Elizabeth Martin, Theresa J. DeMaio, and Pamela C. Campanelli U.S. Bureau of the Census

Key Words: question order, questionnaire design

The survey methodology literature contains much documentation of the effect of question sequencing on response to the items. This work has, in large part, involved attitude items (see Schuman and Presser, 1981; Turner and Martin, 1984). The effect of question order may derive from the context invoked by prior questions, which may influence respondents' frame of reference or suggest differing interpretations of the question.

Context effects can also operate on quasifactual and even factual items, but research in this area has been limited. The potential for this type of effect is especially pronounced when the concept being measured is somewhat unclear, and the respondent really isn't sure what is being asked. In this paper, we explore the effect of question sequence on response to two potentially ambiguous, quasifactual items on the census questionnaire. Specifically, the focus of this paper is the race and Hispanic origin items that are included on both the long and short versions of the decennial census form.

In the next two sections of the paper, we will elaborate on the conceptual problems inherent in the measurement of race and Hispanic origin in general and in the census. We then describe two problems with the race and Hispanic origin data collected in the 1980 census--high levels of reporting of "Other" race, and item nonresponse on the Hispanic origin item. We present specific hypotheses about how context affects responses to the race and Hispanic origin items, and present results of a split-panel experiment involving the sequence of the race and Hispanic origin items on the census form.

THE MEASUREMENT OF RACE

Despite its familiarity, the concept of race is not a simple one. Racial classifications, both popular and scientific, are based on a mixture of principles and criteria: national origin, tribal membership, religion, language, minority status, physical characteristics, and behavior. The criteria and categories for racial classification vary among cultures and over time. In the United States, we are accustomed to think in terms of at least two major races: Black and White.

In this country, we tend to treat race as a biological fact, an objective, fixed characterristic of a person which is biologically inherited. This meaning of race is so ingrained that it may come as a surprise to learn that other cultures have very different conceptions of race (Marshall, 1968, Harris, 1968). For example, the racial categories recognized in Brazil are not the same as those used in the United States, even though its population also includes components with White European and Black African origins. More than 40 basic racial categories are used in Brazil, and these are combined to create hundreds of racial distinctions, based primarily on physical characteristics. Racial descent is not the rule; full siblings whose appearance differs are of different races in Brazil. In addition, race is not a fixed characteristic, and it changes when a person achieves wealth, since socioeconomic status is one of the criteria for race.

The meaning of race in Brazil is clearly different from the meaning of race in the United States. Cultures in Central and South America use different racial categories and/or criteria from those used in the United States. This difference in racial classifications implies that "White" and "Black" are not natural categories in terms of which most Spanish-speaking people think about race. This fact has implications for the consistency and meaningfulness of their answers to the census race question.

Even within the United States, there have been significant changes in how Americans understand and categorize race, and these changes have been reflected in changes in the race categories used in the census. The main changes have been in the classification of persons of mixed Black-White race; a proliferation of Asian categories; and changes in race classification of Spanish-speaking persons (see U.S. Bureau of the Census, 1979, for a description of the race questions that have been used in each census since 1790).

Since we think of race as a stable and enduring characteristic, it is surprising to find that no single set of race categories has been used in more than two censuses, and most were used only once. Of course some changes reflect real change in the composition of the U.S. population due to migration from Asia and Central and South America and expansion of U.S. territory. However, fluctuations in census race cate-gories suggest it is difficult if not impossible to devise a meaningful, objective classification Some anthropologists (for example, of race. Marshall, 1968) argue that all racial classifications are arbitrary and artificial, whether they are based on "scientific" or popular criteria. Problems arise when respondents do not share the race categories used by the Census Bureau. Evidence suggests that this is the case for many Spanish-speaking persons.

Until 1970, the census was collected by personal visit enumeration and race was determined by enumerator observation. In 1970, the Census Bureau began conducting the census by self-enumeration. Census questionnaires were mailed out to households and personal enumeration was conducted only for households that did not mail back a form. In 1970, 60 percent of households were enumerated by mail, and by 1980, 90 percent. (See Bounpane and Jordan, 1978; U.S. Bureau of the Census, 1973.)

