms, participants still needed to locate the correct pieces of information and make some calculations.

Researchers discovered that participants made many unnecessary calculations – responses to many questions on the form could have been calculated using fewer numbers from the records. Though this approach did not cause incorrect answers, it did add burden for the participants by increasing the amount of time it took to complete the form. Researchers expected to find differences in the response strategies employed by the participants, but similar strategies were used for both CES and ASM questionnaires. The data provided by participants generally adhered to the survey definitions. Finally, the researchers discovered that participants consulted the instructions frequently and used them as a reference tool.

Using a vignette in the form of mock records showed researchers how potential respondents would react to the questions and instructions presented in both survey questionnaires. If legal impediments were not a factor, researchers could have conducted these think-aloud and debriefing cognitive interviews with real business respondents using real records. The mock records allowed researchers to create a realistic research situation

**Brass Band Hardware Company**
**Biweekly Payroll Summary Report**

| | | | | | |
|---|---|---|---|---|---|
| Pay period number | 1 | | | | |
| Pay period beginning: | 31-Dec-00 | Pay period ending: | 13-Jan-01 | Pay date: | 20-Jan-01 |

| Current Pay Period | A | B | C | D | E | F | G | H | I |
|---|---|---|---|---|---|---|---|---|---|
| Department | Total staff | Regular hours worked | Overtime hours worked | Hours worked, Total | Hours paid leave, paid holidays | Hours paid, Total | Commissions paid | Total Gross Pay | Total Net Pay |
| Executive | 4 | 216 | 0 | 216 | 104 | 320 | $0.00 | $17,017.42 | $11,742.02 |
| Sales | 5 | 248 | 0 | 248 | 152 | 400 | $0.00 | $10,379.81 | $7,162.07 |
| Admin | 4 | 288 | 0 | 288 | 32 | 320 | $0.00 | $5,119.23 | $3,532.27 |
| Operations | 44 | 3,032 | 0 | 3,032 | 488 | 3,520 | $0.00 | $68,008.80 | $46,926.07 |
| Shipping | 10 | 644 | 0 | 644 | 156 | 800 | $0.00 | $11,234.40 | $7,751.74 |
| Total YTD | 67 | 4,428 | 0 | 4,428 | 932 | 5,360 | $0.00 | $111,759.66 | $77,114.17 |

| Staff Gender | Number | Percent |
|---|---|---|
| Male | 42 | 62.7 |
| Female | 25 | 37.3 |

| Year to date | J | K | L | M | N | O | P | Q |
|---|---|---|---|---|---|---|---|---|
| Department | Regular hours worked | Overtime hours worked | Hours worked, Total | Hours paid leave, paid holidays | Hours paid, Total | Commissions paid | Total Gross Pay | Total Net Pay |
| Executive | 216 | 0 | 216 | 104 | 320 | $0.00 | $17,017.42 | $11,742.02 |
| Sales | 248 | 0 | 248 | 152 | 400 | $0.00 | $10,379.81 | $7,162.07 |
| Admin | 288 | 0 | 288 | 32 | 320 | $0.00 | $5,119.23 | $3,532.27 |
| Operations | 3,032 | 0 | 3,032 | 488 | 3,520 | $0.00 | $68,008.80 | $46,926.07 |
| Shipping | 644 | 0 | 644 | 156 | 800 | $0.00 | $11,234.40 | $7,751.74 |
| Total | 4,428 | 0 | 4,428 | 932 | 5,360 | $0.00 | $111,759.66 | $77,114.17 |

*Fig. 1. Sample mock record used in the cognitive testing of the Current Employment Statistics (CES) Survey and the Annual Survey of Manufactures (ASM)*

without violating information-sharing laws between the agencies. Researchers were able to make comparisons between response processes for both surveys that would not have been possible without the use of mock records. Most important, researchers were also able to take advantage of another one of the benefits of using vignettes – knowledge of the correct answer and the easiest way to arrive at that answer. If actual business records were used, researchers could never know the "true" answer. Not only is there variation in the contents and structure of records across businesses, researchers are not permitted intimate access to business records without exceptional circumstances.

There are many federal statistical agencies that collect the same or similar pieces of information from companies. Oftentimes final estimates for these pieces of information differ when they should be similar or are similar when they should be different. By using mock records, researchers for these two surveys were able to look at the response process as a possible explanation for differences in published statistics.

