

Supplement to:

Flores, D. René, Edward Telles, and Ilana M. Ventura.
2024. “New OMB’s Race and Ethnicity Standards
Will Affect How Americans Self-Identify” *Sociologi-
cal Science* 11: 1147-1169.

Online Supplement for

New OMB's race and ethnicity standards will affect how Americans self-identify

René D. Flores, Edward Telles, and Ilana M. Ventura

René D. Flores
Email: renedf@uchicago.edu

This PDF file includes:

Sampling and survey procedures
Tables S1 to S5
Figures S1 to S5

Sampling and Survey procedures

The survey was fielded by YouGov, a reputable online polling company. The fieldwork took place in two stages: the initial stage spanned from July to September 2023, followed by a second stage conducted between January and March 2024. YouGov maintains an opt-in database of over 24 million people worldwide, across more than 50 countries, to take part in their surveys. It routinely collects demographic, socioeconomic, political, and geographic data on participants. Participation was voluntary and respondents were compensated for their time. Informed consent was obtained on the front page of our online survey. YouGov also has its own consent procedures with the survey panels they recruit. The survey was designed to take under 20 minutes. The median completion time was 17.1 minutes. To ensure quality data, we included two standard attention checks. Respondents who failed any of them were excluded.

Weights

YouGov interviewed 7,806 respondents who were then matched down to a sample of 7,350 to produce the final dataset (1,500 non-Hispanic Whites, 1,000 non-Hispanic Blacks, 2,150 non-Hispanic Asians, 2,400 Hispanics, and 300 respondents who identified in all other racial categories). The respondents were matched to a sampling frame on gender, age, race (overall sample only), and education. The sampling frame is a politically representative "modeled frame" of US adults, based upon the American Community Survey (ACS) public use microdata file, public voter file records, the 2020 Current Population Survey (CPS) Voting and Registration supplements, the 2020 National Election Pool (NEP) exit poll, and the 2020 Cooperative Election Survey (CES), including demographics and 2020 presidential vote. The matched cases were weighted to the sampling frame using propensity scores.

The matched cases and the frame were combined and a logistic regression was estimated for inclusion in the frame. The propensity score function included age, gender, race/ethnicity (overall sample only), years of education, region, and Presidential vote choice. The propensity scores were grouped into deciles of the estimated propensity score in the frame and post-stratified according to these deciles.

The weights were then post-stratified on a three-way (four-way for overall sample only) stratification of gender, age (4-categories), race (3-categories, 4-categories for overall sample only), and education (4-categories), to produce the final weights. As robustness checks, we include unweighted estimates of our main analyzes (See Table S5).

Pre-registration

We registered the research design and analysis plan in October 2023 at Open Researcher and Contributor ID, an online repository of social science experiments and observational studies (osf.io/489sq), before analyzing the data. However, there are some differences between the pre-registered plan and our final paper. First, our final sample size is larger and covers a broader segment of the population. Initially, we planned for a sample of 5,000 respondents representative of individuals identifying as White, Hispanic, Black, and Asian. However, to enhance external validity, we modified our sampling design after the pre-registration to ensure it represented the entire U.S. population, regardless of self-identification. As a result, we conducted a second round of data collection in Spring 2024, which resulted in a final sample of 7,350 cases.

Further, some of the tests outlined in the pre-registration were ultimately not performed. We did not test how the combined question impacted the following traits among ethnracial

groups: U.S. region, self-reported street race, employment status, and co-ethnic concentration in respondents' zip codes.

Lastly, we included some tests in the final paper that were not part of the pre-registration. Specifically, we did not pre-register the analysis shown in Table 2, which examines the ethnoracial identification of individuals who previously identified as Hispanic in YouGov's surveys. Similarly, the analyses in Table S3, which presents logistic regressions predicting Black and Indigenous identification among self-identified Hispanics, were not pre-registered. These analyses emerged during the writing process as we sought to uncover the mechanisms driving the observed effects.

