

Stereotype Threat and Race of Interviewer Effects in a Survey on Political Knowledge

Darren W. Davis Michigan State University
Brian D. Silver Michigan State University

Social desirability is generally thought to underlie the propensity for survey respondents to tailor their answers to what they think would satisfy or please the interviewer. While this may in fact be the underlying motivation, especially on attitudinal and opinion questions, social desirability does not seem to be an adequate explanation for interviewer effects on factual questions. Borrowing from the social psychology literature on stereotype threat, we test an alternative account of the race-of-interviewer effects. Stereotype threat maintains that the pressure to disconfirm and to avoid being judged by negative and potentially degrading stereotypes interferes with the processing of information. We argue that the survey context contains many parallels to a testing environment in which stereotype threat might alter responses to factual questions. Through a series of framing experiments in a public opinion survey and the reliance on the sensitivity to the race of the interviewer, our results are consistent with expectations based on a theory of "stereotype threat." African American respondents to a battery of questions about political knowledge get fewer answers right when interviewed by a white interviewer than when interviewed by an African American interviewer. The observed differences in performance on the political knowledge questions cannot be accounted for by differences in the educational background or gender of the respondents.

The Problem

The presentation of self and the desire to project a positive self-image are important factors influencing how people interact with others. To the extent that the projection of a positive self-image becomes salient through the internalization of societal norms, what may appear to be sincere, honest, and unbiased expressions of political and social beliefs may actually be a conscious attempt to conceal beliefs by giving socially desirable responses (Goffman 1963, 1973) or an unconscious reaction to being perceived in a negative light (Steele and Aronson 1995). Measures of political and social attitudes will tend to reflect more than expressed opinions but also come to reflect an attempt to project a positive self-image.

The research on interviewer effects in surveys has produced compelling evidence of how seriously the pro-

jection of a positive self-image can affect the measurement of political and social attitudes (Anderson, Silver, and Abramson 1988a, 1988b; Davis 1997; Schuman and Converse 1971; Finkel, Guterbock, and Borg 1999). It is generally believed that a norm of social desirability underlies the differences in the information respondents report to interviewers of different races. "Social desirability bias" carries with it the idea that respondents overtly "perform" or "front" during an interview in ways that differ from their true feelings. Respondents try to look better in the eyes of the interviewer by expressing opinions that conform with perceived interviewer expectations or wider societal norms. For example, a respondent who has not actually voted in the last election may say that he or she has voted in order to appear to conform to good citizen norms (Silver, Anderson, and Abramson 1986; Anderson and Silver 1986).¹ Voters may say they support an African

Darren W. Davis is Associate Professor of Political Science, Michigan State University, East Lansing, MI 48824-1032 (davisda@msu.edu). Brian D. Silver is Professor of Political Science, Michigan State University, East Lansing, MI 48824-1032 (bsilver@msu.edu).

The data on which this study is based are from the 20th round of the State of the State Survey, conducted by the Office for Survey Research of the Institute for Public Policy and Social Research at Michigan State University. We thank Karen Clark, Katherine Cusick, and Larry Hembroff for administering the survey. An earlier version of this article was presented at the annual meeting of the Midwest Political Science Association, April 25–28, 2002. We thank Christopher Parker and Paul Sniderman for helpful comments on that version of the article. None of these individuals is responsible for any errors in our analysis or interpretations.

¹The literature has shown that respondents who do *not* endorse good citizens' norms in general are not only less likely actually to vote but if they did not vote they are also less likely to report falsely that they did vote (Silver, Anderson, and Abramson 1986; Anderson and Silver 1986).

American Journal of Political Science, Vol. 47, No. 1, January 2003, Pp. 33–45

©2003 by the Midwest Political Science Association

ISSN 0092-5853

American respondent, because it may be “politically correct,” but they will actually vote for the opposing white candidate (Finkel, Guterbock, and Borg 1999). An African American respondent may also put on a performance by “donning the Black mask” and not revealing his or her true feelings when interviewed by a white interviewer but speak more frankly in the presence of a black interviewer (Davis 1997).