### *3.3. Mock records vignette testing form layout*

Research was conducted using mock records vignettes to test alternative formatting for the 2002 Economic Census (Stettler et al. 2000). The census of manufacturing industries, part of the economic census, provides periodic statistics about manufacturing establishments, activities, and production. Forms for manufacturing industries have traditionally used a tabular format, referred to as a "spanner," to display up to five ordered levels of detail for collecting data about products. Economic census forms for other industries typically used indenting, in the style of a topical outline, to organize successive detailed subcategories. Examples of the "spanner" and indented forms can be found in Figures 2 and 3.

In an effort to increase consistency among all economic census forms, it was proposed that the manufacturing forms use the indented layout. However, subject experts were concerned that respondents would aggregate their data onto the first available line, regardless of the detail requested, producing "first-line bias." The primary objective for the testing was to learn about the potential for this type of bias under indentation.

Cognitive interviews were conducted with respondents from 17 single- and multi-unit establishments in three different industries with past reporting problems. Two methods were used to evaluate both the "spanner" and indented layouts: think-aloud with debriefing, and vignettes. Because of the length and complexity of the questionnaire, e.g., as many as 30 questions covering more than two hundred individual data items, it was unreasonable to expect respondents to complete it in the presence of the researchers. Therefore, researchers tested both layouts with respondents using different methods. Prior to the interview, respondents were sent one of the two versions of the form (tabular or indented). Respondents would test one layout by completing the questionnaire using their company's data. During the cognitive interview, researchers would discuss any issues respondents had with the first layout and then ask them to complete the other layout using mock records.

Unfortunately, despite the fact that some respondents completed a subset of the items before the cognitive interview, none had responded fully. Therefore, the use of mock records was critical to the success of the cognitive interviews for this project. To ensure that the mock records were comprehensible to the respondents, separate records were

**Item 18B.** PRODUCTS AND SERVICES OF THIS ESTABLISHMENT DURING 1997 – Continued

| Line No. | Products and services *Item corresponds to products reported on Current Industrial Reports (A) | | | | Census product code 581 (B) | Unit of measure for quantities (C) | Products shipped and other receipts: Quantity 583 (D) | Value, f.o.b. plant (E): 584 Millions | Thousands | Dollars |
|---|---|---|---|---|---|---|---|---|---|---|
| 63 | **Steel Nails, Staples, Tacks, Spikes, and Brads, Made in Plants That Draw Wire** | Round wire nails, collated, prepackaged * (CIR MA33B, item code 1773, pt) | | | 33152 51 3 | Short tons | | $ | | |
| 64 | | ROUND WIRE NAILS, NOT COLLATED, SMOOTH SHANK * (CIR MA33B, item codes 1773, pt) | Coated, plated, or painted | Galvanized | 33152 52 1 | | | | | |
| 65 | | | | Vinyl, resin, or cement coated | 33152 53 9 | | | | | |
| 66 | | | | Other | 33152 54 7 | | | | | |
| 67 | | | Not coated, plated, or painted | | 33152 55 4 | | | | | |

*Fig. 2. Example of the "spanner" layout used in the cognitive testing of the U.S. 2002 Economic Census*

**Item 18B.** PRODUCTS AND SERVICES OF THIS ESTABLISHMENT DURING 1997 – Continued

| Products and services *Item corresponds to products reported on Current Industrial Reports (A) | Census product code 581 (B) | Unit of measure for quantities (C) | Products shipped and other receipts: Quantity 583 (D) | Value, f.o.b. plant (E): 584 Millions | Thousands | Dollars |
|---|---|---|---|---|---|---|
| **1. Steel nails, staples, tacks, spikes, and brads, made in plants that draw wire** | | | | | | |
| **a.** Round wire nails, collated, prepackaged* (CIR MA33B, item code 1773, pt) . . . . . | 33152 51 3 | | | $ | | |
| **b.** Round wire nails, not collated, smooth shank* (CIR MA33B, item codes 1773, pt) | | | | | | |
| **(1)** Coated, plated, or painted | | | | | | |
| **(a)** Galvanized . . . . . | 33152 52 1 | | | | | |

*Fig. 3. Example of the indented layout used in the cognitive testing of the U.S. 2002 Economic Census*

developed for each of the three industries being studied. These one-page records looked like internal documents from a fictitious company; a sample is provided in Figure 4.

