# A More Informative Set of Race and Ethnicity Questions for the United States Census

MIT Student 17.269 Race and Ethnicity 14 March 2017

## Why Race and Ethnicity?

The United States has had many motives for measuring race and ethnicity in the population across its short history. The three-fifths compromise of the late 1780s called for race to be measured to determine representation in Congress—with slaves (blacks) counting as 3/5 of a white person. As time has passed, races have been added and removed on the census for various reasons ranging from changing concepts of racial makeup (delineating southern Europeans from Northern Europeans in the 1880s to considering all Europeans as white today), racist and mal-intentioned motives (deportation), and even experimentation (measuring fertility of mixed-race individual to determine the truth of the superiority of "biologically pure" races) (Desmond and Emirbayer, 2015; Sollors, 2002).

Today, race is measured not because researchers believe it to be a biological reality (the foundation upon which race was first conceptualized) but rather, as Tukufu Zuberi approaches it, as a social construct which has influenced the lives of Americans through history and which continues to have implications for modern day Americans (Zuberi, 2001). Individuals have been denied opportunities and rights ranging from civil to social to political and educational based on their race and ethnicity (Blank et al., 2004), and therefore asking about race is necessary to understand cumulative discrimination and the lasting impact of these historical events as well as current discriminatory policies, such as New York's "Stop-and-Frisk" policy (Gelman et al., 2007).

## **The Current US Census**

The current United States census employs two questions to gain data about the overall racial and ethnic make-up of the United States population. The first question (Figure 1 question

	Question 9 about race. For this census, Hispanic origins are not races.					
8.	Is Person 1 of Hispanic, Latino, or Spanish origin?					
	No, not of Hispanic, Latino, or Spanish origin  Yes, Mexican, Mexican Am., Chicano  Yes, Puerto Rican  Yes, Cuban  Yes, another Hispanic, Latino, or Spanish origin — Print origin, for example, Argentinean, Colombian, Dominican, Nicaraguan, Salvadoran, Spaniard, and so on.   ✓					
9.	What is Person 1's race? Mark ▼ one or more boxes.  White Black, African Am., or Negro American Indian or Alaska Native — Print name of enrolled or principal tribe. ▼					
	☐ Asian Indian       ☐ Japanese       ☐ Native Hawaiian         ☐ Chinese       ☐ Korean       ☐ Guamanian or Chamorro         ☐ Filipino       ☐ Vietnamese       ☐ Samoan					
	☐ Chinese ☐ Korean ☐ Guamanian or Chamorro ☐ Filipino ☐ Vietnamese ☐ Samoan					
	☐ Chinese ☐ Korean ☐ Guamanian or Chamorro ☐ Filipino ☐ Vietnamese ☐ Samoan ☐ Other Asian — Print race, for example, Hmong, Laotian, Thai, race, for example, Fijian, Tongan,					
	☐ Chinese ☐ Korean ☐ Guamanian or Chamorro ☐ Filipino ☐ Vietnamese ☐ Samoan ☐ Other Asian — Print race, for example, Hmong, Laotian, Thai, Pakistani, Cambodian, and so on. ☐ ☐ Other Pacific Islander — Print race, for example, Fijian, Tongan, and so on. ☐					
	☐ Chinese ☐ Korean ☐ Guamanian or Chamorro ☐ Filipino ☐ Vietnamese ☐ Samoan ☐ Other Asian — Print race, for example, Hmong, Laotian, Thai, race, for example, Fijian, Tongan,					

Figure 1 Current format of United States Census questions on race and ethnicity.

8) asks the census taker about their ethnicity. Rather than giving a wide range of options for this question, the census only allows individuals to identify as having Hispanic, Latino, or Spanish ethnic roots or not. Because Hispanic origins are not classified as races in this census, this is an opportunity for individuals who are typically considered racially white to be able to differentiate themselves from other white populations such as those of Northern European descent. This requires that

researchers must use both of these questions in tandem to form meaningful groups and to compare those groups across a variety of variables such as life expectancy, average income, homeownership, and others.

The second question on this topic (Figure 1, question 9) asks the individual to select one or multiple races. The option of selecting multiple races is a more recent development for the census, originating on the 2000 census. Individuals are given a selection of 12 races with the option of writing in any not listed. A common occurrence on this question is for individuals who marked that they have Hispanic or Latino origins in question 8 to write in Hispanic or Latino for

their race as well. This begs the question of whether it is necessary to have two separate questions on race and ethnicity when individuals seem to consider them as the same.

