



5-24-2023

Questioning Identity: How a Diverse Set of Respondents Answer Standard Questions About Ethnicity and Race

Dana Garbarski
Loyola University Chicago, dgarbarski@luc.edu

Jennifer Dykema
University of Wisconsin-Madison

Cameron P. Jones
University of Wisconsin-Madison

Tiffany S. Neman
University of Wisconsin-Madison

Nora Cate Schaeffer
University of Wisconsin-Madison

Follow this and additional works at: https://ecommons.luc.edu/soc_facpubs



Part of the Faculty Publications and Other Works Commons

Author Manuscript

This is a pre-publication author manuscript of the final, published article.

Recommended Citation

Garbarski, Dana; Dykema, Jennifer; Jones, Cameron P.; Neman, Tiffany S.; Schaeffer, Nora Cate; and Edwards, Dorothy Farrar. Questioning Identity: How a Diverse Set of Respondents Answer Standard Questions About Ethnicity and Race. *Field Methods*, , , 2023. Retrieved from Loyola eCommons, Sociology: Faculty Publications and Other Works, <http://dx.doi.org/10.1177/1525822X231173805>

This Article is brought to you for free and open access by the Faculty Publications and Other Works by Department at Loyola eCommons. It has been accepted for inclusion in Sociology: Faculty Publications and Other Works by an authorized administrator of Loyola eCommons. For more information, please contact ecommons@luc.edu.



This work is licensed under a [Creative Commons Attribution-NonCommercial-No Derivative Works 3.0 License](https://creativecommons.org/licenses/by-nc-nd/3.0/).

Authors

Dana Garbarski, Jennifer Dykema, Cameron P. Jones, Tiffany S. Neman, Nora Cate Schaeffer, and Dorothy Farrar Edwards

Questioning Identity:

How a Diverse Set of Respondents Answer Standard Questions about Ethnicity and Race

*Dana Garbarski

Loyola University Chicago

dgarbarski@luc.edu

Jennifer Dykema

University of Wisconsin-Madison

dykema@ssc.wisc.edu

Cameron P. Jones

University of Wisconsin-Madison

cjones46@wisc.edu

Tiffany S. Neman

University of Wisconsin-Madison

tneman@wisc.edu

Nora Cate Schaeffer

University of Wisconsin-Madison

schaeffe@ssc.wisc.edu

Dorothy Farrar Edwards

University of Wisconsin-Madison

dfedwards@wisc.edu

Abstract

Ethnoracial identity refers to the racial and ethnic categories that people use to classify themselves and others. How ethnoracial identity is measured in surveys has implications for understanding inequalities. Yet how people self-identify may not conform to the categories standardized survey questions use to measure ethnicity and race, leading to potential measurement error.

In interviewer-administered surveys, answers to survey questions are achieved through interviewer-respondent interaction. An analysis of interviewer-respondent interaction can illuminate whether, when, how, and why respondents experience problems with questions. In this study, we examine how indicators of interviewer-respondent interactional problems vary across ethnoracial groups when respondents answer questions about ethnicity and race. Further, we explore how interviewers respond in the presence of these interactional problems. Data are provided by the 2013-2014 Voices Heard Survey, a computer-assisted telephone survey designed to measure perceptions of participating in medical research among an ethnoracially diverse sample of respondents.

Background

Ethnoracial identity in surveys

Ethnoracial identity refers to the racial and ethnic categories that people use to classify themselves and others based on a constellation of factors including ancestry, phenotypic traits, and culture (Omi and Winant 2014; Telles 2014). Ethnoracial classifications are hierarchical in society, such that some groups have higher status than others. Ethnoracial groups and the status accorded to them vary across time and place, supporting the idea that racial and ethnic groupings are socially constructed and defined by the individuals (Schachter et al. 2021) and institutions (Mora 2014) in the society in which they are embedded. Variation in ethnoracial identification can also occur within an individual. For example, 20% of respondents shifted their self-identified ethnoracial identity over a 20-year period (Saperstein and Penner 2012). In the U.S., this within-individual fluidity is more prevalent among individuals who identify as Hispanic/Latina/o/x/e¹ or mixed race (Croll and Gerteis 2019).

Although socially constructed, ethnicity and race are real and objective determinants for well-being and life chances given the status accorded to them. How ethnoracial identity is measured and reported in surveys has implications for understandings of disparities and inequities across a range of outcomes. Yet how people self-identify or are racialized by others may not conform to standardized survey questions used to measure ethnicity and race. For example, the 2020 U.S. Census did not include a Middle Eastern/North African (MENA) category and counted as “white” several groups that are racialized by others as non-White (Maghbouleh 2017; Mathews et al. 2017; Olmstead-Hawala and Nichols 2020; Ortman et al.

