Integrating Early Stage Scoping Techniques into Traditional Pretesting Methods: Inside the Development of a Survey on Small Business Lending

Kenneth Herrell, Michael Purcell, Mary Susan Bucci
kenneth.herrell@census.gov; michael.b.purcell@census.gov;
mary.susan.bucci@census.gov
U.S. Census Bureau, 4600 Silver Hill Road, Washington, D.C. 20233-1912

Abstract
Snijkers and Willimack (2011) have identified a gap in questionnaire development, where questions may be written without a full understanding of the underlying concept, which they refer to as "the Missing Link." Subject experts may have a different understanding of key terms and feasibility of data collection than respondents. Cognitive testing of questions written without understanding of the concept can reveal a need for substantial changes. Stettler and Featherston (2010) discuss early stage scoping (ESS): interviews that learn how respondents understand the survey's key concepts, before the questions are written. In practice, however, subject matter experts often prepare initial questions before the development process begins. While the questions may seem reasonable, testing can reveal that respondents do not understand the conceptual objectives. When this issue is discovered, it may be too late for a full round of ESS. This paper looks at the development of a survey that experienced such issues, and how methodologists and analysts integrated aspects of ESS into subsequent rounds. We also discuss how this may be a way forward in making the practice of early stage scoping a more indelible part of the survey development process.

Keywords: early stage scoping, cognitive interviews, survey methodology, questionnaire development, pretesting

1. Background on Early Stage Scoping

Commonly in survey development, researchers write questions without first exploring the concept that they wish to measure with members of the target population: potential survey respondents. They then test these questions through cognitive interviewing and change the questions based on the results as needed (Snijkers and Willimack, 2011). With cognitive interviews, a respondent sees questions and provides feedback on these questions. The conversation is framed around the question, which may lead the respondent to focus on the question in front of them, rather than considering some of the broader issues related to, but not directly touched by, the specific question itself. Asking a respondent, "in your own words, what is this question asking you?" is different from asking a respondent to talk spontaneously about a certain term, concept, or domain.

Disclaimer: Any views expressed are those of the authors and not necessarily those of the U.S. Census Bureau.

Respondents may misunderstand the question or be unable to provide the data that is being requested. If questions are developed without the input from data providers about how they understand the concepts behind the question, there may be a gap between the concept and the question, which Willimack and Snijkers (2012) refer to as "the Missing Link." There is then a risk of investing time and effort into developing a question or questions that data providers do not understand or are unable to answer.

This issue can be addressed with a process called Early Stage Scoping (ESS). Stettler and Featherston (2010) describe ESS as interviews with potential respondents before survey questions are developed, focusing on the feasibility of the survey objectives, what data the respondents have, how they store the data, and the terminology that they use. This allows survey designers to craft questions around what data can be provided, and in terms native to the data provider. Adding ESS may lead to a faster survey development process overall, even though it is adding a new activity to the beginning of the testing process. When ESS is employed, the first draft of the questionnaire is informed by the feasibility of collecting data on the survey topics and the language that respondents use. Thus, the cognitive interviews using the first draft of the questionnaire are more efficient, because the surveyors have already addressed problems with feasibility and terminology.

Stettler and Featherston (2010) also state that the evidence from early stage scoping can be useful in helping stakeholders limit the scope of the survey. Without early stage scoping, subject matter experts (SMEs) may focus more on the data that they need, and ask for it in terminology with which they are comfortable. In doing so, they may not give due consideration for how respondents think about the concepts, what data the respondents have access to, or how burdensome the survey is. Once SMEs write questions, they may feel more invested in them and may be reluctant to drop questions that are not practical, even if there is strong evidence that they should. With early stage scoping, the feasibility issue is addressed before the questions are developed, allowing SMEs to craft questions that take the respondents' perspective into consideration.

2. Practical Application of Early Stage Scoping

While ESS is grounded in theory and has practical uses and applications, it may not always be possible to incorporate into a testing plan. The foremost reason, mentioned earlier, is that the SMEs may have questions already written prior to meeting with a broader range of survey practitioners, from analysts to methodologists. These questions are often based on the SMEs' knowledge and experience dealing with the subject matter, and created out of a desire to fill gaps in the knowledge and data on a particular topic. SMEs are the drivers behind the survey that provide the funding, backing, and buy-in, and they are generally experts in the topic of the survey, and it is both customary and appropriate to defer to SMEs in these situations.

That approach is not without risk, however, as SMEs could underestimate the difference between their perception of the topic and the perception held by the people who are responsible for providing the data. Imagine the differences in perception between a soldier and a military author, a quarterback and a TV commentator, or a painter and an art critic. In each of these instances, there is a difference in the perception of the person living and creating something and the person analyzing that something, one step removed. The same applies to the relationship between data providers and SMEs. These two parties might have differences in understanding along a wide variety of topics: the importance of the data, the importance of the particular concepts that the SMEs wish to explore, their understanding of the causes and consequences of the concepts, the terminology they use to describe such concepts, the expectations of the record keeping related to these concepts – all topics that fall under the Missing Link.

