Using Iterative Pretesting and Response Analysis Surveys to Develop a Questionnaire for a Multi-Mode Survey July 2012

William Mockovak Office of Survey Methods Research Sharon Stang Office of Employment and Unemployment Statistics

Introduction

The Bureau of Labor Statistics (BLS) received funding beginning in fiscal year 2010 to develop and implement the collection of new data on green jobs in the American economy. The specific objectives of this data collection were to obtain information on (1) the number and trend over time in green jobs, (2) their industrial, occupational, and geographic distribution, and (3) the wages of workers in these jobs.

To address these objectives, two surveys were designed and conducted in 2011 and 2012: the Green Goods and Services (GGS) Survey focused on measuring employment involved in the production of green goods and services, and the Green Technologies and Practices (GTP) Survey focused on collecting information on green jobs at firms that use environmentally friendly production processes. BLS' work developing and refining its definition of green jobs for these surveys is discussed in Stang and Jones (2011).¹ This paper focuses on the steps taken to develop and test the questionnaire used in the Green Technologies & Practices (GTP) survey. Because the schedule and budget for the GTP survey did not allow time to design and conduct large-scale, embedded field experiments, this paper describes how an approach employing small, iterative field testing with mixed-method research approaches was used to develop the questionnaire and to refine field procedures.

Overview

The process of developing the questionnaire for the GTP survey progressed through the following phases, each of which will be discussed in turn:

- 1. Cognitive interviews
- 2. Feasibility study
- 3. Iterative pilot testing

Research Approach: Cognitive Interviews and Feasibility Study

As noted by Stang and Jones (2011), a survey attempting to measure the number of jobs associated with environmentally friendly production processes faces several challenges, including:

- How should environmentally friendly production processes be defined?
- What terminology should be used to describe green processes?
- Do respondents have the information available, and are they willing to provide it?
- Is more than one survey mode necessary to collect the information?

¹ See <u>http://www.bls.gov/green/</u> for more detail on the BLS definition of green jobs.

Since 'green' can be defined in numerous ways, definitional issues were first addressed through cognitive interviews conducted with businesses expected to have green technologies and practices. As anticipated, these interviews uncovered some disagreements about definitions and whether specific business activities qualified as being green. In addition, the cognitive interviews discovered another potential cause of measurement error, social desirability, because some companies valued being seen as green by the public, as well as internally.

The next step in the testing process was a feasibility study of an initial draft of the questionnaire with a select sample of 201 businesses thought to use green practices. This study attempted to answer the following questions:

- Who at the firm is the best respondent for this type of survey? What are the job titles of "best contacts" in establishments?
- Do businesses with known green technologies and practices understand and answer the survey questions as intended?
- Would one respondent be able to answer all of the questions, or would more than one person in the business have to provide the information requested?
- Would respondents be willing to supply detailed data about occupations (title and main activities), along with wages?
- Did respondents understand the purpose of the survey, how did they assess its value, and were they likely to participate in a national survey?
- What were preferred modes for reporting the data?
- What industries are likely to need more help completing the form?

The feasibility study arrived at some clear conclusions.

- 1. A survey about the use of green technologies and practices was generally viewed positively.
- 2. The biggest threat to measurement error was providing a clear definition of green technologies and practices. Moreover, respondents had some difficulty determining how much of a worker's time was devoted to green activities.
- 3. Respondents wanted clear, concrete examples of green technologies and practices so that they could more easily categorize their own activities.
- 4. There was an increased likelihood in larger establishments that more than one respondent would be required. For example, one respondent would be needed to answer questions about the use of green technologies and practices, and another would be needed to provide occupational and wage information for green employees. This finding identified challenges for designing the paper survey form, as well as the web-based data collection instrument, since the electronic form would have to be made accessible to more than one person in a business. The importance of this point was further reinforced when many respondents strongly indicated that they wanted a web-reporting option.

Test, Revise, and Retest Strategy for Developing the Survey Form

Although it can be argued that the use of embedded experiments with random assignment of cases (Fienberg and Tanur, 1989) is the ideal approach for testing a survey questionnaire, there were two reasons why this approach was not feasible in this case. First, since a survey of this nature had not previously been conducted by BLS, there were a wide variety of questionnaire-

design and procedural issues that needed to be addressed, and which could not be resolved in a single field test. Second, the schedule and budget did not allow for an extensive embedded experiment. Therefore, rather than attempt to address all of these issues in a single large field test, an iterative field-testing approach employing five separate panels (pilot tests) with relatively small samples (430–703) was used to test and refine field procedures and the survey questionnaire. In addition, since this was BLS's first green survey of this nature, 50 percent of the samples for Panels 1 and 2 focused on establishments that were expected to have green technologies and practices, so that steps could be taken to ensure the form got to the right person in the establishment and that the questionnaire was usable. Once these objectives were met, random samples based on the Quarterly Census of Employment and Wages (QCEW) was used in Panels 3, 4, and 5 to better simulate what the production survey would encounter.

Since the purpose of the field tests was to develop the paper survey instrument and to maximize the mail response to it, the field tests were conducted with no telephone data collection. Completes for the panels include only responses to the mailed paper form, and in the later panels (4 and 5), the faxable and e-mail fillable forms. The sample size for each field test (panel), the number of completed forms received, and sample sizes for respondents and non-respondents in the Response Analysis Survey (RAS) are shown in the following table:

			RAS		
	Pilot	Test	(number of completed		
			interviews)		
Panel	Ν	Completes	Respondents	Non-Resp	
1	703	252	27	26	
2	430	202	26	25	
3	570	289	26	25	
4	677	288	26	25	
5	677	236	26	25	

Table 1: Sample Sizes for Each Test Panel

Since previous research on green jobs has shown that the number of workers involved is still quite small relative to the size of the workforce, the samples for Panels 1, 2, and 3 were equally divided between expected green businesses and units selected systematically from the QCEW to obtain enough businesses with green jobs so that the green jobs section of the form could be adequately tested. In Panels 4 and 5, the sample was randomly drawn from the QCEW. Each panel lasted only 6 weeks versus the 24 weeks allotted for the production survey. The survey procedures for the test panels are described in the table below.