¹ From this point forward, we use the term “Latine” to denote respondents who identified as Hispanic, Latino, or Latina in our study. Latine is used increasingly as an alternative form of identification that removes the gender binary in Latino or Latina but aligns more with spoken Spanish than the alternative Latinx, with the latter having limited use among the Latine community.

2017; Schachter et al. 2021; Wang 2021). Studies conducted by the U.S. Census Bureau over the last decade examine the distribution of responses to questions about ethnicity and race using separate ethnicity and race questions (Hispanic/Latino ethnicity as one question, race as another) compared to a single question that combines ethnicity and race (Compton et al. 2013; Mathews et al. 2017; Olmstead-Hawala and Nichols 2020; Ortman et al. 2017; Rios, Romero, and Ramirez 2014). One consistent finding from this research is that Latine respondents are more likely to select “some other race” when separate questions about ethnicity and race are included, increasing the heterogeneity of the “some other race” category. In contrast, using a single ethnoracial question allows Latine respondents to select only that category, and “some other race” becomes the residual category it is intended to be. Indeed, even something as simple as the order of questions impacts response distributions, with non-Latine respondents being less likely to answer a question about Hispanic/Latine ethnicity when the question follows rather than precedes a question about race (Martin 2002). Although reframing race and ethnicity questions was supported by ample research done by the Census Bureau, these recommendations were not adopted in the 2020 U.S. Census, as decided by the White House’s Office of Management and Budget in 2018 (Wang 2018).

Interviewer-administered surveys and interviewer-respondent interaction

Despite the potential cost savings of self-administered modes, interviewer-administered surveys continue to be central to the collection of valid and reliable survey data (Olson et al. 2020; Schaeffer, Dykema, and Maynard 2010). Interviewers are able to facilitate the deliberate selection of respondents, administration of complex instruments, and collection of auxiliary measures (e.g., biomeasures, cognitive tests, linkages between survey data and sensitive external records such as Social Security data) that are increasingly incorporated into study designs.

Telephone interviews are a potentially taxing mode of collecting data because respondents may experience problems with complex questions and require extra help (de Leeuw 2005). The motivation of sample members to participate, to work to provide accurate and honest answers, and to consent to providing sensitive information is critical to data quality, and interviewers play a key role in encouraging respondents to fulfill these criteria.

Survey data are overwhelmingly gathered using standardized interviewing, which aims to limit the effect of interviewers on the collected survey data (Hyman 1975 [1954]; O’Muircheartaigh and Campanelli 1998; Schaeffer 1991; Schaeffer et al. 2010; West and Blom 2017). The rules of standardization most commonly referred to are those offered by Fowler and Mangione (1990): read questions as written; probe inadequate answers non-directively; record answers without discretion; and be interpersonally nonjudgmental regarding the substance of answers. If survey questions are clearly written and fit the target population, standardized interviews should consist of a series of “paradigmatic” question-answer sequences (Schaeffer and Maynard 1996, 2008), in which the interviewer reads the question as scripted and the respondent provides an answer that is codable, that is, one of the response options that the interviewer can code (e.g., “yes” for a yes/no question); optionally, the interviewer may acknowledge the respondent’s answer before moving on to the next question.

However, answers to survey questions are interactional accomplishments, and nonparadigmatic question-answer sequences arise for many reasons. These include respondents’ displays of problems comprehending the meanings of questions and the terms they contain, difficulties respondents encounter mapping responses that summarize their attitudes and experiences onto the response categories provided, and a poor fit between the content of

questions and respondents' knowledge or past experiences (Dykema et al. 1997, 2020; Garbarski, Schaeffer, and Dykema 2011, 2016; Holbrook et al. 2006).

The study of interviewer-respondent interaction documents the structure and content of question-answer sequences, and has indicated several behaviors of respondents and interviewers that are associated with lower data quality, such as respondents qualifying responses or saying “don't know,” or interviewers having to ask follow-up questions (see Schaeffer and Dykema 2011 for a review). A detailed analysis of interviewer-respondent interaction can illuminate how respondents from various ethnoracial groups respond to various types of questions, thus informing best practices for writing survey questions and designing survey forms for respondents from multiple backgrounds (Dykema et al. 2020). Such analysis also highlights techniques to improve interviewer training: using evidence from actual interviewer-respondent question-answer sequences to inform decisions about when and how interviewers should intervene to obtain codable answers (Dykema et al. 2020; Garbarski et al. 2016; Schaeffer et al. 2020).

Current study

Overall, researchers know very little about the underlying response process that produces the survey measurement of ethnicity and race in interviewer-administered surveys. While there is some limited qualitative data from cognitive interviews conducted by the U.S. Census (summarized in Ortman et al. [2017]), it is not detailed enough to illuminate the response process. In contrast, a detailed analysis of interviewer-respondent interaction can highlight whether, when, how, and why respondents experience issues with questions on ethnicity and race.