The change in census-taking procedure meant that, after 1970, race was based on self-identification rather than enumerator observation. The change in procedure was associated with dramatic changes in the racial characteristics of the population. The number of Hispanic origin persons classified as "Other race" rose from 700,000 in 1970 to 5.8 million in 1980 (U.S. Bureau of the Census, 1987:100). The transition to self-enumeration contributed to the increase in "Other race" and decline in "White" race reporting among people of Hispanic origin. In addition, Hispanic origin persons were no longer recoded as "White" if they reported themselves as "Other race." Before 1980, persons of Latin descent had been classified by enumerators as White unless they were definitely Black, Indian, or another race.

Additional evidence suggests that selfenumeration is associated with high levels of "Other race" reported by Hispanics. Hispanics reported their race differently in the 1980 census and in reinterviews conducted after the census. While over one-third (35 percent) of Hispanics identified themselves as "Other race" in the census questionnaire, only 10 percent reported "Other race" in the personal visit reinterview (McKenney, Fernandez, and Masamura, 1985). The difference in "Other race" reporting by Hispanics may be a true difference in reporting, or interviewers may have changed "Other race" responses to "White" for respondents who "looked white." In either case, the results suggest that Hispanic respondents do not share the same race categories as Census Bureau enumerators, resulting in responses that appear to be inconsistent.

THE MEASUREMENT OF HISPANIC ORIGIN

In response to pressure from the Hispanic community (Choldin, 1986), the Hispanic origin item was included in the census for the first time in 1970. In 1970, the item appeared only on the long form, but in 1980 it was asked of everyone. Placement near the race item on the page containing all the 100-percent population items may have affected reporting for both items.

Hispanic origin had the highest nonresponse rate of any 100-percent item. The computer allocation rate was 4.2 percent for short forms and 2.3 percent for long forms (U.S. Bureau of the Census, 1986:32). One possible reason is that the item was redundant for respondents who had just reported an "Other race" and written in a Hispanic nationality. In addition, non-Hispanic persons may have left the item blank, thinking that a blank was the same as a "No" response. Other people may have left the item blank because they didn't understand it, or because they did not find a category that fit them.

We hypothesized that the high item nonresponse rate for the Hispanic origin question and the reporting of "Other" in the race question may be the result of context effects. The item that comes first on the census form may create a frame of reference that affects interpretation of the second item. In this case, the abundance of national origin groups listed as categories for race may encourage write-in entries of other nationality groups in that item. The Hispanic origin item then appears redundant, which may lead respondents to leave it blank (see figures 1 and 2).

The reporting of "Other" in the race item may or may not be affected by the order of the two items. The majority of the "Other" races which are written in are Hispanic nationalities. Hispanics may be more likely to report their race as "White" or "Black" rather than "Other" if they have already had a chance to report their Hispanic origin. However, if these Hispanics simply do not think of themselves as "White" or "Black," then their reporting of "Other" race may be insensitive to the order in which the items are asked.

We reasoned that the Hispanic origin item would seem less redundant if it was asked before race. People who think that their answer to the Hispanic origin question can be inferred from their response to the race question would be less likely to leave the Hispanic origin question blank if they answered it first. By reversing the order of the two items, we hoped to decrease item nonresponse for Hispanic origin.

We also thought that giving Hispanics a chance up front to report their Hispanic origin would create a more restricted frame of reference for the race item. We hypothesized that asking Hispanic origin first would reduce the reporting of "Other race" by Hispanics. This would be true, however, only to the extent that Hispanics actually view themselves as "White" or "Black."

METHODOLOGY

This research was conducted as part of an effort to improve the design of the decennial census long form questionnaire. In this redesign effort, typographic and other layout changes were made to increase the consistency of the form's appearance, clearer instructions were provided for respondents on how to complete the census form, questions were reworded to simplify and clarify concepts, and questions were reordered to improve the flow and coherence of the census form (see DeMaio, Martin, and Sigman, 1987, for further discussion of the goals and methods of the research).