Table 2: Survey Procedures for Test Panels

Week	Activity
Week 1	Survey forms mailed
Week 2	Re-mailing activities for postal returns and additional telephone number research

Week	Activity
Week 3	 Non-response prompt postcards mailed Respondent response analysis surveys (RAS) began Non-response prompting (NRP) telephone calls began (for Panels 2-5)
Week 4	Non-response prompting (NRP) telephone calls continue
Week 5	Non-respondent response analysis surveys began
Week 6	Active data collection efforts concluded. Late arriving survey responses are included in the response rates and totals.

The following chart shows the response rates obtained for each panel. The final response rate in the production survey was 70 percent. We hypothesize that the lower response rate in Panel 5 resulted from the timing of the data-collection period, which overlapped with summer vacations, and response was probably also affected by the acerbic 2011 government budget debate that took place at the same time, and which was widely reported in the press.



Chart 1 - Weekly Response Rates for QCEW Samples in each Panel, 1-5²

The following key questionnaire design features were addressed in the separate field tests:

^{*}See footnote on this page for this chart.

 $^{^{2}}$ As noted, Panels 1 and 2 included a half sample of pre-identified green establishments. This graph only shows the response rates for the QCEW portions of the first 3 panels, since they are comparable to the panels drawn for Panels 4 & 5, which were drawn solely from the QCEW. The 'green' portions of Panels 1-3 had response rates that were 7-10 percent higher than the QCEW portions.

- Perceived burden of the survey form.
- Use of a grid or table for obtaining information about green technologies and practices.
- A variety of smaller design changes, such as changing the location of instructions and the use of formatting and highlighting to emphasize questions, examples, and skip instructions.

In addition, each of the five panels addressed practical, operational issues such as determining best practices for cleaning up addresses, reducing undeliverable postal returns, identifying titles of contact persons in businesses, designing effective mailing packages, refining non-response procedures, and determining the impact of offering different modes of response (for example, phone, fax, e-mail, web).

As noted, because both time and resources were limited and it was anticipated that many changes to the questionnaire and procedures would occur from panel to panel, it was not feasible to implement experimental designs to guide design changes at each step of the testing. Instead, an 'action research' model was used (Mockovak, 2010). With this model, researchers systematically study a problem and collect a variety of evaluation data to address issues of interest. Close collaboration between survey researchers and practitioners is an integral part of the approach, and an attempt is made to approach issues scientifically and to rely on quantitative data whenever possible.

However, qualitative research is also used to make decisions, for example, in addition to an itemnon-response analysis, a visual screening of completed paper questionnaires was done to identify completion problems and comments entered by respondents. In addition, a key part of each panel was a follow-up Response Analysis Survey (RAS), which is a structured respondent debriefing conducted over the telephone (Phipps et al., 1995). Response Analysis Surveys of both respondents and non-respondents were conducted as a part of each panel. Although some core questions were kept, the RAS scripts were revised after each panel to address new issues that arose or to more deeply probe issues raised in a previous panel. Using this approach, questionnaire revisions and procedural changes are based on theory, relevant research, and results from current evaluation measures. All these sources of information are interpreted in the context of professional judgment since sample sizes are often too small to reach statistical significance.

The five panels, their key research objectives, and a summary of some key findings are described in Appendix A. An important point about this summary table is that some common objectives were addressed in each panel (for example, response), whereas others changed as respondent comments obtained from the RAS and more objective measures (item non-response rates) pointed out areas of the questionnaire that were causing difficulty. The most significant changes to the form were implemented in Panels 2 and 3.

Perceived Burden - Length of the Survey Form

As noted by Bogen (1996), the relationship between interview length and non-response is "surprisingly weak and inconsistent." Some of the factors that might affect non-response include a respondent's interest in and perceived value of the survey topic; questionnaire length as defined by the number of pages, questions, or time it takes to answer the questions; mode of interview

(self-administered vs. interviewer); presence of financial or other incentives; and format of the questions (Galesic, 2003).

Despite equivocal findings about the impact of the length of a questionnaire on response, many survey methodologists believe that establishment respondents are much more likely to disregard a voluntary survey with a questionnaire that appears too long and, hence, burdensome. The length of the questionnaire tested in Panel 1 was 8 pages, with the RAS results for non-respondents in Panel 1 indicating that about 35 percent of the non-respondents mentioned time or burden when asked why they decided not to complete the survey. For this reason, the questionnaire was shortened to 6 pages, a length that remained constant during the remaining tests (Panels 2-5). RAS respondents were asked in Panels 3 and 4 if the survey form seemed to be of an appropriate length or too long when they first looked at it. These results from the respondent RAS, which are shown in the next table, suggest that efforts to reduce perceived burden were successful.

	Panel 3	Panel 4
Appropriate length	92.3%	84.6%
rippropriate tengui	(24)	(22)
Too long	7.7%	3.8%
loo long	(2)	(1)
Don't know/No opinion	0.0%	11.5%
	(0)	(3)

Table 3a: Did the survey form seem to be an appropriate length or too long when you first looked at it?

Perceived Burden vs. Actual Burden

As noted by Stang and Jones (2011), the perceived burden affecting the respondent's decision to respond, and the actual burden associated with responding (reported time to complete form), did not correlate. While changes to the grid reduced the number of physical pages of the survey and the overall perceived burden, the actual reported time required to complete the form actually increased after Panel 1, while the percentage of respondents reporting that the survey was easy to complete remained relatively unchanged for Panels 3 and 4.

Table 3b: Reported actual time to complete survey, by establishment size

	Reported	Easy to		
	Small	Medium	Overall	complete
Panel 1	9 min	16 min	12 min	80%
Panel 2	13 min	26 min	21 min	96%
Panel 3	17 min	29 min	27 min	77%
Panel 4	20 min	34 min	24 min	81%

	Reported	Easy to			
	Small	Medium	Overall	complete	
	<30 Minutes	>30 Minutes	Cannot Remember		
Panel 5	56%	39%	5%	89%	

To minimize the questionnaire's length, grids (also referred to as a table or matrix) were used in two sections of the form: the green-activities section and the occupations-and-wages section. Since the occupations-and-wages grid has been used successfully in another BLS survey (Occupational Employment Statistics) for many years, the focus of this paper is on steps taken to improve reporting using the green-activities grid. The use of a grid in form design is discussed next.