We examine the following research questions in an ethnographically diverse sample of respondents:

- 1) When respondents answer questions about ethnicity and race, do indicators of interactional problems vary across ethnoracial groups? We address this question using mixed methods: qualitative coding and quantitative enumeration of these differences.
- 2) How do interviewers respond in the presence of these indicators of interactional problems? We examine this question qualitatively, analyzing selected transcripts using conversation analytic and content analysis methods (Dykema et al. 2020; Garbarski et al. 2016).

Methods

Data

Data for this study are from the Voices Heard computer-assisted telephone interview (CATI) survey, which was designed to measure perceptions of barriers and facilitators to participating in medical research studies that collect biomarkers (e.g., saliva and blood) among respondents from different racial and ethnic groups (Black, Latine, American Indian, and white) in Wisconsin. We employed a quota sampling strategy because screening to identify members in non-white groups would have been prohibitively expensive; consequently, we have a non-probability sample for which response rates cannot be calculated. The quota sample consisted primarily of volunteers but also used a targeted list of names provided by a commercial vendor (see Online Supplementary Appendix A). Interviewers conducted 410 usable interviews (in English only) with an average length of 25.21 minutes ($n = 96$ questions) between October 2013 and March 2014. Respondents received a \$20 cash incentive. Interviews were audio recorded, and the current study is restricted to the 375 respondents for which there were usable recordings and the 23 interviewers who interviewed them.

We focus on the features of the interaction that occurred with Questions 89 and 90. Question 89 reads “Are you Hispanic or Latino?” The interviewer is shown the following codable responses on their screen: Yes, No, Don’t Know, and Refused; these are not read to respondents. Question 90 reads “Which one or more of the following would you say is your race: White, Black or African American, American Indian, Alaska Native, Asian, or Native Hawaiian or Other Pacific Islander?” The interviewer is instructed to read all the categories and enter all the categories indicated by the respondent. Importantly, an option for “other (specify)” is listed on the interviewer’s screen, but is not read to respondents.

Measures

Transcription and coding of interviews. Trained transcribers listened to the audio recordings and systematically transcribed each interview. Within a question-answer sequence, interaction was segmented into turns, a unit-of-talk from one actor—the interviewer or respondent—that was not broken up by talk from the other actor. A turn ended when the other actor began talking, either because the original actor’s talk concluded or the current actor interrupted the original actor. In addition to transcribing talk verbatim, transcribers also recorded tokens (e.g., “ah”), coded whether the respondent’s utterance interrupted the interviewer’s initial reading of the question, and coded whether the turn contained overlapping talk, freestanding laughter (laughter that occurs between words), or laugh tokens (units of laughter that occur within words or phrases).

Using Stata’s string coding functions and the electronic transcripts, coders identified strings of text capturing different interactional features, such as codable answers, uncodable answers, requests for clarification or repetition, and conversational elements such as elaborations and mitigators. These string functions allowed us to parse talk into discrete coding units.

Interactional indicators. We examine several behaviors of respondents that index potential problems they experience with the survey questions (see Garbarski et al. [2016] and Schaeffer and Dykema [2011] for a review), which represent separate but related features of the response process and potential breakdowns in cognitive processing:

- 1) Whether the question-answer sequence has more than 3 turns of talk (vs. 3 or fewer).
Question-answer sequences with more than 3 turns indicate an interactional problem requiring multiple turns to resolve.

Whether the respondent's first turn-of-talk contains:

- 2) an immediately codable answer (i.e., the answer repeats or unambiguously paraphrases one of the response categories with no additional talk).
- 3) interruption of the interviewer's initial reading of the question.
- 4) elaborations such as comments about the answer, question, or difficulty of the task.
- 5) requests for clarification or repetition of a word, phrase, or some part of the question or response categories, e.g., "are there other choices for latino because I don't consider myself white."
- 6) "don't know" or equivalent responses.
- 7) mitigating expressions that reduce the exactness, precision, or certainty of another utterance or that itself expresses uncertainty, such as "I guess" or "probably."

And whether the question-answer sequence contains:

- 8) respondent laughter or laugh tokens.
- 9) respondent repaired talk, such as "ver very likely", or respondent-uttered tokens (e.g., "uh" or "um").

Analytic strategy

Each of the interactional indicators are coded as 1 or 0 (i.e., present or absent) within a given question-answer sequence. We examine whether the proportion of question-answer sequences containing an interactional indicator varies across ethnoracial groups using chi-square tests for differences across groups; we also confirm the results are similar using Fisher's exact test given the small number of cases across groups for some of the interactional indicators. We also report post hoc pairwise comparisons across groups.

Results

Question 89: Are you Hispanic or Latino?

