

**COMPARING RESPONSES TO THE RACE QUESTION ASKED
TWO DIFFERENT WAYS**

Barbara Lepidus Carlson

Mathematica Policy Research, P.O. Box 2393, Princeton, New Jersey 08543-2393

KEY WORDS: Survey, Race, Community Tracking Study, Directive 15

Abstract. In preparation for the 2000 decennial Census, the government developed new guidelines for asking about race in federally sponsored surveys. In addition to separating Asians and Pacific Islanders into two categories, the main change was to allow for more than one race category to be given as a response. Although not a federally sponsored survey, the Community Tracking Study Household Survey (sponsored by the Robert Wood Johnson Foundation) uses federal data sources as benchmarks when post-stratifying its weights. In the first three rounds of data collection (1996-97, 1998-99, 2000-01), we asked about the race of respondents allowing for only one race category to be entered. Because federal benchmark surveys, such as the March 2003 Supplement of the Current Population Survey and the 2000 Census, will have asked about race following the new OMB guidelines, we revised the CTS race question accordingly for round four (2003). More than half of the CTS Household Survey round four sample will have also responded in round three. This paper compares the responses of people who were in both rounds to see how the new format of the race question affected their responses to that question.

Background. In 1977, the U.S. government's Office of Management and Budget (OMB) developed guidelines for asking about race in federally funded surveys, called Directive 15. In 1997, OMB updated these guidelines based upon recommendations resulting from a number of studies of proposed changes. Among the most significant changes were that: (1) the racial category of "Asian or Pacific Islander" was to be broken out into two categories ("Asian" and "Native Hawaiian or Other Pacific Islander"), and (2) the respondent could choose more than one racial category to describe himself. Previously, people who considered themselves to belong to more than one racial category either chose the category they identified with most, or responded with "other race" and then provided a verbatim response indicating the relevant racial categories. Because Hispanic ethnicity is not considered to be a racial category according to Directive 15, OMB recommends asking about Hispanic ethnicity separately from (and prior to) asking about race; however, many Hispanic respondents consider "Hispanic" to be their race, and often they too choose

to respond with "other race" and then provide a verbatim response.

At the 2002 Joint Statistical Meetings, David Hubble at the U.S. Bureau of the Census gave a presentation¹ reporting on the Bureau's attempts to develop bridging parameters from responses to the new race question to those of the single-response version of the race question. He presented findings from an experiment using the Census Quality Survey (CQS) that involved initially recontacting a sample of 2000 census respondents in the summer of 2001 (just over a year after the census) by mail with in-person follow-up. Half of the sample was asked the race question the same way as it was asked on the census form (allowing for multiple race categories). Unexpectedly, he and his co-authors found that people in this group often did not respond the same way at two points in time with respect to multiple race categories. In fact, among the roughly 17,000 people in the CQS who had reported two or more races on the census form, only 47 percent reported two or more races one year later (using the same question format and same interviewing mode as in the census). The authors proposed further research to better understand this inconsistency in reporting.

Data. At about the same time as this presentation, the Center for Studying Health System Change (HSC) was proposing changes to the survey instrument for round four of the Community Tracking Study (CTS) Household Survey. The CTS, which is designed and sponsored by HSC, is intended to document how the U.S. health care system is changing over time; to document how those changes affect people; and focuses on local health care markets as well as the nation as a whole. It is comprised of multiple components (several surveys plus site visits). Mathematica Policy Research (MPR) designed the sample, implemented the survey (including data collection), and developed weights and variance estimation parameters for the

¹*Study of Responses to the Census 2000 Race Question Instruction: "Mark One or More Races" and Bridging to Single Race Distributions* (David Hubble, James Poyer, and Michael Bentley, U.S. Census Bureau).

Household Survey. Social and Scientific Systems is responsible for final data processing and file production.