Figures 1 and 2 contain facsimiles of the race and Hispanic origin items for the revised form and the 1986 test census form, which was used as a control. As can be seen, several changes were made to these items, the most important of which is that they are placed in reverse order. The rationale for this change and the hypothesized outcome of the item reversal were outlined in the previous section. In both forms, race and Hispanic origin were separated by other items (age in the revised form; age and marital status in the 1986 form).

Other changes were also made. The response categories for the Hispanic origin item were reordered, as they were throughout the form, so that "Yes" preceded "No." There were also differences in response categories in the revised form, which reflected the thinking within the Census Bureau at that point in time regarding what these items would look like in 1990. In the 1986 form, detailed Hispanic (for Hispanic origin) and Asian (for race) categories were listed separately. In the revised form, categories were combined with an instruction to respondents to write in their specific group.

The revised and 1986 forms were compared in a series of split-panel experiments. Our data were collected in April 1987, in about 30 group sessions organized by the Census Bureau Regional Offices in Boston, Dallas, Chicago, and Philadelphia. Volunteers were recruited to over-represent minority racial and ethnic populations with relatively little education. A total of 515 people filled out forms about themselves and members of their households, providing data on 1,446 persons. Participants included people aged 18 to 80, members of different racial and ethnic groups, and people with various levels of education.

During each session, half of the participants were randomly assigned the 1986 form and the other half were assigned our revised form. Even though these respondents do not represent a sample, the randomization by form type does permit us to make some statistical comparisons between forms. We tested for statistically significant form differences using chi-square tests that take into account the clustering of persons within group sessions, and within households. In the tables that follow, the chi-square values that are labelled X^2 are Pearson chi-squares, calculated on the assumption of simple random sampling; the J's refer to jack-knife statistics that take into account clustering in the data and thus represent a more conservative test. Likelihood ratio chi-squares (identified as L^2 in the tables) were used to test the fit of alternative log-linear models.

RESULTS

Form Differences in Distributions of Race and Hispanic Origin: Despite changes in the questions, the distributions for these two items are very close, as shown in Table 1. Thus, placing the race item last does not appear to affect the data obtained. As shown in the first panel of Table 1, each form identified about the same percentage of individuals in the categories of White, Black, Asian or Pacific Islander, and Other race. The revised form, however, identified a significantly larger percentage of American Indians. This result is due to sampling variability rather than any differences in the forms. Three American Indians, each with large families, were randomly assigned to receive the revised form.

The second panel of Table 1 contains the distribution of responses to the Hispanic origin item, collapsed into the categories on the revised form for comparability. The two forms obtained similar levels of reporting of Hispanic origin. Although based on small samples, this result is encouraging for two reasons. First, it suggests that the increased response to the item in the revised form (as reported below) did not alter the distribution. Second, there was some concern that switching the "Yes" and "No" categories would affect the distribution of the responses given. This concern appears to have been unwarranted.

Item Nonresponse: Examination of the item nonresponse rates in Table 2 suggests that our hypothesis concerning item nonresponse to the Hispanic origin item was supported. We were successful in significantly reducing nonresponse for the Hispanic origin item from 18 percent on the 1986 form to 9 percent on the revised form. In addition, the nonresponse rate for the race item was not detrimentally affected, with a rate of 3 percent for the 1986 form vs. 4 percent on the revised form. Although the number of

changes made to the form precludes us from knowing the exact cause of the difference in item nonresponse rates for Hispanic origin, the pattern of missing data by race suggests that the sequence of the items was responsible. As shown in the first panel of Table 3, reported race was significantly related to nonresponse on Hispanic origin for the 1986 form. For this form, people who reported their race as "Black" or "Asian/Pacific Islander" were far more likely to leave the Hispanic origin item blank than people reporting "White" or "Other race." On the revised form, however, there was no relationship between race and nonresponse on Hispanic origin: all race groups were equally likely to leave the Hispanic origin item blank (see second panel of Table 3). A loglinear model confirms the presence of a three-way interaction (Non-response to Hispanic origin X Race X Type of form). This finding supports our hypothesis that the race item conditions respondents' understanding of the intent of the Hispanic origin item. The pattern of differences is consistent with our reasoning that people who think response to the Hispanic origin item can be inferred from their response to the race item are less likely to leave the Hispanic origin item blank if they answered it first. Very few Blacks (2 percent) and no Asians report being of Hispanic origin. Thus, Blacks and Asians are very likely to think that the Hispanic item is redundant and that their response ("No") can be inferred from their response to the race item. When the Hispanic item is last (1986 form), Blacks and Asians are more likely to perceive the redundancy and not answer the item. When the Hispanic item is first (revised form), nonresponse is lower and unrelated to race.