The mock records were constructed so that respondents would be encouraged to make difficult choices that could possibly result in first-line bias. That is, the data provided in the mock records did not correspond exactly to the categories requested on the form, a common phenomenon in establishment surveys. In some cases, an aggregation of multiple items was provided in the vignette, without appropriate detail breakdowns. Because the researchers were studying the effect of the question format, respondents were required to write numbers on the form, as if they were completing it with their own data. This permitted an evaluation of whether the "spanner" or indented versions were related to tendencies for respondents to enter data incorrectly. Since the economic census is a self-administered questionnaire, and there is no interviewer present to help resolve misunderstandings, it was important to find out about the potential for misreporting on the questionnaire due to formatting or layout.

The mock records used in this research were similar to the "fictional scenarios" used by Schober and Conrad (1997) in two ways. Both research teams used mock records so that potential sources of response errors could be identified. If respondents had used their own records, then it would have been difficult to maintain comparability across respondent interviews. The mock records and "fictional scenarios" were also similar in the degree of complexity involved in mapping the information presented to the appropriate response categories. Both sets of researchers provided some simple, straightforward information that was easily mapped to response categories along with other data that could be mapped to multiple categories or that required additional information.

The use of mock records in Stettler et al.'s research showed how respondents reacted to a form's layout even though the data provided did not fit perfectly into the available answer categories. Although respondents used different methods for determining how to report their data, differences in a respondent's reporting behavior did not appear to be related to the formatting. A few reported aggregated data on the first detailed line, but they did so equally on the "spanner" and indented versions.

Generally, respondents in this study were reluctant to take the time to access their own records to complete the form either during or prior to the cognitive interviews. Using mock records enabled the researchers to observe the respondents using the visual cues of the different layouts for entering their data, which was especially helpful for this research specifically and for establishment surveys in general. Multiple respondents and the time- and labor-intensive process associated with business record retrieval would have made respondent burden prohibitively high during a cognitive interview.

If the researchers had used only think-aloud interviews with debriefings, without mock records, it is unlikely that they would have been able to observe the mapping of data found in records onto the questionnaire. It would have taken too much time and effort for respondents to access their own records and then respond to the lengthy and complex questionnaire. Without the vignette, researchers would also have been unable to observe the communication step. Furthermore, researchers would not have been able to learn what response strategies are used when records do not easily map to the response categories on the questionnaire. In the end, researchers would have been unable to observe possible tendencies toward first-line bias with either the tabular or indented format. Also, because

**ABC Widgets, Inc.**
123 Street
Washington, D.C. 12345
Phone: (202) 123-4567

**PRODUCT SHIPMENT REPORT**

*Instructions: Please take the information below and put it into the correct place(s) on form MC-3621 or form MC-33421(X).*

| Date | Description | Quantity | Value of Shipments | TOTALS |
|---|---|---|---|---|
| 1/97-12/97 | Etch | NA | $ 3,145,000 | $ 3,145,000 |
| 1/97-12/97 | Strip | NA | $ 2,451,000 | $ 2,451,000 |
| 1/97-12/97 | Aligners | NA | $ 987,000 | $ 987,000 |
| 1/97-12/97 | Vapor deposition | NA | $ 1,541,000 | $ 1,541,000 |
| 1/97-12/97 | Voltage ion implanters | NA | $ 2,598,000 | $ 2,598,000 |
| 1/97-12/97 | Polishing machines for semiconductor wafers | NA | $ 751,000 | $ 751,000 |
| 1/97-12/97 | Focused ion beam milling machines | NA | $ 541,500 | $ 541,500 |
| 1/97-12/97 | Machines that cut blank semiconductor wafers | NA | $ 4,212,000 | $ 4,212,000 |
| 1/97-12/97 | Die and wire bonders | NA | $ 3,457,000 | $ 3,457,000 |
| | | | | |
| | | | | |
| | **TOTALS** | | $ 19,683,500 | $ 19,683,500 |

*Fig. 4. Sample mock record vignette used in the cognitive testing of the U.S. 2002 Economic Census*

researchers knew the correct answers for each set of mock records, they were able to evaluate the response errors associated with the formatting and layout of the question. Researchers could not have known this if the respondent's own records were used.

