CAMBRIDGE
UNIVERSITY PRESS

RESEARCH ARTICLE

Biased Interviewer Assessments of Respondent Knowledge Based on Perceptions of Skin Tone

Adam M. Enders[1] and Judd R. Thornton[2]* ●

[1]Department of Political Science, University of Louisville, Louisville, KY, USA and [2]Department of Political Science, Georgia State University, Atlanta, GA, USA
*Corresponding author. E-mail: jrthornton@gsu.edu

Abstract

A rich literature documents the effects of survey interviewer race on respondents' answers to questions about political issues and factual knowledge. In this paper, we advance the study of interviewer effects in two ways. First, we examine the impact of race on interviewers' subjective evaluations of respondents' political knowledge. Second, we substitute measures of respondent/interviewer racial self-identification with interviewer perceptions of respondent skin tone. We find that white interviewers subjectively rate black respondents' knowledge lower than do black interviewers, even controlling for objective knowledge measures. Moreover, we identify a negative relationship between relative skin tone and interviewer's assessment of knowledge. Subsequent analyses show a linear relationship between subjective knowledge assessments and the difference between respondent and interviewer skin tone. We conclude with a discussion of the impact of colorism on survey administration and the measurement of political attitudes and democratic capabilities.

Introduction

A robust literature documents the impact of the race of interviewers on respondents' subjective attitudes about various objects, especially racial issues (e.g., Davis 1997b, Schuman and Converse 1971, Williams Jr. 1964). Simply put, respondents register different attitudes depending on the race of the interviewer. This bias is especially pronounced among black respondents, with attitudes varying considerably by interviewer race. Moreover, this interviewer effect even extends to factual knowledge questions—black respondents answer fewer questions correctly when being interviewed by a white interviewer compared to a black interviewer (Davis and Silver 2003).

In these classic studies of race-of-interviewer effects, biased estimates and underperformance are generally attributed to the climate of the survey, the sensitivity of questions

being asked about, respondent reactions to interviewers of a different race, or a combination thereof. In this paper, we extend this literature by examining interviewers' subjective assessment of respondent knowledge, an item that is frequently used in order to circumvent the aforementioned trappings of "objective" knowledge questions—are black respondents perceived by white interviewers to possess relatively low levels of political knowledge even when they perform well on objective knowledge questions?

To answer this question, we first examine differences in interviewers' subjective assessments of respondent knowledge by self-identified race of interviewers. Next, we extend this analysis by substituting self-identified interviewer race for a social distance measure based on the difference in skin tone between the interviewer and the respondent. A wealth of research shows that relatively dark-skinned blacks experience lower wages (Goldsmith, Hamilton, and Darity Jr. 2006), harsher criminal sentencing (Blair, Judd, and Chapleau 2004, Eberhardt, Davies, Purdie-Vaughns, and Johnson 2006), worse health outcomes (Klonoff and Landrine 2000), and an increased likelihood of school suspension compared to lighter-skinned blacks (Hannon, Defina, and Bruch 2013); likewise, dark-skinned immigrants have less upward mobility than light-skinned immigrants (Han 2020). Does the racial bias—colorism—that underlies these examples of discrimination translate to trained interviewers' evaluations of respondent knowledge?

We find evidence that subjective assessments of black respondents' knowledge are related to both interviewer race and the relative skin tone of interviewers compared to respondents. In particular, white interviewers systematically rate black respondents' knowledge lower than do black interviewers, even controlling for objective measures of political knowledge. Moreover, as the skin tone of the respondent becomes darker compared to that of the interviewer, interviewers' subjective rating of respondent knowledge becomes poorer. Interestingly, an auxiliary analysis of the functional form of this relationship reveals that it is squarely linear—there is a one-to-one relationship between differences in skin tone and subjective knowledge assessment.

Our findings have a number of implications, both substantive and methodological. First, the patterns we observe suggest that scholars should be circumspect in making direct comparisons of perceived political knowledge across racial groups—these can be misleading and should not, without additional analysis, be used to judge the relative civic characteristics of different racial groups. Second, because even "objective" measures of political knowledge show similar biases (e.g., Davis and Silver 2003) and are hardly a panacea for aiding in the comparison of groups more generally (e.g., Perez 2015), political sophistication should be measured in a broader way that encompasses other dimensions of engagement, such as political interest and participation in political activities. Third, that colorism pervades the judgments of even trained interviewers showcases the reach of the phenomenon and suggests that alternative strategies for garnering unbiased estimates of political knowledge—whether in the form of more or different training of interviewers or a strategic assignment of interviewers based on skin tone, for example—may be necessary. We expand upon these ideas in the conclusion.

Background and expectations

Foundational work on the impact of interviewer race on respondent attitudes focused on differences in responses provided by black respondents depending on the race of the interviewer, especially when it came to attitudes about racial issues.

For example, Williams (1964) finds sharp differences in black respondents' stated attitudes about sit-ins, segregated schooling, and even less explicitly racial sentiments regarding "making changes in the way our country is run," depending on interviewer race. Schuman and Converse (1971) extended this work, finding considerable variance in the level of bias—observed differences between attitudes elicited by white and black interviewers—across various racial issue attitudes.

Fundamental to the majority of work in this area is an assumption about the mechanism by which observed biases are created. Minimally, the literature holds that bias is a product of social tension between racial groups when it comes to salient racial policies (Davis 1997b, Schuman and Converse 1971). Many others have also made the case that some version of "threat" drives observed biases (e.g., Williams Jr. 1964). Davis and Silver (2003), like Steele and Aronson (1995) before them, hypothesize that stereotype threat—"the pressure to disconfirm and to avoid being judged by negative and potentially degrading stereotypes" (pg. 33)—is behind biases regarding even questions of fact about politics or educational topics. To evade this threat, black respondents "don the black mask" as Davis (1997a) puts it, thereby strategically altering responses in order to conform with perceived expectations or to reduce potential tension, more generally.

