

## Cognitive testing of racial and ethnic questions for the CPS supplement

*Extensive cognitive testing aided the development of a new questionnaire that is readily understood, no longer evokes negative emotional responses, and allows analysts to pinpoint the circumstances associated with multiracial reporting*

Ruth B. McKay  
and  
Manuel de la Puente

**R**ace” and “ethnic origin” are two of the most emotionally charged and conceptually ambiguous terms in U.S. contemporary social discourse. Surveys with racial and ethnic questions that have been thoroughly researched, sensitively tested, and carefully structured should enable researchers to collect reliable data and make sound conclusions. Thus, a study was conducted on the May 1995 CPS supplement to provide questions that would allow respondents to identify themselves most appropriately by race, ethnic origin, and multiracial groups.

This article describes the development of the Current Population Survey Supplement on Race and Ethnic Origin that could be easily understood and would no longer evoke negative emotional responses. (See the article by Clyde Tucker and Brian Kojetin in this issue for a description of the supplement.)

An interagency team of behavioral scientists representing anthropology, psychology, and sociology was closely involved in the development and testing of the questions for the May 1995 CPS supplement.<sup>1</sup> The team continued to be involved during and after collection of the supplement, analyzing live and taped interviews, and contributing insights from the analysis to help interpret findings from the statistical analysis of the data.

### Pretesting by cognitive research

To test successive versions of the supplement, cognitive research interviews were conducted by a

multiracial and multi-ethnic team of behavioral scientists from several Federal agencies, with additional support from survey specialists in academia. The research protocol called for individual, face-to-face interviews in which the respondent answered all of the questions in one of the four panel questionnaires to be tested.<sup>2</sup> After responding to each question, the respondent was asked to paraphrase the question, that is, to tell the interviewer what the question meant in his or her own words. For questions containing terms of special interest to the research, for example, race, ethnicity, Latino, the respondent was also asked to provide a definition of these terms.<sup>3</sup>

The research plan called for matching the race and Hispanic origin of the respondent and researcher. This was achieved for all groups except for the American Indian respondents, who were interviewed by an Asian-American researcher. All of the materials used in the cognitive research, including the four panel questionnaires of the supplement and the research interview protocols, were translated into Spanish for respondents who were monolingual in Spanish. The cognitive interviews were carried out in three phases between November 1994 and January 1995. Each successive phase tested a version of the questionnaire that had been revised in light of problems found in earlier rounds of cognitive research.

A total of 82 cognitive interviews were carried out in the following locations with the groups indicated: Albuquerque (American Indians); Chicago (blacks); Houston (Hispanics, whites); New Orleans (Creoles); New York (Hispanics,

Ruth McKay is a behavioral scientist in the Statistical Methods Division, Bureau of Labor Statistics. Manuel de la Puente is chief of the Ethnic and Hispanic Statistics Branch, Bureau of the Census.

whites); Rural California (Hispanics); Rural Mississippi (blacks); Rural West Virginia (whites); San Francisco (Asians and Pacific Islanders, Hispanics, multiracials); and Washington, DC (Asians and Pacific Islanders, blacks, Hispanics, multiracials, whites). For each of the racial groups, it was possible to interview respondents who had a high school education or less, as well as respondents with some years of college.

The phase one interviews revealed a multitude of problems in the four panel questionnaires of the supplement. These problems included vague or imprecise questions, sensitive questions, abstract questions, vocabulary problems, and order effects. The following examples illustrate the types of problems that were identified by the cognitive interviews.

*Vague or imprecise questions.* After naming a race from a list of racial categories, the respondent was asked, "Do you also have a more specific group that you belong to?" Because the question did not specifically call for an ethnic group *within* that race, responses to the question ranged from "Christian," "Masons," "Black Muslims," to "rebellious teenager."

*Sensitive questions.* Many respondents were uncomfortable answering any question about race, because they feared that the questionnaire was really about racism, and the supplement, a covert attempt to learn about their racial attitudes. Asking questions about their comfort level with using certain racial terms to describe themselves only increased their discomfort. For example, after a respondent selected a category from the list of races, a subsequent question asked, "Please tell me if using this category to describe your race would make you feel: very comfortable; uncomfortable; somewhat comfortable; very comfortable."

*Abstract questions.* Earlier versions of the supplement included conceptual questions designed to learn about the criteria which respondents use to categorize themselves and others into racial, ethnic, and ancestral groups. Questions such as, "Do you think there is any difference between race, ethnicity, and ancestry?" were perceived by respondents as being too abstract and difficult. Some believed that the questions were a test of their intelligence.

