

# THE COGNITIVE EVALUATION OF THE NATIVITY QUESTIONS FOR THE CURRENT POPULATION SURVEY

Tracy Wellens, U.S. Bureau of the Census, Center for Survey Methods Research  
Bureau of the Census, CSMR, WPB#1 Rm 433, Washington, DC.

**KEY WORDS:** Cognitive Interviewing, Nativity

## **INTRODUCTION**

In recent years there has been an increase in the use of cognitive pretesting methods for survey questions. This paper represents one such cognitive evaluation that was particularly interesting because the population of interest consisted of respondents who were difficult to recruit and for whom the topic of the questions was potentially sensitive. In addition, the interviewing situation was expected to be hampered by the respondents' language skills. This paper will focus on the method used to locate these respondents and the flexibility of the cognitive interviewing methodology employed in this evaluation.

This evaluation focused on a series of questions known as the nativity questions. These questions are used to collect information on the foreign-born population in the United States. Respondents are asked for their country of birth, their mother's country of birth, their father's country of birth, whether they are citizens of the United States, their citizenship type, and their year of immigration. The nativity questions were designed to allow comparisons between different foreign-born groups and between different immigrant cohorts based on their citizenship status and their length of exposure to U.S. culture. This data is sought by demographers and policy makers in order to assess the impact of immigrant populations on the country's economy and on broader societal issues.

The Census Bureau was asked to evaluate these nativity questions before their proposed inclusion in the Current Population Survey (CPS). Including these questions in the CPS is important because it provides the only large scale data collection on the foreign-born population in the United States between decennial censuses. The purpose of this evaluation was to: a) improve the quality of data elicited from the nativity questions by evaluating how respondents interpreted and comprehended the questions; b) to make the questions easier for respondents to answer and c) to determine if the nativity questions were perceived as "sensitive" in the context of the CPS interview.

In terms of questionnaire design, these questions are very interesting methodologically because the targeted population consists of foreign-born persons, many of whom are non-native English speakers. Further, the potentially sensitive nature of these

questions posed interesting questions in terms of how respondents should be recruited and where interviews should be conducted. Cognitive interviews were completed in two phases to determine how respondents interpreted and responded to the nativity questions. The first phase of interviewing allowed for the evaluation of problems respondents encountered when answering the nativity questions. The questions were then revised and the second phase of interviewing provided an opportunity to assess the effectiveness of the proposed question revisions.

This paper will focus on a subset of the nativity questions. Questions related to citizenship, citizenship type and year of immigration into the U.S. provided interesting insights about the respondents' comprehension of various terms and will be the focus of this paper. The remainder of this paper will discuss foreign-born respondent recruitment for potentially sensitive questions, the methodological issues involved in conducting cognitive research with a non-native English speaking population, and the results of the evaluation for the nativity citizenship question series and the year of immigration question.

## **RESPONDENT RECRUITMENT ISSUES**

As stated earlier, foreign-born persons were the target-population for these questions. We decided to select respondents from Hispanic, Asian, White European and other Non-Hispanic categories of foreign-born populations. It was also decided that both bilingual (Spanish-English) and monolingual (Spanish) respondents would be recruited from the Hispanic population in order to evaluate the Spanish translation of the nativity questions.

Two initial dilemmas surfaced early in the planning stages of this research. The first centered on determining where we could locate and then recruit this population for interviewing. We wanted to include respondents from a variety of countries with a range of English language skills and citizenship statuses. The second issue focused on where we could conduct the interviews. We were somewhat concerned that the nature of these questions might be sensitive to the population of interest and we wanted to ensure that our respondents would agree to participate in the research and would feel comfortable in the interview situation.

Fortunately, both of these dilemmas were easily resolved with the help of a community organization.

We were able to coordinate respondent recruitment through a community center that had ties to the Hispanic community and to another community organization which specialized in providing services for foreign-born Asian populations. Having this introduction into the community was a true asset. First, the community organization provided access to people that we would not have had access to without the community center's help. Second, using the community center provided non-threatening locations for conducting the interviews. Respondents were interviewed in the community centers. These organizations were deemed trustworthy by respondents and were known for providing much needed services. This facilitated the cooperation required for our cognitive interviews.

The community organization that arranged respondent recruitment was unable to locate any community centers that targeted their services toward White Europeans. Therefore, the White European respondents were recruited through their local restaurant employers and were interviewed in those restaurants.