<u>Content Differences--Race</u>: Our second major hypothesis concerned the level of reporting of "Other race" by persons of Hispanic origin. For both questionnaire versions, the vast majority of write-in entries in "Other race" were Hispanic (over 90 percent for both forms). This finding is consistent with prior research.

Further examination shows that the form of the questionnaire did affect Hispanic individuals' response to the race item. As shown in Table 4, respondents who reported Hispanic origin were more likely to report their race as "White" in the revised form (39 percent) than in the 1986 form (25 percent) although this difference is only marginally significant (p = .12). However, the extent of reporting "Other race" write-ins by Hispanics was substantial on both forms (61 percent for the revised form vs. 75 percent for the 1986 form). Thus, our goal of reducing reporting of "Other race" by Hispanics was partially successful; however, these data suggest that there are large numbers of persons of Hispanic origin who do not believe they fit into any of the major racial classifications measured in the census.

Further analysis shows that the effect of context was restricted to Hispanics who were born in the United States. Table 5 presents responses to the race item separately by form for people born in a U.S. state or outside the United States; the table includes data for respondents of Hispanic origin only. The first panel shows that Hispanic respondents who were born outside the U.S. are unaffected by the version of the questionnaire. For both forms, the vast majority (over three-quarters) identified their race as "Other" and wrote in a Hispanic nationality. However, questionnaire form had a very large effect on race-reporting by Hispanic respondents who were born in a U.S. State. U.S.-born Hispanics were much more likely to report their race as "White" in the revised form (74 percent) than in the 1986 form (22 percent). For this group, race-reporting was apparently quite dependent on the context of the question. (The three-way interaction between form, place of birth, and race is significant.) This result is consistent with our initial hypothesis that giving Hispanic respondents a chance to report their Hispanic origin before asking race would reduce the extent of "Other race" reporting.

It is interesting to speculate why our revised form reduced "Other race" reporting for U.S.-born Hispanics, but not for Hispanics who migrated to the mainland U.S. The difference may be due to U.S.-born Hispanics' greater assimilation and understanding of the meaning of the U.S. racial categories "White" and "Black." Hispanics who were born and raised in the mainland United States may find these terms more natural and acceptable than Hispanics who migrated from another cultural setting. The latter group appear unwilling to report themselves as either "White" or "Black," regardless of whether they have first reported Hispanic origin or not. The former group appears more variable in their race-reporting; like other Hispanics, they tend to report their race as "Other," but if they are first given a chance to report their Hispanic origin, they identify themselves as "White.'

These results imply that the measurement properties of census items about race and Hispanic origin are affected by question order effects, and that the measurement properties of these items vary systematically over the population. This result is consistent with Johnson's (1987) finding that the measurement properties of different indicators of Hispanicity are not constant over the population, but vary between first- or second-generation Hispanic immigrants versus others.

DISCUSSION AND CONCLUSIONS

Overall, these findings suggest that the revisions to the form tested in these experiments resulted in improvements in data quality. We increased the item response rate to the Hispanic origin item, which was one of the major hypotheses guiding this research on question sequence. This was done without affecting the distribution of substantive responses, and without affecting the response rate for the race item.

Our second hypothesis, that our changes would decrease reporting by persons of Hispanic origin in the "Other" category of the race item, was partially supported. While we did increase reporting by Hispanics in the "White" category, more than half of the persons who reported being of Hispanic origin still reported being "Other race." This suggests that perhaps the reporting of Hispanics in the "Other race" category is not an error, but simply a real perception by these