The first way we seek to extend this literature on race-of-interviewer effects is to move beyond consideration of how respondents alter their behavior in response to interviewers—whether that be in the form of stereotype threat, social desirability bias, discomfort of the social climate or question topic, or something else—toward a consideration of how interviewers react to respondents. In particular, while past work has focused on biases in the assessment of knowledge and subjective attitudes, we investigate how interviewers subjectively rate the knowledge of subjects of color, including how well those subjective evaluations comport with objective measures of respondent knowledge.

We also extend previous work by examining the effect of both discrete racial self-identification of interviewers (e.g., black, white) and differences between interviewer and respondent skin tone. As noted earlier, a substantial body of work demonstrates that relatively dark-skinned blacks experience worse outcomes across a variety of domains. Simply put, skin tone is correlated with a wide variety of social and economic outcomes—perhaps it impacts interviewers' evaluations of survey respondents as well.

We have good reason to expect as much. Hagiwara, Kashy, and Cesario (2012) find, using classical implicit bias tests, that whites feel more negatively toward darker-skinned blacks than lighter-skinned blacks. Keith et al. (2017) show that darker-skinned blacks were more likely to experience discrimination in the form of disrespect/condescension and high-level microaggressions. Moreover, Foy and Ray (2019) find that sports announcers are more likely to evaluate lighter-skinned basketball players in terms of performance and mental ability, while darker-skinned players are more likely to be discussed in terms of physical characteristics even controlling for objective performance. The findings of this study are echoed very closely in our study. To complete the theoretical puzzle, Hannon and Defina (2014) show that light-skinned interviewers are apt to perceive blacks as being darker-skinned than they really are; indeed, skin tone is simply the primary way that people categorize others (Campbell et al. 2020, Feliciano 2016). Taken together,

previous work shows that (1) interviewers, like anybody, categorize others into racial groups by perceptions of skin tone and (2) lighter-skinned people tend to discriminate against those with darker skin, on average. These patterns undergird our primary expectations:

H₁: White interviewers will subjectively rate the knowledge of black respondents lower than will black interviewers, even controlling for performance on objective knowledge measures.

H₂: Interviewers with light skin will subjectively rate the knowledge of respondents with comparatively dark skin lower than will darker-skinned interviewers, even controlling for performance on objective knowledge measures.

If discrepancies in respondent knowledge persist when interviewer-driven, subjective assessments of respondent knowledge and skin tone are substituted for objective assessments of knowledge and discrete racial categories, we will possess evidence that, minimally, interviewers hold systematic racial biases—albeit, perhaps, implicit ones—against out-group respondents. Such a finding would traverse past work showing only a tendency for people to rate the skin tone of racial out-groups much lighter or darker than their own (e.g., Hannon and Defina 2014), whereas a tendency to misperceive could potentially prove innocuous—the product of standard psychological mechanisms, such as out-group homogeneity bias (Hill 2002)—we expect to observe more than mere difference. Instead, systematically *poorer* performance evaluations (rather than merely "different" ones) would be more accurately interpreted as the product of a racial bias of some sort, whether it be overt prejudice or implicit bias.

   If our expectations bear out, they will also provide some additional clarification of the mechanisms at play in past work on race-of-interviewer effects that have been attributed to the likes of stereotype threat and respondent discomfort. Indeed, even implicit biases can frequently still be perceived by their targets. For example, implicit biases may behaviorally manifest as microaggressions (e.g., Sue 2010) or other subtle verbal and behavioral cues that respondents might perceive as hostile, dismissive, or critical.[1] In other words, just because the biases held by one individual may not be felt consciously does not mean those biases are not perceptible to their targets. As Weaver (2012) notes, there is an "absence of a strong societal norm to avoid" skin tone bias. Likewise, while people are capable of actively suppressing racial bias, suppression of skin tone bias is more difficult (Blair, Judd, and Chapleau 2004).[2] Thus, observing that interviewers systematically evaluate particular respondents more poorly than others provides further supporting evidence that there is, indeed, some threat or social discomfort to be avoided when black respondents are interviewed by white interviewers.

---

[1]Perceptions of what others think of them are oftentimes referred to as "metaperceptions," which is a burgeoning avenue for research on intergroup interactions and relationships, especially when it comes to race (e.g., Frey and Tropp 2006).

[2]Interviewer assessments of respondent knowledge also appear to be impacted by respondent appearance (e.g., attractiveness), more generally (e.g., Palmer and Peterson 2015).

Data

We utilize the 2012 American National Election Study (ANES). We begin by noting that in presidential years, the ANES includes both a pre-election (conducted after the nominating conventions through the election) and a post-election wave (conducted in the weeks following the election). The 2012 ANES has two desirable features, to test our hypotheses. First, the 2012 ANES oversampled black respondents ($n = 554$ for the face-to-face sample; 205 of whom were interviewed by a white interviewer and 255 by a black interviewer), allowing a well-powered examination of effects by race. Second, it includes interviewers' assessments of their own skin tone, as well as that of the respondent, in the pre-election portion of the survey. This allows for a deeper examination of the mechanism by which racial differences in subjective assessments of respondent knowledge might come about.

