

Getting your money's worth! Targeting resources to make cognitive interviews most effective

Jaki S. McCarthy

USDA's National Agricultural Statistics Service,
3251 Old Lee Highway, Fairfax, VA 22030

Abstract

Cognitive interviewing has long been hailed as an effective technique to evaluate and improve survey questions. However, cognitive interviews are typically resource intensive and thus conducted on limited sets of questions and with limited sets of respondents. To be most effective, questions that are most likely to have adverse impacts on data quality should be targeted. Respondents most likely to exhibit problems with these questions should likewise be selected for testing. One way to target a subset of questions is to use available information from previous data collections to identify questions with the greatest number of quality problems (e.g. high edit or item imputation rates, greater numbers of requests for assistance answering these questions, etc.) Once a subset of questions has been identified as good candidates for cognitive testing, respondents must also be selected. Again, information from existing data sets can be used to identify characteristics of respondents most likely to exhibit problems. Data mining techniques, such as classification trees, can be used to determine the type of respondents most likely to contribute to low quality responses. These criteria can be used to select respondents for cognitive interviews. Knowing the pertinent characteristics of these respondents may also suggest useful probes that can be included in the cognitive interviews. Once questions have been revised based on the cognitive interviews, the same indicators of quality can be used to measure the improvement in data collection using the new questions. This approach has been employed in making revisions to questions on the Census of Agriculture; a case study provided will illustrate how this is an effective use of scarce testing resources.

Key Words: Cognitive Interview, Pretest, Questionnaire Testing

1. Introduction

Cognitive interviews are widely recognized as rich sources of information on questionnaire performance (see for example, Madans, Miller, Maitland and Willis, 2011 and Willis, 2005). They are widely used by survey researchers to evaluate survey questions and uncover potential problems with their use and administration. However, cognitive interviews are resource intensive, and thus typically include small numbers of respondents. Cognitive interview studies with more than 25 respondents would be considered large. In addition, respondents will likely provide an extensive think aloud narrative while answering and then answer additional follow up probe questions for items

of interest. This lengthens the interview time substantially, thus only a small number of questions can be included. An important, but often unanswered question is: how do you target resources to the right people and focus on the right questions? That is, who should be included in a small cognitive interview sample and for a large questionnaire, which questions should be the focus of investigation?

Typical practice in cognitive interview studies is to try to include a diverse pool of respondents. For NASS surveys we will typically try to include different types of farm operations (perhaps some with livestock, some with crops), different size operations, and in different locations. If we know other characteristics may be relevant to the topic of the questionnaire, we may also try to take that into consideration. For example, when testing a questionnaire about agricultural workers, farms known to have different types of workers can be included. However, cognitive interview samples are typically ad hoc and often simply convenience samples.

Of course, a convenience sample in no way ensures that any of the cognitive interview sample will provide useful insight into potential problems with the tested questionnaire. Imagine a survey in which errors on a particular question occur 25% of the time. A typical cognitive interview sample might be 10 respondents. If we select those respondents randomly, we might expect 2 or 3 of them to have these reporting errors. However, this is only if the errors are also distributed randomly. Useful cognitive testing rarely seeks to find random error; most errors of interest are clustered.

Imagine a hypothetical survey asking questions of Justin Bieber concert goers. Perhaps we are interested in improving questions about whether concert tickets are worth the money paid for them, or the quality of the sound at Mr. Bieber's concerts. A typical cognitive interview sample might be selected by picking a fan at a Midwest concert, another fan on the East Coast or in another part of the country, then we might pick a girl who appeared to be crying in the crowd, as well as one who was screaming and then one who seemed subdued. However, it is likely that these (and even a majority) of people we might pick because they are convenient, will say the tickets are worth the money and the sound is great, regardless of how we word these questions. Instead, we might need to test questions with fan parents who've paid for tickets, or less "fanatical" fans (for lack of a better adjective!). Perhaps these respondents will be the only ones without consistent interpretations of the questions, or problems reporting in our answer categories.

Similarly, in a NASS survey with questions about agricultural labor, questions may only be problematic for operations with seasonal labor, migrant workers, or some other particular type of worker. If we include mostly operations that do not have these types of workers, our cognitive interviews will yield little insight into potential problems. Is it really an effective use of resources to include a majority of interviews that show that people can understand and answer the questions correctly?

