Statistics of Income

2007 SOI Paper Series

SOI Develops Better Survey Questions Through Pretesting

by Diane Milleville and Tara Wells

SOI Develops Better Survey Questions Through Pretesting

Diane Milleville and Tara Wells, Internal Revenue Service

Within the Statistics of Income (SOI) Division of the Internal Revenue Service (IRS) is a small section called Statistical Support, which consists mainly of mathematical statisticians who provide statistical consulting services to areas within the IRS and occasionally other branches of the Federal Government. This group works on various projects, including sampling, quality and performance measurement, surveys and questionnaires, focus groups, usability studies, and Web site analysis. Depending on the needs of the customer, Statistical Support may work on the design of the project, the analysis, or both. Statistical Support (SS) has several primary customers, but also takes on ad hoc projects from new customers, who usually learn about the group through word of mouth.

The Servicewide Policy, Directives, and Electronic Research (SPDER) Division of the IRS is responsible for all electronic research tools available to IRS employees. SPDER currently has contracts with two main electronic research companies: LexisNexis and Westlaw and provides direct access to both of these tools to any IRS employee who may benefit from the information provided. With the current electronic research contracts expiring at the end of 2008, SPDER needed to determine a set of requirements in order to write a statement of work for the new contract that would be instituted in 2009.

A statement of work is the first step required when trying to establish a contract with an outside vendor and is used to find potential companies that can provide the service(s) as described. It is a document that defines the services and requirements to be delivered and/or the outcomes to be achieved through a contract with a vendor. This document must include a well-detailed description of all necessary requirements.

To ensure that all critical research requirements would be included in the statement of work, SPDER decided to survey a group of IRS researchers who use the tools frequently and are extremely competent navigating through them. However, with little experience in survey development, SPDER sought additional help for this task. In the fall of 2006, in order to ensure that this survey was well-designed and would capture all of the necessary information needed to obtain specific, user-based requirements that would satisfy customer needs, SPDER contacted SS for assistance. After an explanation of the project we agreed to help design the survey for SPDER.

▶ Electronic Research Tools, etc.

IRS employees have available a comprehensive suite of tax, legal, news, and business research services under their current electronic research contract. As previously stated, the two main electronic research tools are LexisNexis and Westlaw. Employees use LexisNexis to access a comprehensive assortment of primary and secondary tax research sources, while Westlaw provides employees with access to tax and legal sources, business, and news information.

Specific IRS employees, known as research specialists, serve primarily as inhouse consultants to assist other IRS employees in the use of commercial research services. These employees have received advanced training and are able to answer LexisNexis or Westlaw related questions, which made them the ideal respondents for our survey.

While we did not have any experience in using the electronic research tools, SPDER was able to provide a general overview of both LexisNexis and Westlaw. Prior to meeting with SS, SPDER developed a small set of questions for the survey. In addition to these, SS also created an introduction and a list of possible questions. The phrasing of these questions was finetuned in order to have them as objective and easy to comprehend as possible.

Besides some basic demographics, the questions in the survey covered specific details of the research tools, what aspects of each tool is important to users, if one research tool offers a specific option that they would

like offered in another research tool, if the research specialists go elsewhere for info, and how many times they are contacted for help.

In conjunction with our goal of creating a set of requirements for the next electronic research contract, we believed that the questionnaire responses would be able to verify whether one electronic research tool was preferred over the other or if each electronic research tool exhibited individual benefits of use.

▶ Pretesting

As SOI did not have any prior knowledge regarding this survey topic, we decided to pretest our draft questionnaire in order to validate its accuracy in targeting the correct information. Pretesting, by definition, is showing a product to a sample of the target audience during its development to check whether it is conveying the desired message, in the desired way, with the desired effect. There are many methods of pretesting survey questions, such as cognitive interviews, expert reviews, focus groups, and observational interviews.

After researching different methods of pretesting, we believed that cognitive interviewing would be the most beneficial method to use for our project. Cognitive interviewing consists of one-on-one interviews using a draft questionnaire in which respondents are asked questions about the survey questions. These interviews, which are usually recorded, can identify problems in question comprehension, memory recall, selecting responses, interpreting reference periods, and reactions to sensitive questions. During cognitive interviewing, probing questions are asked, and respondents are encouraged to verbalize their thought processes as they answer. In addition, these interviews can yield information about major problems in the survey design as various respondents repeatedly identify the same questions and concepts as sources of confusion.