By using these procedures, we were able to recruit the range of respondents we desired. In the first phase of cognitive research we conducted 24 cognitive interviews. The second phase of cognitive research consisted of 23 cognitive interviews. The age of our respondents ranged from 18 to 64 and their level of education ranged from no formal education to a MBA. Eleven of the interviews were conducted with monolingual Hispanics in Spanish with the Spanish version of the questionnaire.

#### **METHODOLOGY**

The nativity questions were evaluated within the broader context of the CPS demographic questions. The cognitive interviews were conducted using concurrent think-aloud procedures. The researchers followed a protocol to probe respondents about their understanding of the questions. Specific probes were used to evaluate how various key terms were comprehended and understood by respondents.

It should be noted that these procedures were sometimes difficult to employ with this population of respondents. We were quite concerned about attempting to conduct this type of interview with non-native English speakers in English. We thought that it might be difficult for respondents to think-aloud when they were trying to translate concepts from their native language to English. However, since for most of these respondents the actual survey interview would be conducted in English, we decided to attempt this procedure with the knowledge that at times it would be difficult. We believed that obtaining linguistic information from respondents was the most effective way of evaluating the terminology employed in the questions

and the types of information respondents had available to them when answering the questions.

Many of the respondents were from low income areas with little formal education and it seemed that the task of thinking aloud was unfamiliar and difficult. These interviews often required the use of specific question probes. Thus, much of the information we obtained from respondents was in response to specific questions and not from an uninterrupted flow of thoughts.

There were also additional problems from the interviewers' perspective. It was difficult to conduct these interviews because of respondents' accents and their limited familiarity with spoken English. On occasion we could not decipher what was being said and at other times unfamiliar words were being incorporated from other languages. In these cases non-verbal communication, such as smiles, nods and gestures, helped to facilitate the interaction.

At the end of the cognitive interview, respondents were also probed on sensitivity issues. These probes were designed to determine how individual respondents felt about answering these questions and how they thought other people may feel when they were asked these questions.

#### **COGNITIVE INTERVIEWING RESULTS**

##### **A. Citizenship Question: (Are/Is) ... a CITIZEN of the United States?**

The purpose of the citizenship question is to determine whether the household member is a citizen of the United States. This question is only asked if citizenship status cannot be determined from the country of birth questions.

In general, respondents tended to understand the citizenship question. They knew what a citizen was and whether or not they were citizens. Respondents also seemed to know the difference between being a citizen and a variety of other statuses (i.e., political refugee, permanent resident, legal resident, legal alien, holding a green card, and having an illegal status).

Of course, their beliefs were not always consistent with the legal reality of these statuses but they clearly knew what a citizen was and whether they had that status. As one respondent illustrates:

*I: Are you a citizen of the United States?*

*R: No. I think not yet but in the future. If I have been here they say for five or six years I can be a citizen. I heard somebody say that. You learn about America or something.*

The majority of individuals with legal statuses felt that there was not a major difference between being a citizen and their status. Many mentioned that they also pay taxes and that the only real differences were in terms of being qualified for government jobs and being

eligible to vote. Respondents with illegal statuses often mentioned that they had fewer opportunities and rights than citizens or legal aliens. Many respondents knew that one can acquire U.S. citizenship through birth in this country or through completing some process. As one respondent said:

*R: Well, I think you have two kinds of citizenships, two kinds of citizens. Either by birth or by document, paper.*

Given the lack of comprehension problems with the citizenship question, changes to this question were not warranted for the second phase of cognitive interviewing.

During the second phase of cognitive interviewing, respondents were again certain of their own citizenship status and were able to make distinctions between being a citizen and other types of statuses. Respondents were also able to explain why they were or were not citizens of the United States. For example:

*I: Are you a citizen of the United States?*

*R: No.*

*I: In your own words, what does the term "citizen" mean to you?*

*R: I don't ... I don't understand.*

*I: I just want to know what you think the definition of the word "citizen" is.*

*R: A citizen is a person who belongs to this country. That person has the right to vote or join the government.*

*I: Can you tell me more about why you are not a citizen of this country?*

*R: A person must live here for five years, take a test, obtain history course, then can person become citizen.*

In general, their reasons included whether they had been in the country for the required length of legal residence before application, whether they had studied for and passed a test, whether they had been interviewed by INS and sworn their allegiance and whether they carried a U.S. passport. Thus, there appear to be clear markers our respondents used for determining citizenship status. In addition, respondents knew that the citizenship status was reserved for people who were born in this country or who went through some process. As illustrated below, this distinction between being born a citizen and becoming a citizen through some process was often discussed.