Table 1 shows how the distribution of the interactional indicators of data quality varies across ethnoracial groups when answering the question "Are you Hispanic or Latino?" Just over one third of Latine respondents had an interaction involving more than 3 turns-of-talk, while the other three groups (Black, American Indian, and white) had relatively shorter interactions. While less than 80% of Latine respondents answered with an immediately codable answer, close to or over 90% of respondents did so in the other three groups.

Table 1. Distribution of Interactional Indicators by Ethnoracial Group for Question 89 “Are you Hispanic or Latino,” Voices Heard Survey 2013-2014

	Latine	Black	American Indian	White	Group Difference	Pairwise Difference
More than 3 turns (vs. 3 or fewer)	36.5 %	7.6 %	6.5 %	5.3 %	***	abc
Immediately codable answer (vs. any other talk)	78.1 %	88.0 %	95.7 %	92.6 %	***	bc
Interruption of question reading (vs. none)	0.0 %	0.0 %	0.0 %	0.0 %		
Elaborations (vs. none)	5.2 %	8.7 %	0.0 %	4.3 %	*	bdf
Requests for clarification or repetition (vs. none)	4.2 %	0.0 %	1.1 %	2.1 %		a
“Don't know” responses (vs. none)	2.1 %	0.0 %	1.1 %	1.1 %		
Mitigating expressions (vs. none)	2.1 %	0.0 %	0.0 %	0.0 %		
Laughter or laugh tokens (vs. none)	7.3 %	3.3 %	2.2 %	1.1 %		c
Repaired talk (vs. none)	9.4 %	0.0 %	1.1 %	3.2 %	**	ab
Tokens (vs. none)	19.8 %	4.4 %	1.1 %	3.2 %	***	abc
N	96	92	93	94		

Notes. Chi-square tests for differences across group, + $p < .1$, * $p < .05$, ** $p < .01$, *** $p < .001$. Levels of significance are the same when using Fisher’s exact test.

Pairwise comparisons ($p < .05$): a: Latine and Black, b: Latine and American Indian, c: Latine and white, d: Black and American Indian, e: Black and white, f: American Indian and white

There were no occurrences of interruptions during question reading.

The question “Are you Hispanic or Latino” is a list-item question, that is, one that lists a set of objects linked with the words “or” or “and.” Although question writers treat this question as a “yes/no” question (given the listed response options), list-item questions have multiple correct grammatical answers: for this question, responses of “Hispanic” or “Latino” would be synonymous with a “yes” response and would be acceptable answers in terms of grammar and conversational practice.

When we examine the transcripts, we see that the Black, American Indian, and white respondents largely treat this question as a “yes/no” question, answering with “no.” However, the Latine respondents are more likely to engage with the “either/or” suggestion in this question; that is, they do not provide an immediately codable answer, which leads to follow-up by the interviewer and thus a question-answer sequence with more than 3 turns-of-talk. For example, in

Excerpt 1, after the reading of the question, the Latine respondent repeats one of the categories listed in the question: “Hispanic.” The interviewer follows up with a directive probe “so you would say yes,” turning what the respondent reports into a codable answer.

Excerpt 1
I: are you hispanic or latino?
R: hispanic
I: so you would say yes is that correct?
R: oh yes sorry {Laughter}
I: that's ok

In Excerpt 2, the respondent requests clarification, and the interviewer follows up with a “means to you” statement before defining the parameters of the question as “either of those.”

Excerpt 2
I: are you hispanic or latino?
R: uh I think both are the same aren't they?
I: I'm sorry what was that?
R: are they both the same?
I: are they the same?
R: yes
I: um you know I'm not sure but um it's the question is whatever it means to you if you consider yourself either of those
R: um hispanic
I: ok

Not answering with a codable “yes” or “no” leads to follow-up from the interviewer, extending the length of and variability within the interview, with implications for measurement error if variation in interviewer behavior influences the responses obtained from the respondent (Olson et al. 2020). Interviewer follow-up also has implications for rapport: the respondent in Excerpt 1 has provided a grammatical answer to the survey question, yet the interviewer’s follow up indicates a correction that is both in line with the rules of standardization but may conflict with conversational practices (Garbarski et al. 2016). Thus, the contextualization provided by the transcripts helps to illustrate why and how this question would benefit from revision.

We also find some evidence of group differences in elaborations, with Black respondents having higher levels than other groups and American Indian respondents having lower levels than other groups. Excerpts 3 and 4 show examples of elaborations: each respondent gives a codable answer “no” and then elaborates on their answer with “African American.” The excerpts show how interviewers might vary in their approach to dealing with these elaborations. The interviewer in Excerpt 3 acknowledges the elaboration with the neutral “thank you,” while the interviewer in Excerpt 4 provides a more expansive account for why they are about to read the next question that was anticipated by the respondent’s elaboration.