One of the changes that was being proposed to the CTS questionnaire in 2002 was to modify the race question to conform to the new OMB guidelines; in particular, to allow the respondent to choose more than one race category. Because the CTS Household Survey is funded by the Robert Wood Johnson Foundation, and not the federal government, it was not bound by the new guidelines; however, because the CTS survey weights are poststratified to demographic totals provided by census estimates, we decided to revise the CTS weight question according to the new guidelines.

The three prior rounds of data collection for the CTS Household Survey took place in 1996-97, 1998-99, and 2000-01. Round four is taking place in 2003. In round one, the main sample component was a nationally representative list-assisted random-digit dial sample of phone numbers. Starting in round two, the sample consisted of a combination of newly selected phone numbers and phone numbers from the prior round (both respondents and nonrespondents). Because we had phone numbers in common from round three to round four, this means that we are likely to have many of the same people responding in both rounds. This provided us an opportunity to investigate how the same people responded to the race question asked two different ways two-and-a-half years apart.

Because the CTS Household Survey sampling unit is the phone number (and not the household), we do not attempt to find households who responded in the prior round if they no longer have the same phone number in the current round. In addition, household composition can change over the course of a couple of years. To carry out this investigation, we made use of the household enumeration section of each round's questionnaire to match people across the two rounds as best we could. One other caveat is that the race question, although asked about each household member, is answered by one respondent per family unit within the household. This respondent is supposed to be the person most knowledgeable about the family's health care, but may not be the same person from round to round.

We took a snapshot of the data as of May, 2003, about four months into data collection. Because the bulk of the initial sample releases happened to be "re-interview" cases (prior round completes), we had a fairly large number of completed interviews to use for this investigation. By the time we limited our sample to those people thought to be in common between the two rounds, we still had a fairly large number of people (about 13,000). One caveat is that these people may not represent the U.S. population as

a whole, because they tend to be among the more stable in terms of mobility and household structure.

The Questions. In round three, the Hispanic ethnicity and race questions were asked as follows:

--Do you consider yourself to be of Hispanic origin, such as Mexican, Puerto Rican, Cuban, or other Spanish background?

--What race do you consider yourself to be?
INTERVIEWER: CODE MIXED RACE IN OTHER

- WHITE**
- AFRICAN AMERICAN OR BLACK**
- NATIVE AMERICAN (AMERICAN INDIAN) OR ALASKA NATIVE**
- ASIAN OR PACIFIC ISLANDER**
- OTHER [SPECIFY]**

MPR telephone interviewers are generally instructed not to read aloud text appearing in capital letters, so they would tend not to read the race categories unless the respondent was hesitating and/or unsure what was being asked. Note the interviewer instruction to code mixed race responses in "other."

In round four, the comparable questions were:

-- Do you consider yourself to be of Hispanic origin, such as Mexican, Puerto Rican, Cuban, or other Spanish background?

--I'm going to read you a list of five race categories. Please choose one or more races that you consider yourself to be.

- White**
- African American or Black**
- American Indian or Alaska Native**
- Asian**
- Native Hawaiian or Other Pacific Islander**
- OTHER [SPECIFY]**

Note that the same Hispanic ethnicity question was used in both rounds. But the race question differed in a few respects. First, the Asian and Pacific Islander racial categories are separated in round four. Second, the interviewer tells the respondent that he or she can choose one or more races. Third, the interviewer is reading the race categories aloud (except for "other")—although they were told they did not have to read the entire list of race categories if the person responded before they finished reading them (unlike standard interviewer procedures for similarly structured questions). Due to the sensitive nature of the question for some people, we did not want to unnecessarily provoke people who already provided their response by continuing to read down the list.

Coding. In both rounds three and four, a number of respondents were coded under the “other” race category, which allowed the interviewer to record the verbatim response. For this exercise, I attempted to code each verbatim response into one of the existing categories, if possible. I used as my guideline the OMB classification standards for racial categories. Some responses turned out to be uncodable in terms of racial categories, such as “American,” “Indian,” “Muslim,” and flippant responses such as “human,” and “UFO.” I created a specific category for people who responded that they were multiracial, but who did not specify the races.