There are several pros and cons to the current way that the census asks about race and ethnicity. The current census assumes that only Hispanic, Latino, or Spanish ethnicities are significant enough to be worth asking about. Every other ethnic group is lumped into the category of "not of Hispanic, Latino, or Spanish origin" and those individuals are assumed to be primarily defined by their race—whatever that means to them. This means that any variations in income, home ownership, or other information gathered by the census cannot be delineated by any other ethnic lines. This strongly limits the usefulness of the ethnic data by itself and does not allow us to meaningfully answer how ethnicity in general affects the lives of Americans.

The current census questions are also vague in terms of what they are asking for. What makes up race and racial identity is an ever-changing and ill-defined subject. Race has ties to geography, personal identity, skin color, cultural identity, and more. Asking a single question about race leaves the respondent to define the term on their own. Each census taker may take different variables into account as they select which race they belong to. Because of this, it is questionable to what degree one person's answer can be meaningfully compared to another's as individuals are unlikely to place the same weight on each component that can make up racial identity.

These questions also present race and ethnicity as facts, rather than malleable concepts that have the potential to change over time. Since the census has been shown to have broad implications as to how Americans view race (Sollors, 2002), this wording has the potential to

give individuals the idea that race is a biological fact or well-defined truth—a historically dangerous concept.

## **Proposed Modifications to the Census**

Questions

This paper proposes replacing the current census questions with six alternative questions, presented in full in Appendix 1. Each of the proposed questions is formulated to ask about a slightly different aspect of race and ethnicity. When combined, these questions will form a more holistic characterization of not only what racial and ethnic groups make up the US, but how Americans themselves view race and ethnicity. These proposed questions are:

**Question 1** Person 1 considers themselves as from... (select one or more country)

**Question 2** Person 1 has ancestry in... (select one or more region)

**Question 3** Person 1 describes their race as... (select one or more race)

**Question 4** Others have (rightly or wrongly) described Person 1's race as... (select one or more race)

**Question 5** Person 1 describes their ethnicity as... (select one or more ethnicity)

**Question 6** Of the previous questions (1-5), Person 1's experiences are most strongly shaped by... (select one answer)

Each of these questions approaches race and/or ethnicity from a slightly different perspective. Because race and ethnicity are such vague concepts, asking about specific aspects of race and ethnicity allows us to gather data that is more self-consistent in that respondents know more clearly what the question is asking. This approach is less ambiguous in what it is asking from the respondent in each question and allows a more expansive exploration of the individual's identity by isolating many possible definitions of race and ethnicity.

The first of the proposed questions is designed to touch on an individual's cultural upbringing. Which present-day country(s) the individual most strongly identifies with is explored here. This question probes an aspect of ethnicity—one that seeks to capture broad present-day influences on the individual that stem from the society(s) that they consider most formative for them as an individual.

The second question asks about the individual's ancestry. Companies such as ancestry.com and 23 andme are particularly drawn to this conception of race as being delineated by early populations of different geographic regions. This question may be useful to track the implication of having a certain geographic ancestry in the United States. What is important to note about this question, however, is that it is not necessarily a reflection of an individual's physical appearance. There may be an individual who marks that they have Native American roots who has light skin and freckles. Additionally, the term "ancestry" suggests a long time scale, but exactly how long is up to each respondent to determine. One individual who puts Northern American ancestry may mean that they have Native American (Navajo for example) ancestry, while another may mean their parents or grandparents lived in the United States.

The third question explicitly asks about race. This question allows the individual to answer based upon their personal conception of race. This question differs from the current census question on race in that it follows two other questions that touch on racial ideals.

Because the individual has already provided two answers that could contribute to their definition of race—geographic ancestry and geographic identification—they do not need to wonder if their definition of race "needs" to be rooted in either of these things. If their personal

conception of race lies in their geographic ancestry, their answer to this question will be the same as their answer to question 2. However, if race for the individual is more a function of their personal experiences or other factors, they can feel free to answer the question in that manner as they have already provided their geographical ancestry and will not feel like they need answer in that way because it might be what the surveyor means.

The fourth question attempts get at the individual's visual appearance and possible discrepancies in life outcome that may be a result of explicit cultural and systematic racism. Asking what race others have classified the person as allows researchers to explore whether race as self-identification is most important for life outcome or if race as others perceive it is most important. Does being classified by others as black or middle eastern put individuals at a fiscal disadvantage as compared to individuals who tend to be classified by others as white?

The fifth question allows the individual to provide their ethnic identity. This question will include more response options than the current census' ethnicity question (see Appendix 1), but the list of ethnic identities provided is certainly not exhaustive. Several options are given to allow the survey respondent to understand what could be considered an identity. If the respondent feels strongly about having a cultural identity that is not included on the list, they are free to write it in the "other" option. If they do not feel strongly about their ethnicity, they can select "none" or simply "American." This question varies significantly from the current ethnicity question in that it is not a binary choice between Hispanic or Latino and not, allowing for a more fine-grained analysis of the possible effects of cultural identification in the United States.