### *3.4. Mock questionnaire testing electronic instrument design*

Testing of edits within electronic instruments was conducted as part of the development of the U.S. Census Bureau's "Style Guide for the 2002 Economic Census Electronic Forms" (Anderson et al. 2001). A style guide is a list of rules for developers to use when designing electronic forms. One section of the U.S. Census Bureau's style guide focused on edit messages, which usually consist of a couple of sentences telling the respondent that data he or she entered has failed an edit (a means of detecting potential inconsistencies in the reported data). Issues included how many edit messages to present, how they should be phrased, and when they should be shown.

Due to time and budget constraints, the proposed edit messages were tested with eight internal U.S. Census Bureau employees, rather than actual business respondents. The testing focused primarily on how edit messages should be phrased. The researchers were concerned about participants' high level of familiarity with economic census forms, so they developed a vignette in the form of a partially completed mock on-line mortgage application (a mock questionnaire) that easily incorporated the proposed edits from the style guide.

This example and those discussed previously differ in one important respect. Earlier, mock records or traditional vignettes were used to test questions in the context of the vignette. Here, the edit wording was tested in the context of the mock questionnaire vignette, which was presented as a prototype. (See Figure 5 for a portion of the mock questionnaire presented to participants.)

Prototypes used in the electronic instrument design process can vary from low-fidelity, like the one used here, to high-fidelity (a near-final version of the instrument). Researchers give respondents the prototype and a set of tasks to complete, so that the interaction between respondent and prototype can be examined (Hix and Hartson 1993). Prototypes can be used to test the general look and feel of a system or instrument, as well as specific pieces or parts.

Cognitive testing of the edit messages was done in the U.S. Census Bureau's cognitive laboratory. All participants were asked to think aloud. Subjects were initially presented with printed screen shots of the mock questionnaire vignette and asked to read them in order to familiarize themselves with the form, its pre-printed data, and the language used on mortgage applications. Three edit messages were presented to the subjects as they read through the vignette; fifteen edit messages were presented to participants at the end of the vignette. Each message was no longer than three sentences. They varied in terms of the inclusion and placement of the question number, the question topic, and the respondent's entry of data. As a result, some messages were presented less clearly than others. Three examples of the edit messages that were tested are shown below:

- Entry exceeds maximum length allowed. Only "5" characters are allowed.
- Item 29 is $10,000, which does not equal $3,000 in item 30. Please check for typing mistakes. If entries are correct, explain discrepancy in the remarks section.