3. Concepts of Early Stage Scoping in the Development of the Small Business Lending Survey

In early 2015, the Federal Deposit Insurance Corporation (FDIC) and the U.S. Census Bureau commenced work on the development of the FDIC-funded Small Business Lending Survey (SBLS). FDIC commissioned the survey to gain a better understanding about the current state of lending to small businesses, focusing on topics such as:

- The working definition of "small business" that banks use;
- How the size of the bank relates to the types of business loan customers;
- The eligibility requirements that small businesses must meet to obtain a loan;
- The collateral that these businesses must provide to obtain a loan;
- The types of loans available to small businesses;
- The typical size of banks' geographic markets;
- The ways banks work with existing customers;
- The strategies and challenges for reaching new customers, and;
- Competition between banks for small business loans.

When FDIC and Census began work on this project, the FDIC's team of subject matter experts came with fully-developed questions in hand. The questions were detailed and thorough, and the questionnaire as a whole was ambitious in its attempt to gather as much information as possible from respondents. Since FDIC had fully-prepared questions in hand from the outset of the working relationship, it was not feasible to conduct early stage scoping.

Research on this project took place in three distinct phases: an expert review, a series of cognitive interviews, and usability testing. In the spring of 2015, Census methodologists conducted an expert review of the first draft of the SBLS, which aimed to identify questions that may produce inaccurate data or be too burdensome. Some problems were identified in the expert review, such as a reliance on complex matrices that covered several dimensions of a single topic. The methodologists recommended in the expert

review that matrices like this be broken down into individual questions in order to make them easier for respondents to process and answer.

Another issue from the expert review, which would continue throughout testing, was terminology that could be confusing to respondents. Some terminology used in the initial draft could have posed problems for respondents, and SMEs worked with methodologists to use different terminology that might work better for respondents. In some cases, where it was not clear how respondents would interpret certain terms, the SMEs and methodologists agreed to ask about these terms during cognitive interviews.

Following the expert review and subsequent revision, Census Bureau methodologists conducted three rounds of cognitive interviews, attended by FDIC staff observers. The interviewers conducted 16 interviews each in Rounds 1 and 2, and 8 interviews in Round 3, for a total of 40 interviews. FDIC stratified potential participants to reflect the diversity of banks in terms of assets and geography. Table 1 shows the stratification of banks and the number of interviews in each stratum. "Asset range" refers to the amount of assets held by the bank as a whole, which was used as a measure of the size of the bank. "Location of Majority of Deposits" refers to whether the majority of the assets held by the bank were owned by account holders within Metropolitan Statistical Areas (MSAs) or outside of them; this was used as a measure for whether the clientele served by the banks were more urban or more rural. The "Interviews" column is a count of how many of the 40 interviews were conducted with banks that fell into that stratum.

Table 1: Bank stratum identified by FDIC, and number of cognitive interviews conducted with banks in each stratum

Stratum #	*Asset Range*	*Location of Majority of Deposits*	*# of Interviews*
1	Less than $250 million	Inside MSAs	7
2	Less than $250 million	Outside MSAs	4
3	$250 million through less than $1 billion	Inside MSAs	10
4	$250 million through less than $1 billion	Outside MSAs	6
5	$1 billion through less than $10 billion	Inside MSAs	10
6	$1 billion through less than $10 billion	Outside MSAs	2
7	$10 billion through less than $50 billion	All	0
8	$50 billion and more	All	1

The rounds of interviews were iterative, meaning that after each round of testing, the sponsors and methodologists made modifications to the questionnaire, which were then tested in the next round. This allowed the protocol to be modified between rounds, as well.

Some problems persisted through multiple rounds of cognitive testing, but were eventually resolved. For example, several smaller banks said that they could not easily retrieve some of the data that the survey was asking for (such as total of loan origination values across various sizes of companies), because they did not maintain databases

containing these types of data. These banks said that these questions would be very burdensome, and some banks hinted that they might not be able to, or willing to, complete the survey. Through the process of cognitive testing, FDIC recognized the need to allow smaller banks that are unlikely to have these records (Strata 1 through 4) to skip these questions entirely. FDIC further modified the survey so that banks in Strata 5 and 6 will only see these questions if they indicate that they are able to find the data in their record-keeping systems.

Cognitive interviewing also identified and resolved a number of terminology issues. In the first round of cognitive interviewing, it became apparent that simply identifying loans as "commercial" was too vague for respondents, and that specific types of commercial loans should be broken out by the most common types: Commercial and Industrial (C&I) loans and Commercial Real Estate (CRE) loans. Cognitive interviewing also demonstrated that many banks do not offer loan products that were specifically designed for small businesses, but instead offered loans that were tailored to the needs of each individual customer.