Many of the responses were people who reported their country or countries of European ancestry (for example, “part Irish, part French”). I coded such people in the “white” racial category. As per the guidelines, people who responded that they were “Arab,” “Middle-Eastern,” or from a North African country were coded as “white.” Pakistani and [East] Indian responses were coded as “Asian.”

A large number of “other” responses were from Hispanic respondents, who responded that their race was “Hispanic,” “Latino,” “Spanish,” or identified their race in terms of their country of origin (for example, Colombian, Dominican, Puerto Rican, or Mexican). I created a Hispanic racial category for such responses.

Results. Note that all of the results shown in this paper are unweighted sample counts, and so do not reflect any population distributions.

Table 1a shows the cross-tabulation of race by Hispanic ethnicity in round three, with the “other” category not yet coded, and Table 1b shows the same cross-tabulation for round four.² The Hispanic classification here is based solely on the response to the Hispanic ethnicity question that preceded the race question. Note that the round four responses allow for more than one racial category to be reported without resorting to the “other” category.

As you can see from Table 1a, the “other” category comprised 2.5 percent of the non-Hispanics in round three. In round four (Table 1b), the percent of non-Hispanics coded as “other” decreased significantly (to 0.5 percent) as expected, with 2 percent specifying more than one race.

Table 1a. Round 3 Race/Ethnicity

	Non-Hispanic	Hispanic	Total
White	10,127 (86.7%)	167 (30.4%)	10,294 (84.2%)
Black	1,035 (8.9%)	11 (2.0%)	1,046 (8.6%)
Native Am	56 (0.5%)	3 (0.5%)	59 (0.5%)
Asian/Pac	169 (1.4%)	1 (0.2%)	170 (1.4%)
“Other”	288 (2.5%)	367 (66.8%)	655 (5.4%)
Total	11,675 (100.0%)	549 (100.0%)	12,224 (100.0%)

Table 1b. Round 4 Race/Ethnicity

	Non-Hispanic	Hispanic	Total
White	10,425 (86.9%)	275 (48.2%)	10,700 (85.1%)
Black	1,039 (8.7%)	16 (2.8%)	1,055 (8.4%)
Native Am	54 (0.5%)	18 (3.2%)	72 (0.6%)
Asian	168 (1.4%)	1 (0.2%)	169 (1.3%)
Haw/Pacific	14 (0.1%)	1 (0.2%)	15 (0.1%)
“Other”	62 (0.5%)	234 (41.0%)	296 (2.4%)
White + Native	143 (1.2%)	2 (0.4%)	145 (1.2%)
White + Black	14 (0.1%)	0 (0.0%)	14 (0.1%)
White + “Other”	24 (0.2%)	17 (3.0%)	41 (0.3%)
Black + Native	20 (0.2%)	0 (0.0%)	20 (0.2%)
Other pairs	22 (0.2%)	4 (0.7%)	26 (0.2%)
Three or more	13 (0.1%)	3 (0.5%)	16 (0.1%)
Total	11,998 (100.0%)	571 (100.0%)	12,569 (100.0%)

While one would not necessarily expect the racial distribution among Hispanic respondents to change because of the new question format, there were some interesting shifts between rounds three and four. The percent of Hispanics classifying themselves as “White” increased from 30 percent in round three to 48 percent in round four (and to 51 percent if we include “White + Native American” and “White + Other”). The percent classifying themselves as “Native American” increased from 0.5

² The numbers do not match precisely between the two rounds (Tables 1a and 1b) because there are different missing value patterns in the two surveys. The numbers in these tables exclude responses for people who have a missing value for either the race or the Hispanic ethnicity question in that round.

percent in round three to 3 percent in round four. The percent classifying themselves as “other” decreased from 67 percent in round three to 41 percent in round four. Almost 5 percent of Hispanics chose to report more than one racial category in round four using the new question format.