Use of a Grid to Display Questions

Grids have long been viewed as being more difficult to complete than individual questions on printed self-administered questionnaires (DeMaio et al., 1987), a concern that has carried forward to web surveys (Callegaro et al., 2009; Dillman et al., 2009). However, other research has shown that grids can be effective when compared to alternative approaches (Moyer, 1996). Moreover, despite their potential drawbacks, grids do allow for an efficient use of space, and they have been widely, and successfully, used in both business and household surveys.

However, the development of a grid for this survey posed new challenges because of the length and complexity of the definition of green technologies and practices and the need to present concrete examples, which had been identified in the feasibility study.

Figure 1a shows a portion of the initial version of the green-activities grid, as well as the final version that resulted after the pilot testing. This section of the survey form presented a list of green technologies and practices and asked the respondent to indicate if any of these were used in the sampled establishment.

The entire grid asked about six specific green technologies and practices and provided space for the respondent to enter two additional green practices if the pre-selected list was not adequate. If a respondent indicated that a green technology or practice was used and some workers devoted at least half of their time to it, details about the occupation (title, chief duties, and wages) were collected in the occupations and wages grid shown in Figure 2a. As mentioned previously, the grid shown in Figure 2a is very similar to a grid currently used in forms completed by respondents to the BLS Occupational Employment Statistics (OES) survey.

Figure 1a: Original and Final Versions of a Part of the Green-Activities Grid*

Panel 1 Version of a Section of the Green Technologies and Practices Grid



*Note: This grid was preceded by completion instructions, which are not shown.

Panel 4 Version of a Section of the Green Technologies and Practices Grid



Figure 2a: Original and Final Versions of Part of the Occupations-and-Wages Grid



Panel 1 Version of a Section of the Occupations-and-Wages Grid

Panel 4 Version of a Section of the Occupations-and-Wages Grid



The preceding figures illustrate some questionnaire design features that were added to the form during the pilot testing. These include the use of arrows, brackets, and instructions presented in text boxes. The directional arrows, brackets, and text boxes were used to draw attention to skip patterns, as well as to relevant instructions. The importance of these types of visual cues for reducing item non-response are described in Dillman et al. (1999) and Dillman and Christian (2002).

Both grids underwent a series of other changes during the pilot tests (for example, slight question wording changes, formatting of examples, use of bolding, revised instructions) during the pilot testing, which will not be described in this paper. These changes were influenced by itemnonresponse rates, as well as feedback provided by respondents in the response analysis surveys.

Item Non-Response in the Green Technologies & Practices Grid

As shown in the next table, item non-response for the pre-listed green technologies and practices ranged from a low of 1.0 percent to a high of 6.9 across the first four panels. As an aside, non-response for Column A in the production survey was consistently lower. The grid used in Panel 1 was originally four columns, which was expanded to five columns in Panel 2 when examples of green technologies and practices were placed in a separate column to provide better visual separation and to reduce the overall length of the grid.

	Co	lumn A	, Non-F	Respons	e			Colun	ın B, N	on-Res	ponse	
Item	P1	P2	P3	P4	P5	Prod	Item	P1	P2	P3	P4	P5
1	1.6% (4)	4.0%	2.4%	1.0%	*	.7% (61)	1	1.2% (3)	4.5% (9)	3.1% (9)	1.4% (4)	*
2	(1) 1.6% (4)	5.4% (11)	(.) 1.4% (4)	(c) 1.0% (3)	*	.9% (77)	2	5.2% (13)	8.9% (18)	5.2% (15)	5.9% (17)	*
3	1.2% (3)	4.0% (8)	2.1% (6)	1.7% (5)	*	1.0% (83)	3	2.0% (5)	6.9% (14)	4.2% (12)	2.4% (7)	*
4	3.2% (8)	6.9% (14)	1.7% (5)	2.1% (6)	*	1.2% (102)	4	4.4% (11)	7.4% (15)	3.1% (9)	2.4% (7)	*
5	2.0% (5)	4.0% (8)	2.4% (7)	1.4% (4)	*	1.0% (84)	5	5.6% (14)	8.4% (17)	8.7% (25)	5.2% (15)	*
6	3.6% (9)	5.9% (12)	3.5% (10)	2.1% (6)	*	1.4% (116)	6	4.4% (11)	8.4% (17)	5.5% (16)	4.8% (14)	*

Table 1.	Itom Non Ros	nonce for the	Groon Activ	itias Grid
1 abie 4.	Item Non-Kes	polise for the	Green Activ	mes ond

*P5 ---Panel 5 is the panel that offered use of the Internet, and there was no item non-response.

If businesses reported using a green activity, they were then asked to indicate in Column B of the grid whether or not they had any employees involved in researching, developing, maintaining, using, or installing technologies or practices to lessen the environmental impact of their establishment, or if they were involved in training the establishment's workers in these

technologies or practices (the criterion was that a worker had to spend more than 50 percent of his/her time on this activity).

If a respondent indicated a business used an activity by checking yes in Column A, but failed to enter a response in Column B, it was considered an item non-response in Column B. Across all grid items and panels, item non-response rates for Column B ranged from a low of 1.2 in Panel 1 to a high of 8.9 percent in Panel 2. Also, non-response for Column B was generally higher than for Column A. The data in Table 4 are shown graphically in the next figure.



Figure 3: Column A Item Non-Response Across Panels

Figure 4: Column B Item Non-Response Across Panels



As can be seen in the preceding figures, there was a consistent rise in Column A's non-response in Panel 2 when another column was added to the grid and the grid was condensed, with item non-response more than doubling in Column A for all but Item 6. This increase in non-response also occurred for Column B. A conjecture at this point is that the more condensed grid format may have contributed to the increase in item non-response.

