

DEVELOPING AND TESTING AGGREGATE REPORTING FORMS FOR DATA ON RACE AND ETHNICITY

Susan Schechter
National Center for Health Statistics

Linda L. Stinson
Bureau of Labor Statistics

Luann Moy
U.S. General Accounting Office

Introduction

In October 1997, the Office of Management and Budget (OMB) published "Standards for Maintaining, Collecting, and Presenting Federal Data on Race and Ethnicity" (*Federal Register*, 62 FR 58781 – 58790). Federal agencies are now required to offer respondents the option of selecting one or more of a minimum of five racial categories. These categories are the minimum set for data on race for Federal statistics, program administrative reporting, and civil rights compliance reporting: American Indian or Alaska Native; Asian; Black or African American; Native Hawaiian or Other Pacific Islander; White. With respect to ethnicity, the standards provide for the collection of data on whether or not a person is of Hispanic or Latino origin. The standards do not permit a multiple response that would indicate an ethnic heritage that is both Hispanic or Latino and non-Hispanic or non-Latino.

The 1997 standards specify new ways in which data collection of race and Hispanic or Latino origin information should be obtained. First, self-report or self-identification using two separate questions, one asking Hispanic or Latino origin and one asking race, is the preferred method for collecting data on race and ethnicity. When race and ethnicity are collected separately, the ethnicity question should be placed first followed by the race question. Second, respondents should be offered the option of selecting one or more racial designations. Third, Native Hawaiian or Other Pacific Islander is a separate category from Asian.

As a result of the change in policy for collecting data on race and ethnicity, the reporting categories used to present these data must similarly reflect this change. Agencies are expected to provide as much detail as possible on all responses including multiple race responses, consistent with agency confidentiality and data quality procedures. The 1997 standards specify that, at a minimum, the number of individuals who marked one of the five race categories and the number who marked more than one race category are to be reported and that the race of those indicating Hispanic or Latino origin be reported if available. In addition to providing the number of people who marked one of the five racial categories, the standards strongly encourage data producers to provide detailed distributions of multiple responses--at a minimum, the total number of respondents reporting more than one race should be reported.

An interagency committee was formed to assist in developing, among other things, a reporting form that would aggregate data on race and ethnicity for a given population. The committee decided to use cognitive research methods to develop and test these forms. While cognitive interviews designed to pretest establishment forms are typically done on-site at establishments, the committee extended these methods and designed a research protocol which included both on-site and laboratory interviews (see Edwards and Cantor, 1991, for a discussion of using cognitive methods to test administrative forms).

Use of Cognitive Research Methods to Develop and Test Administrative Forms

For more than 15 years, the interdisciplinary efforts of survey methodologists, cognitive psychologists, and behavioral scientists have stimulated considerable interest in establishing cognitive pretesting of questionnaires as a standard component of survey research. By applying cognitive psychology techniques to develop and test data collection instruments, survey researchers continue to improve and expand methods used to interview small numbers of subjects in a laboratory environment in order to identify questionnaire problems (Lessler, Tourangeau, and Salter, 1989; Willis, Royston, and Bercini, 1991). Cognitive interviews, like the survey questionnaires they test, can be administered in a variety of ways and researchers must carefully consider pretest design decisions (Beatty and Schechter, 1994; DeMaio, Mathiowetz, Rothgeb, Beach, and Durant, 1993; Forsyth and Lessler, 1991). Most of the research to develop cognitive methods has centered on face-to-face interviews for household surveys. Less attention has been paid to developing methods to design and test administrative reporting forms, although evidence is appearing more frequently in the survey literature (e.g., Cantor and Phipps, 1999; Edwards and Cantor, 1991; Goldenberg, Butani, and Phipps, 1993; Gower and Nargundkar, 1991; Jenkins and Dillman, 1994; Martin and Tucker, 1999; Willimack, Nichols, and Sudman, 1999).