Excerpt 3

I: are you hispanic or latino?
R: no um african american
I: uh thank you

Excerpt 4

I: are you hispanic or latina?
R: no african american
I: alright this next question's going to ask about that again I do have to read the whole thing so bear with me but I will keep your answer in mind

Overall, when the respondent’s answer anticipates the answer to a subsequent question, there are several ways in which interviewers can and do respond that vary in their responsiveness and adherence to the rules of standardization (Garbarski et al. 2016)—increasing variability and potential measurement error. In addition, the order of the questions—Hispanic/Latino identity, then racial identity—stems from research that shows that the Hispanic/Latino question is more likely to be skipped when it follows rather than precedes the question about race (Martin 2002). However, the issues shown in Excerpts 3 and 4 demonstrate the current order of the questions may produce other interactional problems.

Relatively few or no occurrences of the following indicators occurred with Question 89, so we do not examine them in more detail: interruption of question reading, requests for clarification or repetition, “don’t know,” mitigating talk, or laughter or laugh tokens (Table 1). Finally, Latine respondents are also more likely to produce repaired talk or disfluency tokens (Table 1; Excerpt 2 is an example), potentially indicating a greater level of cognitive processing required for them to understand the intent of the question and provide a response compared to respondents of other groups.

Question 90: Which one or more of the following would you say is your race?

Table 2 shows how the distributions of the interactional indicators vary across ethnoracial groups when answering the question “Which one or more of the following would you say is your race?” Over half of Latine respondents have question-answer sequences with more than 3 turns, while this only occurs for 20 to 30% of respondents in other groups. Just over 40% of Latine respondents answer with an immediately codable answer, while over 80% of respondents in the other groups do so.

Table 2. Distribution of Interactional Indicators by Ethnoracial Group for Question 90 “Which one or more of the following would you say is your race? ,” Voices Heard Survey 2013-2014

	Latine	Black	American Indian	White	Group Difference	Pairwise Difference
More than 3 turns (vs. 3 or fewer)	51.0 %	28.3 %	26.9 %	19.2 %	***	abc
Immediately codable answer (vs. any other talk)	44.8 %	87.0 %	78.5 %	88.3 %	***	abc
Interruption of question reading (vs. none)	7.3 %	26.1 %	20.4 %	16.0 %	**	ab
Elaborations (vs. none)	7.3 %	1.1 %	3.2 %	6.4 %		a
Requests for clarification or repetition (vs. none)	16.7 %	3.3 %	2.2 %	1.1 %	***	abc
“Don't know” responses (vs. none)	12.5 %	2.2 %	0.0 %	0.0 %	***	abc
Mitigating expressions (vs. none)	8.3 %	1.1 %	1.1 %	3.2 %	*	ab
Laughter or laugh tokens (vs. none)	14.6 %	6.5 %	1.1 %	5.3 %	**	bc
Repaired talk (vs. none)	20.8 %	15.2 %	10.8 %	5.3 %	*	ce
Tokens (vs. none)	51.0 %	9.8 %	14.0 %	11.7 %	***	abc
N	96	92	93	94		

Notes. Chi-square tests for differences across group, +p<.1, *p<.05, **p<.01, ***p<.001. Levels of significance are the same when using Fisher’s exact test.

Pairwise comparisons (p<.05): a: Latine and Black, b: Latine and American Indian, c: Latine and white, d: Black and American Indian, e: Black and white, f: American Indian and white

We again examine the transcripts to contextualize the discrepancy between the Latine respondents and the other groups with respect to these first two indicators. Rather than immediately codable answers to the survey question, we observe several uncodable answers in the first turns of Latine respondents. In Excerpts 5 and 6, the respondent initially responds with “no” or “none.” In Excerpt 5, the interviewer offers “other” as a response (which is listed on their screen but not read to the respondent), and then asks the respondent to specify. In Excerpt 6, the interviewer offers “other” as a directive probe, and then asks the respondent what they would say, which is potentially a bit clearer than the “specify” in Excerpt 5.

Excerpt 5

- I: which one or more of the following would you say is your race white black or african american american indian alaska native asian or native hawaiian or other pacific islander?
R: uh no
I: or you could say other
R: I'm other
I: ok and then could you specify for me please?
R: I'm mexican

Excerpt 6

- I: which one or more of the following would you say is your race white black or african american american indian alaska native asian or native hawaiian or other pacific islander?
R: none of those
I: and you would say other is that correct?
R: other correct
I: and what would you say for other?
R: uh mexican
I: you said mexican is that correct?
R: yes

Similarly, Excerpts 7 and 8 illustrate elaborations that offer candidate answers that are beyond the scope of response options as they have been listed: “no it’s Spanish.” In Excerpt 7, the interviewer first confirms “none of those” and then “Spanish”; in Excerpt 8, the interviewer confirms “Spanish.”