Interviewers provide subjective assessments of respondents' level of "political information" after the interview has concluded. This is our dependent variable. These assessments are made on a five-point scale ranging from "very low" (1) to "very high" (5). Interviewers are not provided any specific instructions on how this assessment is to be made. Still, this subjective assessment is correlated with an index of responses to political knowledge questions, as we would expect.[3] There is, however, a slight discrepancy in the correlations among white ($r = 0.540$, $p < 0.001$) and black respondents ($r = 0.407$, $p < 0.001$),[4] which is suggestive of the patterns we anticipate—if objective and subjective knowledge are less aligned for black respondents than white ones, some other factor may be at play.

Interviewers completed the assessment of respondents' skin tone directly after the face-to-face interview was completed (though it is unclear when they completed the assessment of their own skin tone) using a ten-point graphical scale ranging from the whitest possible skin tones (1) to the darkest (10) developed by Massey and Martin (2003).[5] The graphic interviewers used to make these assessments, as well as the relevant excerpt from their training manual, appear in the Supplemental Appendix. In the analyses below, we employ the *difference* in perceived skin tone measures as a measure of relative racial (dis)similarity. We do this because skin tone should primarily motivate racial categorization and cue (implicit) racial biases if there are differences between the respondent and interviewer. This relative measure accounts for both the fact that (1) most bias comes in the form of differences in perceptions of racial categorization and (2) that biases can be observed even within discrete racial groupings (e.g., whites' perceptions of the skin tone of other whites, blacks' perceptions of other blacks—see recent work by Adams et al. 2016, Feliciano 2016, Yadon and Ostfeld 2020, for example).

None of the interviewers used the extremes of the scale (1, 10) to describe themselves. As such, subtracting the interviewer-assessed respondent skin tone from

---

[3]Interviewers are not told the correct answers to the political knowledge questions in their training—some may know the answers, others may not. To account for this, we control for objective knowledge and cluster standard errors by interviewer, as we describe below.

[4]We used the average subjective and objective knowledge assessments across pre- and post-election waves, though the pattern is identical for each wave separately.

[5]While previous work finds low inter-coder reliability of the Massey-Martin measure (Hannon and DeFina 2016), we are not interested in evaluations of respondents' skin tone alone, but rather *interviewer-perceived differences* between the interviewer's and respondent's skin tone. Whether these perceptions are an accurate reflection of reality is of less importance if interviewers are nevertheless acting on them.
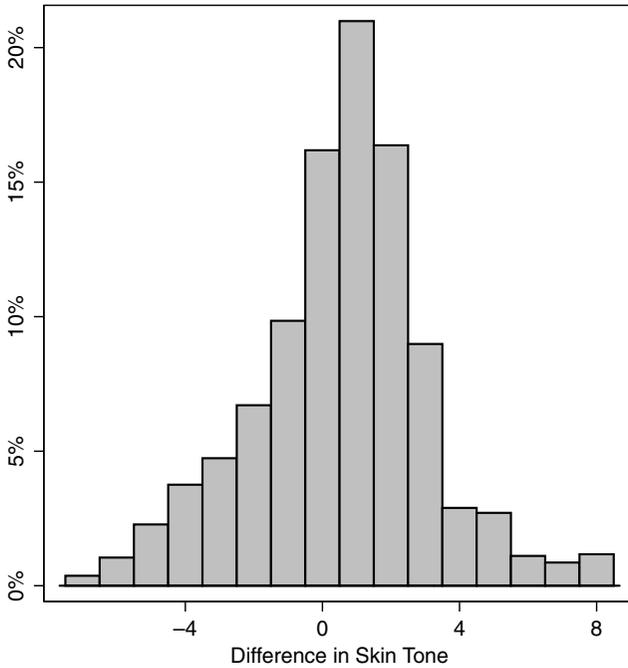
Figure 1. Distribution of difference between perceived skin tone of respondents and self-reported skin tone of interviewer. Positive values represent respondent being perceived as lighter-skinned than interviewer, negative values correspond to respondents being perceived as darker-skinned than interviewer, and 0 represents no difference

the interviewer's perception of their own skin tone results in a measure that ranges from –7 (respondent much darker than interviewer) to 8 (respondent much lighter than interviewer). We display the distribution of our measure of skin tone difference in Figure 1. This and a dichotomous measure of interviewer race (black/white) serve as our primary independent variables.

In our analyses, we cannot strictly consider the race of interviewer to be a semi-random treatment as it is related to several demographic variables; that is, an assumption of random assignment is dubious.[6] Consequently we control for age, sex, education, marital status, and whether the respondent resides in the South in all of our models. To avoid post-treatment bias, we do not control for variables that might also be influenced by the race of the interviewer, such as subjective political attitudes and orientations. We also include objective information in the model as it assuredly influences how informed the interviewers perceive the respondent to be, at least to some extent (details of the factual questions are included in the Supplemental Appendix).[7] We estimate the models

---

[6]We provide results of a balance test in the Supplemental Appendix.

[7]Since objective information may be influenced by the race of the interviewer (e.g., Davis and Silver 2003), including it may influence our estimate of interviewer race on subjective assessments. We find that race of interviewer is unrelated to objective knowledge in the pre-election wave and shares only a weak relationship in the post-election wave. As such, we re-estimated our primary post-election model presented in Table 1,

with linear regression and cluster standard errors by interviewer. Our results are substantively identical when estimating the models with ordered logit (these results are presented in the Supplemental Appendix).

Results

We begin by examining the relationship between race of interviewer and subjective assessments of knowledge among black respondents. Results of this analysis are presented in Table 1; the first and second columns include model results from the pre- and post-election waves, respectively. Figure 2 displays the model-predicted interviewer assessments of black respondents' level of information for both the pre- and post-election surveys. In each wave, we observe that black respondents interviewed by a black interviewer are rated higher than black respondents interviewed by a white interviewer.