To address the increase in item non-response between Panels 1 and 2, as well as the observation that some respondents were writing "not applicable" or "does not apply" on the form by a green technology or practice, a "does not apply" response option was added in Panel 3 to the yes/no choices in Column A. In addition, brackets and arrows were used to reinforce the instruction that if the answer was yes in Column A, an answer was desired in Column B (see Figure 1a).

As can be seen in the table and figures, with one exception (Item 5, Column B), this led to a consistent drop in item non-response across all items and for both Columns A and B of the grid in Panel 3. Moreover, item non-response remained steady or decreased for Panel 4 when a random sample of QCEW establishments was first used. Results for Panel 5 are not reported because this panel offered a web-reporting option, along with the paper questionnaire, and the web-reporting option did not use a grid to present the questions nor did it allow incomplete data submissions.

RAS Results

As mentioned previously, response analysis surveys were conducted with a small sample (about 26) of both respondents and non-respondents in each panel. Because respondents were expected to be more cooperative, their interviewing script was more ambitious and asked a variety of questions about the survey procedures, contact strategies, materials, the number of people required to complete the form, and specific features of the form. On the other hand, based on past experience, we knew that conducting a RAS with non-respondents would be challenging so the interviewing script had limited objectives. Its key objective was to engage the respondent in conversation and encourage open-ended responses so that the interviewer could probe, if possible, to determine reasons for non-response and what the respondent's view was toward the survey and form.

Respondents were selected for the RAS in Panels 1 & 2 based on whether they had returned the survey immediately, or waited very long in the six week collection cycle to respond. Respondents were selected for the RAS in Panels 3 & 4 if they appeared to have difficulty completing the form or provided inconsistent answers. RAS samples for Panels 1-3 were split between the expected green businesses and the randomly sampled units.

Modifications were made to the RAS for both respondents and non-respondents after each panel. This was done to address changes in the form and to fine tune the questions to address procedural questions or issues that arose during data collection (for example, how the form was handled in the establishment). The small sample size of both RASs (respondent and nonrespondent) precludes making any generalizations about the larger population. Instead, the RASs were used to confirm that procedures were working as intended, as well as to uncover potential problems in form design, instructions, procedures, or materials. When possible, results from the RAS were interpreted in the context of other sources of information (for example, a visual review of the questionnaires, item non-response analyses, and professional judgment). Some selected results from the RAS follow.

Did the Respondent Read the Cover Letter?

Because it was important to clearly explain the purpose of the survey, one question asked if the respondent read the cover letter. The relatively poor response reflected in Panel 1 for reading the cover letter led to some changes in the formatting and appeal used, which appeared to have a positive impact in Panel 2, but the improvements did not hold in Panels 3 and 4.

Panel	Read	Skimmed	Did not read	Total
1	44.0%	48.0%	8.0%	100.0%
1	(11)	(12)	(2)	(25)
2	77.8%	22.2%	0.0%	100.0%
2	(14)	(4)	(0)	(18)
2	29.4%	58.8%	11.8%	100.0%
3	(5)	(10)	(2)	(17)
1	52.6%	26.3%	21.1%	100.0%
4	(10)	(5)	(4)	(19)

Table 5: Did you read the cover letter?

RAS results from Panel 1 for the cover letter also indicated that the mailing package should appear more official and that the cover letter should be revised to emphasize that all businesses, even those without any green technologies and practices should respond.

Perceived Importance of Survey

Because of the importance of salience, or interest in the survey topic as a moderating variable, a RAS question asked the respondent how important it is to measure the growth of green jobs. As can be seen in the table that follows, respondents generally viewed the survey as important with between 42 and 52 percent saying it was very important, and between 23 and 54 percent saying it was somewhat important. Of course, a selection bias is at play here since these are respondents who chose to respond to the survey, and who also participated in the follow-up RAS.

Panel	Very Important	Somewhat important	A little important	Not at all important	Total
1	48.0%	28.0%	12.0%	12.0%	100.0%
1	(12)	(7)	(3)	(3)	(25)
2	42.3%	53.8%	0.0%	3.8%	100.0%
Ζ.	(11)	(14)	(0)	(1)	(26)
2	53.8%	23.1%	19.2%	3.8%	100.0%
5	(14)	(6)	(5)	(1)	(26)
4	52.2%	43.5%	4.3%	0.0%	100.0%
4	(12)	(10)	(1)	(0)	(23)

Table 6 [.]	How im	portant is	it to	measure the	growth o	f green	iobs?
1 uoie 0.	110 w IIII	portant 15	11 10	measure me	giowin o	'i green	1005.

Initial Reaction to the Form and Ease of Completing the Questionnaire

In Panels 1 and 2, respondents were asked if the form was easy or hard to complete. As shown in the next table, the Panel 1 form was viewed as hard to complete by a relatively large number of respondents (19.2 percent), but this figure dropped substantially in Panel 2 to only 3.8 percent.

Table 7: Please tell me your general reactions to the form. Did you think it was easy or hard to complete?

	Easy	Hard	Sample Size
Panel 1	80.8% (21)	19.2% (5)	26
Panel 2	96.2% (25)	3.8% (1)	26

In Panels 3 and 4, the decision was made to first ask an open-ended question to obtain more general reactions to the form, followed by an "ease of completing" question with more scale values than used in Panels 1 and 2.

As shown in the next table, the open-ended question resulted in a less favorable impression of the questionnaire, with 27 percent of respondents offering a negative comment in Panel 3, a figure which was reduced to 8 percent in Panel 4. In Panel 4 though, it's worth noting that a sizeable number of respondents either could not recall their initial reactions or reported no reactions. In addition, as shown in Table 9, a sizeable number of respondents (23% in Panel 3; 15% in Panel 4) did not think that the survey form looked easy to complete (the grid appeared on Page 2).

	Panel 3	Panel 4
Appeared simple/Straightforward/Easy	34.6%	23.1%
	(9)	(6)
Appeared overwhelming or confusing/	26.9%	7.7%
Appeared long or time consuming	(7)	(2)
Displeased about receiving yet another	7.7%	19.2%
survey to fill out	(2)	(5)
Don't recall/No reaction	11.5%	34.6%
	(3)	(9)
Other	19.2%	15.4%
	(5)	(4)
Sample Size	26	26

Table 8: What was your general reaction when you first looked at the green jobs survey form?