▶ Preparation of Pretest

The goal of the pretest was to identify sources of potential nonsampling error, specifically measurement error and nonresponse error, and to identify areas of confusion due to wording of the introduction and questions or sequencing of questions. To ensure that the pretest accomplished these goals, we spent time planning each part of the process.

We decided to select eight research specialists as participants in the pretest; two would be interviewed as part of a practice run to ensure that our process worked properly.

SPDER suggested using a communications application called Centra for the pretesting. This computer application is available to all IRS employees and can easily unite people from all over the nation. It allows the meeting leader to share his or her screen with the participants regardless of their geographic location and control exactly what they see and when they see it. Centra also has a record function, which makes it easy to replay the session to capture more notes.

Once we determined the participants and method of data collection, we developed the materials for the pretest, which included a script that detailed each step of the one-on-one session, and a presentation file that included the survey introduction and questions, exactly as we wanted the participants to view them.

Once the pretesting materials were ready, we set up individual 1-hour appointments with the participants and reserved a private room to ensure the proper setting for leading the pretests.

▶ Pretest Details

Half of the eight research specialists chosen to participate in the pretest were designated to answer questions concerning LexisNexis while the other half were more knowledgeable in using Westlaw. The pretest participants were given a specific time to expect a phone call from SS to begin the pretest. Once the respondent was contacted, introductions were made, and a brief synopsis of the project was provided to the respondent, as well as, a description of what was expected during the pretest. Because we utilized Centra communications software, the participant regardless of location was able to view the actual slides from which SS was reading on his or her own computer screen. This use of Centra also enabled SS to record the interviews, in

case we needed clarification or further detail on the respondents' answers. The approximate time spent on the pretest interview was 45 minutes; however, the pretest gradually became less time-consuming as the interviewers became more familiar with the script and the format of the survey. Two members of SS took handwritten notes during each of the interviews, while a third member conducted the pretesting.

▶ Probing

Interviewers often ask probing questions when conducting a pretest. Probes are used to obtain information concerning the survey content. Probes can be preplanned surrounding specific content and worded in a specific way, but they can also be spontaneous, allowing the interviewer to determine the exact cause of confusion and how the confusion can be alleviated. Both types of probes should be asked in such a way that the interviewer does not influence the participant.

While developing the script for our pretest, we identified several words and phrases in the introduction and questions that we thought might be difficult to understand. We made note to ask each participant what the word or phrase meant to them to see if they interpreted each as intended.

In the introduction, words and phrases such as "requirements," "confidentiality statement," and "statement of work" stood out as potentially troublesome, while, in the survey questions themselves, "content" and "functionality" seemed too vague. We made sure to probe about each of these words or phrases if any of these items was, in fact, difficult to interpret or if the participants were familiar with them and could understand them in the correct context. If participants indicated that they did not understand the item, we would ask them for another word or phrase that would communicate the message we intended to convey.

Comments from these specific probes confirmed that almost all of the items were too difficult or vague for our audience. All participants understood "confidentiality statement" and "requirements," but the other highlighted items caused confusion. Some participants had heard of a "statement of work" but did not have a full understanding of the term, while others had never heard of it. Both "content" and "functionality" were deemed ambiguous, and participants interpreted these terms much differently than intended. When probed, our participants made various suggestions for how to revise the wording to capture the desired information. This is illustrated in more detail in Evolution of a Question.

In addition to specific, preplanned probing, we also used spontaneous probing when a participant showed signs of struggling with the information provided. To do this, we used generic questions, such as: What do you think this question means? Is this word clear to you? How would you reword this question so that it would be clearer? We probed as soon as we recognized that the participant was having trouble understanding the question, which allowed us to determine what caused the confusion while the initial reaction was still present.

Through the use of probes, both preplanned and spontaneous, we identified areas of the survey that were problematic and learned how to rephrase questions to make them easier to understand. The information gained through probing had a positive impact on the refinement of the survey.

▶ Evolution of a Question

Throughout the project, the survey questions evolved, based on comments given by the pretest participants. While some of the question refinement involved merely tweaking the questions or response options, other survey questions were significantly altered, which greatly benefited the survey. The pretest results showed that the original questions did not capture the information we intended to obtain. Below are two examples of how our survey questions evolved and why.

Question 1:

Original Question: Of the content and functionality provided in Westlaw/LexisNexis, which is most important to you?

First Revision: What content provided in Westlaw/LexisNexis is most important to you?

Second Revision: In your opinion, what is the most important feature provided by Westlaw/LexisNexis?