In the pre-election wave, we find that black respondents are rated 0.438 ($p = 0.022$) points lower by white interviewers than black interviewers; the discrepancy is even larger at 0.670 points ($p = 0.006$) in the post-election wave. This corresponds to a shift of between 0.402 and 0.624 standard deviations or about 11–17% of the five-point scale. This difference is similar in magnitude to other known correlates of knowledge. For example, consider the relationship between education and interviewers' assessments: movement from less than high school education to an advanced degree is associated with an increase of about 0.5 points on the subjective assessment scale. We also note that women are rated about 0.2 points lower than men, controlling for other factors—this suggests that gender bias may also factor into interviewers' assessments.

Taken together, the observed relationships between interviewer/respondent race and knowledge are meaningful. Not only is racial bias systematic, but it is seemingly capable of overriding other considerations that should guide interviewers' assessments of political knowledge. Indeed, the interviewers appear not to be accurately translating information they may already possess about respondents' levels of political knowledge into their subjective assessments of respondent knowledge. This is a particularly troubling finding given the frequency with which the interviewer's knowledge assessment is employed as a proxy for political knowledge (e.g., Lupia 2016).

Substituting skin tone for self-identification

Next, we substitute racial self-identification for skin tone in our examination of interviewers' subjective assessments of respondent knowledge.[8] Are darker-skinned respondents rated more poorly than lighter-skinned ones? As a descriptive initial examination of the impact of skin tone, we find that subjective assessments of respondent knowledge by white interviewers are significantly negatively correlated with absolute skin tone of the respondent (i.e., the perceived shade of the

---

excluding objective information. Results are consistent across strategies; hence, we control for objective knowledge below. Details of these analyses are included in the Supplemental Appendix.

[8]As skin tone was only assessed in the pre-election wave, we restrict our analysis to all pre-election respondents as well as those post-election respondents with the same interviewer across waves.

Table 1. The relationship between race of interviewer and subjective knowledge among black respondents

	Pre-election	Post-election
Black interviewer (vs. White)	0.407**	0.597**
	(0.178)	(0.207)
Objective information	0.090**	0.261**
	(0.044)	(0.040)
Age	0.007	-0.002
	(0.004)	(0.003)
South	0.005	-0.106
	(0.137)	(0.168)
Female	-0.194**	-0.132*
	(0.087)	(0.079)
Single	-0.393**	-0.231**
	(0.141)	(0.115)
Widowed	-0.209	0.056
	(0.223)	(0.195)
Divorced	-0.026	0.078
	(0.114)	(0.161)
Separated	-0.217*	0.175
	(0.125)	(0.161)
Education	0.213**	0.075*
	(0.031)	(0.043)
Income	0.026**	0.017**
	(0.008)	(0.008)
Constant	1.993**	2.217**
	(0.281)	(0.294)
n	411	398
R^2	0.275	0.348

Standard errors, clustered by interviewer, in parentheses.
** $p < 0.05$, * $p < 0.1$.

respondent's skin tone, coded such that greater values reflect darker skin tones) at −0.102 ($p < 0.001$) in the pre-election survey, and −0.128 ($p < 0.001$) in the post-election survey. This simple test showcases how darker-skinned respondents are the subject of colorism by even trained interviewers.
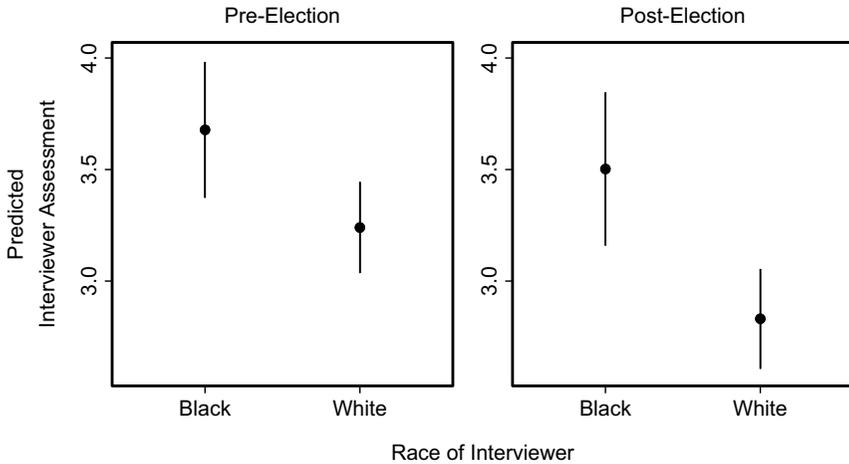
Figure 2. Predicted subjective assessments of political knowledge by race of interviewer with 95% confidence bars. Black respondents only. Note: Predicted values calculated from models in Table 1

To provide a more robust test of this phenomenon, we examine the relationship between subjective assessments of respondent knowledge and the skin tone of interviewers relative to that of respondents. Our expectation is that respondents with darker skin tones relative to interviewers (negative values on the skin tone difference measure) will correspond to poorer knowledge ratings by interviewers—in other words, we should observe a positive relationship between skin tone difference and subjective interviewer assessments. Results from a regression model controlling for the factors we previously discussed are displayed in Table 2; the first column contains estimates from the pre-election model, the second column corresponds to the post-election model.