Excerpt 7

- I: which one or more of the following would you say is your race I know we just kind of asked but um white black or african american american indian alaska native or excuse me or native hawaiian or other pacific islander?
R: no it's spanish
I: so none of those is that correct?
R: yeah
I: and did you say spanish?
R: yeah
I: thank you ma'am
R: mhmm

Excerpt 8

I: which one or more of the following would you say is your race I know we just kind of asked but um white black or african american american indian alaska native or excuse me or native hawaiian or other pacific islander?

R: no it's spanish

I: so none of those is that correct?

R: yeah

I: and did you say spanish?

R: yeah

I: thank you ma'am

R: mhmm

Overall, Excerpts 5 through 8 illustrate that when respondents have issues answering questions in a way that invites interviewer follow-up, the follow-up behavior may vary, with implications for measurement error.

All the remaining interactional indicators vary across ethnoracial groups for Question 90 (Table 2). Interruptions are least likely to occur for the Latine respondents; based on the transcripts reviewed above, Latine respondents are likely waiting for the interviewer to present a category that aligns with their self-identity. The remaining indicators are more likely to occur for Latine respondents, with significant differences across groups for all indicators other than elaborations.

Discussion

This study examines how a diverse set of respondents answer questions about ethnoracial identity in an interviewer-administered survey. First, we find varying types of answers for question 89 (“Are you Hispanic or Latino”) across respondents: Latine respondents are more likely to interpret the response categories to be “Hispanic” or “Latino,” while other groups interpret the question as “yes/no,” that is, as “Hispanic or Latino” or “Not Hispanic or Latino,” with the latter being how the researchers intend for the question to be answered. When respondents provide answers that are not immediately codable, such as how the Latine respondents answer question 89, or offer information in addition to a codable answer, such as an anticipatory answer, interviewers are more likely to follow up. As shown by the excerpts, the ways in which interviewers follow up vary, systematically increasing interviewer variability and measurement error for those ethnoracial groups.

Second, for the race question (question 90), we observe both increased use of the hidden “other” category and more interactional problems for Latine respondents. While question order communicates meaning, such that answering a question about race after Hispanic/Latino origin theoretically communicates a contrast effect, in practice it often does not—Latine respondents often want to indicate that component of their identity when answering about race. Indeed, given the racialization of Latine groups in the U.S., identifying as white may not map onto their lived experiences with discrimination (Roth 2012). The race question is also cumbersome to administer in interviewer-administered modes due in part to the number of response categories: we observe interruptions when respondents hear the category that applies to them and requests for repetition or clarification when they do not. Overall, these results support the idea that ethnicity and race should be ascertained using a single survey question on ethnoracial identity

(Compton et al. 2013; Mathews et al. 2017; Olmstead-Hawala and Nichols 2020; Ortman et al. 2017; Rios, Romero, and Ramirez 2014; Wang 2018).

This study has limitations. First, the survey took place over the phone; the interactional behaviors we observe may differ in face-to-face modes of data collection. The interviews were conducted in English, and although the respondents in this study agreed to participate in a study done in English, it is plausible that the questions posed difficulty for respondents for whom English is not their first language, an issue that cannot be explored with these data but should be considered in future research. The interviewer-respondent interaction illustrates some of the issues respondents may have with these questions regardless of mode; however, self-administered modes may lead to other issues with responding not captured here. The study results derive from one set of 23 interviewers at one research organization; more studies or studies with many interviewers and multiple organizations are needed to begin to trace the boundaries and contours of interviewers' impact on both self and reflected appraisals of respondents' ethnoracial identity. Finally, the sample is a non-probability sample, although the benefit of this quota sample is that it intentionally recruited across ethnoracial groups such that differences across groups are made apparent here in ways that were not feasible in prior studies.

In conclusion, this study illuminates the interactional accomplishment of producing answers to survey questions on ethnoracial identity. Our examination of transcripts shows variation in how interviewers respond to interactional troubles. The lack of standardization for interviewers' follow-up behaviors has implications for data quality, in that it increases the effect of interviewers on the quality of the data obtained. The issues of data quality also indirectly influence costs—all of that interaction is expensive in terms of time spent on the phone. Future research should continue to use transcripts of and real experiences in interviews to develop

trainings for interviewers to respond to interactional troubles deriving from ubiquitous yet complicated questions on ethnicity and race.