2. THE CENSUS OF AGRICULTURE CASE STUDY

For many surveys, particularly longitudinal or ongoing data collections, there may be information from previous administrations of the survey or information from similar data collections available. This can be used to identify what errors have already occurred in data collection and the types of respondents that make these errors. This is the approach taken for identifying cognitive interview respondents in order to test the Census of Agriculture (COA) report form. The COA is conducted every five years collecting data for the reference year ending in 2 or 7 (for example, 2007, 2012). It includes questions on land use, agricultural inventory and production, farm economics and operator demographics. The report form is quite long, 24 pages, so think aloud protocols and follow up probes clearly cannot be administered for the complete questionnaire. Any cognitive interviews will have to be conducted on select questions. In addition, the COA mail list is huge, comprising over 3 million names and addresses, so again, we cannot hope to cover all types of operations with a cognitive interview sample. So how do we decide where to focus our efforts to making improvements to the 2012 COA report form?

In order to answer this question for the 2012 COA, we began with a review of the previous 2007 COA. For every COA there is a toll free telephone help line available. The number is listed on all of the COA materials, including the report form. For anyone who calls that line and asks for help completing their report form, we record what section of the form they asked for help with and collect a comment about their request. For the 2007 COA, the majority of calls requesting help to complete the report form were about the land sections of the report form. Of the 99,993 calls for help with the form, 23,780 (24%) were for help with the land sections. In addition, we calculated for each variable on the report form the number of times it had been edited or imputed, either by an automated edit or by an analyst. This way items could be identified that had been changed a high percentage or number of times. Four of the items in the land section were among those identified as particularly problematic. This review also showed that items in the Acreage and Land sections were edited more often than many other items on the form. The “total acres operated” was edited 28.1% of the time it was reported, cropland pasture was edited 59% of the time it was reported, permanent pasture was edited 35.5% of the time and woodland pasture was edited 33.5%. This identified these questions as the focus for our cognitive interviews. The 2007 COA land sections appeared on facing pages in the form and are shown in Figures 1 and 2 below.

Report land owned, rented, or used by you, your spouse, or by the partnership, corporation, or organization named on the front of this form. Include ALL LAND, REGARDLESS OF LOCATION OR USE - cropland, Conservation Reserve Program (CRP) and Wetlands Reserve Program (WRP) land, pastureland, rangeland, woodland, idle land, farmsteads, etc.


	None	Number of Acres
1. All land owned. 0043	<input type="checkbox"/>	
2. All land rented or leased from others , including land worked by you on shares, used rent free, in exchange for services, payment of taxes, etc. Include Federal, State, and railroad land leased on a per-acre basis. Exclude land (i.e. private, Federal, State, railroad, etc.) used on a per-head or animal unit month (AUM) basis under a grazing permit. . . . 0044	<input type="checkbox"/>	
3. All land rented or leased to others , including land worked on shares by others and land subleased. 0045	<input type="checkbox"/>	
4. TOTAL ACRES in this operation for this census - Add items 1 and 2, then subtract item 3. If the entry is zero, please refer to the enclosed Instruction Sheet, section 1. These acres are referred to as THIS OPERATION for the remainder of this report. 		0046