Had it been used as an independent stage of testing, ESS may have identified the issues of data retrieval and terminology prior to cognitive interviewing, and the questions that went out for cognitive interviewing could have been adjusted to address these concerns. Since ESS was not planned for, though, several aspects of ESS crept into the initial round of cognitive interviews, via "emergent" probes that were reacting to the responses of the participant (Willis, 2005). Pre-written protocol questions in Round 1 focused on respondents' cognition of the question as it was asked, and how the respondent would answer the question. Once researchers noticed that the questions touched on concepts that were often misunderstood, there was a broader look at the questions and the underlying concepts. The protocol was modified for Rounds 2 and 3 to focus on terminology and data sources for almost every item, a change from Round 1, which asked these questions for only a select few items. There was also an increased focus on whether the respondents' had the systems in place to retrieve the data being asked for, and on the appropriateness of asking certain questions to all banks.

4. Integrating Early Stage Scoping Techniques into Cognitive Interviews

Although the integration of ESS techniques into the cognitive interviews was not planned, we view the cognitive testing as a success. We were grateful that the subject matter experts at FDIC provided the flexibility in the testing plan to take a step back and address conceptual issues, and start anew with questions that were not connecting with respondents as expected. Because of this support from FDIC, the survey ended up with improved questions that are expected to collect higher-quality data.

Learning from this experience, we propose a way to integrate early stage scoping methodology into cognitive interviews, for use in instances where it is simply not feasible to have a dedicated round of early stage scoping.

The first step we recommend, prior to interviewing any participants, is having a detailed discussion with the SMEs who initially prepared the questions and commissioned the study, and other key stakeholders (such as primary data users). Topics that we suggest covering in this discussion include:

- the measurement objective (the expected publication data or the analytical outcome they wish to address);
- the key concepts of the area of study;
- respondents' expected reporting capabilities;
- respondents' native terminology; and,
- expected survey respondents.

In addition to this discussion, it may be helpful to review the survey with SMEs to discover underlying concepts. If the survey is broken out into sections, a natural starting point would be reviewing the sections to see if they cover a single concept. If the survey is not broken into sections, it may be useful to group questions with similar themes together, which would hint at the underlying concepts of the survey.

When conducting the cognitive interviews, researchers can ask respondents to describe the concepts of the survey (as identified during the review) in broad terms, using their own terminology. The researchers may also want to discuss some specific terms that they plan to use, to see if their understanding of the terms matches that of the respondents. Researchers should also focus on if, and how, respondents track the information related to these concepts. When showing the draft questions, researchers should inform the respondents that these are indeed draft questions, and that there may be necessary changes to the questions based on their feedback. This will allow the respondent to elaborate on specifics of where the draft questionnaire falls short of meeting some of the concepts that it wishes to measure. It will also give the respondent the opportunity to address whether or not they have access to the requested data, or if it might be more appropriate to ask for another metric that they track.

There are drawbacks to using this method. Allotting time for scoping-style questions potentially means less time available for studying some of the more specific issues in the questions. In addition, the draft questions may need significant revisions, or not apply at all, after a thorough discussion of the concept. Lastly, the findings for the cognitive aspects of the questions themselves would be less pure than if they were studied independently, as the coverage of the topic may have interfered with their reactions to the questions.

Despite these drawbacks, we believe a broader focus by way of integrating these early stage scoping strategies would be beneficial to the testing process. Conducting cognitive interviews under this format will allow respondents to more easily address their concerns with conceptual issues by providing the opportunity to do so outside of the framework of the question. Further, respondents are still provided the opportunity to review the questions themselves, and address the primary purpose of the cognitive interviews.

Respondents will also have a fuller understanding of the concepts, and will be able to think through the questions in that context.

Acknowledgements

The authors would like to thank Ken Pick, formerly of the Census Bureau, for his efforts in guiding this testing process through its initial stages, and for inspiring this paper. We would also like to thank Karyen Chu, Yan Lee, Smith Williams, Jack Reidhill, Kris Rengert, and Claire Brennecke of FDIC for their flexibility and support throughout the testing process. Thanks to Lynda Lee of the Census Bureau for her help in developing survey content, and to the Census Bureau's Kristin Stettler, Amy Anderson Riemer, Diane Willimack, Carma Hogue, Carol Caldwell, Aryn Hernandez, Rebecca Keegan, Krysten Mesner, Temika Holland and Alda Rivas for their helpful comments and insights. Lastly, we would like to thank the participants in the cognitive interviews; their contributions to the final survey product were invaluable.

References

Snijkers, G., and Willimack, D.K. (2011, September). *The Missing Link: From Concepts to Questions in Economic Surveys*. Paper presented at the 2nd European Establishment Statistics Workshop (EESW11), Neuchâtel, Switzerland.

Stettler, K., and Featherston, F. (2010, May). *Early Stage Scoping: Building High Quality Survey Instruments Without High Costs.* Paper presented at the American Association for Public Opinion Research (AAPOR), Chicago, Illinois.

Willimack, D.K., and Snijkers, G., (2012, January). *The Missing Link: From Concepts to Questions in Economic Surveys*. Paper presented at the Federal Committee on Statistical Methodology (FCSM), Washington, D.C.

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