Hand coding of responses

Respondents who provided detail in their answers to the race question were coded into its corresponding major 1997 OMB race group, according to the Census's 2020 Hispanic Origin and Race Code List, following procedures outlined by the Census (Marks and Ríos-Vargas 2021). Importantly, respondents who provided a Hispanic origin write-in in the "Some other race" category were left as "Some other race" and those who provided a Middle Eastern origin were coded as White for the separate questions format. We created a MENA flag from write-in's from this treatment for the sake of our MENA estimates only. However, in the combined question format, respondents who provided a write-in of a Hispanic origin were coded as racially Hispanic, and those who provided a write-in of a MENA origin were coded as racially MENA.

Index Supplementary Tables and Figures

1. Table S1. Comparison of PERI sample with the 2021 American Community Survey
2. Table S2. Randomization Check: Predicting Assignment to Treatment Condition
3. Table S3. Logistic regressions predicting Black and Indigenous Identification among self-identified Hispanics
4. Table S4. Unweighted Counts of SOR Write-in's by Type
5. Table S5. Ethnic and racial Identification by experimental condition (unweighted estimates)
6. Figure S1. Ethnic and racial Identification by experimental condition
7. Figure S2. Comparing Treatment Effects on Identifying as Black by Selected Variables (Latino respondents only)
8. Figure S3. Comparing Treatment Effects on Identifying as American Indian/Alaska Native by Selected Variables (Self-identified Latino Respondents Only)
9. Figure S4. Screenshots of the web survey questions. Separate Question Format (*control* condition)
10. Figure S5. Screenshots of the web survey questions. Combined Question Format (*experimental* condition)

Table S1. Comparison of PERI sample with the 2021 American Community Survey

	Non-Hispanic Whites		Hispanics		Non-Hispanic African-Americans		Non-Hispanic Asians	
	ACS	PERI	ACS	PERI	ACS	PERI	ACS	PERI
Population size	64.02 (0.03)	64.09 (0.56)	16.9 (0.02)	16.4 (0.43)	12.68 (0.02)	12.56 (0.39)	6.7 (0.02)	5.89 (0.27)
Educ: HS	58.59 (0.04)	58.02 (1.17)	60.32 (0.08)	72.36 (0.9)	67.59 (0.09)	70.25 (1.42)	39.59 (0.11)	40.79 (1.07)
Educ: College	36.37 (0.04)	37.96 (1.15)	17.79 (0.06)	18.03 (0.78)	23.03 (0.08)	23.33 (1.32)	51.96 (0.12)	55.32 (1.09)
Female	50.78 (0.04)	51.06 (1.18)	49.72 (0.08)	50.05 (1.01)	52.89 (0.1)	53.22 (1.55)	52.79 (0.12)	53.59 (1.09)
Age: 18-29	17.83 (0.03)	18.11 (0.91)	27.11 (0.07)	27.86 (0.9)	24.42 (0.08)	24.57 (1.34)	21.45 (0.1)	21.25 (0.89)
Age: 30-44	23.42 (0.03)	22.65 (0.99)	31.56 (0.08)	31.47 (0.94)	27.98 (0.09)	27.13 (1.38)	31.57 (0.11)	30.6 (1.01)
Age: 45-64	33.06 (0.03)	33.83 (1.12)	29.78 (0.08)	29.36 (0.92)	31.19 (0.09)	32.36 (1.45)	30.71 (0.11)	31.54 (1.01)
Age: 65+	25.7 (0.03)	25.4 (1.03)	11.54 (0.05)	11.31 (0.64)	16.41 (0.07)	15.94 (1.14)	16.27 (0.09)	16.62 (0.81)

Given that all PERI respondents are adults, the data from the American Community Survey only includes respondents aged 18 and above.