Tables 2a and 2b show how the “other” verbatim responses to the race question were coded in rounds three and four, respectively.³

Table 2a. Round 3 “Other” Races – Coded

	Non-Hispanic	Hispanic	Total
White	176 (61.3%)	2 (0.5%)	178 (27.3%)
Black	3 (1.0%)	0 (0.0%)	3 (0.5%)
Native American	3 (1.0%)	1 (0.3%)	4 (0.6%)
Asian	7 (2.4%)	0 (0.0%)	7 (1.1%)
Hispanic	3 (1.0%)	333 (91.5%)	336 (51.6%)
Multiracial	11 (3.8%)	8 (2.2%)	19 (2.9%)
Uncodable	57 (19.9%)	0 (0.0%)	57 (8.8%)
White + Native	14 (4.9%)	0 (0.0%)	14 (2.2%)
White + Black	3 (1.0%)	0 (0.0%)	3 (0.5%)
White + Asian	3 (1.0%)	0 (0.0%)	3 (0.5%)
White + Uncodable	3 (1.0%)	0 (0.0%)	3 (0.5%)
Hispanic + White	0 (0.0%)	10 (2.7%)	10 (1.5%)
Hispanic + Black	0 (0.0%)	4 (1.1%)	4 (0.6%)
Hispanic + Native	1 (0.3%)	2 (0.5%)	3 (0.5%)
Other pairs	1 (0.3%)	2 (0.5%)	3 (0.5%)
Three or more	2 (0.7%)	2 (0.5%)	4 (0.6%)
Total	287 (100.0%)	364 (100.0%)	651 (100.0%)

³ Note that these figures do not precisely match the “other” race figures in Tables 1a and 1b due to missing values in the verbatim text file used for coding. Also, Table 1b contains some “other” race responses in the categories of “Other pairs” and “Three or more”; these cases are found in Table 2b.

Table 2b. Round 4 “Other” Races – Coded

	Non-Hisp	Hispanic	Total
White	60 (60.6%)	1 (0.4%)	61 (17.3%)
Black	6 (6.1%)	0 (0.0%)	6 (1.7%)
Native American	4 (4.0%)	1 (0.4%)	5 (1.4%)
Asian	6 (6.1%)	0 (0.0%)	6 (1.7%)
Pacific	2 (2.0%)	0 (0.0%)	2 (0.6%)
Hispanic	4 (4.0%)	228 (90.1%)	232 (65.9%)
Multiracial	1 (1.0%)	1 (0.4%)	2 (0.6%)
Uncodable	11 (11.1%)	9 (3.6%)	20 (5.7%)
White + Native	1 (1.0%)	0 (0.0%)	1 (0.3%)
White + Black	1 (1.0%)	0 (0.0%)	1 (0.3%)
White + Asian	3 (3.0%)	0 (0.0%)	3 (0.9%)
Hispanic + White	0 (0.0%)	8 (3.2%)	8 (2.3%)
Hispanic + Black	0 (0.0%)	3 (1.2%)	3 (0.9%)
Other pairs	0 (0.0%)	1 (0.4%)	1 (0.3%)
Three or more	0 (0.0%)	1 (0.4%)	1 (0.3%)
Total	99 (100.0%)	253 (100.0%)	352 (100.0%)

There are a few things to note in Tables 2a and 2b. First, in both rounds, over 90 percent of Hispanics who were coded as “other” race gave as their verbatim response something that was coded as Hispanic. The next most common response for these Hispanic respondents in round three was coded as “White and Hispanic,” followed by the unspecified “Multiracial” category. In round four, the next most common response (other than “Uncodable”) was “Hispanic and White.”

Overall, combining Hispanics and non-Hispanics, 52 percent of those coded as “other” race in round three had their verbatim response coded to Hispanic; 29 percent were coded to one of the original race categories; 8 percent were coded as multiracial (either with specified or unspecified races); and 3 percent as Hispanic plus some race category.

In round four, 66 percent had their verbatim response coded to Hispanic; 23 percent to one of the original race categories; 3 percent as multiracial; and 3 percent as Hispanic plus some race category. Even though the round four race question allowed for more than one race category to be reported, nine people were coded as “other” who then went on to report more than one race category in their verbatim response.