In looking at the guidelines this literature has to offer, one frequently sees use of examples of both directed and non-directed probing questions used successfully to study respondents' understanding of and reaction to the content of questions (e.g., DeMaio and Rothgeb, 1996; Hess and Singer, 1995; Esposito, Campanelli, Rothgeb, and Polivaka, 1991). Likewise, there has been much discussion about the reliability of these methods, including reviews of how and when they are used, debates over the use of cognitive techniques versus other pretest methods, and descriptions of specific protocols used in laboratory and on-site establishment settings (DeMaio and Rothgeb, 1996; Presser and Blair, 1994; Tucker, 1997; Willis, DeMaio, and Harris-Kojetin, 1999).

Testing administrative forms on-site at the establishment is grounded in the belief that the information retrieval processes can and will affect the quality of the responses and that it is critical for the researcher to observe this process at work. Turning specifically to the literature of establishment surveys, the emphasis is often placed on the various protocols used to identify and gather information from knowledgeable officials. For example, Edwards and Cantor (1991) developed and described a response model for establishment surveys that conceptualizes the completion of a form as an information system in which the respondent is asked to respond on behalf of the establishment. One can easily see the types of testing problems which could arise, not the least of which is that the information required by the form may not be collected by the agency or may be in a substantially different format than the reporting form. Processes such as information retrieval, selection of the most knowledgeable respondent (or respondents), and the working environment is thought to have

as much or more influence on the quality of establishment responses than simply whether the respondent fills out the form correctly. As pointed out by Edwards and Cantor (1991), the impact of such environmental forces (which also may be at work in regard to household surveys) are magnified for establishment surveys and may have even more effect on the survey response process.

Test Forms

The development of forms to report aggregate data on race and ethnicity was a collaborative effort among the committee members, experts in questionnaire design and survey research, and policy and statistical analysts from the federal government who were involved in the review of standards for data on race and ethnicity. In developing test forms, a decision was made to only use the minimum race categories specified in the 1997 standards. Thus, the test forms only asked for aggregate numbers of American Indians or Alaska Natives, Asians, Blacks or African Americans, Native Hawaiians or Other Pacific Islanders, and Whites and did not ask for aggregate subgroups such as Chinese, Japanese, Samoan, and so forth. This design was not to imply that agencies should only require aggregate reports of populations for these five groups but rather, was selected for ease of testing.

Three different types of forms were tested (Figures 1 and 2 contain portions of the first two forms). All of the forms required reporting of an individual's Hispanic or Latino origin status and separately, reporting of that individual's race. The committee recognized that many organizations collect and maintain data at the individual level that includes reporting Hispanic or Latino origin within other race categories (sometimes referred to as a combined format). However, the design and testing of the forms was an attempt to see how data reporters would approach the task of aggregating separate Hispanic or Latino origin counts with the expectation that in the future, agencies will gradually modify the ways in which individual race and Hispanic or Latino origin data are collected.

One form, referred to as Form RH-1 was designed to collect the specific reports of race and record these by the Hispanic or Latino origin responses. There were 31 reporting lines representing every combination of both single and multiple race responses for the five minimum race categories. Total numbers for each race group were then suppose to be entered under an appropriate Hispanic or Latino origin status column. This form conceptualized an automated format and was designed for easy expansion depending on desired reporting categories.

Two other forms were tested, referred to as Forms RH-2 and RH-3. Both of these forms had two parts--one part for the aggregate number of individuals who reported each single race, the number of individuals who reported more than one race, and the number of individuals for whom race information was missing, and a second part to count the number of times each race was included in a multiple race response. All numbers were also to be reported by Hispanic or Latino origin status. The difference between RH-2 and RH-3 was that RH-3 was designed to report crosstabulated data (e.g., race by Hispanic or Latino origin by gender).

Laboratory and On-Site Interview Methods

Cognitive research methods were used both for laboratory and on-site interviews. The interview protocol for the laboratory and on-site interviews were the same. Subjects were first shown the test form and asked for their overall impressions and reactions to the form. They were not given much

information about the 1997 standards but were told that the test form allowed for multiple race responses and that some categories might be different from those they had seen before. Subjects were then given two tasks. First, they had to review individual data and aggregate or tally in some way the race and Hispanic or Latino origin responses. Second, they had to record these tallies onto the test form. Interviewers observed the completion of these tasks and then probed for additional information on how participants understood and interpreted the various sections of the forms. Interviewers asked participants to explain the meanings of the terms used and to identify parts of the forms that were confusing, offensive or problematic.