While a social desirability explanation may apply to attitudinal and opinion questions, it does not adequately explain the bias that interviewer effects may impart on objective indicators. More importantly, social desirability is only *believed* to underlie interviewer effects on attitudinal questions.

In this article, we provide an alternative account of the race-of-interviewer effects on survey responses, drawing on the social psychological literature on race differences in test performance. Steele and his colleagues (Aronson et al. 1999; Spencer, Steele, and Quinn 1999; Steele 1997; Steele and Aronson 1995) show that the awareness of threat of being stereotyped and confirming a negative stereotype interferes with the processing of information by groups about whom the relevant negative stereotype exists, even when they themselves do not believe the stereotype. For example, stereotype threat may impair the test performance of African Americans when taking an apparently standardized test of intellectual ability, or it may impair the test performance of women when taking a standardized test of mathematical ability. The activation of a stereotype about a group’s lower intellectual ability can negatively affect their test performance. But when the stereotype threat is removed, such as by changing the reported nature of the test, test performance improves.

Although most of the research on stereotype threat occurs in the context of a testing environment in which individuals are made to feel self-conscious about race, we explore the extent to which a stereotype threat approach can explain race-of-interviewer effects in an attitude survey. We argue that a survey can heighten respondents’ sensitivity to race and that it parallels the situation that occurs in testing. In addition to the structure of the questions and format of response options, which conveys a sense of formality, respondents often think there are right and wrong answers.

A more important test parallel exists in the fact that attitude surveys do, indeed, often ask respondents about their own “objective” behavior, background, or attributes (e.g., voting, alcohol consumption, work activity, sexual behavior, income, education, family composition, race, and gender).² The answers to such questions can be said

²We are not intending to say that race is simply an objective characteristic, but instead that race is often likely to be perceived as such

to be true or false or (less judgmentally) accurate or inaccurate. Surveys often directly test a respondent’s knowledge of other people, issues, events, and the broader environment by asking questions that clearly have right and wrong, as well as readily verifiable, answers. What are the names of your U.S. Senators? Which political party holds a majority in the state legislature? How many justices are there on the U.S. Supreme Court? Can you place a candidate on the liberal/conservative scale? How many things do you like or dislike about George Bush? Such questions have a potential to be threatening to respondents who do not want to appear uninformed or ignorant. For this reason, survey researchers sometimes even tell the respondents that the questions are “not a test,” to convey the idea that right or wrong answers to the questions will not be used to form a judgment about the respondent’s intelligence, knowledge, or honesty.

We use question wording experiments in a public opinion survey to examine the role of stereotype threat on survey responses. We hypothesize that African American respondents to a set of questions on political knowledge will give fewer correct answers to white interviewers than to African American interviewers. Furthermore, we expect African American respondents to give fewer correct answers to knowledge questions that are explicitly described as a test than to the same questions that are explicitly said *not* to be a test.

The answers to basic questions about political institutions, such as how many Supreme Court justices there are, do not have an inherent racial meaning or links to issues of group identity or solidarity. Also, if the respondents give different answers when the interviewers are members of one race than when they are members of another race, the *difference* in the number of correct answers is unlikely to be due to social desirability. So we need another theory to account for such a pattern of responses to political knowledge questions. The central concept of the theory that we shall rely on is that of “stereotype threat.”

Stereotype Threat

People are motivated to appear competent and to take control rather than being merely victimized by negative stereotypes (Oyserman and Swim 2001). According to Steele (1997), when a widely known negative stereotype (e.g., poor intellectual ability) exists about a group which jeopardizes their appearance of competence, it creates for

both by interviewers and respondents, and hence in principle respondents may assume that it can be verified by observation. Even in telephone surveys, the respondent’s perception of the race of the interviewer agrees with the interviewer’s own self-description by race most of the time.

some members of that group a burden of suspicion that acts as a threat. This threat arises when a person's behavior could be judged (usually by members outside the group) in terms of a stereotype. Such a threat is cued by the mere recognition that a negative group stereotype could apply to oneself in a given situation. The person's anxiousness to disconfirm a negative stereotype and a potentially degrading label may interfere with performance of the task. For members of the group to which the stereotypes apply, the threat of stereotypes may lead to inefficient information processing by redirecting attention needed to perform a task to irrelevant concerns, which may lead to self-consciousness, overcautiousness, and frustration (Steele and Aronson 1995). Through these mechanisms, stereotype threat is expected to undermine the processing of information, and hence performance on tests or other assigned tasks.