Table S2. Randomization Check: Predicting Assignment to Treatment Condition

VARIABLES	(1) Combined question format
Age (<i>ref=18-29</i>)	
30-44	-0.105 (0.066)
45-64	-0.108 (0.067)
65+	-0.003 (0.08)
Female	-0.083 (0.047)
Multilingual	0.038 (0.056)
Skin Color (<i>ref= Light</i>)	
Medium	-0.017 (0.053)
Dark	-0.056 (0.064)
U.S. born	0.118 (0.062)
Finished college	0.057 (0.053)
Home Owned	0.016 (0.051)
High income (\$100k+)	-0.012 (0.061)
Republican	0.052 (0.061)
Constant	-0.038 (0.092)
Observations	7,350
Pseudo R-squared	0.0027

Source: Project on Ethnicity, Race, and Immigration (PERI). Standard errors in parentheses. Log odds reported. *** p<0.001, ** p<0.01, * p<0.05

Table S3. Logistic regressions predicting Black and Indigenous Identification among self-identified Hispanics

VARIABLES	(1) Black ID	(2) Black ID	(3) Black ID	(4) Indig ID	(5) Indig ID	(6) Indig ID
Combined question	-0.302 (0.523)	-0.324 (0.619)	0.155 (0.639)	-0.652* (0.329)	-2.215* (0.696)	-0.908* (0.436)
National origin (ref= Mexican)						
Puerto Rico	1.431* (0.461)			0.155 (0.406)		
Cuba	1.312* (0.569)			-0.669 (0.772)		
Other	1.015* (0.499)			0.456 (0.336)		
PR*combined question	-0.271 (0.743)			-0.276 (0.728)		
Cuba*combined question	-0.275 (0.952)			-0.388 (1.518)		
Other*combined question	0.039 (0.732)			-1.448* (0.738)		
U.S. born		0.542 (0.475)			-0.117 (0.31)	
U.S. born*combined question		-0.053 (0.702)			1.538* (0.754)	
Skin Color (ref= Light)						
Medium			1.155* (0.569)			0.694* (0.321)
Dark			2.628*** (0.551)			0.769† (0.414)
Medium*combined question			-0.444 (0.797)			-0.237 (0.588)
Dark*combined question			-1.091 (0.798)			-0.115 (0.727)
Constant	-3.686*** (0.338)	-3.359*** (0.425)	-4.139*** (0.468)	-2.401*** (0.197)	-2.217*** (0.257)	-2.703*** (0.24)
	0.055	0.010	0.093	0.045	0.044	0.047
Observations	2460	2460	2460	2460	2460	2460
Max-rescaled R-Square	0.0553	0.0108	0.0938	0.0454	0.0448	0.0473

Source: Project on Ethnicity, Race, and Immigration (PERI). Standard errors in parentheses. Log odds reported. *** p<0.001, ** p<0.01, * p<0.05, †p<0.1.

Table S4. Unweighted Counts of SOR Write-in's by Type

Self-identified Hispanic (2020 Ethnicity or 2030 Race Question)	SOR Write-In	Separate		Combined	
		N	%	N	%
No	Non-Hispanic	50	11.3	62	93.94
No	Blank	2	0.5	1	1.52
Yes	Hispanic	325	73.7	0	0
Yes	Indigenous Latin American	8	1.8	1	1.52
Yes	Non-Hispanic	26	5.9	2	3.03
Yes	Blank	30	6.8	0	0.0

Source: Project on Ethnicity, Race, and Immigration (PERI). All numbers in this table are unweighted and post-processing.

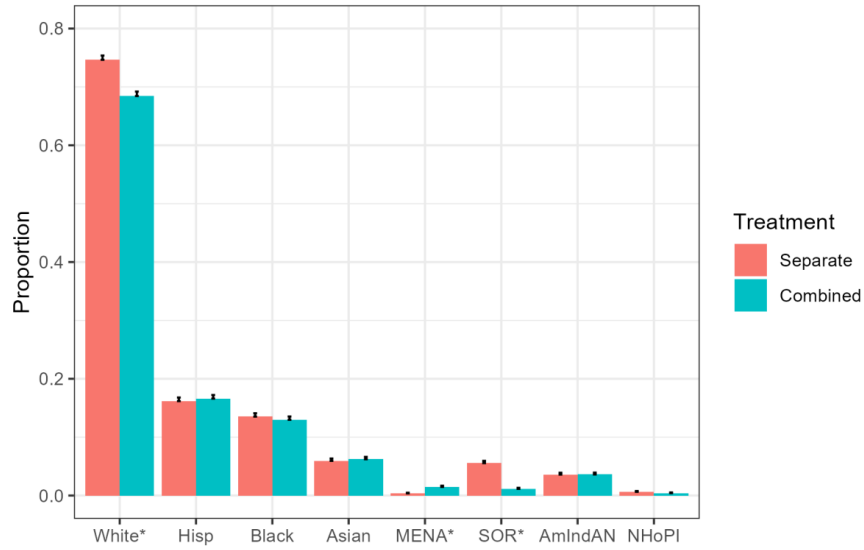
Table S5. Ethnic and racial Identification by experimental condition (unweighted), *percentages*