Lab Interviews

Nine interviews were conducted in government cognitive laboratories. Participants included both Federal and private sector employees whose jobs typically involved completion and use of administrative forms. In order to test the process of completing the forms in the laboratory, subjects were given stacks of fictitious applications that included race and Hispanic or Latino origin as well as other demographic information. These “dummy” records were used to see how participants would complete the forms using different kinds of source data (see Figure 3). One set of dummy records contained a single race question which listed among the race groups the category Hispanic or Latino origin. The second set of dummy records contained a separate Hispanic or Latino origin question followed by a race question.

On-Site Interviews

On-site interviews included a variety of establishments including government agencies, correctional facilities, schools, and private sector businesses. In total, nine interviews were conducted on location and in most cases, actual data collection forms or computer-generated summaries rather than dummy records were used.

Results

Both laboratory and on-site subjects indicated some difficulty grasping the concept of multiple race reporting. Some participants perceived and counted reports of Hispanic as a race. Other subjects mentioned their uncertainty in reporting missing race or Hispanic or Latino origin information. Several participants were confused with the different column headings on the forms used for reporting Hispanic or Latino origin status. For RH-2 and RH-3, subjects had considerable difficulty understanding the requirement to report the number of times a race was reported among multiple race responses.

Laboratory subjects were especially cooperative in offering suggestions for ways to improve the forms. For example, some suggested that a worksheet be attached to the form to assist in tallying responses. A number of subjects noted that detailed instructions with examples would have been very helpful. Some subjects pointed out that the acronyms used on the forms (e.g., AIAN for American Indian or Alaska Native) were misleading and confusing.

Discussion

The findings from both the lab and the field were unexpectedly similar. Obvious, observable problems with the design of the forms were evident in both locations. Clearly, the forms appeared visually complex with many lines and blanks for entering data, so that subjects had to interpret the forms

before they could complete them. Thus, the process of understanding how the individual interacted with the form was not dependent upon the location where the interview occurred. This may have been due largely to the finding that most of the issues were cognitive and conceptual; they were not the result of system retrieval problems.

A significant benefit of conducting the on-site interviews was that actual data processing systems and details of the data collection processes were available for review. Going into the field and using actual establishment records allowed for a critical discovery-- the chasm between the draft forms and the original source data with which these participants worked. In many cases, the test sites were using a combined reporting format which included Hispanic as a race. Some laboratory subjects stated that their agency was using a combined format, but it was easier and clearer for the on-site subjects to discuss the implications of this issue while looking at their agency data. While this finding was critical, it is relevant to note that this could have also been determined with laboratory subjects had they been asked to bring agency data with them to the lab.

Apart from the issue of conceptualizing Hispanic or Latino origin as a race, subjects appeared to vary in their interpretation of actual categories of race. Subjects did not consistently offer the same definitions for race terms and, in some cases, subjects clearly misinterpreted terms. For example, one subject said that if Hispanic was the only item indicated on a corrections department intake form, she would enter Asian as the race because the two categories are similar.

Further, it was clear from these visits that race and ethnicity information is not always collected through self-reports. For example, the race and ethnicity assignments for prisoners in one facility were made by arresting officers, most typically on the basis of appearance (or surname if it appeared to be Spanish). In another case, an administrative secretary reported that she changed the self-reports of her clients if the information did not correspond with information she knew about the family and its history (e.g., she changed self-reports of White to Hispanic when she knew the individuals were born in Mexico, spoke Spanish, and had a Spanish surname). While this information did not directly impinge upon the design and revision of the reporting forms, it did provide further insight into data quality issues.