respondents of their "correct" place in the racial classification structure--neither "White" nor "Black." This interpretation is supported by our finding of an interaction effect between place of birth, form type, and reported race for respondents of Hispanic origin. Hispanics born in a U.S. State were influenced by the context of the race item in the hypothesized direction: they were more likely to report their race as experiment illustrates this point nicely. On one of the questionnaires (revised form), the respondent reported "Spanish" in the "Other "White" if they first answered the Hispanic origin question. Hispanics who were not born in a U.S. State were not influenced by question context: most of them reported "Other race" for both forms. The difference suggests that U.S.born Hispanics have begun to assimilate, to some "Black." Some anecdotal evidence from our race" category of the race item, and "Dominican" in the Hispanic origin item, for both herself and her husband, but reported their child's race as "White" and wrote in "American" for the This observation illus-Spanish origin item. trates perfectly how racial identity changes when one starts a new life in the United States. Clearly, this respondent was not thinking of race as a biological attribute which follows a rule of descent. This observation suggests that a differentiation is made within families concerning racial classification, in ways that don't correspond to our native ideas about how that classification should be made. While this may not be the way we expect that classification to be made, this respondent's answers were clearly not "errors" but were quite logical and consistent, in terms of her frame of reference.

Unfortunately, due to the number of experi-mental manipulations that were introduced in these experiments, it is not possible to firmly conclude that the observed results were due to the change in the sequence of the race and Hispanic origin items. Some of our results are marginal in this relatively small sample. However, we are currently conducting a large, carefully controlled experiment, and we hope to have more conclusive evidence in the near future.

REFERENCES

- Bounpane, Peter, and C. Jordan. 1978. "Plans for Coverage Improvement in the 1980 Census." Paper presented at the Annual Meeting of the American Statistical Association, San Diego, California.
- Choldin, Harvey M. 1986. "Statistics and Politics: The 'Hispanic Issue' in the 1980 Census." <u>Demography</u>, Volume 23, Number 3. DeMaio, Theresa J., and Elizabeth Martin. 1987. "Documentation of Revisions to 1986 Test
- Census Long Form." Unpublished paper. March 16, 1987.
- Harris, Marvin. 1968. "Race." In <u>International</u> <u>Encyclopedia of the Social Sciences</u>, Encyclopedia of the Volume 13, pp. 263-268.
- "Evaluating U.S. Johnson, Robert A. 1987. Census Bureau Measures of Hispanic Origin: An Approach Based on Latent Structure Analysis." Paper presented at the Annual Meeting of the American Statistical Association, San Francisco, California.

- Marshall, Gloria. 1968. "Racial Classifications: Popular and Scientific." pp. 149-64 in Science and the Concept of Race, edited by M. Mead, T. Dobzhansky, E. Tobach, and R. Light. New York: Columbia University Press.
- York: Columbia University Press. McKenney, Nampeo R., Edward W. Fernandez, and Wilfred T. Masamura. 1985. "The Quality of the Race and Hispanic Origin Information Reported in the 1980 Census." <u>Proceedings</u> of the Survey Research Methods Section (American Statistical Association), pp. 46-50.
- Schuman, Howard and Stanley Presser. 1981. Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording, and Context. New York: Academic Press.
- Turner, Charles F. and Elizabeth Martin. 1984. Surveying Subjective Phenomena. New York: Russell Sage Foundation/Basic Books.
- U.S. Bureau of the Census. 1973. <u>The Coverage</u> of <u>Housing in the 1970 Census</u>. Census of Population and Housing: 1970. Evaluation and

Research Program PHC (E)-5. Washington, DC: Government Printing Office.

- U. S. Bureau of the Census. 1979. <u>Twenty</u> <u>Censuses: Population and Housing Questions</u>, <u>1790-1980</u>. Washington, D.C.: U.S. Government Printing Office.
- U.S. Bureau of the Census. 1986. <u>Content Rein-</u> terview Study: <u>Accuracy of Data for Selected</u> <u>Population and Housing Characteristics as</u> <u>Measured by Reinterview</u>. 1980 Census of Population and Housing Evaluation and Research Reports, PHC 80-E2.
- U.S. Bureau of the Census. 1987. <u>1990 Planning</u> <u>Conference Series</u>. No. 12, Race and Ethnic Items, July 1985. Washington, D.C.: U.S. Bureau of the Census.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the statistical assistance of Robert Fay and comments by Robert Johnson and Irv Schreiner.