References

- Compton, Elizabeth, Michael Bentley, Sharon Ennis and Sonya Rastogi. 2010. "Census Race and Hispanic Origin Alternative Questionnaire Experiment." *US Census Bureau*.
- Croll, Paul R and Joseph Gerteis. 2019. "Race as an Open Field: Exploring Identity Beyond Fixed Choices." *Sociology of Race and Ethnicity* 5(1):55-69.
- de Leeuw, Edith D. 2005. "To Mix or Not to Mix Data Collection Modes in Surveys." *Journal of Official Statistics* 21(2):233-55.
- Dykema, Jennifer, James M. Lepkowski and Steven Blixt. 1997. "The Effect of Interviewer and Respondent Behavior on Data Quality: Analysis of Interaction Coding in a Validation Study." Pp. 287-310 in *Survey Measurement and Process Quality*, edited by L. Lyberg, P. Biemer, M. Collins, E. de Leeuw, C. Dippo, N. Schwarz and D. Trewin. New York, NY: Wiley.
- Dykema, Jennifer, Dana Garbarski, Nora Cate Schaeffer, Tiffany Neman and Dorothy Farrar Edwards. 2020a. "Parenthetical Phrases and Respondent Motivation, Interview Efficiency, and Interviewer Variation." Paper presented at the annual meeting virtual conference of the American Association for Public Opinion Research.
- Dykema, Jennifer, Nora Cate Schaeffer, Dana Garbarski and Michael Hout. 2020b. "The Role of Question Characteristics in Designing and Evaluating Survey Questions." Pp. 449-70 in *Advances in Questionnaire Design, Development, Evaluation, and Testing*, edited by P. Beatty, D. Collins, L. Kaye, J. Padilla, G. Willis and A. Wilmot. Hoboken, NJ: Wiley.
- Fowler, Jr., Floyd J. and Thomas W. Mangione. 1990. *Standardized Survey Interviewing: Minimizing Interviewer-Related Error*. Newbury Park: Sage.
- Garbarski, Dana, Nora Cate Schaeffer and Jennifer Dykema. 2011. "Are Interactional Behaviors Exhibited When the Self-Reported Health Question Is Asked Associated with Health Status?". *Social Science Research* 40(4):1025-36. doi: 10.1016/j.ssresearch.2011.04.002
- Garbarski, Dana, Nora Cate Schaeffer and Jennifer Dykema. 2016. "Interviewing Practices, Conversational Practices, and Rapport: Responsiveness and Engagement in the Standardized Survey Interview." *Sociological Methodology* 46(1):1-38. doi: 10.1177/0081175016637890.
- Holbrook, Allyson, Young Ik Cho and Timothy Johnson. 2006. "The Impact of Question and Respondent Characteristics on Comprehension and Mapping Difficulties." *Public Opinion Quarterly* 70(4):565-95. doi: 10.1093/poq/nfl027.
- Hyman, Herbert H. 1975 [1954]. *Interviewing in Social Research*. Chicago, IL: The University of Chicago.
- Maghbouleh, Neda. 2017. *The Limits of Whiteness*. Stanford, CA: Stanford University Press.
- Martin, Elizabeth. 2002. "The Effects of Questionnaire Design on Reporting of Detailed Hispanic Origin in Census 2000 Mail Questionnaires." *Public Opinion Quarterly* 66(4):582-93.
- Mathews, Mary Beth Swetnam. 2017. *Doctrine and Race: African American Evangelicals and Fundamentalism between the Wars*: University of Alabama Press.
- Miller, Kristen, Rory Fitzgerald, José-Luis Padilla, Stephanie Willson, Sally Widdop, Rachel Caspar, Martin Dimov, Michelle Gray, Cátia Nunes, Peter Prüfer, Nicole Schöbi and Alisú Schoua-Glusberg. 2011. "Design and Analysis of Cognitive Interviews for Comparative Multinational Testing." *Field Methods* 23(4):379-96. doi: 10.1177/1525822x11414802.