*I: Can you tell me the difference between a resident alien and a citizen?*

*R: A resident alien is a citizen of another country who is permitted to live and work in the United States under U.S. immigration laws. A citizen is somebody who was either born in the U.S. or chose to become a citizen of the U.S. under U.S. naturalization laws.*

**B. Citizenship Type Question: (Were/Was) ... born abroad of an American parent or parents, or**

**(are/is) ... a citizen by naturalization?**

After a respondent indicates that he/she is a U.S. citizen, the citizenship type question is used to determine whether the person with U.S. citizenship was born with citizenship or whether the individual acquired citizenship through the process of naturalization. Only those people who are born in the United States, Puerto Rico, or an Outlying U.S. Area or Territory and those who are born abroad of American parents are born with U.S. citizenship. Everyone else must go through the process of naturalization to acquire citizenship.

The citizenship type question was somewhat problematic and difficult to evaluate. Many of our respondents were not citizens and did not have citizens in their households; therefore, we did not find an overwhelming number of respondents who would be asked this question. In order to more fully evaluate the terms employed in this question we asked all respondents specific follow-up probes which requested definitions of the following terms: "born abroad," "American parent," and "citizen by naturalization."

We found that the term "born abroad" was very difficult for the majority of respondents to comprehend. Respondents' definitions included such things as being born on a ship and being first-generation born in the United States. There were several respondents who understood the term but they tended to be better educated with better English language skills.

The term "American parent" out of context is somewhat difficult to understand. Most respondents chose to define the two words separately. The majority of our respondents defined American as being born in the United States. Several respondents expressed that "American" could refer to anyone from one of the "Americas" and felt that this term referred to a geographical location that included the United States but was not limited to it.

"Citizen by naturalization" was understood by those respondents who had either completed the process or were contemplating undergoing the process. Of these respondents, most understood that there was a process, although the specific requirements for that process varied among respondents. On the other hand, this term was problematic for the majority of our respondents who were not citizens. These respondents tended to say that a "citizen by naturalization" was someone who was born here. Perhaps they were confusing the term "naturalized" with "natural." This is potentially problematic for proxy responses where the individual respondent is not a citizen but other members of the household are. In addition, most of our respondents seemed to realize that you could either be born with citizenship or you could go through a

process to attain citizenship. However, this question did not allow them to make this distinction because all terms in this question were not clear.

To rectify these problems, a two-step question was proposed as displayed below.

- a. (Were/Was) ... born a citizen of the United States?
- b. Did ... become a citizen of the United States through naturalization?

These questions did not employ the confusing terms (e.g., "born abroad," "American parent," or "citizen by naturalization") from the original citizenship type question. Moreover, these questions were shorter and require simple yes or no answers instead of making a distinction between two ambiguous choices as in the first phase citizenship question.

During the second phase of cognitive interviewing, all respondents who were U.S. citizens answered these questions correctly. We did not encounter anyone who was born with U.S. citizenship and none of our respondents wrongly reported that they were U.S. citizens through birth. Thus, it seems that respondents were able to answer the "born a citizen" question correctly. Once again, we did not find an overwhelming number of respondents or household members who were U.S. citizens during these interviews. In order to ensure that the terminology in these questions was thoroughly evaluated, we again asked specific questions about the terms used in both the "born a citizen" and "become a citizen" questions.

When probed for the meaning of being "born a citizen of the United States," the vast majority of respondents stated that it meant being born in the United States. A few respondents included being born to U.S. citizen parents but this was not discussed by most respondents. However, for this population of foreign-born respondents it was more important for them to realize that they were not born with U.S. citizenship than to demonstrate that they knew individuals born abroad to American parents were also born with U.S. citizenship. Answers to this question were fast and even respondents with poor language skills did not indicate a need for clarification of the words used in the question. In addition, when probed for the meaning of being "born a citizen of the United States," several respondents spontaneously asked whether the interviewer wanted to know the difference between being born a citizen and being a naturalized citizen.