The sixth question helps to identify American attitudes towards race and ethnicity. If asked every census, a clearer perception of what components of racial identity are significant for individuals at different time points in the United States and across different racial and ethnic groups will be available.

### **Answers**

The answer choices respondents are given can affect both response rate and quality of data. Having more choices has been shown to decrease the response rate on questions (Tugend, 2010). Question 1 errs on the side of more options rather than fewer. The hope is that the nature of the survey (the US Census) will encourage individuals to respond despite that large amount of questions. However, if a lower response rate is observed, fewer options should be given on future iterations. With the multitude of options, more fine-grained data can be obtained. Additionally, simple data regressions can still be performed by grouping answers into larger categories (such as using North American instead of American and Canadian separately).

For question 1, countries are divided into subregions based on the UNSD definition. Only countries with populations over 1 million are included as possible answers, however individuals are free to write in other countries in a region under the "other" option in that region. The answers to question 2 are divided into the same subregions as are present in question 1.

## Conclusion

The current questions on race and ethnicity on the US Census, though elegantly simple, miss potentially important and worthwhile aspects of race and ethnicity in the United States. By using a more thorough set of questions, the proposed census questions will be able to create a more accurate picture of the United States population without compromising the potential for

simple data. This set of data will not only allow research into how life outcomes vary with race, but also will allow them to answer what aspects of race contribute most strongly to disparities in outcomes. With race and ethnicity being hazy subjects in their exact definition, collecting more detailed data will allow researchers to understand how Americans conceptualize race and ethnicity, and how this changes over time.

### Sources

Blank, Rebecca M. Dabady, Marilyn. Citro, Constance F (2004). *Measuring Racial Discrimination*. National Research Council, 2004.

Desmond, Matthew. Emirbayer, Mustafa (2015). Race in America. Norton, New York, 2015.

Gelman, Andrew. Fagan, Jeffrey. Kiss, Alex (2007). An Analysis of the New York City Police Department's "Stop-and-Frisk" Policy in the Context of Claims of Racial Bias. Journal of the American Statistical Association, September 2007, Vol 201, No. 479.

Tugend, Alina (2010). *Too Many Choices: A Problem That Can Paralyze*. New York Times 26 February, 2010.

Sollors, Werner (2002). What Race are You?. The New Race Question: How the Census Counts Multiracial Individuals Ch. 11. Russell Sage Foundation, 2002.

Zuberi, Tukufu (2001). *Thicker than Blood: How Racial Statistics Lie.* Regents of the University of Minnesota, 2001.

# **APPENDIX 1**

Proposed Census Questions

## 1) Person 1 considers themselves as from... (select one or more)

#### North America Northern Europe Eastern Africa Central Asia Bermuda Denmark Burundi Kazakhstan Canada Estonia Eritrea Kyrgystan Greenland Finland Ethiopia Tajikistan **United States** Ireland Turkmenistan Kenya Uzbekistan Other\_\_\_ Latvia Madagascar Other Lithuania Malawi Central America Mauritius Norway Eastern Asia Costa Rica Sark Mozambique China Sweden Rwanda Clipperton Island **United Kingdom** Somalia Japan El Salvador Other\_\_\_\_ South Sudan Macau Guatemala Tanzania Mongolia Honduras North Korea Uganda Mexico Western Europe Zambia South Korea Nicaragua Austria Taiwan Zimbabwe Panama Belgium Other Other Other France Germany Southern Asia South America Middle Africa Netherlands Afghanistan Argentina Switzerland Angola Bangladesh Bolivia Cameroon Other India Brazil Central African Republic Iran Chile Chad Eastern Europe Nepal Colombia Democratic Republic of Belarus Pakistan Ecuador the Congo Bulgaria Sri Lanka **Paraguay Equatorial Guinea** Czech Republic Other Peru Gabon Hungary Uruguay Republic of the Congo Poland Southeastern Asia Venezuela Other Republic of Moldova Cambodia Other Romania Indonesia Russia Laos Caribbean Northern Africa Slovakia Malaysia Cuba Algeria Ukraine Myanmar Dominican Republic Egypt Other **Philippines** Haiti Libya Singapore Jamaica Morocco Southern Europe **Thailand** Puerto Rico Sudan Albania Timor-Leste Trinidad and Tobago Tunisia Bosnia and Herzegovina Viet Nam Other\_\_ Other Croatia Other Greece Oceana Italy Western Asia Southern Africa Australia Macedonia Armenia Botswana Fiji **Portugal** Azerbaijan Lesotho Micronesia Serbia Bahrain Namibia New Zealand Slovenia Georgia South Africa Papua New Guinea Spain Iraq Samoa Swaziland Other Israel Other \_\_\_\_\_ Solomon Islands Jordan Tonga Kuwait Other Svria Lebanon Polynesia Turkey Oman Micronesia **United Arab Emirates** Qatar Melanesia