	2023 PERI (unweighted)	
	Separate Questions (Control)	Combined Question (Treatment)
White	44.0	32.9*
	(0.8)	(0.8)
Hispanic	33.4	33.6
	(0.8)	(0.8)
Black	16.3	15.3
	(0.6)	(0.6)
Asian	29.6	29.0
	(0.8)	(0.7)
MENA	0.3	1.7*
	(0.0)	(0.2)
SOR	12.2	1.8*
	(0.5)	(0.2)
AIAN	5.9	4.4*
	(0.4)	(0.3)
NHoPI	1.0	0.8
	(0.2)	(0.1)
Multiracial	7.2	7.3
	(0.4)	(0.4)

Source: Project on Ethnicity, Race, and Immigration (PERI).

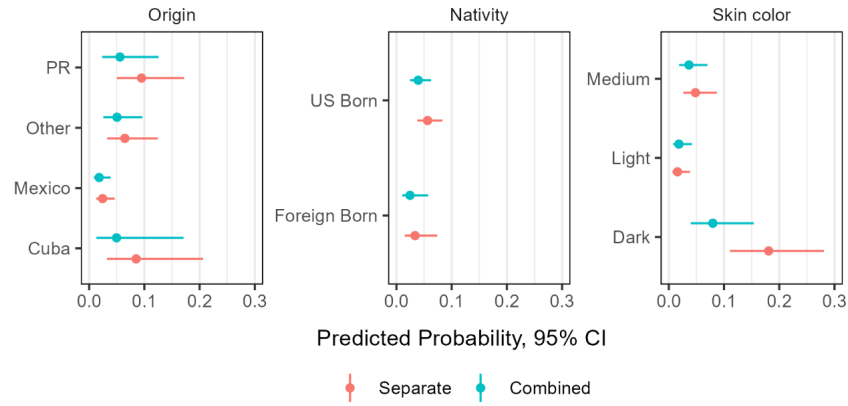
Note: * $p < 0.05$ with T-test between separate questions (control) and combined question (treatment) formats, in PERI, only. Multiracial is defined as individuals selecting more than one box except for Hispanic or MENA. MENA estimates in the Separate Questions are from the write-ins only.

Figure S1. Ethnic and racial Identification by experimental condition



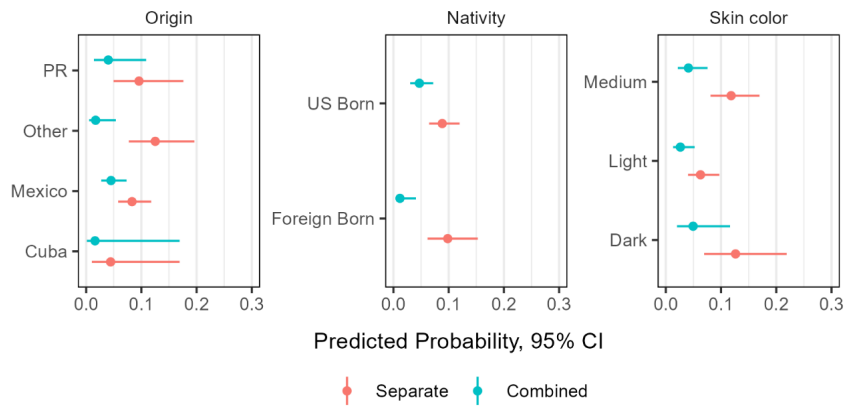
Source: Project on Ethnicity, Race, and Immigration (PERI). Weighted estimates.
 Note: * $p < 0.05$ with T-test between separate questions (control) and combined question (treatment) formats, in PERI, only. MENA estimates in the Separate Questions are from the write-ins only. Weighted estimates.