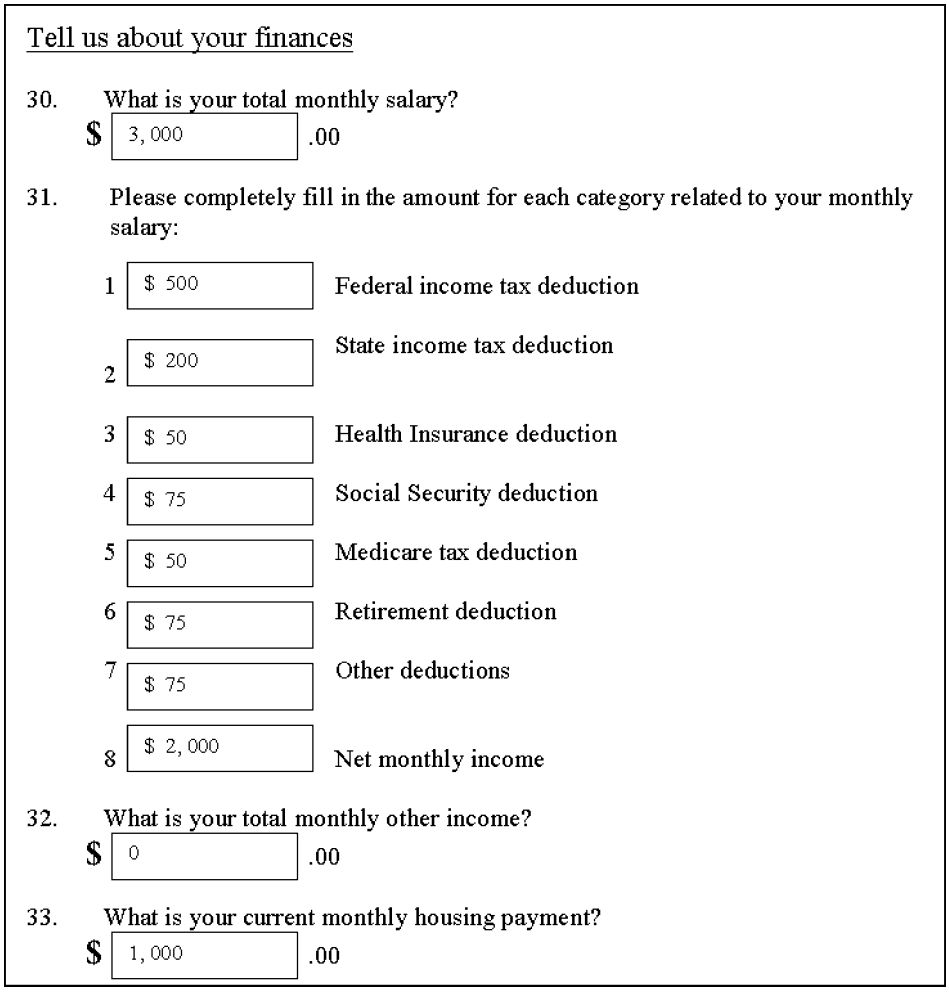


*Fig. 5. Portion of the mock questionnaire vignette used in the testing of electronic instrument design for the U.S. Economic Census*

- ⟨VARNAME1⟩ should be greater than or equal to ⟨VARNAME2⟩. If values are correct, explain discrepancy in the remarks section.

For each edit, participants were asked a series of probes about what they thought the edit message meant, what they would need to do in order to correct or resolve the error, and where on the mock questionnaire they needed to go in order to make their corrections. Participants were encouraged to explain any confusion they experienced with the edit messages. Researchers watched and listened as participants used the messages to fix the data problems. This allowed researchers to pinpoint current edit message limitations and outright problems.

The findings from this research had several implications for the style of edit messages presented to respondents on economic census electronic forms. Respondents frequently cited a need for the question number, a description of the problem, and an action to take.

As a result, researchers recommended that these three elements be incorporated into each edit message, where feasible.

Incorporating a mock questionnaire into the development of the style guide was beneficial to the project. Without it, researchers would have been making decisions regarding edit message wording without input from people behaving in a manner expected of respondents, and it may have been difficult to pinpoint what sort of information was useful and what was not useful. Since testing was limited to internal staff, using a vignette in the form of a mock questionnaire mimicked the situation survey respondents would likely find themselves in – having to interpret and decipher a questionnaire, an instrument, and edit messages they would be unfamiliar with.

In contrast to most household and social surveys, the mainstay of many establishment surveys is to collect data that consists of figures or numbers (e.g., employment, payroll, sales). Establishment surveys frequently require respondents to refer back to previous questions, split out totals reported elsewhere in the questionnaire, and match figures from one question to another. Therefore, the inter-relationships and consistencies among reported figures must be maintained. Edit message design elements that make it easy for respondents to understand the nature of any data problems and take appropriate action are essential to good data quality.

## 4. Summary and Implications

To this point, this article has provided background for the cognitive response process and the use of vignettes in the social and establishment survey settings, and has described four examples of how vignettes have been used in the testing of establishment survey instruments at the U.S. Census Bureau. In addition, the utility and effectiveness of vignettes as a tool that augments the core cognitive interviewing techniques of retrospective debriefing, concurrent think-aloud, and probes have been demonstrated. The primary advantages that vignettes offer to the researcher are their ability to be tailored to specific uses, and their systematic application that lends comparability across respondents.

Two different forms of vignettes have also been discussed. Traditional vignettes are brief narratives that describe hypothetical situations. Mock records are similar but take a different form in order to accommodate the restrictions associated with record retrieval during an establishment survey cognitive interview. These mock records can be tailored for different types of respondents, or the same records can be given to all respondents, depending on the researcher's needs.

The establishment survey setting has unique characteristics that require traditional vignette methodology to be broadened to include mock records and mock questionnaires. The goals are the same as with traditional vignettes – to identify sources of error arising from the cognitive response process. However, because of the nature of records and their importance in the establishment survey response process, using vignettes in the form of mock records can help researchers gain a better understanding of potential response errors during questionnaire development and testing. Likewise, mock questionnaires can be used to test aspects of the survey instrument other than the questionnaire itself that could affect data quality, as in the development of electronic instruments.

In each example provided, the vignette took one of three forms: traditional, mock record, or mock questionnaire. Vignettes added to the knowledge gained by the researchers during the course of their cognitive interviews by exploring extra dimensions of survey response. Vignettes have been shown to provide information about survey concepts, questionnaire layout, and electronic instrument design.