*Number of respondents is shown in parentheses

	Panel 3	Panel 4
Very easy to complete	15.4%	26.9%
	(4)	(7)
Easy to complete	61.5%	53.8%
	(16)	(14)
Not very easy to complete	23.1%	15.4%
r in the second s	(6)	(4)
Not easy at all	0.0%	3.8%
	(0)	(1)
Don't know/No opinion	0.0%	0.0%
	(0)	(0)
Sample Size	26	26

Table 9: Would you say that the green jobs survey form looked (read responses)...?"

How Clear Were the Descriptions of Green Technologies and Practices?

In Panels 3 and 4, questions were asked to determine how clear the descriptions of green technologies and practices were to respondents. As shown in the next table, over 80 percent of the respondents in each panel found the definitions to be clear, with between 8 and 12 percent not being able to recall them.

Table 10: Were the descriptions of green activities clear or unclear?

	Panel 3	Panel 4
Clear	80.8	88.5
	(21)	(23)
Unclear	7.7	3.8
	(2)	(1)
Do not recall	11.5	7.7
	(3)	(2)
Sample Size	26	26

Did Respondents Read Each Green Activity Before Responding?

Since we hoped to minimize item non-response in the grid, in Panels 3 and 4, respondents were asked if they read each green activity before answering the questions.

As can be seen in the next table, respondents reported that they read the descriptions of green technologies and practices.

	Panel 3	Panel 4
Yes	96.2%	100.0%
105	(25)	(26)
No	0.0%	0.0%
Do not recall	3.8% (1)	0.0%
Sample Size	26	26

Table 11: Did Respondents Read Each Green Activity Before Responding?

Non-Respondent RAS

The main reason non-respondents gave for not completing the survey was that it would take too much time to complete or the perceived burden was too high. As can be seen in the next table, this response remained fairly consistent across the four panels, ranging from a low of about 35 percent in Panel 1 to a high of 44 percent in Panels 2 and 3. Other reasons for not participating that remained fairly consistent across the panels were the expression of anti-government attitudes, existence of a company policy not to participate in conjunction with the policy of only completing mandatory surveys, and lack of relevance of the survey topic to the business.

Getting the form to the right location was an issue in Panels 1 and 2, but changes in procedures seemed to be effective because this disappeared as a reason for non-response in Panels 3 and 4.

Reason Given	Panel 1	Panel 2	Panel 3	Panel 4	Overall
Time/burden – would take too	34.6%	44.0%	44.0%	36.0%	39.6%
much time to complete	(9)	(11)	(11)	(9)	(40)
Anti-government attitudes	11.5%	8.0%	12.0%	12.0%	10.9%
	(3)	(2)	(3)	(3)	(11)
Company policy not to participate	11.5%	32.0%	20.0%	52%	28.7%
	(3)	(8)	(5)	(13)	(29)
Only do mandatory surveys	3.8%	0.0%	4.0%	4.0%	3.0%
	(1)	(0)	(1)	(1)	(3)
Topic not relevant to this business/Do not use green technologies or practices	7.7% (2)	12.0% (3)	20.0% (5)	8.0% (2)	11.9% (12)
Forwarded to another location	19.2%	16.0%	0.0%	0.0%	8.9%
	(5)	(4)	(0)	(0)	(3)

Table 12: Can you tell me why you decided not to complete the survey?

*Respondents could have given more than one reason for non-response.

Other Changes of Interest Made to the Form

As noted by Griffin and Clark (2009), changes that appear to be superficial on a selfadministered survey form (for example, adding white space or removing horizontal lines between questions) can result in a significant increase in item non-response. Therefore, several other changes that were made during the testing of the survey form were evaluated. 'Before' and 'after' versions for some of these changes are shown in the next series of illustrations.

One of these changes included a question that asked if the description of the business's main products or services was accurate.

Figure 5: Revisions to the Question Wording and Formatting for Q3 - Main Products or Services

Original Version (Panel 1)

Enter	Renting consumer electronics equipment and appliances,
corrections>	such as televisions, stereos, and refrigerators. Included in this
here	industry are appliance rental centers.

Final Version (Panels 2-4)

Our records show that your main products or services in the space p	vices are related rovided and cont	to those listed below. If they are not, please inue with the rest of the survey.	
Renting consumer electronics equipment and appliances, such as televisions, stereos, and refrigerators. Included in this industry are appliance rental centers.	Enter corrections		

As shown in the next table, about 45 percent of respondents changed the pre-printed entry in Panel 1. However, an analysis revealed that many of these text entries were unnecessary, and probably resulted because the respondent did not carefully read either the instruction that followed the bolded question or the existing description that appeared to the right of the entry field/box (see Figure 3).

Therefore, with the objective of reducing unnecessary changes to the products or services description, the order of the description and location of the entry field was changed in Panel 2, so that assuming a respondent engaged in normal left-to-right reading behavior on the form, the respondent would first encounter the description and, hopefully, read it. In addition, the question was expanded so that what had been a separate instruction in Panel 1 was directly incorporated into the question.

As can be seen in the table below, these changes resulted in statistically significantly fewer corrections to the description in Panel 2. The percentage bounced up in Panel 3, but then dropped again in Panel 4. For the production survey, 11.7 percent of usable responses had text entered in the changes box, and many of these were simply providing a finer level of detail for the industrial activity not a wholesale change.

Table 13: Percentage of Respondents Changing the Description of the Main Products or Services

Panel	Percentage of Respondents Changing the NAICS Description	No. of Completes
1	44.7%	244
2	22.7%	194
3	36.9%	268
4	29.8%	265

Number of Green Employees

Another important question on the form asked about the number of green employees, defined as those employees who spent more than half of their time on green activities. As noted in the next figure, the question about the number of green employees was located at the bottom of the page in Panel 1, a location which some research has suggested could conceivably lead to increased error rates (Dillman et al., 1999). In Panels 2-4, this question was moved to the top of the next page, but instead of item non-response improving, as previous research might suggest, it increased slightly, and then remained in a relatively narrow range.