In both the pre- ($p = 0.045$) and post-election ($p = 0.074$) models, we observe that as respondents are perceived to have lighter skin relative to the interviewer, assessed levels of information increase. For example, results from the model in column 1 indicate that moving from the minimum (very dark-skinned respondents compared to interviewers) to the maximum (very light-skinned respondents compared to interviewers) on the skin tone difference variable results in a change of 0.370 in knowledge, from 3.440 to 3.803, corresponding to 0.340 standard deviations (about 9% of the scale). Moving from one standard deviation below the mean to one standard deviation above is associated with a shift from 3.223 to 3.475, a change of 0.252. Predicted values from both models across the range of our measure of skin tone difference are presented in Figure 3. We observe similar estimates across both models, though the coefficient for skin tone difference for the post-election model is estimated less precisely, perhaps due to the slightly smaller sample size.

Finally, we explore the functional form of the relationship between relative skin tone and subjective knowledge assessment. One could imagine that the strength of the relationship between these two variables exponentially increases as interviewers perceive respondents' skin tone to darker compared to their own; likewise, the relationship might be very weak in cases where the interviewer perceives the respondent

Table 2. The relationship between perceived skin tone difference and subjective knowledge among all respondents

|	Pre-election	Post-Election
Difference in skin tone	0.025**	0.030*
	(0.012)	(0.017)
Objective information	0.152**	0.324**
	(0.026)	(0.021)
Age	0.008**	0.002
	(0.002)	(0.002)
South	0.178*	0.169
	(0.095)	(0.119)
Female	-0.200**	-0.102**
	(0.047)	(0.051)
Single	-0.101	-0.109
	(0.066)	(0.075)
Widowed	-0.122	-0.023
	(0.114)	(0.107)
Divorced	-0.115	0.029
	(0.080)	(0.077)
Separated	-0.158	0.129
	(0.102)	(0.137)
Education	0.197**	0.111**
	(0.018)	(0.020)
Income	0.015**	0.011**
	(0.004)	(0.004)
Constant	1.750**	1.863**
	(0.148)	(0.147)
n	1486	1,217
R^2	0.241	0.381

Standard errors, clustered by interviewer, in parentheses.
** $p < 0.05$, * $p < 0.1$.

to look similar to them, or even lighter than them. Knowing something about the nature of the relationship between relative skin tone difference and subjective knowledge assessments can potentially reveal nuance to theories of racial prejudice and intolerance, especially those founded in social and evolutionary psychological approaches that take perceptions of others' physical characteristics as a foundational mechanism by which prejudice unfolds (e.g., Blair et al. 2002).
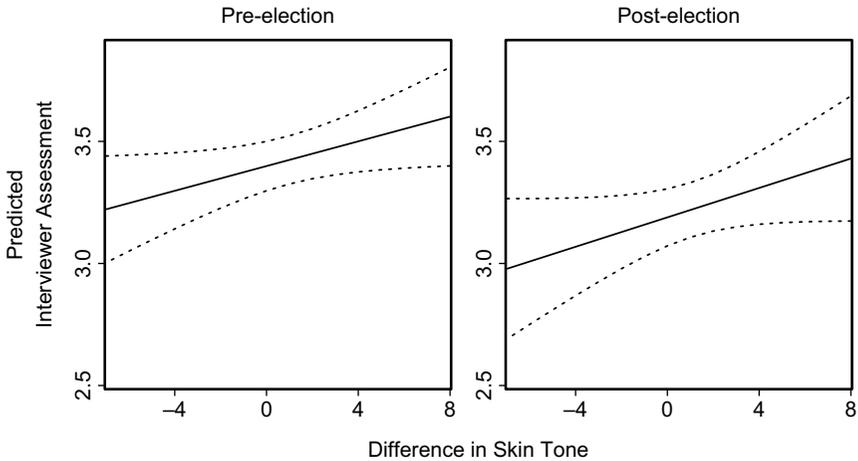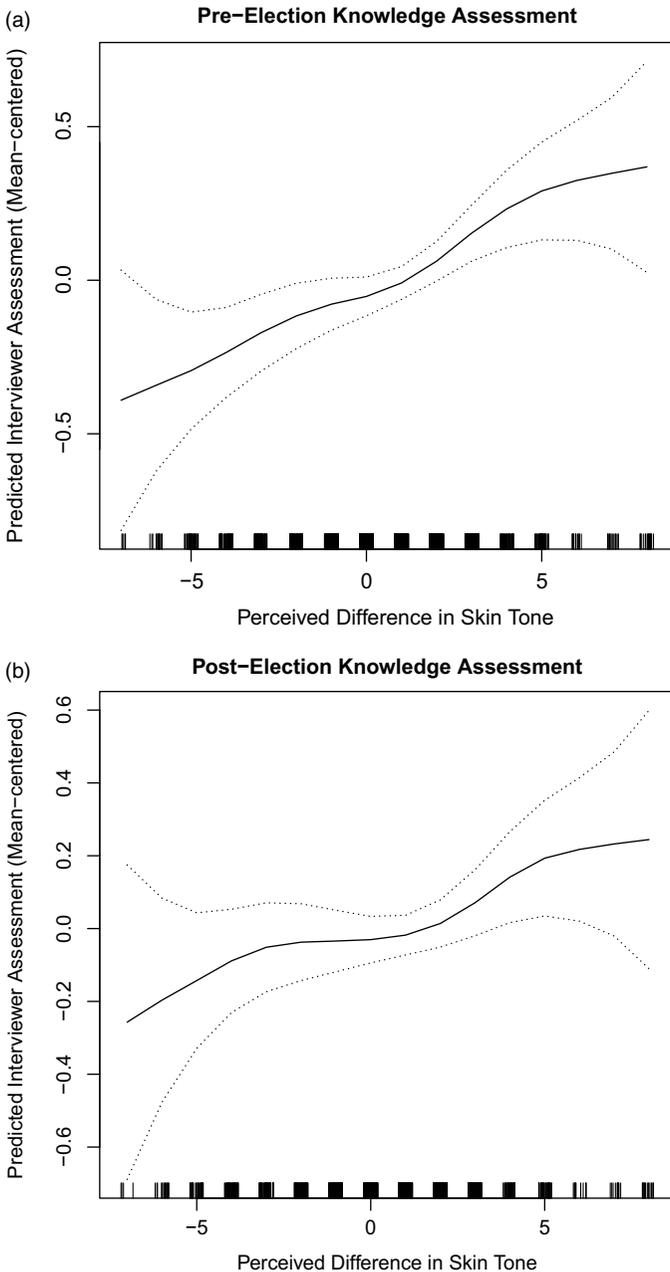
Figure 3. Predicted subjective assessments of political knowledge by difference in skin tone with 95% confidence bars. All respondents. Higher values represent respondent is perceived as lighter than interviewer. Note: Predicted values calculated from models in Table 2
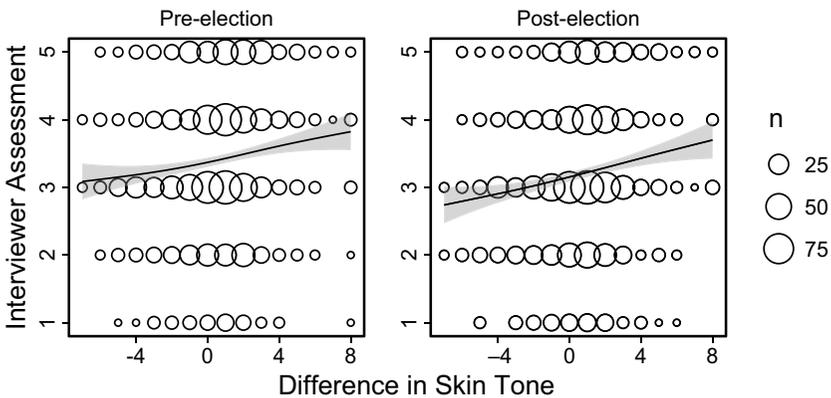
Figure 4. Relationship between subjective assessment of knowledge and perceived skin tone difference. Plotted points are sized proportionally to the number of observations. LOWESS curve with 95% confidence interval is included. Higher values represent respondent is perceived as lighter than interviewer