Stereotype threat may apply to any stigmatized group (African Americans, women, whites, and students from low socio-economic status) and is usually treated as context-dependent or situationally specific (cf. Marx, Brown, and Steele 1999). Its activation does not require overt references to a group's inferiority; however, such references may enhance the effect. Steele and Aronson (1995) examined performance on the most difficult verbal items in the GRE among African American and white college students. Activating a stereotype threat among one group by introducing the exam as "a test of intellectual ability" led African American participants to score significantly worse than whites who were given the same instructions. African Americans in the nondiagnostic groups performed equally to whites. In another experiment, the mere recording of their race on a demographic section of the exam, just before taking the exam, impaired the performance of African American students. Blascovich et al. (2001) found impaired performance on a cognitive task among African Americans high in stereotype threat, and they also revealed that African Americans who had high stereotype threat exhibited higher blood pressure than whites.

Walsh, Hickey, and Duffy (1999) examined the extent to which a negative stereotype threat might explain gender differences in mathematical problem solving. Instructing participants that the SAT "has been found to show gender differences in math performance and that males score higher than females" had the effect of producing significantly lower test scores among the women. The researchers argued that the threat of a negative stigma associated with lower math ability creates a threat to self-characterization and an emotional reaction that interferes with performance (Walsh, Hickey, and Duffy 1999, 236). Research by Spencer et al. (1997), Stangor, Carr, and Kiang

(1998), and Quinn and Spencer (2001) revealed similar findings: women significantly underperformed in comparison to equally qualified men on a difficult math test, but women performed just as well as equally qualified men on an advanced literature test.

Extending the concept of stereotype threat to social class, Croizet and Claire (1998) introduced stereotype threat to low-SES participants by informing them the GRE was intended to "assess your intellectual ability for solving verbal problems." The hypothesis that low-SES participants would do less well because stereotype anxiety would disrupt their performance was strongly supported.

Aronson et al. (1999) explored the extent to which even white males can experience stereotype threat. In an experiment involving a group of white male college students with high math aptitude, a stereotype threat was activated by introducing a math section on the SAT by asking participants to read articles about the phenomenal math achievement of Asians and informing the test takers that Asians outperform other students on tests of math ability. Making salient the Asian stereotype depressed the performance of white males, even though they reported exerting more effort.

Stone et al. (1999) explored stereotype threat on athletic performance. They argued that if an athletic performance were framed as an indication of natural athletic ability—one's genetically determined physical capabilities—white athletes might infer that they are being evaluated on the basis of a negative stereotype and consequently perform more poorly than if the stereotype were not made salient. When a negative stereotype was activated, white participants did worse than control participants when a golf task was framed as diagnostic of "natural ability." Likewise, African American participants performed worse than did control participants when performance on a golf task was framed as diagnostic of "sports intelligence."

Survey Design

The activation of stereotype threat seems to be affected by how tasks are framed. If a task is framed in such a way that individuals feel that their performance or responses may be judged against or evaluated in the context of a negative group stereotype, the pressure to disconfirm the negative stereotype produces anxiety that interferes with the processing of information or the ability to solve problems.

We take a similar framing approach to questions tapping political knowledge. Though there are many survey questions that parallel test questions, such as asking respondents to place candidates and parties on ideological

and issue dimensions and what types of things respondents like and dislike about political parties and candidates, political knowledge questions come closer in form to an actual test. A battery of seven “political knowledge” questions was included:

1. How many years is the term of office of a United States Senator? (correct answer: 6)
2. How many justices (judges) are there on the U.S. Supreme Court? (correct answer: 9)
3. What is the minimum voting age in Michigan? (correct answer: 18)
4. Do you happen to know how many times an individual can be elected president? (correct answer: 2)
5. How much of a majority is needed for the Senate and House of the U.S. Congress to override a presidential veto? (correct answer: two-thirds majority)
6. Do you happen to know what political office William Rehnquist holds? (correct answer: Chief Justice, Justice, or judge of the Supreme Court)
7. Which political party holds a majority in the Michigan legislature? (correct answer: Republican)