Note also a few people in each round who responded “no” to the Hispanic ethnicity question, but whose racial classification (coded from the verbatim response) was “Hispanic.” Four of these are people who reported their race as “Brazilian” (which is generally classified as “Hispanic” even though Brazil is not a Spanish-speaking country) and one as “South American.” (The two others reported “Mexican” and “Hispanic.”) One of the people who reported “Brazilian” in round three responded “Indian from Brazil” in round four, and so was coded in round four as “Native American.” While these numbers are very small, they may illustrate broader self-classification issues.

Tables 3a and 3b examine who reported multiple races in each of the two rounds. In round three, the only way multiple races could be reported was through the “other” category. In round four, multiple races could be reported directly by the respondent choosing more than one of the given race categories; in addition, some multiple races were reported via the “other” category.

Table 3a. Round 3 Multiple Races (Coded from “Other”)

White + Native American	14	20.9%
White + Black (+ Uncodable)	4	6.0%
White + (Asian and/or Pacific)	4	6.0%
White + Uncodable	3	4.5%
Black + Native American	1	1.5%
Hispanic + White	10	14.9%
Hispanic + Black	4	6.0%
Hispanic + Native American	3	4.5%
Hispanic + Asian	1	1.5%
Hispanic + Uncodable	1	1.5%
Hispanic + White + Asian	2	3.0%
Multiracial (unspecified)	20	29.9%
Total	67	100.0%

In round three, almost one-third of the 67 people reporting multiple races reported “Hispanic” as one of the races. About 21 percent reported the combination of White and Native American. In round four, 13 of the 20 reporting multiple races under “other” reported “Hispanic” as one of the races; but the vast majority (93 percent) of those reporting multiple races in round four did so by choosing from among the racial categories in the

question. Once again, the most common combination of race categories was White and Native American (51 percent) in round four.

Table 3b. Round 4 Multiple Races

	Selected multiple race categories	Coded as multiple from “other” verbatim	Total
White + Native	145 (54.9%)	1 (5.0%)	146 (51.4%)
White + Black	16 (6.1%)	1 (5.0%)	17 (6.0%)
White + (Asian or Pac)	8 (3.0%)	3 (15.0%)	11 (3.9%)
White + “Other”	41 (15.5%)		41 (14.4%)
Black + Native	20 (7.6%)		20 (7.0%)
Black + (Asian or Pac)	3 (1.1%)		3 (1.1%)
Black + “Other”	5 (1.9%)		5 (1.8%)
Native + “Other”	2 (0.8%)		2 (0.7%)
Asian + (Pacif or “Other”)	8 (3.0%)		8 (2.8%)
White+ Black + Native	4 (1.5%)		4 (1.4%)
White+Native + “Other”	2 (0.8%)		2 (0.7%)
Black+Native + “Other”	2 (0.8%)		2 (0.7%)
Wh+Black+ Asian+ “Oth”	2 (0.8%)		2 (0.7%)
Other three or more	6 (2.3%)		6 (2.1%)
Multiracial (unspec)		2 (10.0%)	2 (0.7%)
Hisp + White (+Uncodable)		9 (45.0%)	9 (3.2%)
Hispanic + Black		3 (15.0%)	3 (1.1%)
Hispanic+ Uncodable		1 (5.0%)	1 (0.4%)
Total	264 (100.0%)	20 (100.0%)	284 (100.0%)

The prior tables showed results separately for rounds three and four, although limited to those persons in common between the two rounds. Tables 4 through 6 below show the linked results for these persons.

Table 4. Comparing Race in Rounds 3 and 4

Round 3	Round 4				
	Selected race		Coded from "other" verbatim		Total
	1 race	2+ races	1 race	2+ races	
1 race (selected race category)	11,139	167	125	7	11,438
1 race coded from "other" verbatim	390	15	182	5	592
2+ races coded from "other" verbatim	24	9	8	5	46
Total	11,553	191	315	17	12,076

The bolded figures in Table 4 are those reflecting a change from single to multiple races, or multiple to single races, between rounds three and four. While the numbers are very small here, 32 of the 46 people who reported more than one race in round three (via the "other" category) reported only one race when given the multiple-race opportunity in round four. Table 5 shows the responses given in the two rounds by these 32 people.