Thus the main difficulties were how to convey the various race and ethnicity categories that were being requested and how to provide a conceptual bridge between agency categories and those requested by the form. Consequently, the location for testing the form was as portable as the individual respondent with knowledge of the organization's racial categories. The work could be done in the lab or in the field. Doing the work in the field simply made it easier to review copies of the organization's forms and racial categories and identify the appropriate person(s) who typically would be filling out such a form.

Conclusion

Research is still continuing in order to revise the test forms, with more attention to developing instructions that are easy to understand and follow. One problem that can only be overcome in time is the inconsistencies among agencies in the ways data on race and Hispanic or Latino origin are collected. A system that collects every multiple race combination along with Hispanic or Latino origin information will allow the maximum flexibility for an agency in further reporting and analysis.

However, this option will not always be feasible and therefore, further efforts to design a variety of forms for agencies to choose from should continue.

Acknowledgments

Much appreciation is extended to the interagency committee and especially to Scott Fricker and Kathleen Sorensen. Portions of this paper were presented at the 1999 Joint Statistical Meetings in Baltimore, MD and can be found in the forthcoming Provisional Guidance on the Implementation of the 1997 Standards for the Collection of Federal Data on Race and Ethnicity to be issued by OMB.

Any opinions expressed in this paper are those of the authors and do not constitute policy or views of the National Center for Health Statistics, the Bureau of Labor Statistics, or the US General Accounting Office.

References

Beatty, P. and Schechter, S. (1994). An examination of mode effects in cognitive laboratory research. 1994 Proceedings of the Section on Survey Research Methods, American Statistical Association, 1275 - 1280.

Cantor, D. and Phipps, P. (1999). Adapting cognitive techniques to establishment surveys. In M. Sirken, T. Jabine, G. Willis, E. Martin, and C. Tucker (Eds.) A new agenda for interdisciplinary survey research methods: Proceedings of the CASM II seminar. Hyattsville, MD: US Department of Health and Human Services, 74-78.

DeMaio, T. J., and Rothgeb, J. (1996). Cognitive interviewing techniques: In the lab and in the field. In S. Sudman and N. Schwarz (Eds.) Answering questions. San Francisco: Jossey-Bass.

DeMaio, T., Mathiowetz, N., Rothgeb, J., Beach, M., and Durant, S. (1993). Protocols for pretesting demographic surveys at the Census Bureau. Report of the Pretesting Committee. Washington DC: Bureau of the Census.

Edwards, W. S. and Cantor, D. (1991). Toward a response model in establishment surveys. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman (Eds.), Measurement errors in surveys. New York: Wiley, 221-236.

Esposito, J. Campanelli, P. Rothgeb, J. and Polivaka, A. (1991). Determining which questions are best: Methodologies for evaluating survey questions. Proceedings of the Section on Survey Methods Research, American Statistical Association, 46-55.

Forsyth, B. and Lessler, J. (1991). Cognitive laboratory methods: A taxonomy. In P. Biemer, R. Groves, L. Lyberg, N. Mathiowetz, and S. Sudman, (Eds.), Measurement error in surveys. New York: Wiley.

Goldenberg, K. L., Butani, S., and Phipps, P.A. (1993). Response analysis surveys for assessing response errors in establishment surveys. Proceedings of the International Conference on Establishment Surveys, American Statistical Association, 290-299.

Gower, A. and Nargundkar, M (1991). Cognitive aspects of questionnaire design: Business surveys versus household surveys. Proceedings of the 1991 Annual Research Conference, Washington DC: Bureau of the Census, 299-312.

Hess, J. and Singer, E. (1995). The role of respondent debriefing questions in questionnaire development. No. 021, Working Paper Series, Survey Methodology Program, Institute for Social Research, University of Michigan.

Jenkins, C. and Dillman, D. (1994). The language of self-administered questionnaires as seen through the eyes of respondents. Presented at the 1994 Council of Professional Associations on Federal Statistics, Bethesda, MD.

Lessler, J., Tourangeau, R., and Salter, W. (1989). Questionnaire design research in the cognitive research laboratory. Vital and Health Statistics, 6(1), Hyattsville, MD: US Department of Health and Human Services.