Figure 6: Revisions to the Location and Formatting of Instructions for Question 5

Panel 1 Version (bottom of Page 3 of the survey form is shown)

						Yes	No	[Yes	No	
5 How Please cover of are inc more th	many emplo provide the tota If this form work luded in Column han half of their	yees spent I number of em ing on the gree B more than or time on green a	more than ployees at this n activities for nce. Please do activities please	establishment which you ans not include co e enter zero.	eir time er that spent m wwered Yes i n ontractors or	ore than half Column B. Consultants no	green activ of their time in Please count ot on your estal	i ties at t the pay p each em lishment's	your est eriod includ iployee on i payroll. I	ablishment ded in item 2 or I ly once , even f no employees	the if they spent



Panel 4 Version (bottom of Page 3 and the top of Page 4 are shown)

Item Non-response

As can be seen in the following table, despite making a wide variety of changes to the formatting of the question and instructions, item non-response fluctuated very little among the panels.

	Pa	nel 1	Panel 2		Panel 3		Panel 4	
	%	N	%	Ν	%	Ν	%	Ν
Missing/Blank	3.2%	8	4.0%	8	5.2%	15	3.8%	11

Table 14: Item Non-Response for Question 5 (Number of Green Employees)

Contact Information

Accurate contact information was important for sample cases that required follow-up calls. Since item non-response for contact information was higher than desired in Panel 1, the order of the contact information was changed in Panel 2, so that contact information for the respondent providing information about green technologies and practices appeared on the left. However, this change did not significantly reduce non-response (see Table 15).

Therefore, in Panels 3 and 4, the question wording was changed slightly to make it more conversational, and the formatting was changed to make the entry fields stand out better. This was done by right justifying the labels for the entry fields and moving them outside the entry field, bolding the labels, and then blocking and aligning the entry fields (see Figure that follows).

Figure 7: Revisions to the Contact Information

Panel 1 Version

Primary Contact:	Secondary Contact:
E-Mail:	E-Mail:
Department;	Department;
Job Title;	Job Title;
Business Website:	Business Website:
Phone number: ()	Phone number: ()

Panel 2 Version

s who completed this form. For Occupation and Wage Information:
Contact Name:
E-Mail:
Department:
Job Title:
Business Website:
Phone number: ()

Panels 3 & 4 Version

7 Please tell us who provided the information on this form for:	
Green Technologies and Practices:	Occupations and Wages:
	Check here, if the same person as green technologies & practices
Contact Name:	Contact Name:
E-Mail:	E-Mail:
Job Title:	Job Title:
Business Website:	Business Website:
Phone Number:	Phone Number:

Item Non-response

As shown in the next table, these changes led to a significant reduction in item non-response in Panels 3 and 4 for name, title, and phone number.

Panel	Ν	Contact Name	E-mail	Title	Website	Phone
1	244	14.3%	25.0%	19.3%	38.5%	18.0%
2	194	18.0%	24.2%	28.9%	37.6%	25.8%
3	268	7.8%	19.4%	12.3%	29.5%	10.4%
4	265	10.2%	24.9%	14.0%	43.0%	11.3%

Table 15: Item Non-response for Contact Information Listed in Left Column

Discussion

Because time and resources precluded the use of large, embedded field experiments, the research described in this paper relied heavily on a development process known as "action research" (Lewin, 1946). As the following diagram indicates, this model relies on an iterative process where survey methodologists and content specialists work together using a variety of sources of quantitative and qualitative information to determine if research objectives are being met.

Figure: The Action-Research Model



Key objectives in this effort were to develop a usable questionnaire that minimized measurement error and maximized response. To accomplish these objectives, five separate panels were used to test and refine the questionnaire and field procedures.

The decision to use five field-test panels once a questionnaire had been drafted versus some other number reflected the unique demands of this research. Since BLS had no prior experience measuring the concept of jobs associated with green technologies and practices, we knew there would be many operational and procedural issues that would need to be resolved, in addition to measurement issues associated with the questionnaire.

In addition to standard measures such as item non-response, an important tool used in this approach was the Response Analysis Survey (RAS), which provided timely feedback about the success of changes made to the form, field procedures, and data quality. Quantitative tools such as item non-response may point out a problem, but the reasons for that non-response may remain unclear. The RAS, along with the judgment and experience of the survey team, helped to uncover possible reasons as well as possible corrective actions. For maximum effectiveness, the timing of a RAS is important. The effort was made in each panel of this study to minimize the lag time between completion of the interview and the RAS so that a respondent's recall was as fresh as possible.

Although larger sample sizes for a RAS are generally recommended (Phipps et al., 1995), a lack of resources and time led to the use of small sample sizes in this effort. These small sample sizes turned the RAS into a qualitative research tool whose purpose was to identify potential sources of response and measurement error. In addition, as noted previously, although some core items were kept across all panels, the RAS scripts were revised after each panel to address new issues that arose or to more deeply probe issues raised in a previous panel. Moreover, acknowledging the resistance that would occur from non-respondents, the non-respondent RAS used a highly streamlined approach that relied heavily on a conversational approach with open-ended questions and follow-up questioning by interviewers who probed reasons for non-response as deeply as possible.

An effective use of the action-research model is characterized by open communication among content specialists and survey methodologists, the use of both quantitative and qualitative research, and reliance on past research, or when that is lacking, personal experience. The current research is an example of the use of sequential mixed-methods research (Driscoll, et al. 2007), as well as triangulation (Jick, 1979), which in this case refers to the use of multiple sources of information to provide different perspectives and, hopefully, improved validity of results.

Although the mixed-methods approach resulted in a variety of positive changes that could be documented using quantitative measures (for example, reductions in item non-response), relying on qualitative research is not without risks. For example, a message that seemed to come across loud and clear during the feasibility study and associated RAS was that respondents wanted an Internet-reporting option. For example, in debriefings, respondents frequently wondered why a survey about green jobs relied on a paper instrument.