*Vocabulary problems.* Very few of the respondents knew the meaning of the word "ethnicity," but several respondents thought that the question was asking about the ethical character of races. One woman thought that the word "characteristic" meant that the interviewer was asking her about her character. Vocabulary problems were found in several of the questions.

*Order effects.* In panels 1 and 2, the question asking if the respondent would have liked to have had a multiracial category

from which to further specify race came after the question on the respondent's ancestry or ethnic origin. For those respondents who had recalled a remote ancestor of a different race on the ancestry question, the later question about liking a multiracial category led them to think aloud about whether they should say they were multiracial, in light of how they had answered the ancestry question. This did not happen for panels 3 and 4, in which the multiracial question *preceded* the ancestry question.

Following phase one, there were major revisions made to the questionnaire to correct problems identified by the cognitive research. The test of the revised instrument in phase two revealed that many of the problems identified in phase one had been corrected. However, other questions that had been revised continued to present problems. For example, the first question in panels 1 and 2 asked, "Are you Hispanic, Latino, or of Spanish origin?" Because government surveys such as the decennial census always ask race first, respondents, out of habit, paraphrased this question as "What is your race?" or "What race are you?" and responded by stating their race, for example, "I'm white."

The key findings of the phase two interviews were that respondents found some of the questions to be confusing or redundant, and that the supplement continued to evoke a negative emotional response. The intermingling of self-descriptive questions, such as, "What is your race?", with preference questions asking, "Which term do you prefer for your racial group?" was found to contribute to the apparent redundancy of the questions.

The questionnaire was revised again in light of the findings of the phase two research. Questions that had evoked a negative emotional response were eliminated. The problematic question on Hispanic origin and the conceptual questions were revised. Finally, the supplement was reorganized to group the questions in each of the panel questionnaires into three distinct sections: 1) self-identification questions; 2) preference questions; and 3) conceptual questions.

The new version used in phase three was tested in a small sample of non-Hispanic whites in West Virginia and a small group of Hispanics in suburban Maryland. Respondents in both groups had much less difficulty with the supplement in phase three. The only remaining problem for both groups was the conceptual questions, which were found to be too difficult. The non-Hispanic whites still responded to the first question in panels 1 and 2, "Are you Hispanic, Latino, of Spanish origin, or none of these?" by naming their race or ethnicity, for example, "I'm white," "I'm American."

After the phase three interviews, the interagency team decided to abandon any attempt to salvage the conceptual questions. The one remaining problem was to reword the Hispanic origin question in such a way as to lead respondents to identify as either Hispanic or not Hispanic. The final wording of the question, "First, are you one of the following: Hispanic,

**Table 1. Multiracial breakdown by race, 1995 CPS supplement**

Category	Panel 2	Panel 4
Total multiracial .....	1.65	1.55
No race; don't know; not ascertained .....	.02	.0
"Something else" as only race .....	.51	.22
Only one race .....	.53	.15
White and black or black and white .....	.09	.16
American Indian, Eskimo, or Aleut plus one race .....	.20	.28
Asian or Pacific Islander plus one race .....	.07	.28
One race plus "something else" .....	.16	.07
Other two races .....	.0	.20
Three or more races .....	.08	.21

NOTE: Panels 1 and 3 are not shown because multiracial questions were not asked in these panels.

Latino, of Spanish origin?", was tested with 48 non-Hispanic and 6 Hispanic respondents in Washington, DC. In all cases, the question elicited either "Yes—Hispanic/Latino," or "No" responses.

### Interpreting the findings

Once the final supplement instrument was constructed, additional cognitive research was conducted to help interpret the results from the supplement. On one day during the CPS collection week of May 14th through May 20th, cognitive researchers monitored supplement interviews in computer-assisted telephone interviewing (CATI) facilities located in Hagerstown, MD, and Tucson, AZ. Researchers also conducted focus groups with interviewers in both facilities on the following day to learn about their experiences in conducting the supplement interviews. Also during that week, researchers accompanied CPS interviewers in Tucson and in Miami, FL, to observe how the interview was conducted in the field. Four hundred CATI interviews, two hundred in each CATI facility, were taped for subsequent behavior coding.

Following collection of the May 1995 CPS supplement, open-ended answers to supplement questions on multiracial status and ancestry or ethnic origin were analyzed for content. The findings from the several phases of cognitive research provided the basis for greater accuracy in interpreting the results of the statistical analysis of the supplement data, especially the ambiguous findings relating to multiracial reporting.