Table 5. Who Reported 2 or More Races in Round 3 and Only 1 Race in Round 4

Round 3 (coded "other")	Round 4	Number
White + Native Amer	White	5
	Native Amer	2
White + Black	White	1
	Black	1
White + Asian (+ Pacific)	White	1
	Asian or Pac	2
White+Uncodable	White	3
Black +Native Amer	Black	1
Hispanic + White	White	5
	"Other": White	1
	"Other": Hisp	3
Hispanic + Black	Black	1
Hispanic + Native Amer	White	1
	"Other": Hisp	1
Hispanic + Asian	White	1
Hispanic + Uncodable	White	1
Hisp+White + Asian	"Other": Hisp	2
Total		32

Note that 16 of these 32 people involve "Hispanic" being reported as one of the multiple coded races in round three.

Table 6 shows the 194 people who reported one race in round three and more than one race in round four (see bolded figures in Table 4).

Table 6. Who Reported 1 Race in Round 3 and 2 or More Races in Round 4

Round 3	Round 4	Number
White	White + Native	116
	White + Black	5
	Another pair	5
	"Other": Hisp+White	3
Black	White + Black	5
	Black + Native	16
	Another pair	1
	Three or more	5
Native Amer	White+Black	1
	White + Native	9
	Black + Native	2
	Another pair	1
Asian/Pac	"Other": White+Native	1
	Another pair	2
	Pair w/"oth": Wh+Asn	1
	"Other": White+Asian	1
"Other": Hisp	White + Native	1
	Wh+ "Oth": Hsp+Wh	1
	Another pair	1
	"Other": Hisp+White	3
"Other": Multi	"Other": Hisp+Uncod	1
	White + Native	3
	White + Black	1
	Three or more	1
"Other": White	White + Native	1
	White + Black	2
	Another pair	1
"Other": Native	Black + Native	1
"Other": Uncod	White + Native	2
	White + Black	1
Total		194

The most common race category added was Native American. More than two-thirds of these 194 people reported only White or Black in round three, and then added Native American as a second race in round four.

Summary of Results. Adding the ability to report multiple races reduced the percent classified as "other" by 80 percent among non-Hispanics.

Among Hispanics, more reported their race (one or more) and fewer were coded as "other" under the newer format of the race question.

Hispanics made up half of the "other" race reports in round three, but two-thirds in round four,

due to the reduction of reports of multiple races under “other” in round four.

Multi-racial reports made up 7.5 percent of the “other” race reports in round three, but only 2.6 percent in round four.

Of those who reported more than one race in round three, 70 percent reported only one race in round four, when given the opportunity to report more than one. Half of these cases involved reporting “Hispanic” as one of the races in round three.

Most of those who reported more than one race in round four, but only one race in round three, added Native American as a second race.

While this exercise was not intended to replicate the design used by Hubble et al. in their experiment, we did see some similar results in terms of how consistently people report multiple races when given the opportunity at two points in time. Hubble et al. found that less than half of those people who reported more than one race on the census form did so again

when asked the same question one year later. We found that, among the 46 people in our matched sample who reported more than one race (using a verbatim response) in round three of the CTS, there were only 14 who reported more than one race in round four. Among these 14, only 9 made use of the explicit instruction that they could report more than one race category.

Future work. I plan to re-examine these data with the complete round four data. At that time, we may use a more detailed algorithm to match persons between rounds three and four. Similar analysis should be carried out on any other longitudinal surveys that have updated the race question format to be in compliance with the new federal guidelines.

Acknowledgment. The author would like to acknowledge the assistance of Michael Watts of MPR, who prepared the matched data file used in this paper.