This type of factual item has been used widely to measure political knowledge. Political knowledge has been related to political tolerance, political participation and voting, ideological self-identification, acquisition of information, and perceptions of issues (e.g., racial issues, abortion; Delli Carpini and Keeter 1993, 1996; Jennings 1996; Mondak 2001). Unfortunately, these measures are not without criticism. A great deal of this criticism involves the handling of the “Don’t Know” response. Delli Carpini and Keeter (1993) maintain that the “Don’t Know” response should be encouraged in order to reduce the amount of guessing. However, Mondak (2001) argues that discouraging the “Don’t Know” response does not eliminate guessing, but rather such discouragement creates a validity problem because it activates a differential propensity to guess. In addition, Mondak (2001) argues that the open-ended response format of knowledge questions makes it difficult to handle partially correct responses.³ Following this logic, we attempted to discourage “Don’t Know” responses by not offering it as a response option.

After each question was read, the interviewers could either record that the respondent gave the correct answer (preprogrammed and coded on the screen) or they were to enter the verbatim answer. If a respondent refused to

answer the question, the answer was entered as “refused.” If a respondent at first answered “don’t know” or “not sure,” the interviewers prompted the respondent with the statement “Just give me your best guess.” After prompting, the respondents’ final answers were recorded as either a correct answer, a different answer (recorded verbatim), or a “don’t know.” Thus, without browbeating the respondents the interviewers did prod those respondents who at first responded “don’t know” or “not sure” to answer the question if they could.

The seven questions were preceded by an introduction, which was offered in two variations.

Nonthreatening: Now I have a few more questions concerning public figures and government affairs. These questions are not a test of any sort. Instead, we want to see how much information about them gets out to the public from television, newspapers, and the like.

Threatening: Now I have a few more questions concerning public figures and government affairs. These questions are a kind of test. We want to see how much information about them gets out to the public from television, newspapers, and the like.

The two versions of the introduction were programmed to be offered randomly to respondents. Although we anticipated that placing any tests of knowledge in a survey could be perceived as threatening, we hypothesized that an introduction that explicitly labeled questions as “a kind of test” would represent greater threat than one that explicitly stated that the questions were “not a test of any sort.” The “nonthreatening” version was modeled on the pilot of the 1985 National Election Study as reported by Zaller (1986).

We did not make explicit references to racial differences in our framing, but instead we relied on more subtle cues based on the race of the interviewer. Explicit statements about racial differences in intellectual ability, like the ones used in laboratory experiments, would likely have been met with a great deal of apprehension among our respondents and risked termination of the interview. Instead, we relied on the racial cuing and sensitivity to the race of the interviewer. Research by Danso and Esses (2000) suggests that the race of test administrator alone is enough to create a stereotype threat. Public opinion surveys are often replete with questions that explicitly deal with racial stereotypes that require great forbearance for African American respondents to sit through. Although such questions can activate a stereotype threat, the race of

³In addition to discouraging a “Don’t Know” response, Mondak (2001) prefers a multiple choice format to the political knowledge questions. He argues that a multiple choice format in knowledge questions is not susceptible to response sets and they will tend to measure knowledge more reliably than open-ended questions.

the interviewer provides another element (or perceived expectation) that should be heeded or avoided (Davis 1997). It is not just the framing of survey questions, or the general social stereotypes about the groups to which the respondents belong, that induces stereotype threat but also *who is asking the questions*.

Hence, we hypothesized that respondents would offer fewer correct answers if they were given the threatening version of the questions than if they were given the nonthreatening version. And we hypothesized that black respondents would offer fewer correct answers to the political information questions when they interviewed by white interviewers than by black interviewers.

Data

A random digit dialing (RDD) telephone survey of the Michigan adult population was conducted during January and February 2001. Using disproportionate sampling by region of the state, 954 interviews were completed.⁴ The interviews averaged 23 minutes in length. The interviewers were assigned telephone numbers using standard RDD methodology. The assignment of telephone numbers was not based on either the characteristics of the interviewer or the characteristics of the respondents (e.g., region, “expected” race). A sufficient percentage of the interviewer staff were either African American or white (by self-description) that it was not necessary to try to assign interviewers to particular (potential) respondents by race of respondent.