To address this issue, BLS developed an Internet-reporting option, which was tested in Panel 5. However, when this option was offered along with a paper option, only 3.4 percent of respondents chose to complete the form using the Internet mode. As noted in Jones and Stang (2012) ten percent of usable survey responses in the production survey were completed via Internet with significant differences seen by size of establishment.

One could argue that the low take-up rate for the Internet mode reflects the increased burden associated with using the Internet such as the requirements to create an account and complicated passwords, but as Mockovak (2010) discovered, a low take-up rate can occur even when the Internet-reporting system does not require a password and a simple, short form is used.

Although establishment respondents can be successfully pushed to use the Internet to complete long or complex questionnaires, the cost even in a mandatory survey may be an increase in overall unit non-response (Downey et al. 2007). Given an initial choice of multiple response options that include a web-reporting mode in a voluntary survey, many establishment respondents will still opt to use a paper survey form.

As discussed previously, a key objective of this research was the development of an effective, usable grid for obtaining information about a respondent's use of green technologies and practices. As noted by Christian and Dillman (2004), respondents' answers to questions in self-

administered surveys are influenced by more than words, and it's important to take the visual presentation of questions into account when designing surveys. Our revisions to the grid across the iterations of field testing showed that adding a "does not apply" response option to Column A and directional arrows to highlight the subsequent skip pattern led to a significant drop in item non-response for both columns in the grid (see Table 4).

References

Bogen, K. (1996). The Effect of Questionnaire Length on Response Rates – A Review of the Literature. Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association, pp. 1020-1025.

Callegaro, Mario; Shand-Lubbers, Jeffrey; and Dennis, J. Michael (2009) "Presentation of a Single Item Versus a Grid: Effects on the Vitality and Mental Health Subscales of the SF-36v2 Health Survey. Paper presented at European survey research association conference 2009, ESRA, Warsaw. WebSM, Web Survey Methodology, <u>http://www.websm.org</u>.

Christian, Leah M. and Dillman, Don A. (2004), "The Influence of Graphical and Symbolic Language Manipulations on Responses to Self-Administered Questions," *Public Opinion Quarterly*, 68, No. 1, 57-80.

DeMaio, T. J., Martin, E. A. and Sigman, E. P. (1987). "Improving the Design of the Decennial Census Questionnaire." Paper presented at the Annual Meeting of the American Statistical Association.

Dillman, D.A. (1978) Mail and Telephone Surveys: The Total Design Method. John Wiley & Sons.

Dillman, D. A., Sinclair, M. and Clark, J. (1993), "Effects of Questionnaire Length, Respondent-Friendly Design, and a Difficult Question on Response Rates for Occupant-Addressed Census Mail Surveys," *Public Opinion Quarterly*, 57, 289-304.

Dillman, D. A., Redline, C. D., Carley-Baxter, L. R. (1999). "Influence of Type of Question on Skip Pattern Compliance in Self-Administered Questionnaires. Paper presented at the Annual Meeting of the American Statistical Association.

Dillman, D. A. and Christian, L. (2002) "The Influence of Words, Symbols, Numbers, and Graphics on Answers to Self-Administered Questionnaires: Results from 18 Experimental Comparisons," WebSM - Web Survey Methodology: http://www.websm.org/.

Dillman, D.A., Smyth, J.D. & Christian, L.M. (2009). *Internet, Mail, and Mixed-Mode Surveys. The Tailored Design Method*. (Third ed.) Hoboken: John Wiley & Sons, Inc.

Downey, K; McCarthy, D. and McCarthy, W. (2007) "Encouraging the Use of Alternative Modes of Electronic Data Collection: Results of Two Field Studies." Paper presented at the Annual Meeting of the American Statistical Association.

Driscoll, D., Appiah-Yeboah. A., Salib, P., and Rupert, D. (2007). "Merging Qualitative and Quantitative Data in Mixed Methods Research: How To and Why Not," *Ecological and Environmental Anthropology*, Vol. 3, No. 1, 19-28.

Featherstone, F. and Moy, L. (1990) "Item Nonresponse in Mail Surveys." Paper presented at the International Conference of Measurement Errors in Surveys, Tucson, Arizona.

Fienberg, Stephen E. and Tanur, Judith M. (1989) "Combining Cognitive and Statistical Approaches to Survey Design" *Science*, 24 February 1989: Vol. 243 No. 4894, 1017-1022.

Galesic, M. (2003). Effect of questionnaire length on response rates. *Social Research (Drustvena istrazivanja)*, 12, 5, 807-824.

Griffin, Deborah h and Clark, Sandra L. (2009) "Respondent Effects Associated with Questionnaires Designed to Accommodate Survey Processing." Presented at the annual conference of the American Association for Public Opinion Research, Hollywood, FL.

Jenkins, Cleo R. and Dillman, Don A. (1997) "Toward a Theory of Self-Administered Questionnaire Design." In Lyberg, et al. (eds.), *Survey Measurement and Process Quality*, 165-198, John Wiley & Sons, Inc.

Jick, T. D. (1979). "Mixing Qualitative and Quantitative Methods: Triangulation in Action," *Administrative Science Quarterly*, Vol. 24, No. 4, Qualitative Methodology (Dec., 1979), 602-611.

Jones, C. K. and Stang, S. (2012). "Counting Green Technologies and Practices (GTP) Jobs." Paper presented at the Fourth International Conference on Establishment Surveys.

Lewin, K. (1946) Action Research and Minority Problems, Journal of Social Issues, Vol.2.

Mockovak, W. (2010). "Using an Action-Research Model to Move from Conversational to Hybrid Standardized Interviewing: A Case Study." Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association.

Mockovak, W. (2011) "The Impact of Visual Design in Survey Cover Letters on Response and Web Take-up Rates." Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association.

Moyer, Laureen H. (1996) "Which Is Better: Grid Listing or Grouped Questions Design for Data Collection in Establishment Surveys?" Proceedings of the Section on Survey Research Methods, Alexandria, VA: American Statistical Association, pp. 986-991.

Phipps, Polly A; Butani, Shail J.; and Chun, Young I. (1995) "Research on Establishment-Survey Questionnaire Design," *Journal of Business & Economic Statistics*, Vol. 13, No. 3 (Jul., 1995), pp. 337-346.