The original question addressed two aspects of the tools: content and functionality. During the first round of pretesting, it became clear that the word "functionality" caused too much confusion and that participants did not understand how to answer the question. They indicated that they correctly understood what we meant by "content," which is why the first revision still contains the word.

During the second round of pretesting, we recognized that "content" was still an issue. Through probing the participants, we learned that the word "feature" was easier to understand, which is why we made that change. Participants also suggested adding the phrase "in your opinion" to make it clear that the answer was related to the survey respondent's own experience.

Question 2:

Original Question: For job-related research, do you sometimes find it easier to perform an Internet search (e.g., Google), instead of using Westlaw or LexisNexis, to locate information? If so, which search engine do you use and why?

First Revision: For job-related research, do you find it easier to use an Internet search engine, instead of using Westlaw or LexisNexis, to locate information?

Second Revision: This question was eliminated from the survey, based on further information provided by pretest participants.

The first group of pretest participants thought that the phrase "perform an Internet search" was not as clear as the phrase "use an Internet search engine." Both participants also felt that asking for the actual search engine would not provide any additional value to the survey. We agreed with both of these observations and revised the question.

During the second round of pretesting, the participants expressed concern with this question because they felt that almost all employees would need to use an Internet search engine to locate information. They named several work-related databases which are outside of the IRS, but are used by IRS employees frequently. Another concern was that IRS employees use Westlaw and/or LexisNexis for one kind of research, but turn to the Internet for another kind of research. This type of Internet usage does not mean that the current electronic research tools are not valuable; it just indicates that they do not cover all information needed. Based on the information gathered through the second round of pretesting, we decided that this question would not gather the type of information that SPDER wanted to collect. We decided to delete this question from the survey.

These two questions changed drastically thanks to the pretesting process. Prior to pretesting the survey, both SS and SPDER thought that these questions were well-worded and would be understood by survey participants; however, the results proved otherwise. These and other results from our pretest greatly improved the quality of our survey.

▶ Pros and Cons of Pretesting in our Survey

In this project, pretesting proved to benefit our survey question development and our understanding of the survey. However, we also found that sometimes pretesting can have certain negative effects on the survey design.

On a positive note, our results from pretesting revealed the differences in question comprehension and aided as a guide to how questions should be revised. Pretesting allowed for questions to be developed to ensure the quality of response and the data collected. As the reviewers, we were given the ability to know the exact areas where the respondents had confusion in the survey. This aspect of pretesting gave SS a sense of foresight that we otherwise would not have had. Another benefit of pretesting was that the survey could be altered again and again until the final version was created. This process of continued editing, with the help of the research specialists, throughout the design phase reduces time needed to be spent on the survey during the

implementation phase. When the time spent on revising is reduced, ultimately total cost is reduced, as well.

However, we also found that, sometimes, pretesting had certain negative effects on survey questionnaire design. For an individual who has full understanding of a statement or question, the probing could have encouraged overanalysis of questions, diction, and semantics. As in any sample, we ran the risk of sampling error with our pretest sample of research specialists inaccurately representing the population. Our pretest was performed using Centra instead of having the respondents face to face with the interviewer. By having the pretest performed via telephone, we were not able to see the respondents' visual cues and reactions to the questions.

Nevertheless, SS was able to benefit from the experience of pretesting a survey. Pretesting helped to give SS and SPDER a firm grasp on how to word questions in order to attain the most important requirements for the statement of work. SS believes that the positive effects of pretesting our survey greatly outweigh the few negative aspects. The generous amount of helpful feedback that aided us in the survey question development would never have been gathered without the use of pretesting.

▶ Conclusion

We believe all the information gathered during the pretesting process helped us design two quality surveys, one for Westlaw and a separate one for LexisNexis, that would garner the type of responses sought by SPDER.

SPDER decided against our recommendation to use the two surveys that we developed and instead created a new survey that combined the two surveys into one. Unfortunately, we are unable to analyze any results from the pretested LexisNexis and Westlaw surveys and how they compared to the responses from the pretested surveys. Some of the questions in the combined survey remained the same, while other questions were changed or deleted and new questions were added. These new questions were not pretested. As consultants, we do not have control over the final product. We can provide our educated opinions and recommendations to our customers, but, ultimately, the final decision lies in their hands.

SS plans to experiment with other types of pretesting and to continue to practice pretesting while using both types of probing. In the future, we hope to incorporate pretesting as an integral part of SS's survey question development.