### Multiracial reporting

In the May 1995 CPS supplement, the multiracial response category was listed in the race question for panels 2 and 4. Respondents who chose the multiracial response category were asked a follow-up question: "Which of the following list (do/

does) (you/proxy) consider (yourself/himself/herself) to be: white; black; American Indian, Eskimo, or Aleut; Asian/Pacific Islander; or something else?" Those who answered, "something else," were asked to specify what that was.

The breakdown of racial identities for those who chose the multiracial response option in panels 2 and 4 may be seen in table 1. Some respondents who identified themselves as "multiracial" in panels 2 and 4 selected only one race in the follow-up question. The finding of a large group of "one-race multiracials" led the team of analysts to develop "multiracial" and "indeterminate multiracial" categories for analyzing the CPS supplement data. These are defined as follows:

- *Multiracial*: persons who report belonging to two or more of the racial groups, or a racial term that encompasses two or more races, listed on the race question for that panel;
- *Indeterminate multiracial*: persons who do not report belonging to two or more of the racial groups listed on the race question for that panel.

A very large proportion of the "one-race multiracials" in table 1 selected "something else" as their single race, while a portion of the two plus multiracials chose a single racial category and "something else." From the cognitive research interviews, it had been observed that multiracial respondents who were asked to respond to race questions that did not offer the multiracial category most often chose "something else," and listed their several racial identities in the open-ended follow-up question. An analysis of the "open-ended" responses to the follow-up question for those who identified as "something else" under multiracial revealed a range of racial and ethnic designations.

*Further breakdowns.* The open-ended answers from respondents who indicated "one-race equals 'something else'" included such diverse entries as: Creole; Eurasian; Cape Verdian; and German and Irish. Some of the entries, such as Creole or Eurasian, although single terms, do represent multiracial groups. Reporting two ethnicities, such as German and Irish, illustrates the overlapping of the semantic categories of race and ethnicity observed during the cognitive research interviews.

A wide range of open-ended answers was found for the "something else" follow-up question for multiracials who identified as belonging to one of the races on the list, for example, black, white, and as "something else." The following list illustrates these combinations specifically: white, Mexican, American Indian, and German; black, Puerto Rican, German, and African American; white and Armenian; white, Italian, Dutch, and Irish. Thus, the pattern of equating ethnicity and racial groups was common in these open-ended responses as well.

From the analysis of the "something else" entries it became apparent that some of the "one-race equals 'something else'

**Table 2. Percentage of multiracials and indeterminate multiracials by Hispanic origin status, 1995 CPS supplement**

Hispanic origin status	Panel 2		Panel 4	
	Multiracial	Indeterminate multiracial	Multiracial	Indeterminate multiracial
Total .....	36.73	63.47	80.40	19.60
Named one race:				
Hispanic .....	2.24	10.74	.0	.71
Non-Hispanic .....	4.88	45.78	5.15	17.02
Named two or more races:				
Hispanic .....	.58	4.61	22.79	.0
Non-Hispanic .....	28.03	2.34	52.46	1.88

NOTE: Panels 1 and 3 are not shown because multiracial questions were not asked in these panels.

multiracials," such as Creole, *did fit* the definition of multiracial constructed for the statistical analysis. The analysis also revealed that some of the "two races equals race plus 'something else' multiracials" *did not fit* this definition, and should be classified as "indeterminate multiracials."

Following this analysis, the percentages of "multiracials" and "indeterminate multiracials" were calculated. Table 2 presents the percentages of both types of multiracials for Hispanics and for non-Hispanics. Persons who reported two or more of the racial categories other than "something else" are also included in the "multiracial" category.

The initial racial breakdown of the multiracial group displayed previously in table 1 indicated that about half of the one-race multiracials on both panels had identified themselves as belonging to one of the racial groups other than "something else." Observations made in the course of monitoring CPS supplement interviews provided a lead for the investigation of factors contributing to one-race reporting by some multiracial reporters. It had been observed that, occasionally, a CPS respondent who identified as "multiracial" and listed only one race, such as white, would include a second race in answering the later ancestry/ethnic origin question.

An analysis of entries to the ancestry/ethnic origin question for the one-race "indeterminate multiracials" who had listed a single race other than "something else" revealed that 54 percent of the 152 individuals in this group had listed a second race under ancestry. To learn how widespread this phenomenon was, an analysis of entries under ancestry was carried out for a random sample of 2,000 panel 2 and panel 4 respondents, drawn proportional to the racial distribution for those two panels, who had not identified as multiracial and had selected a racial response option other than "something else." This latter group was designated "single race respondents." For the single race respondents, only 7 percent of the entries under ancestry included a second race not previously named. Although the samples are small, there is a basis to suggest that the existence of a second racial group in their heritage contributed to the se-

lection of the "multiracial" designation for many of the "one-race indeterminate multiracials."