The simplest way to examine functional form is to impose a non-parametric smoother, like LOWESS, on a scatterplot of the relationship in question. In Figure 4, we present such a scatterplot, depicting the relationship between skin tone difference and interviewer assessment for both the pre- and post-election waves. The black curve is the LOWESS smoother, the shaded region corresponds to a 95% confidence interval, and the circumference of each circular plotting symbol represents the sample size for a particular set of scores for the two variables.

(a)

Pre−Election Knowledge Assessment

(b)

Post−Election Knowledge Assessment

Figure 5. Predicted subjective interviewer assessment of respondent knowledge across levels of perceived skin tone difference. From a generalized additive model including controls for age, gender, education, income, marital status, residence in the South, and performance on factual knowledge questions. Higher values represent respondent is perceived as lighter than interviewer

Remarkably, the relationship between relative skin tone and interviewers' assessment of political knowledge is squarely linear across both waves.[9] As the perceived skin tone of the respondent, compared to the interviewer, becomes lighter, interviewers linearly rate respondents as more knowledgeable. Does this result hold when we control for other known sources of perceived information?

To answer this question, we estimated a generalized additive model (GAM), which allows us to specify a multivariate regression model without imposing assumptions about the functional form of relationships. The models we estimate—one for pre-election interviewer assessments, one for post-election—look identical to those reported in Table 2, except the relationship between perceived skin tone difference and subjective interviewer assessments of knowledge is non-parametric (i.e., allowed to be non-linear).[10] Because we have no particular interest in, or theory about, non-linearity in the relationships with any other control variables, we restrict them to the standard linear specification.

We begin by noting that for neither pre- ($p = 0.474$) nor post-election ($p = 0.434$) subjective assessments does the ANOVA test for non-parametric effects suggests significant non-linearity. The ANOVA tests for parametric effects are, however, significant in both instances ($p = 0.001$). As such, the GAM results do not significantly improve on the OLS results presented in Table 2—the relationship between perceived skin tone differences and subjective assessments of respondent knowledge is linear. Model-based predictions, which indicate that the relationship between the two variables is linear in both cases, are presented in Figure 5.

That the results from a more flexible model that does not require any assumptions about the functional form of the relationship between skin tone differences and subjective information assessment are entirely consistent with the results presented in our parametric models, and the simple bivariate relationships depicted in Figure 4 is striking. This firmly suggests that there is a one-to-one relationship between skin tone and subjective knowledge assessment: as the perceived skin tone of the respondent becomes darker compared to that of the interviewer, interviewers' subjective rating of respondent knowledge becomes more negative. Such a pattern has considerable implications not only for the assessment of respondents' political knowledge but potentially other subjective attitudes, as well. Moreover, that highly trained interviewers operate in such a manner showcases how relentless and pernicious racial biases can be.

Conclusion

Our findings extend previous work in two ways. First, we consider race-of-interviewer effects on subjective assessments of political knowledge, rather than objective ones. Second, we expand the operationalization of "racial differences" between interviewers and respondents by using perceptions of skin tone on the 2012 ANES, in addition to the standard racial self-identification measures. These shifts in focus and operationalization revealed that white interviewers are, at least on average, outwardly biased toward respondents of other races. Previous work intuited this (e.g., Davis and Silver 2003, Steele and Aronson 1995), but tended to focus more on

---

[9]That is to say, a linear regression line fits comfortably within the 95% confidence bands; we did not include this in the figure so as not to clutter the image.