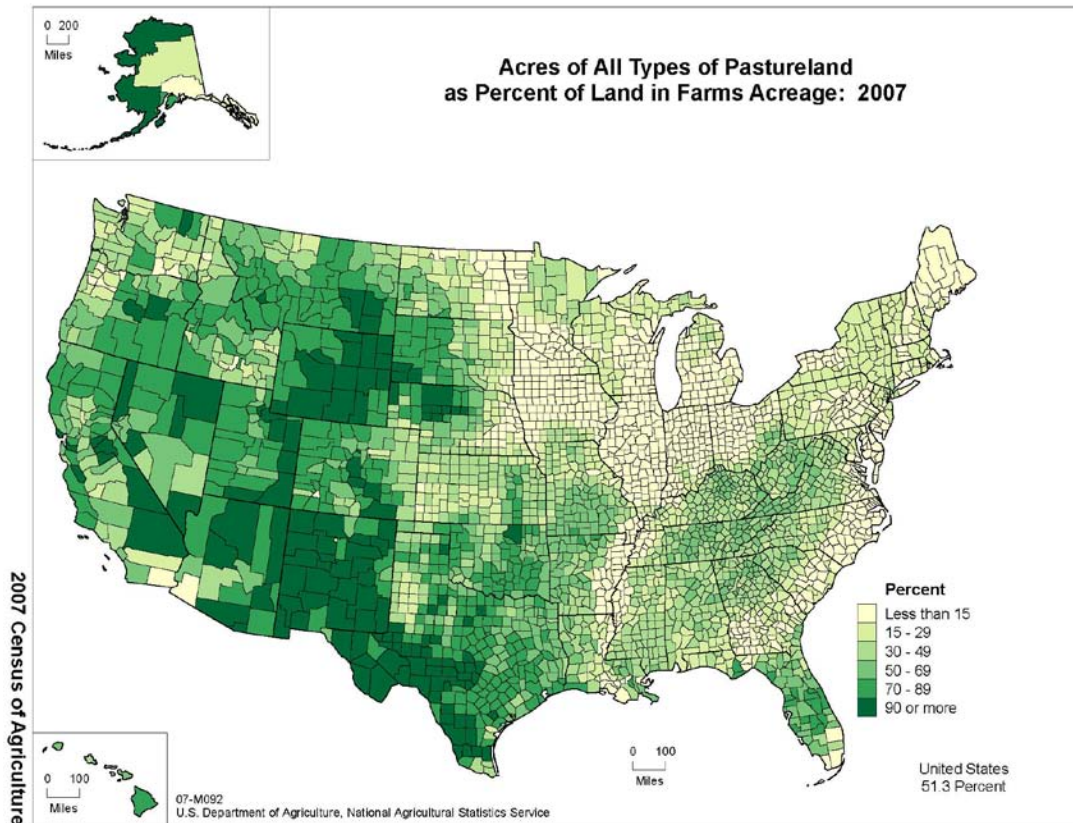
Figure 1. Excerpt of 2007 COA, Section 1

1. Cropland - Exclude cropland pasture.		
a. Cropland harvested - Include all land from which crops were harvested or hay was cut, all land in orchards, citrus groves, vineyards, berries, and nursery and greenhouse crops, Christmas trees, and short rotation woody crops. 0787	None <input type="checkbox"/>	Number of Acres
b. Cropland on which all crops failed or were abandoned - Exclude land in orchards and vineyards. 0790	<input type="checkbox"/>	
c. Cropland in cultivated summer fallow. 0791	<input type="checkbox"/>	
d. Cropland idle or used for cover crops or soil-improvement but not harvested and not pastured or grazed. 1062	<input type="checkbox"/>	
2. Pasture		
a. Permanent pasture and rangeland - Exclude cropland pasture. 0796	<input type="checkbox"/>	
b. Woodland pastured. 0794	<input type="checkbox"/>	
c. Cropland used only for pasture or grazing - Include rotation pasture and grazing land that could have been used for crops without additional improvements. 0788	<input type="checkbox"/>	
3. Woodland not pastured - Include woodlots, timber tracts, and sugarbush. . 0795	<input type="checkbox"/>	

Figure 2. Excerpt of 2007 COA, Section 2

This data review helped narrow the questions of interest for the cognitive interviews, but we still needed to identify the most productive types of respondents to include. One of the most common errors in the land sections is that the total number of acres reported does not match the total of the reported acres in the subpieces (cropland, woodland pastured, etc.). We used a data mining approach to identify smaller subgroups of 2007 COA respondents most likely to have this error in their reported “total acres operated”.

In this analysis, classification trees identified those operations with rented land or with pasture as much more likely to report total acres operated which did not agree with the reported subpieces of total acres operated (McCarthy and Earp, 2009). We then selectively reviewed data reported by these types of operations rather than all operations. This review suggested that some operations may have reported the same acres in multiple pasture categories that should be mutually exclusive and incorrectly excluded specific types of acres. We also made sure to go to areas of the country with higher percentages of pasture, such as Colorado and Arizona, rather than states where little of the agricultural land is in pasture. The areas with respondents most likely to have pasture can be seen in the map below, where the darker shaded counties have the highest percent of pastureland. Certainly, had a completely random (or merely population representative) sample been selected, it would be quite likely that few, if any, respondents would have pasture land as part of their operations. Of course, even if respondents with pasture are targeted, we are interested in the subset of those that will display problems reporting their land.



In addition, follow up probes were specifically designed for the cognitive interviews that focused on rented land and pasture. These included asking respondents to define the different types of pasture listed on the form, asking them whether it was clear that land should be reported only in one category, and general difficulties answering these questions.