As stated earlier, none of our respondents who were U.S. citizens were born with that citizenship and they were able to correctly answer the "born a citizen" question. In accordance with the skip pattern, they were then asked if they became a citizen of the United States through naturalization. All respondents correctly answered this question. For example:

*I: Were you born a citizen of the United States?*

*R: No.*

*I: What does being born a citizen of the United States mean?*

*R: Well, I think it means that you are born in the United States.*

*I: Did you become a citizen of the United States through naturalization?*

*R: Correct. Yes.*

In addition, many of those respondents who were not citizens but were asked about the terms used in this question knew that if they themselves became citizens this is how they would be classified. They associated this with a process although it was not always easily verbalized. Defining the term "naturalization" was still problematic for many respondents. This difficulty was once again more pronounced for those individuals who were not naturalized citizens but even those people who knew that they were naturalized citizens sometimes had trouble defining the term. However, these respondents did not have trouble describing the process.

Further, these questions were often accompanied by quick responses which indicated that the answers to these questions were readily available to the respondent.

Thus, the new citizenship type question series was a definite improvement over the original question. This series allowed respondents to differentiate between being born with U.S. citizenship and undergoing a process to become a citizen by naturalization. The shorter sentence structure and simpler word choice which was adopted in these new questions facilitated responding. Although one could infer that those individuals who are not born with U.S. citizenship must be naturalized, we feel it is still necessary to ask respondents to provide that information directly. Providing an answer to this question allows the respondent to feel that all information has been obtained. In addition, it seems that this two-part question series creates the context which is needed to make the distinction between the two types of citizenship.

### **C. Year of Immigration: When did ... come to the United States to stay?**

The purpose of the year of immigration question is to measure the length of permanent residence in the United States in order to determine the length of exposure to U.S. culture. This serves as an important factor in evaluating other demographic data on foreign born individuals.

The year of immigration question produced an overwhelming number of actual dates. The majority of respondents could not only provide the year very quickly but they could also provide a month and

sometimes even a day. When asked about this level of detail many responded that this was a very meaningful date, one that they could never forget. This date was associated with major life changes and also used in "official" paperwork (i.e., immigration, social service assistance) quite often.

On the other hand, respondents had a very hard time explaining how they interpreted this question. We were able to ascertain through the use of very specific additional questions that most people gave the date that they entered the United States for the first time. This may be due to the fact that many of our respondents had never left the U.S. since the date they reported or had only left for short periods of time. They seemed to have a strong feeling of the differences between a "vacation or a visit" and "living or staying." For those who had been in the U.S. before the date that they provided, most stated that it was only for a short time (i.e., ranging from a few weeks to one year) and many had even forgotten about it until they were probed for such details. Although this question tended to be associated with a clear date for most respondents, it is troublesome that they had difficulty explaining the meaning of the question. Many people stated that this question was asking them "when did I come here to live." For example:

*I: What does come to the United States to stay mean to you?*

*R: To live.*

*I: To live?*

*R: Yeah, to stay. When you stay somewhere you live there. That is how I look at it. How I understand it. When you live somewhere you ... You said stay right? You used the word stay?*

As a result of these findings, we suggested revising the question wording to incorporate the term "come to live" which was used by respondents. We also wanted to ensure that this revised wording would take into account the natural process respondents used when answering the original question. It was important for the revised question to reflect the recall strategy of providing an exact year so we proposed the following question: When did you come to live in the United States?

The revised question wording of the year of immigration question seemed to be easily understood by respondents. All but three respondents answered this question with an exact year for their self-response to this question. Many respondents also spontaneously included a specific month and day in their answer to this question. However, when estimating an answer to this question for other household members when the specific information was not known, respondents sometimes provided an answer in terms of the number

of years. It seems that it was easier for respondents to provide their estimates of length of time based on the number of years rather than on a specific year. Several respondents indicated that when they were uncertain of the arrival time of a household member, they based their response on an estimate which focused on the amount of time they themselves had been in the country. Thus, their own arrival time served as the comparison point for estimating the amount of time other household members had resided in the United States. For example:

*I: When did Joe come to live in the United States?*

*R: Oh, you got me. I don't know. I think eight years ago.*

*I: Eight years ago?*

*R: Yeah, 'cause when he told me, I was thinking that I have been here like a couple years before him.*

When probed about the meaning of the phrase "coming to live in the United States," respondents tended to explain a time when they "settled here," "moved here" or "started a life here." There seemed to be some permanence to this decision as indicated by describing this as a time of establishing themselves here and intending not to leave. Respondents were much more verbal about the meaning of this revised question than they were about the original year of immigration question.