[10]We achieve this using a spline. Models are estimated using the gam R package.

reasons why respondents would react to interviewers in a particular way—a line of reasoning that often lead to an assumption about the impact of respondent threat perception or discomfort on survey response. Our findings, however, reveal some level of conscious or unconscious hostility toward black respondents among white interviewers, not merely a potential for the perception of one among black respondents. In other words, black respondents may not feel discomfort only because of the obvious racial tensions that permeate social interaction, but also because white interviewers are sending perceptible signals to that effect. Simply put, even highly trained interviewers are capable of engaging in colorism, thereby biasing our understanding of black public opinion and political behavior.

These findings have several implications for the literatures on survey methodology and racial bias. Most obviously, our findings suggest that interviewer race is a factor that potentially biases all manner of survey items. Indeed, we now possess evidence for interviewer effects on objective measures of fact (e.g., Davis and Silver 2003), subjective attitudes (e.g., Schuman and Converse 1971, Williams Jr. 1964), and the subjective assessments of factual knowledge that our study revealed—these three categories cover most of the broad question types utilized on surveys like the American National Election Study. Strategies to account for these effects should span both best practices for appropriately utilizing rich datasets that already exist (e.g., the ANES, GSS) and measures for circumventing these problems in the future collection of data. As for the former, interviewer race should probably be adjusted for in models of political attitudes; in some instances (e.g., attitudes about racial issues), researchers might even consider the conditional effect of interviewer race or incorporation of interviewer random effects.

As for strategies for more accurately collecting survey data in the future, researchers might reconsider best practices when it comes to interviewer assignment and training. Evidence for the efficacy of implicit bias training is decidedly mixed, though a recent meta-analysis finds a positive average effect (Bezrukova et al. 2016). Moreover, learning about structural issues (Pritlove et al. 2019) and the impact of social contexts (Payne and Vuletich 2018) can have a impact on the core foundations of implicit biases even when training programs designed to address specific beliefs might occasionally prove ineffective. Thus, programs designed to help interviewers become aware of their biases in evaluating and interacting with respondents —in some form—should not be ruled out.

In addition to altering interviewers' training protocols, researchers might also consider manipulating interviewer assignment based on congruence between interviewer and respondent race. We find that both discrete racial categories (i.e., white interviewers rating black respondents) and differences in skin tone (i.e., lighter-skinned interviewers rating darker-skinned respondents) resulted in discrepancies in evaluations of political knowledge. In this light, interviewer assignment based on either discrete racial identification or skin tone could result in an improvement (i.e., less bias) in subjective assessments. That said, previous work also shows that there are intra-racial group differences of various sorts based on skin tone (e.g., Adams et al. 2016, Ostfeld and Yadon, Forthcoming, Yadon and Ostfeld 2020). Therefore, interviewer assignment based on skin tone might prove more efficacious in reducing bias. These strategies for reducing bias may pose some practical and statistical

challenges, but an accurate understanding of the civic interests and capabilities of all strata of citizens strikes us a worthy cause.

Finally, our results have implications for the measurement of political sophistication. The results presented here, coupled with evidence that traditional objective measures of information tend to ask questions that whites are more likely to know the answer to (Cohen and Luttig 2020, Perez 2015), suggest that when scholars seek to capture political sophistication (variously referred to as "political awareness" or "expertise"), they should rely on more than traditional political information items as neither objective nor subjective measures allow for straightforward comparisons across racial (or gender, per previous work) groups. Minimally, we recommend that measures of information be included alongside variables that capture other dimensions of political engagement, such as voter registration, participation (Zingher and Flynn 2019), and interest in politics (Jones 2020); combining all of these variables into a single index has been a fruitful strategy for reducing measurement error (e.g., Enders and Armaly 2019, Lau and Pomper 2001, Lupton, Myers, and Thornton 2015). A more sophisticated approach might entail adjusting political knowledge questions for differential item functioning (or measurement invariance) by respondent and interviewer race (e.g., see Enders 2021). Evaluating citizens' democratic capabilities is a critical function of political scientists (Achen and Bartels 2017), one we must take care in properly executing.

Supplementary material. For supplementary material accompanying this paper visit https://doi.org/10.1017/rep.2021.40

References

Achen CH and Bartels LM (2017) *Democracy for Realists: Why Elections Do Not Produce Responsive Government*, Vol. **4**. Princeton, NJ: Princeton University Press.

Adams EA, Kurtz-Costes BE and Hoffman AJ (2016) Skin tone bias among african americans: antecedents and consequences across the life span. *Developmental Review* **40**, 93–116.

Bezrukova K, Spell CS, Perry JL and Jehn KA (2016) A meta-analytical integration of over 40 years of research on diversity training evaluation. *Psychological Bulletin* **142**, 1227–1274.

Blair IV, Judd CM and Chapleau KM (2004) The influence of afrocentric facial features in criminal sentencing. *Psychological Science* **15**, 674–679.

Blair IV, Judd CM, Sadler MS and Jenkins C (2002) The role of afrocentric features in person perception: judging by features and categories. *Journal of Personality and Social Psychology* **83**, 5–25.

Campbell ME, Keith VA, Gonlin V and Carter-Sowell AR (2020) Is a picture worth a thousand words? An experiment comparing observer-based skin tone measures. *Race and Social Problems* **12**, 266–278.

Cohen CJ and Luttig MD (2020) Reconceptualizing political knowledge: race, ethnicity, and carceral violence. *Perspectives on Politics* **18**, 805–818.