Redline, Cleo D; Dillman, Don A.; Carley-Baxter, Lisa; Creecy, Robert (2003) "Factors that Influence Reading and Comprehension in Self-Administered Questionnaires," Paper presented at the Workshop on Item-Nonresponse and Data Quality, Basel Switzerland, October 10, 2003.

Stang, S. and Jones, C.K. (2011). "Counting Green Jobs: Developing the Green Technologies and Practices (GTP) Survey." Paper presented at the Annual Meeting of the American Statistical Association.

Panel	Objectives	Key Findings
Panel 1	 Determine response rate, completion time, and general reactions to form and definition of green activities from a sample of 'green' businesses and a random selection of establishments from the QCEW. Determine success of address correction and follow-up efforts, and effort required in follow-up to boost non-response. Assess effectiveness of mail-out methodology and materials. Identify improvements to form and mail-out methodology. Determine number of respondents in each establishment who provided information. Assess quality of information provided. Determine main reasons for non-response. 	 Respondents generally agreed with the definition of green activities, but encountered some problems deciding when workers spent more than 50% of their time on green activities. Form was perceived as too long (perceived burden was high). Item non-response in the green activities grid was too high. Large establishments had a lower response rate. Some sections of the form led to respondent confusion (e.g., reporting location, grid instructions for green activities, examples for green activities grid). Visual review of completion errors indicated some confusion with occupations and wages grid. Mailing should appear more official. Cover letter should be revised to address key respondent concerns. Stress that response is desired from all establishments, even non-green. Item non-response for contact information was high.
Panel 2	 Continue to study reactions from green subsample and a random selection of establishments. Determine if form is getting to right person. Determine if questionnaire-design changes from Panel 1 had desired impact. Reduce survey length to 6 pages and determine impact on response rate, completion time, and general reactions to form. Assess effectiveness of mail-out methodology and materials. 	 Respondents generally agreed with the definition of green activities, but some disagreement remained about the definition of a green worker. Overall response improved by 13%. Some confusion remained about which locations to include. Large establishments had a lower response rate. Still some difficulty getting form to the right person. There was a drop from Panel 1 in the number of corrections reported in pre-identified products or services (location of correction box had been changed from Panel 1). Item non-response in the green activities grid for Column A increased. Visual review of forms indicated some continued difficulty with the occupation and wages grid. 4.0 percent of respondents left the question about the number of green employees blank. Item non-response for contact information was lower than in Panel 1, but still too high at 18 percent.
Panel 3	 Continue to study reactions from green subsample and a random selection of establishments. Determine if questionnaire-design changes from Panel 2 had desired impact Assess effectiveness of adding a "does not apply" response in the green activities grid. Add brackets and directional arrows to highlight skip pattern in the green activities grid. Move instructions about reporting location 	 Survey response rate showed a slight improvement (2%) over Panel 2. The number of "return to sender" envelopes was reduced to less than 9 percent versus 16 and 13 percent in Panels 1 and 2. As in previous panels, larger establishments had lower response. Revisions to the NAICS descriptions dropped. Item non-response in the green activities grid dropped, which indicated the changes made (addition of NA response and directional arrows) had the

Appendix A – Summary of Panel Objectives and Key Findings

Panel	Objectives	Key Findings
	 to top of page. Change formatting of instructions in the occupation and wages grid to address completion problems. In the question asking for the number of green employees, add a directional arrow pointing to instructions. RAS for respondents was revised to collect more detail about respondents' reactions to form and steps taken to complete it. 	 desired effect. Item non-response to the question about the number of green employees was still higher than desired. A small number of respondents continued to have difficulty providing the information requested in the occupations and wages grid. For example, 51 percent of respondents correctly entered the total number of employees in an occupation versus about 33 percent in Panel 2. The format for contact information was changed and the amount of item non-response was reduced compared to Panels 1 and 2. 77 percent of RAS respondents said the form looked easy or very easy to complete. 32 percent of the completed surveys could be directly attributed to the non-response follow-up effort.
Panel 4	 Did pre-contacting establishments to verify a mailing address or identify a contact person hurt or help? To more closely mirror the production survey, study response and reactions only from randomly selected QCEW establishments (no special sample of pre- identified green establishments). Continue to monitor response to the green activities and occupations and wages grids. 	 Response rate was 13 percent higher for establishments that had been pre-contacted. The number of "return to sender" envelopes was just over 9 percent versus slightly less than 9 percent in Panel 3, so the maximum benefits of the address cleaning operation appears to have been met. 18 percent of total response was by fax, which was lower than Panel 3. Item non-response in the green activities grid generally showed a decline, and was comparable with the levels observed in Panel 3. Adding the label 'Job Description' under 'Job Title' in the occupations and wages grid resulted in fewer respondents not giving a job description. A larger percentage of respondents (58%) entered the total number of employees per job in the occupations and wages grid. This was an improvement over Panel 2 (33%) and Panel 3 (51%). Item non-response for the contact information was comparable to Panel 3.
Panel 5	 This panel offered a web-reporting option, in addition to the paper form, a fax-reporting option, and an Adobe-fillable form that could be returned via e-mail. There were no revisions to the Panel 4 form. An experiment was conducted with the instructions for web reporting. Because there was concern that the length of the web-reporting instructions could add to perceived burden and discourage response, Group 1 received a cover letter that referred respondents to enclosed (separate document) web-reporting instructions. These separate web-reporting instructions were not included in the mailing to Group 2. Instead, respondents in Group 2 were referred to a web page (URL) for the instructions. 	 Overall response was down significantly in this panel, but there was no difference between the two experimental groups. A conjecture was that this panel's response was negatively affected by the harsh political climate and budget conflicts in Washington, DC, which dominated media coverage for weeks. Most establishments responded using paper. Depending on the group, between 3.2 and 3.6 percent chose the web-reporting option. This result was not expected given the large percentage of respondents who said they wanted a Web option in the feasibility study. The percentage returns by fax or Adobe-fillable form were even lower.

Panel	Objectives	Key Findings
	• The sample consisted of randomly selected establishments (no special sample of pre-identified green establishments).	