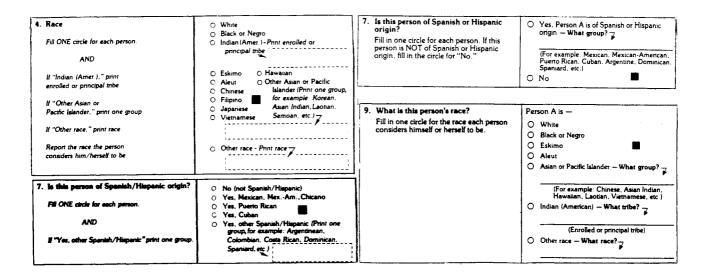


FIGURE 1. Race and Spanish Origin Items on the 1986 Form (reduced) FIGURE 2. Spanish Origin and Race Items on the Revised Form (reduced)

TABLE 1: Distribution of Race and Hispanic Origin by Type of Form

	<u>Ra</u> 1986 Form	<u>ce</u> <u>Revised Form</u>]	Hispanic 1986 Form	<u>c Origin</u> Revised Form
White Black Asian or Pacific Islander	53% 30 3	53% 31 2 3	Hispanic Non-Hispanic Total	23% 77 100	19% <u>81</u> 100
American Indian Other Race Total	$\frac{14}{100}$	$\frac{11}{100}$	Number of persons with complete dat on Hispanic origi		678
Number of persons with complete data on Race	684	715			
the second s	X2 = 3.37, d			X2 = 2.49, 0	df = 1, n.s.
*American Indian category Less than 1 percent	excluded from	test			

TABLE 2: Item Nonresponse Rates for the Race and Hispanic Origin Items by Type of Form

	Response	Response to Race Item		panic Origin Item		
	1986 Form	Revised Form	1986 Form	Revised Form		
No response	3%	4%	18%	9%		
Valid response	97	96	82	91		
Total	100	100	100	100		
Number of persons	703	743	703	743		
	$\chi^2 = 1.31$,	df = 1, n.s.	$\chi^2 = 25.13, d$ J = 2.78, d	f = 1, p < .001 f = 1, p < .01		

TABLE 3: Item Nonresponse to Hispanic Origin Item by Race and Type of Form

	1986 Form				Revised Form					
	White	Black	Asian or <u>Pacific</u>	Other Race	White	Black	Asian or Pacific	Other Race		
No response on Hispanic origin	7%	36%	41%	4%	8%	11%	0%	5%		
Valid response on Hispanic origin	93	64	59	96	92	89	100	95		
Total	100	100	100	100	100	100	100	100		
Number of persons with complete data on Race, excluding American Indians	361	204	22	96	380	220	16	80		
	x2 ر	= 98.40, = 5.80,	df = 3, p < df = 3, p <	.001 .001	χ2	χ^2 = 4.83, df = 3, n.s.				

Three-way interaction: $L^2 = 25.63$, df = 2, p < .001 J = 1.93, df = 2, p < .05

TABLE 4: Reporting of White vs. Other Race Among Hispanics by Type of Form

	1986 Form	Revised Form
White Other Race Total	25% 75 100	39% <u>61</u> 100
Number of persons with complete data on Hispanic Origin and Race	122	111
Nate	X ² = 4.76, c J = .81, c	lf = 1, p < .05 lf = 1, p = .12

TABLE 5:	Reporting	of	White	vs.	0ther	Race	Among	Hispanics	bу	US	Birth	and	Туре	of	Form
----------	-----------	----	-------	-----	-------	------	-------	-----------	----	----	-------	-----	------	----	------

	Born Outs 1986 Form	ide the US Revised Form	Born in 1986 Form	in a US State Revised Form			
White Other Race Total	17% <u>83</u> 100	18% 82 100	22% <u>78</u> 100	74% <u>26</u> 100			
Number of Hispanic persons with complete data on Place of Birth		44	32	38			
	$\chi^2 = 0.$	01, df = 1, n.s		df = 1, p < .01 df = 1, p = .01			
Three-way interaction:	$L^2 = 8.34, J = 1.04,$	df = 1, p < .01 df = 1, p = .08					