Yemen

Other\_

Saudi Arabia

Palestine

The Americas	Europe	Africa	Asia
(Native Populations)	Northern Europe	Northern Africa	Central Asia
North America	Western Europe	Middle Africa	Eastern Asia
Central America	Eastern Europe	Eastern Africa	Southern Asia
South America	Southern Europe	Southern Africa	Southeastern Asia
Caribbean	Unknown Region	Unknown Region	Western Asia
Unknown Region	Oceana		Unknown Region

# 2) Person 1 as ancestry in... (select one or more)

## 3) Person 1 describes their race as...

White Black or African American American Indian or Alaska Native Latino/a Middle Eastern Asian Indian Chinese Japanese Korean Filipino Vietnamese Native Hawaiian Guamanian or Chamorro Samoan Other Asian\_ Other Pacific Islander \_\_\_ Some other race \_

# (select one or more)

American	Kurd
American Indian or	Norwegian
Alaskan Native	Oromo
Arab	Pashtun
Armenian	Persian
British	Punjabi
Bulgarian	Russian
Catalan	Scottish
Cornish	Serb
Chinese	Slovak
Croat	Spaniard
Czech	Swede
Dutch	Tibetan
French	Ukranian
Greek	Uzbek
Hispanic or Latino	Welsh
Indian	Zulu
Irish	Other
Italian	
Japanese	None
Jewish	

# 4) Others have (rightly or wrongly) described Person 1's race as ...

White
Black or African American
American Indian or Alaska Native
Latino/a
Middle Eastern
Asian Indian
Chinese
Japanese
Korean
Filipino
Vietnamese
Native Hawaiian
Guamanian or Chamorro
Samoan
Other Asian
Other Pacific Islander
Some other race

(select one or more)

# 5) Person 1 belongs describes themselves as which ethnicity... (select one or more)

Question 1: geographic affiliation

Question 2: ancestry Question 3: race

Question 4: mistaken race Question 5: ethnicity Questions 1-5 equally None of these questions

6) Of the previous questions (1-5) Person 1's experiences are most strongly impacted by... (select one)

## **APPENDIX 2** Theoretical Comparison of Question Sets

To see how the proposed system compares to the current one, take for example an Indian individual who was born in the UK and moved to the United States when they were young. Let's call this individual P1. In answering the current census questions, P1 would likely respond:

Table 1. Possible answers to current census question by P1.

**Question 1** No, not of Hispanic, Latino, or Spanish origin

**Question 2** Asian Indian

When asking many individuals with a similar background these two questions, each one would likely respond the same way (perhaps also identifying as racially white). This is positive as it allows for consistency between individuals of similar backgrounds. However, it also lumps together individuals that may have significantly different experiences even though they share the same race—what if race is not the most influential factor? A racially Indian individual born in India would mark the same responses as P1, but may not have the same experiences and life outcome.

When confronted with the proposed set of six questions, P1 may answer in a variety of ways, a few of which are explored in Table 2. In the current format, the individual is defined solely by their race with no regard for their cultural upbringing (except for Hispanic populations). In column A (Table 2), P1 describes their race as Indian, but their ethnicity as British. In column B, they describe their ethnicity as American. These different ways of answering give researchers insight into how Americans identify in terms of race and ethnicity and how that might affect their life. Having multiple questions allows the individual to express multiple facets of their identity, something that a single-question response does not allow for.

Table 2. Possible answers to proposed census questions from P1

	Α	В	С
Question 1	The UK	The US	The UK and the US
Question 2	India and Europe	India and Europe	India and Europe
Question 3	Indian	White	White and Indian
Question 4	Indian	Indian	Middle Eastern
Question 5	British	American	Indian
Question 6	Question 1	Question 3	Question 4

Unlike the current questions, which are most accurate when used together, the proposed questions can each stand alone as an individual lens through which to view other data. More information can be explored by using the questions together and comparing the magnitude of the effect of each of the questions.

# MIT OpenCourseWare https://ocw.mit.edu

17.269 Race, Ethnicity, and American Politics Spring 2017

For information about citing these materials or our Terms of Use, visit: https://ocw.mit.edu/terms.