Our approach lead to focused cognitive interviews conducted with targeted respondents. These produced much useful information. For example, several respondents reported that they did not understand what “cropland pasture” was. To these respondents, if land was being used as pasture, then it was not cropland, so the phrase “cropland used only for pasture” did not make sense. Land that the respondents did not intend to ever use for anything but pasture was classified as permanent pasture, but could also “have been used for crops without additional improvements” as stated on the form. Several respondents then stated that they might then report those acres in both items with comments such as, “I’ll just put it in both and you can figure it out.”

Based on the results of these cognitive interviews and additional review and testing by NASS staff, several changes were proposed for this section and implemented in the 2012 COA form, shown below.

1. Cropland – Exclude cropland pasture.			
a. Cropland harvested			
<u>INCLUDE</u>		None	Number of Acres
<ul style="list-style-type: none"> • land from which crops were harvested or hay was cut • orchards and vineyards • nursery and greenhouse crops 	<ul style="list-style-type: none"> • Christmas trees • citrus groves • berries • short rotation woody crops 0787 	<input type="checkbox"/>	
b. Cropland on which all crops failed or were abandoned –			
Exclude land in orchards and vineyards 0790		<input type="checkbox"/>	
c. Cropland in cultivated summer fallow 0791			
		<input type="checkbox"/>	
d. Cropland idle or used for cover crops or soil-improvement but not harvested and not pastured or grazed 1062			
		<input type="checkbox"/>	
2. Pasture			
a. Permanent pasture and rangeland – Exclude cropland pasture 0796		<input type="checkbox"/>	
b. Woodland pastured 0794		<input type="checkbox"/>	
c. Other pasture and grazing land (including rotational pasture) that could have been used for crops without additional improvements 0788		<input type="checkbox"/>	
3. Woodland not pastured			
<u>INCLUDE</u>			
<ul style="list-style-type: none"> • woodlots • timber tracts • sugarbush 0795 		<input type="checkbox"/>	

The fonts in this section were modified to make it clearer that this is really 3 categories of land. In addition, the wording for the pasture items was changed, so that none of the pasture items were also called “cropland”.

Using the information from the 2007 COA helped target efforts in cognitive testing to make revisions to the form. Results (as of August 2013) from the 2012 COA indicate that our changes to this section improved the data we collected. As mentioned earlier, for the 2007 COA, 23.8% of the calls for help completing the form (23,780 of 99,993) were

for help with the land sections. For the 2012 COA, with a similar number of calls for help on the form, this was drastically reduced with only 13.3% (14,340¹ of 107,664 calls) asking for help on the land sections. (However, this is still the area of the form where help is most likely to be requested.) At the time of this writing, 2012 COA data are not available to make comparisons to the 2007 COA error rates, but we are hopeful that this will also show improvements in the reported data.

3. Discussion

Cognitive interviews are a staple of questionnaire evaluation, but there is little guidance on how to select the best cognitive interview respondents. Many cognitive interview samples are chosen with the idea that they should represent the population of interest. However, it could be argued that this is the wrong approach, particularly for testing questionnaires that have been fielded before and are known to produce reporting errors. If only a handful of people will be interviewed out of a population of millions, it is NOT a good use of resources to interview people who have no problems reporting and who will report correctly. In order to identify and reduce reporting problems, we need to interview and ask questions of respondents with those problems. For surveys which have been conducted in the past, past data can be reviewed and used to identify items with higher levels of data quality problems (e.g. high edit or imputation rates, unreasonable estimates, etc.). In addition, analysis of the characteristics of the respondents displaying those problems can be used as selection criteria for our cognitive interview respondents. Doing this initial work to target both the items to test and the types of respondents to test them with increases the likelihood the limited resources devoted to cognitive interviews will be productive.

This is not to say that questionnaires, once modified, should not be tested on a wide range of people, including those who reported without error in the past. But to fix problems in the questionnaires, limited resources available for cognitive interviews should be targeted to the biggest problems and those most likely to have them. Those with “Bieber fever” need not be included.

References

Madans, J., Miller, K., Maitland, A. and Willis, G. (2011). *Question Evaluation Methods: Contributing to the Science of Data Quality*. NY: Wiley.

Willis, G. (2005). *Cognitive Interviewing: A Tool for Improving Questionnaire Design*. London: Sage Publications.

McCarthy, J. and Earp, M. (2009). Who Makes Mistakes? Using Data Mining Techniques to Analyze Reporting Errors in Total Acres Operated. US Department of Agriculture, National Agricultural Statistics Service. RDD Research Report Number RDD-09-02, Washington, DC.

¹ As of August 28, 2013