Davis DW (1997a) The direction of race of interviewer effects among African-Americans: Donning the Black Mask. *American Journal of Political Science* **41**, 309–322.

Davis DW (1997b) Nonrandom measurement error and race of interviewer effects among African Americans. *Public Opinion Quarterly* **61**, 183–207.

Davis DW and Silver BD (2003) Stereotype threat and race of interviewer effects in a survey on political knowledge. *American Journal of Political Science* **47**, 33–45.

Eberhardt JL, Davies PG, Purdie-Vaughns VJ and Johnson SL (2006) Looking deathworthy: perceived stereotypicality of black defendants predicts capital-sentencing outcomes. *Psychological Science* **17**, 383–386.

Enders AM (2021) A matter of principle? On the relationship between racial resentment and ideology. *Political Behavior* **43**, 561–584.

Enders AM and Armaly MT (2019) The differential effects of actual and perceived polarization. *Political Behavior* **41**, 815–839.

Feliciano C (2016) Shades of race: how phenotype and observer characteristics shape racial classiffication. *American Behavioral Scientist* **60**, 390–419.

Foy SL and Ray R (2019) Skin in the game: colorism and the subtle operation of stereotypes in Men's College Basketball. *American Journal of Sociology* **125**, 730–785.

Frey FE and Tropp LR (2006) Being seen as individuals versus as group members: extending research on metaperception to intergroup contexts. *Personality and Social Psychology Review* **10**, 3265–3280.

Goldsmith AH, Hamilton D and Darity Jr. W(2006) Shades of discrimination: skin tone and wages. *American Economic Review* **96**, 242–245.

Hagiwara N and Kashy Joseph Cesario DA (2012) The independent effects of skin tone and facial features on whites' affective reactions to blacks. *Journal of Experimental Social Psychology* **48**, 892–898.

Han JH (2020) Does skin tone matter? Immigrant mobility in the U.S. labor market. *Demography* **57**, 705–726.

Hannon L and Deffna R (2014) Just skin deep? The impact of interviewer race on the assessment of African American respondent skin tone. *Race and Social Problems* **6**, 356–364.

Hannon L and DeFina R (2016) Reliability concerns in measuring respondent skin tone by interviewer observation. *Public Opinion Quarterly* **80**, 534–541.

Hannon L, Defina R and Bruch S (2013) The relationship between skin tone and school suspension for African Americans. *Race and Social Problems* **5**, 281–295.

Hill ME (2002) Race of interviewer and perception of skin color: evidence from the multi-city study of urban inequality. *American Sociological Review* **67**, 99–108.

Jones PE (2020) Partisanship, political awareness, and retrospective evaluations, 1956–2016. *Political Behavior* **42**, 1295–1317.

Keith VM, Nguyen AW, Taylor RJ, Mouzon DM and Chatters LM (2017) Microaggressions, discrimination, and phenotype among African Americans: a latent class analysis of the impact of skin tone and BMI. *Sociological Inquiry* **87**, 233–255.

Klonoff EA and Landrine H (2000) Is skin color a marker for racial discrimination? Explaining the skin color-hypertension relationship. *Journal of Behavioral Medicine* **24**, 329–338.

Lau RR and Pomper GM (2001) Effects of negative campaigning on turnout in U.S. senate elections, 1988–1998. *The Journal of Politics* **63**, 804–819.

Lupia A (2016) *Uninformed: Why People Know so Little About Politics and What We Can do About It.* Oxford University Press.

Lupton RN, Myers WM and Thornton JR (2015) Political sophistication and the dimensionality of elite and mass attitudes, 1980–2004. *Journal of Politics* **77**, 368–380.

Massey DS and Martin JA (2003) *The NIS Skin Color Scale.* Princeton, NJ: Princeton University Press.

Ostfeld MC and Yadon N (forthcoming) ?Mejorando La Raza?: The Political Undertones of Latinos' Skin Color in the United States. *Social Forces.*

Palmer CL and Peterson RD (2015) Halo effects and the attractiveness premium in perceptions of political expertise. *American Politics Research* **44**, 353–382.

Payne BK and Vuletich HA (2018) Policy insights from advances in implicit bias research. *Policy Insights from the Behavioral and Brain Sciences* **5**, 49–56.

Perez EO (2015) Mind the gap: why large group deficits in political knowledge emerge and what to do about them. *Political Behavior* **37**, 933–954.

Pritlove C, Juando-Prats C, Ala-leppilampi K and Parson JA (2019) The good, the bad, and the ugly of implicit bias. *The Lancet* **393**, 502–504.

Schuman H and Converse JM (1971) The effects of black and white interviewers on black responses in 1968. *Public Opinion Quarterly* **35**, 44–68.

Steele CM and Aronson J (1995) Stereotype threat and the intellectual test performance of African Americans. *Journal of Personality and Social Psychology* **69**, 797–811.

Sue DW (2010) *Microaggressions in Everyday Life: Race, Gender, and Sexual Orientation.* New York, NY: Wiley and Sons.

Weaver VM (2012) The electoral consequences of skin color: the 'hidden' side of race in politics. *Political Behavior* **34**, 159–192.

Williams Jr. JA. (1964) Interviewer-respondent interaction: a study of bias in the information interview. *Sociometry* **27**, 338–352.

Yadon N and Ostfeld MC (2020) Shades of privilege: the relationship between skin color and political attitudes among white Americans. *Political Behavior* **42**, 1369–1392.

Zingher JN and Flynn ME (2019) Does polarization affect even the inattentive? Assessing the relationship between political sophistication, policy orientations, and elite cues. *Electoral Studies* **57**, 131–142.