The main topics of the survey concerned quality of life in communities, evaluation of governmental performance, and electoral reform. The political knowledge questions appeared approximately midway in the interview.

The survey also included an oversample of telephone exchanges that had a higher than average percentage of African-American households, in an effort to complete a

minimum of 200 interviews with African-American respondents. In all, 212 self-identified African Americans were interviewed, as well as 643 whites and 89 others. For purposes of administering the political knowledge questions, the “African American” respondents are those who identified themselves as only black or African American but no other racial category, and those who are “white” in the political knowledge study identified their race as “white” but selected no other racial category.⁵

The very last question in the survey asked the respondents to report on their perception of the race of the interviewer: “*Finally, what do you think is my racial background?*” We use the answers to this question as well as the interviewer’s own racial self-identification in the analysis of race-of-interviewer effects. Because we are especially concerned about the effects of perceived racially related “expectations” on the respondents’ performance on the political knowledge questions, we expect to find stronger correlations between perceived race of the interviewer and test performance than between the actual (self-reported) race of the interviewer and test performance.

Eighty-five percent of the self-identified African American respondents identified the interviewers either as black or as white, while 11 percent answered Don’t Know or Refused. In contrast, 64 percent of the self-identified White respondents identified the interviewers either as black or white (66 percent among those who answered the political knowledge question battery), while 26 percent answered “Don’t Know” and 4 percent refused. We speculate that the differential reflects sensitivity about the race issue among whites. Below, we replicate the analysis of race-of-interviewer effects using the interviewer’s own racial identification, not only the respondent’s reported perception of the race of the interviewer.

For the African American respondents, who are the main concern of this study, the fact that the race of some of the interviewers was perceived to be ambiguous—that is, the respondents reported that they did not know or they refused to state the race of the interviewer—raises another possibility for analysis. Given the theoretical expectation that African American respondents will offer fewer

⁴The regional strata are the six regions of the Michigan State University Extension service (MSUE), with an additional break-out of Detroit city, so that in all seven regional strata were employed. For statewide analyses, weights are constructed to make the overall results representative of the state adult population (see Hembroff and Silver 2001). For purposes of this article, the data are used in *unweighted* form, because there is no attempt to generalize to the statewide population. However, the survey procedures provided for random selection of respondents using the RDD methodology as well as random the assignment of the alternative tasks or question variants to respondents (see the text for further explanation). Using the American Association for Public Opinion Research (AAPOR) standard, the overall response rate (RR4) was 46.4 percent; the refusal rate (REF2) 16.4 percent; and the cooperation rate 73.9 percent. For the current AAPOR standards, see <http://www.aapor.org/ethics/stddef.html>.

⁵Following recent practice suggested by OMB Directive 15 (1997) and followed in the U.S. Census of 2000, this survey asks respondents first whether they are Hispanic/Latino, then “What is your race?” The racial categories offered are White/Caucasian, Black or African American, Hawaiian or other Pacific Islander, Asian, American Indian, or Alaska Native. The respondents are allowed to make multiple selections. In this survey, if they selected “Black or African American” *but no other racial category*, they were later administered the battery of political knowledge questions. If they selected “White or Caucasian” but no other racial category, then a random one-third of such respondents were later administered the battery of political knowledge questions. This approach was chosen to produce approximately equal numbers of black and white respondents.

TABLE 1 Percent of Correct Answers to Political Knowledge Questions, by Respondent's Race