Martin, E. and Tucker, C. (1999). Toward a research agenda: Future develop and applications of cognitive sciences to surveys. In M. Sirken, D. Herrmann, S. Schechter, N. Schwarz, J. Tanur, and Tourangeau, R. (Eds.). Cognition and survey research. New York: Wiley, 363-381.

Office of Management and Budget (1997). Revisions to the standards for the classification of federal data on race and ethnicity. Federal Register, Vol. 62 (210), 58781-58790.

Presser, S. and Blair, J. (1994). Survey pretesting: Do different methods produce different results? Sociological Methodology, 73-104.

Tucker, C. (1997). Methodological issues surrounding the application of cognitive psychology in survey research, Bulletin de Méthodologie Sociologique, 67-92

Willimack, D., Nichols, E., and Sudman, S. (1999). Understanding the questionnaire in business surveys. Paper presented at the Annual Meetings of the American Association for Public Opinion Research, St. Petersburg, FL.

Willis, G. B., DeMaio, T. J. and Harris-Kojetin, B. (1999). Is the bandwagon headed to the methodological promised land? Evaluating the validity of cognitive interviewing techniques. In M. Sirken, D. Herrmann, S. Schechter, N. Schwarz, J. Tanur, and R. Tourangeau, (Eds.). Cognition and survey research. New York: Wiley, 133-153.

Willis, G., Royston, P., and Bercini, D., (1991). The use of verbal report methods in the development and testing of survey questionnaires. Applied Cognitive Psychology, 5, 251-267.

Figure 1: A portion of Form RH-1

FORM RH-1		Individuals who marked YES, Hispanic Origin	Individuals who marked NO, Hispanic Origin	Individuals who did NOT provide Hispanic Origin information
Individuals who marked ONLY ONE race	1 White			
	2 Black/African American.....			
	3 Asian.....			
	4 American Indian/Alaska Native			
	5 Native Hawaiian/Other Pacific Islander.....			
Individuals who marked TWO races	6			
	7 White + Asian.....			
	8 White + Am Indian/Alaska Nat.....			
	9 White + Nat Hawaiian/OPI.....			
	10 Black/African Am + Asian.....			
	11 Black/African Am + Am Indian/Alaska Nat.....			
	12 Black/African Am + Nat Hawaiian/OPI.....			
	13 Asian + Am Indian/Alaska Nat			
	14 Asian + Nat Hawaiian/OPI			
Race missing	32 Individuals who DID NOT provide race information			
Total	33 Sum of rows 1 through 32			

NOTE: Form RH-1 contains rows 15-31 which are rows for individuals who marked three, four, and five race groups. For space reasons, only the first third of the form is shown above.

Figure 2: A portion of Form RH-2

FORM RH-2 REPORTING MULTIPLE RACES

Count of TIMES each race was marked for individuals who marked MORE THAN 1 race	Hispanics	NON Hispanics	Separate Hispanic Origin Question with no answer given
Number of times WHITE was marked			
Number of times BLACK/AFRICAN AMERICAN was marked			
Number of times ASIAN was marked			
Number of times AMERICAN INDIAN/ ALASKA NATIVE was marked			
Number of times NATIVE HAWAIIAN / OTHER PACIFIC ISLANDER was marked			

Figure 3: Examples of Race and Ethnicity Questions from the Dummy Records

Example 1 - Combined format used on dummy records

Race: Mark one or more

- | | |
|---|---|
| 01 <input type="checkbox"/> White | 04 <input type="checkbox"/> American Indian or Alaska Native |
| 02 <input type="checkbox"/> Black or African American | 05 <input type="checkbox"/> Asian |
| 03 <input type="checkbox"/> Hispanic or Latino | 06 <input type="checkbox"/> Native Hawaiian or Other Pacific Islander |

Example 2 - Two question format used on dummy records

9. Are you Spanish,
Hispanic or Latino?

- 01 Yes
02 No

10. Race: Mark one or more

- 01 White
02 Black or African American
03 American Indian or Alaska Native
04 Asian
05 Native Hawaiian or Other Pacific Islander