Figure S2. Comparing Treatment Effects on Identifying as Black by Selected Variables (Self-identified Latino Respondents Only)



Source: Project on Ethnicity, Race, and Immigration (PERI). Predicted margins are based on a logistic regression model from Table S3.

Figure S3. Comparing Treatment Effects on Identifying as American Indian/Alaska Native by Selected Variables (Self-identified Latino Respondents Only)



Source: Project on Ethnicity, Race, and Immigration (PERI). Predicted margins are based on a logistic regression model from Table S3.

Figure S4. Screenshots of the web survey questions. Separate Question Format (*control* condition).**Are you 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, for example, Salvadoran, Dominican, Columbian, Guatemalan, Spaniard, Ecuadorian, etc.*

What is your race?

Select all that apply **AND** enter additional details in the spaces below.

- White - Specify, for example, German, Irish, English, Italian, Lebanese, Egyptian, etc.*

- Black or African Am. - Specify, for example, African American, Jamaican, Haitian, Nigerian, Ethiopian, Somali, etc.*

- American Indian or Alaska Native - Specify name of enrolled or principal tribe(s), for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat, Nome Eskimo Community, etc.*

- Chinese*

- Vietnamese*

- Native Hawaiian*

- Filipino*

- Korean*

- Samoan*

- Asian Indian*

- Japanese*

- Chamorro*

- Other Asian - Specify, for example, Pakistani, Cambodian, Hmong, etc.*

- Other Pacific Islander - Specify, for example, Tongan, Fijian, Marshallese, etc.*

- Some other race - Print race or origin*

Figure S5. Screenshots of the web survey questions. Combined Question Format (*experimental condition*).

What is your race or ethnicity?

Select all that apply **AND** enter additional details in the spaces below. Note, you may report more than one group.

WHITE

German
Italian
Irish
Polish
English
French
Other, for example, Scottish, Norwegian, Dutch, etc.

HISPANIC, LATINO, OR SPANISH

Mexican or Mexican American
Salvadoran
Puerto Rican
Dominican
Cuban
Colombian
Other, for example, Guatemalan, Spaniard, Ecuadorian, etc.

BLACK OR AFRICAN AMERICAN

African American
Nigerian
Jamaican
Ethiopian
Haitian
Somali
Other, for example, Ghanaian, South African, Barbadian, etc.

ASIAN

Chinese
Vietnamese
Filipino
Korean
Asian Indian
Japanese
Other, for example, Pakistani, Cambodian, Hmong, etc.

AMERICAN INDIAN OR ALASKA NATIVE

Enter, for example, Navajo Nation, Blackfeet Tribe, Mayan, Aztec, Native Village of Barrow Inupiat, Tlingit, etc.

MIDDLE EASTERN OR NORTH AFRICAN

- Lebanese
- Syrian
- Iranian
- Moroccan
- Egyptian
- Israeli

Other, for example, Algerian, Iraqi, Kurdish, etc.

NATIVE HAWAIIAN OR OTHER PACIFIC ISLANDER

- Native Hawaiian
- Tongan
- Samoa
- Fijian
- Chamorro
- Marshallese

Other, for example, Palauan, Tahitian, Chuukese, etc.

SOME OTHER RACE, ETHNICITY, OR ORIGIN

Then, once general group or groups were selected, respondents were asked to mark and/or provide the subgroups. In this example, White and Hispanic were marked.

What is your race or ethnicity?

Select all that apply **AND** enter additional details in the spaces below. Note, you may report more than one group.

WHITE

- German
- Italian
- Irish
- Polish
- English
- French
- Enter, for example, Scottish, Norwegian, Dutch, etc.

HISPANIC, LATINO, OR SPANISH

- Mexican or Mexican American
- Salvadoran
- Puerto Rican
- Dominican
- Cuban
- Colombian
- Enter, for example, Guatemalan, Spaniard, Ecuadorian, etc.