	Minimum Voting Age	Number of Terms President Can Serve	Majority Party State Legislature	Pct. to Override Veto	Number U.S. Supreme Court Justices	Term U.S. Senator	Who Is William Rehnquist
Black Respondents							
Correct	87%	85%	69%	25%	22%	19%	15%
Wrong	11	10	28	54	66	73	16
Don't know	1	4	3	10	10	7	32
Refuse	1	1	1	1	2	1	37
Total Pct.	100%	100%	101%	100%	100%	100%	100%
(Base N)	(212)	(212)	(212)	(212)	(212)	(212)	(212)
White Respondents							
Correct	85%	90%	52%	44%	32%	34%	30%
Wrong	15	8	34	47	63	63	10
Don't know	0	2	12	9	5	2	19
Refuse	1	0	2	1	0	1	42
Total Pct.	101%	100%	100%	101%	100%	100%	101%
(Base N)	(221)	(221)	(221)	(221)	(221)	(221)	(221)

correct answers to the political knowledge questions to white interviewers than to black interviewers, what effect might an *ambiguous* race of the interviewer be expected to have? In our later analysis we will argue that ambiguity in this context might create even greater anxiety than knowing the race of the interviewer for certain.

Findings

Race differences in political knowledge. Table 1 summarizes the distribution of responses to each of the seven political knowledge questions. The questions are arrayed from the easiest to the hardest, based on the percentage of correct answers offered by the African American respondents. For all but one question, "correct" and "wrong" answers comprised the majority of answers. Only on the "Who is William Rehnquist" question were most of the answers either "don't know" or "refused." On the whole, the patterns of answers do not differ much between African American and white respondents, though on five items a larger percentage of whites answered correctly, and on two items (minimum voting age, which party holds a majority in the state legislature) a larger percentage African Americans answered correctly.

On average (Table 2), African American respondents answered 3.05 questions correctly, while white respondents answered 3.83 correctly, a statistically significant difference ($p < .001$). Although there is a strong gradient

in the number of correct answers by educational level of the respondents, the race differences cannot be accounted for by differences in the educational attainments of whites and African Americans.⁶

Blacks and whites did not differ in the tendency to "refuse" to answer the questions ($p = .738$). They did differ significantly, however, in the number of correct answers, wrong answers, and don't know answers that they offered (Table 2).

Race of the interviewers. Half of the black respondents were interviewed by persons whom they identified as black (see top panel of Table 3), compared to only one-fourth of the white respondents. About 15 percent of black respondents did not identify the interviewers as either black or white, compared to 34 percent of the white respondents. The larger percentage among white respondents may reflect sensitivity on the race issue. When we examine the interviewers' own racial self-identification

⁶The educational breakdowns are not shown in a table. To illustrate the conclusion above: among those with a completed college degree, African Americans answered 4.14 questions correctly on average, while whites answered 4.68 correctly; and among those with a high school education, African Americans answered 2.61 correctly while whites answered 2.92 correctly. An ANOVA comparing the mean number of correct answers for whites and blacks at each educational level (college graduate, some college, high school graduate, and less than high school graduate) finds that both the within racial group and between racial group differences are statistically significant at $p .001$.

TABLE 2 Distribution of Answers to Political Knowledge Questions

	Percent Correct	Percent Wrong	Percent Don't Know ^a	Percent Refused
Black Respondents				
0	2	9	52	60
1	8	10	31	36
2	30	24	6	2
3	29	28	6	2
4	13	20	2	0
5	8	6	1	0
6	4	1	2	0
7	4	0	1	0
Mean	3.05	2.64	0.86	0.46
Std. Dev.	1.54	1.40	1.31	0.63
(Base N)	(212)	(212)	(212)	(212)
White Respondents				
0	0	11	71	57
1	6	14	24	42
2	18	29	3	0.0
3	21	25	2	1
4	23	16	1	0
5	15	4	1	0
6	8	0	0	0
7	9	0	0	0
Mean	3.83	2.34	0.39	0.44
Std. Dev.	1.65	1.32	0.75	0.52
(Base N)	(221)	(221)	(221)	(221)
ANOVA Black vs. White Respondents				
F	25.829	5.166	20.883	.112
Sig.	<.001	.024	<.001	.738
df	1 & 431	1 & 431	1 & 431	1 & 431

^aRespondents who initially stated “don’t know” or “not sure” were asked to “just make your best guess.” Respondents who “refused” (said they did not want) to answer a question were not asked again.

(bottom panel of Table 3), we find far larger percentages who identify as black or white than the respondents reported.