Researchers' use of vignettes, in addition to the core cognitive techniques, are important to the establishment survey setting because they allow of insight into response process issues that are unique to the establishment survey setting. The examples above demonstrated the utility and effectiveness of vignettes in three specific areas: 1) to explore technical concepts with complex definitions, 2) to analyze response errors despite a lack of access to respondents' actual business records, and 3) to investigate inter-relationships and consistencies among data items.

Each of the establishment survey vignettes presented here shares the following five characteristics with vignettes in the household survey setting.

(1) Vignettes are brief: In order to take advantage of the limited time in a cognitive interview, vignettes are constructed so that respondents can grasp the situation in a few words or sentences.
(2) Vignettes use respondents' words and language: Using the same words that respondents employ to describe a given situation is imperative. Otherwise, the researcher risks respondent confusion, and the respondent's focus shifts from answering questions about the situation to understanding the situation itself. Using respondents' own words and language also helps increase the realism of the situation or mock record presented.
(3) Vignettes are intentionally ambiguous: to some degree, vignettes are slightly unclear to the respondent with regard to particular features of survey questionnaires. Researchers are able to find out how ambiguous situations are likely to be interpreted; there is little reason to make them straightforward. Since researchers are interested in the effectiveness of a particular survey design, vignettes constitute an attempt to realistically create the "gray areas" that the survey is likely to encounter.
(4) Vignettes are useful in shedding light on the judgment and communication steps of the cognitive response process: They are used to find out how respondents make decisions, what they base their answers on, and how they communicate their answers on the survey (to an interviewer or on a self-administered form). The systematic variation of vignette details and the vagueness of vignettes also facilitate the discovery of a respondent's decision criteria as they relate to a specific question or questions. Furthermore, with the provision of "facts" on which to base answers, the respondent can more easily answer survey questions concretely.
(5) Vignettes efficiently use limited resources: Vignettes enable researchers to take better advantage of the cognitive interviews with respondents they already have. Rather than searching for respondents with special characteristics, researchers can gain information about special characteristics through respondents who may be familiar with or have knowledge of them, though perhaps not personal experience.

Regardless of the survey setting, these similarities provide guidance for the appropriate construction and use of vignettes, and the circumstances under which they should be used.

If researchers are interested in identifying potential sources of error associated with the cognitive response process, especially the judgment and communication steps, vignettes provide a common and constant stimulus across respondents, along with "correct" answers. They are most effectively used when only limited resources are available, or if the researcher is interested in relatively rare situations that need to be accounted for in the survey instrument. They should be created so that they are brief, ambiguous, and use respondents' own words and language. In this way, researchers can make the most effective use of the restricted time within the cognitive interview.

The establishment survey researcher also has to consider additional factors that are unique to the establishment setting, especially record look-up. Though some social surveys ask respondents to look back through their calendars or expense records, asking business respondents for data from records is a routine practice in establishment surveys. However, business respondents are frequently unable or unwilling to review their records in the context of a cognitive interview because of the labor-intensive and burdensome response process. Thus, vignettes provide an effective and efficient tool for obtaining information about the cognitive response process in establishments, potential sources of response error, and user requirements for electronic instrument design.

Vignettes, often used by household survey developers, are useful to the survey researcher who wants to study the decision-making process of business respondents. Regardless of the form they take, vignettes are a valuable addition to the establishment survey designer's toolbox.

## 5. References

Alexander, C.S. and Becker, H.J. (1978). The Use of Vignettes in Survey Research. Public Opinion Quarterly, 42, 93–104.

Anderson, A.E., Nichols E.M., and Pressley, K.D. (2001). Usability Testing and Cognitive Interviewing to Support Electronic Forms Development for the 2002 U.S. Economic Census. Paper prepared for presentation at the Statistics Canada Symposium 2001: Achieving Data Quality in a Statistical Agency, Hull, Quebec.

Cox, B.G. and Chinnappa, B.N. (1995). Unique Features of Business Surveys. In Business Survey Methods, B.G. Cox, D.A. Binder, B.N. Chinnappa, A. Christianson, M.J. Colledge, and P.S. Kott (eds). New York: Wiley.

Edwards, W.S. and Cantor, D. (1991). Towards a Response Model in Establishment Surveys. In Measurement Errors in Surveys, P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (eds). New York: Wiley.