While the overwhelming majority, 91 out of 130 (70 percent), of the single race respondents reporting a second race under ancestry are whites reporting an American Indian, Eskimo, or Aleutian racial component, the one-race indeterminate multiracials are more varied in their racial composition. Whites constitute fewer than 50 percent of the latter group. Blacks at 36 percent, American Indians, Eskimos, and Aleuts at 11 percent, and to a lesser extent, Asians and Pacific Islanders at 6 percent, are overrepresented in this group, compared with the proportional racial distribution for the two panels. Future publications from BLS and Census will report the results of additional demographic analysis of these two groups.

## Conclusions

Testing of the initial version of the Current Population Survey Supplement on Race and Ethnic Origin indicated that many questions contained unfamiliar terms, such as "ethnicity," and presupposed an unrealistic level of sociological sophistication, for example, the ability to differentiate race, ethnicity, and ancestry. Some respondents were suspicious of the survey's purpose. After extensive cognitive research, a questionnaire was developed that was readily understood and no longer evoked a negative emotional response.

Cognitive research also allowed analysts to pinpoint the circumstances associated with multiracial reporting, leading to the recognition of true multiracial status for a smaller portion of the multiracials who reported a single race, that is, "something else." Similarly, this analysis led to the recognition of "indeterminate multiracial" status for a portion of multiracials who reported two races—a race plus "something else."

Cognitive research conducted in conjunction with the supplement thus contributed to solving problems associated with developing a questionnaire on race, ancestry, and ethnic origin.<sup>4</sup> □

## Footnotes

<sup>1</sup> The team consisted of: Adalberto Aguirre (University of California, Riverside); Ade Costa-Cash, Manuel de la Puente, and Eleanor Gerber (Bureau of the Census); Patricia Bell (Oklahoma State University); Lu Ann Moy (General Accounting Office); Ruth McKay (Bureau of Labor Statistics); Jorge Nakamoto (Aguirre International, Inc.); and Jackie Stanley (Department of Agriculture).

<sup>2</sup> Panel 1 contained separate Hispanic origin and race questions with no multiracial category; panel 2 contained separate Hispanic origin and race questions with a multiracial category; panel 3 contained a combined Hispanic origin and race question, with "Hispanic" a category on the list of races, and no multiracial category; panel 4 contained a combined Hispanic origin and race question, with "Hispanic" a category on the list of races, that also included a multiracial category.

<sup>3</sup> For more information on this and other cognitive methods, see Barbara H.

Forsyth and Judith T. Lessler, "Cognitive Laboratory Methods: A Taxonomy," in Paul P. Biemer and others, eds., *Measurement Errors in Surveys* (New York, John Wiley & Sons, Inc., 1991), pp. 393-418.

<sup>4</sup> For a more complete discussion of the cognitive research findings, see: Ruth B. McKay and Manuel de la Puente, "Cognitive Research in Designing the cps Supplement on Race and Ethnicity," *Proceedings of the Bureau of the Census' 1995 Annual Research Conference*, Rosslyn, VA, pp. 435-45; Manuel de la Puente and Ruth B. McKay, "Developing And Testing Race and Ethnic Origin Questions for The Current Population Survey Supplement On Race And Ethnic Origin," *Proceedings of the 1995 American Statistical Association Meetings*, Orlando, FL, pp. 19-28; and Ruth B. McKay, L.L. Stinson, M. de la Puente, and B. A. Kojetin, "Interpreting the Findings of the Statistical Analysis of the cps Supplement on Race and Ethnicity," *Proceedings of the Bureau of the Census' 1996 Annual Research Conference*, Rosslyn, VA (In press).

---

### LABSTAT available via World Wide Web

LABSTAT, the Bureau of Labor Statistics public database, provides current and historical data for many BLS surveys as well as numerous news releases.

LABSTAT Public Access has introduced a new production Internet service over the World Wide Web. BLS and regional offices programs are described using hypertext pages. Access to LABSTAT data and news releases is provided by a link to the BLS gopher server. The URL is:

<http://stats.bls.gov/blshome.html>

If you have questions or comments regarding the LABSTAT system on the Internet, address e-mail to:

[labstat.helpdesk@bls.gov](mailto:labstat.helpdesk@bls.gov)

---