- Mora, Marie T and Alberto Dávila. 2014. "Gender and Business Outcomes of Black and Hispanic New Entrepreneurs in the United States." *American Economic Review* 104(5):245-49.
- O'Muircheartaigh, Colm and Pamela Campanelli. 1998. "The Relative Impact of Interviewer Effects and Sample Design Effects on Survey Precision." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 161(1):63-77.
- Olmsted-Hawala, EL and EM Nichols. 2020. "Usability Testing Results Evaluating the Decennial Census Race and Hispanic Origin Questions Throughout the Decade: 2012-2020 (Research and Methodology Directorate, Center for Behavioral Science Methods Research Report Series [Survey Methodology# 2020-02])."
- Olson, Kristen, Jolene D. Smyth, Jennifer Dykema, Allyson Holbrook, Frauke Kreuter and Brady T. West, eds. 2020. *Interviewer Effects from a Total Survey Error Perspective*. Boca Raton, FL: CRC Press, LLC.
(<https://www.taylorfrancis.com/books/9781003020219>).
- Omi, Michael, and Howard Winant. 2014. *Racial Formation in the United States*, 3d ed. New York: Routledge.
- Ortman, Jennifer, Megan Rabe, Jennifer Reichert, Michael Bentley, Patrick Cantwell, Asaph Young Chun, Sarah Heimerl et al. 2017. "MEMORANDUM FOR Victoria Velkoff Chief, American Community Survey Office From: Karen Battle Chief, Population Division (POP) Prepared by: Jacquelyn Harth."
https://www.census.gov/content/dam/Census/library/working-papers/2017/acs/2017_Harth_01.pdf
- Ríos, Merarys, Fabián Romero and Roberto Ramírez. 2014. "Race Reporting among Hispanics: 2010." *Population Division Working Paper* 102.
- Roth, Wendy. 2012. *Race migrations: Latinos and the cultural transformation of race*. Stanford University Press.
- Saperstein, Aliya and Andrew M Penner. 2012. "Racial Fluidity and Inequality in the United States." *American Journal of Sociology* 118(3):676-727.
- Schachter, Ariela, René D. Flores, and Neda Maghbouleh. 2021. "Ancestry, color, or culture? How whites racially classify others in the US." *American Journal of Sociology* 126(5): 1220-1263.
- Schaeffer, Nora Cate. 1991. "Conversation with a Purpose--or Conversation? Interaction in the Standardized Interview." Pp. 367-92 in *Measurement Errors in Surveys*, edited by P. P. Biemer, R. M. Groves, L. E. Lyberg, N. A. Mathiowetz and S. Sudman. New York: John Wiley & Sons.
- Schaeffer, Nora Cate and Douglas W. Maynard. 2008. "The Contemporary Standardized Survey Interview for Social Research." Pp. 31-57 in *Envisioning the Survey Interview of the Future*, edited by F. G. Conrad and M. F. Schober. Hoboken, NJ: Wiley.
- Schaeffer, Nora Cate, Jennifer Dykema and Douglas W. Maynard. 2010. "Interviewers and Interviewing." Pp. 437-70 in *Handbook of Survey Research, Second Edition*, edited by P. V. Marsden and J. D. Wright. Bingley, UK: Emerald Group Publishing Limited.
- Schaeffer, Nora Cate and Jennifer Dykema. 2011. "Response 1 to Fowler's Chapter: Coding the Behavior of Interviewers and Respondents to Evaluate Survey Questions." Pp. 23-39 in *Question Evaluation Methods: Contributing to the Science of Data Quality*, edited by J. Madans, K. Miller, A. Maitland and G. Willis. Hoboken, NJ: John Wiley & Sons, Inc.

- Schaeffer, Nora Cate, Jennifer Dykema, Steve M. Coombs, Rob K. Schultz, Lisa Holland, and Margaret Hudson. 2020. "General Interviewer Training: Developing Evidence-Based Practices for Standardized Interviewing." Pp. 33-46 in *Interviewer Effects from a Total Survey Error Perspective*, edited by K. Olson, J. D. Smyth, J. Dykema, A. L. Holbrook, F. Kreuter, and B. T. West. Boca Raton, FL: CRC Press.
<https://www.taylorfrancis.com/books/9781003020219/chapters/10.1201/9781003020219-5>
- Telles, Edward. 2014. *Pigmentocracies: Ethnicity, Race, and Color in Latin America*: UNC Press Books.
- Wang, Hansi Lo. 2018. 2020 Census To Keep Racial, Ethnic Categories Used In 2010. NPR. <https://www.npr.org/2018/01/26/580865378/census-request-suggests-no-race-ethnicity-data-changes-in-2020-experts-say>
- Wang, Hansi Lo. 2021. What The New Census Data Can — And Can't — Tell Us About People Living In The U.S. NPR. <https://www.npr.org/2021/08/12/1010222899/2020-census-race-ethnicity-data-categories-hispanic>
- West, Brady T. and Annelies G. Blom. 2017. "Explaining Interviewer Effects: A Research Synthesis." *Journal of Survey Statistics and Methodology* 5(2):175-211. doi: 10.1093/jssam/smw024.

Acknowledgments

Support for this study was provided by the: National Science Foundation (SES-1853094 to J. Dykema and D. Garbarski; “Effects of Interviewers, Respondents, and Questions on Survey Measurement”); University of Wisconsin-Madison Office of the Vice Chancellor for Research and Graduate Education with funding from the Wisconsin Alumni Research Foundation; University of Wisconsin Survey Center (UWSC); facilities of the Social Science Computing Cooperative and the Center for Demography and Ecology (NICHD core grant P2C HD047873). Voices Heard data collection was funded by NIMHD grant P60MD003428 (PD: A. Adams; PI: D. Farrar Edwards). Opinions expressed here are those of the authors and do not necessarily reflect sponsors or related organizations.

Declaration of Conflicting Interest

The Authors declare(s) that there is no conflict of interest.