Eisenhower, D., Mathiowetz, N.A., and Morganstein, D. (1991). Recall Error: Sources and Bias Reduction Techniques. In Measurement Errors in Surveys, P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (eds). New York: Wiley.

Forsyth, B.H. and Lessler, J.T. (1991). Cognitive Laboratory Methods: A Taxonomy. In Measurement Errors in Surveys, P.P. Biemer, R.M. Groves, L.E. Lyberg, N.A. Mathiowetz, and S. Sudman (eds). New York: Wiley.

Gerber, E. (1994). Hidden Assumptions: The Use of Vignettes in Cognitive Interviewing. Working Papers in Survey Methodology no. SM94/05. U.S. Census Bureau: Washington, D.C.

Gerber, E., Wellens, T.R., and Keeley, C. (1996). Who Lives Here? The Use of Vignettes in Household Roster Research. Working Papers in Survey Methodology no. SM96/02. U.S. Census Bureau: Washington, D.C.

Goldenberg, K.L. (2002). Personal Communication with D.K. Willimack, September, U.S. Census Bureau.

Goldenberg, K.L. (1996). Using Cognitive Testing in the Design of a Business Survey Questionnaire. Paper prepared for presentation at the annual meeting of the American Association for Public Opinion Research, Salt Lake City, UT.

Goldenberg, K., Willimack, D.K., Fisher, S.K., and Anderson, A.E. (2002). Measuring Key Economic Indicators in U.S. Government Establishment Surveys. Paper prepared for presentation at the International Conference on Improving Surveys, Copenhagen, Denmark.

Gower, A.R. and Nargundkar, M.S. (1991). Cognitive Aspects of Questionnaire Design: Business Surveys Versus Household Surveys. Proceedings of the Annual Research Conference, Washington, D.C., U.S. Census Bureau, 299–312.

Hix, D. and Hartson, H.R. (1993). Developing User Interfaces: Ensuring Usability Through Product and Process. New York: Wiley.

Kydoniefs, L. (1993). The Occupational Safety and Health Survey: Instrument Redesign. Proceedings of the International Conference on Establishment Surveys, American Statistical Association, Alexandria, Virginia, 99–106.

Martin, E. and Polivka, A.E. (1995). Diagnostics for Redesigning Survey Questionnaires: Measuring Work in the Current Population Survey. Public Opinion Quarterly, 59, 547–567.

Moy, L. and Stinson, L.L. (1999). Two Sides of a Single Coin: Dimensions of Change Suggested in Different Settings. Proceedings of the American Statistical Association, Section on Survey Research Methods, 44–53.

Rivière, P. (2002). What Makes Business Statistics Special? International Statistical Review, 70, 145–159.

Schechter, S., Stinson, L.L., and Moy, L. (1999). Developing and Testing Aggregate Reporting Forms for Data on Race and Ethnicity. Paper prepared for presentation at the Federal Committee on Statistical Methodology Research Conference, Arlington, VA.

Schober, M.F. and Conrad, F.G. (1997). Does Conversational Interviewing Reduce Survey Measurement Error? Public Opinion Quarterly, 61, 576–602.

Stettler, K., Morrison, R.L., and Anderson, A.E. (2000). Results of Cognitive Interviews Studying Alternative Formats for Economic Census Forms. Paper prepared for presentation at the Second International Conference on Establishment Surveys, Buffalo, NY.

Stettler, K., Willimack, D.K., and Anderson, A.E. (2001). Adapting Cognitive Interview Methodology to Compensate for Unique Characteristics of Establishments. Paper prepared for presentation at the annual meeting of the American Association for Public Opinion Research, Montreal, Quebec.

Tourangeau, R. (1984). Cognitive Sciences and Survey Methods. In Cognitive Aspects of Survey Methodology, T.B. Jabine, M.L. Straf, J.M. Tanur, and R. Tourangeau (eds). Washington, DC: National Academy Press.

Willimack, D.K., Lyberg, L., Martin, J., Japec, L., and Whitridge, P. (2004). Evolution and Adaptation of Questionnaire Development, Evaluation and Testing Methods in Establishment Surveys. In Questionnaire Development, Evaluation, and Testing, S. Presser, M.P. Couper, J.T. Lessler, E. Martin, J. Martin, J.M. Rothgeb, and E. Singer (eds). New York: John Wiley and Sons.

Received February 2003
Revised December 2003