From specific additional questions about the answer given in the year of immigration question, we learned that several respondents had been in the U.S. before the date they stated in this question. These respondents made a clear distinction between those earlier entries and the entry they gave in response to the year of immigration question. This distinction was based on the difference between a "visit" or "vacation" and "moving to a country" or "coming to a country to live." In addition, several respondents had been back to their country of birth for a period of a few weeks to several months after the date given in response to the year of immigration question. Again, these absences were categorized by the respondents in terms of a temporary situation such as a vacation or visit.

The revised question wording of the year of immigration question also proved to be an improvement over the original question wording. This was evidenced by an increased ability for respondents to explain in their own words what the revised question meant to them. The majority of responses reflected the use of a recall strategy based on an exact year which was associated with a major event. In both interview phases, respondents indicated that this date represented an important time in their lives that they would not be likely to forget. However, not all respondents first generated an exact year and we felt

that this question should allow for respondents to utilize their own natural recall strategies. This was accomplished by designing response categories which accept either an exact year or the number of years with an exact year verification procedure.

#### **D. Sensitivity Probes**

As stated earlier, during both phases of cognitive interviewing respondents were probed on the sensitivity of the questions. Several respondents indicated that they were suspicious and questioned the true use of the information they had provided. It became clear that several respondents thought the information may be used for deportation purposes. Respondents often verbalized their concern that many people would not want to answer these questions because they would be afraid based on their immigration status. Other respondents stated that these questions were sensitive because of the situations which lead them to leave their country.

On the other hand, many respondents indicated that these questions were not at all sensitive. The main reason stated for this lack of sensitivity was that the content of the questions was based on information that was factual. Apparently, foreign-born individuals are asked often about where they are from and when they arrived in the United States. This topic is common in official situations (e.g., food stamp applications or immigration applications) as well as in social situations (e.g., conversations with friends and neighbors). On the other hand, the majority of respondents did state that the information requested in the interview was sensitive in nature. They felt that the questions were personal and noted that people may not always answer them truthfully. This was particularly the case for individuals with illegal or problematic immigration statuses or those who knew others with immigration difficulties.

When probed for which particular questions were sensitive, a variety of topics were deemed personal or sensitive. Many respondents had difficulty verbalizing which specific questions were particularly sensitive but focused on the content of the questions and it was clear that the nativity series was noted by many. It may be the case that respondents found a wide variety of questions in the interview to be sensitive because of the level of in-depth probing which is required in a cognitive interview.

It should be pointed out that although respondents said that these questions were sensitive, they did not refuse to answer them. We have no way to evaluate the truthfulness of responses, however; several respondents indicated that they were in the United States illegally and this is not the expected socially desirable response. In addition, we did not locate an over-

whelming number of U.S. citizens, a protected legal status, leading us to believe that we did receive many truthful responses. So although this information may be deemed sensitive, it does not necessarily imply that it will not be provided.

#### **DISCUSSION**

In terms of meeting our research goals, the cognitive evaluation was successful. We feel confident that we have questions which are better comprehended by respondents (e.g., the two part citizenship type question series). In addition, the improved question format allows respondents to provide information in a way that reflects their understanding of concepts and their recall strategies (e.g., the year of immigration question).

In terms of sensitivity, we determined that the content of the nativity question series may be perceived as sensitive. The degree of perceived sensitivity varied among respondents and this was probably related to their immigration status (e.g., legal versus illegal) and other consequences of their particular situations. However, we also learned that respondents were willing to talk about this potentially sensitive topic in great detail. We feel that the involvement of the community organization in respondent recruitment and the location for conducting the interviews facilitated respondents' willingness to discuss this topic.

This evaluation also demonstrates that cognitive interviewing methods can be fruitfully applied in situations where respondents have low levels of English proficiency. The flexibility of this technique allowed changes in question probing strategies which resulted in a wealth of information about the types of relevant knowledge the respondents' possessed. Overall, our respondents tended to provide more information when responding to specific probes than to the general request to "think-aloud." Thus, through the use of specific probes we were able to obtain information about how respondents' interpreted the questions even though they could not always verbalize that information directly. Although language skills create some problems when conducting these interviews (i.e., accents and foreign vocabulary) these problems were not insurmountable.

This evaluation suggests that the standard concurrent think-aloud procedure can be successfully modified to meet the needs of special populations. Our methodology employed a modified concurrent think-aloud procedure by incorporating specific probing questions. Future research may suggest other tools (e.g., vignettes or visual aids) or questioning strategies which in combination with concurrent think-aloud procedures would produce the data we need to more fully evaluate survey questions.