For purposes of testing the effects of race of the interviewers on the political knowledge scores of the respondents, it is tempting to rely on the interviewer’s self-identification as likely to be more accurate. However, race of the interviewer effects are more likely to be filtered through the respondents’ *perceptions* of the interviewer. Also, if stereotype threat underlies the pattern of correct responses that we find among African American respondents, then the “threat” is likely to be induced by the *perception* that the interviewers are white, not directly

TABLE 3 Race of Interviewer, by Race of Respondent

Race of Interviewer	Race of Respondent ^a (self-identified)	
	Black	White
A. As Perceived by Respondent		
Black	50%	25%
White	36	42
Other	3	6
Don’t know	8	25
Refuse	4	3
Total Percent (Base N)	101% (212)	101% (221)
B. Interviewer Self-Identified		
Black only	67%	55%
White only	26	34
Black and white ^b	5	10
Other ^c	2	1
Total Percent (Base N)	100% (212)	100% (221)

^aBased only the respondent’s *first* answer to questions on own race or ethnicity. Only respondents who self-identified either as black or as white were administered the political knowledge battery.

^bInterviewer classified self as black and as white in separate questions.

^cIncludes those who say Native American, Asian, or combination of one of these categories with black or white.

by whether the interviewers are *actually* white (by their self-description). Furthermore, since the focus of this research is on the black respondents, it is important that for most black respondents the race of the interviewer was not ambiguous—only 15 percent of black respondents answered “other,” “don’t know,” or “refuse” when asked to report the race of the interviewer.

Race of the interviewer and political knowledge. Among white respondents, the mean number of correct answers is not associated with either the respondent’s perceived race of the interviewer or the interviewer’s self-identified race (Table 4, panel A). The differences in the number of correct answers by perceived race of interviewer are small and not statistically significant ($p = .485$). Nor does the number of correct answers given by white respondents differ significantly by the self-identified race of the interviewers ($p = .922$).

In contrast, among black respondents, the perceived race of the interviewer matters a great deal. When interviewed by a black interviewer (perceived), black

TABLE 4 Mean Number of Correct Answers to Political Knowledge Questions, by Race of Respondent and Race of Interviewer

Race of Interviewer	Race of Respondent (self-identified)	
	Black Mean (N)	White Mean (N)
A. As Perceived by Respondent		
Black	3.42 (105)	4.00 (55)
White	2.80 (76)	3.87 (90)
DK, Refuse, Other	2.39 (31)	3.66 (76)
ANOVA		
By Respondent-Perceived		
Race of Interviewer		
F	7.258	.725
Sig.	.001	.485
df	2 & 209	2 & 215
B. Interviewer Self-Identified		
Black only	3.13 (142)	3.88 (121)
White only	2.87 (55)	3.80 (74)
Black and white	... (11)	3.78 (23)
ANOVA		
By Self-Identified		
Race of Interviewer		
F	.640	.081
Sig.	.528	.922
df	2 & 205	2 & 215

respondents answer an average of 3.42 political knowledge questions correctly. When interviewed by a white interviewer (perceived), black respondents answer an average of 2.80 questions correctly. When interviewed by an interviewer whose race is not perceived as either black or white (DK, Refuse, or Other), black respondents answer an average of 2.39 questions correctly—one fewer correct answers than those who are interviewed by black interviewers. The differences are both large and statistically significant ($p = .001$). However, as with the white respondents, we find no significant difference in the number of correct answers associated with *the self-identified* race of the interviewers ($p = .528$). This finding reinforces the conclusion by Jackson, Hatchett, and Gurin (1990) that it is generally valuable to include perceived race of interviewer, not just self-reported race of interviewer in studies of race differences in attitudes.

Thus, how the interviewers classify themselves by race is not associated with different levels of performance on the political knowledge test. But the respondents' *percep-*

tion of the interviewers' race makes a great deal of difference for black respondents and no difference for white respondents. This result is highly consistent with a stereotype threat interpretation. When black respondents identify the test-giver as black, they do much better on the test than when they identify the test-giver as white or when the race of the interviewer is ambiguous, that is, the respondents are unable to put a black or white label on the interviewer.

The findings with respect to the "ambiguous" category were unanticipated in our original design. Though based on only a small number of cases, they are intriguing. They suggest that even greater anxiety may occur when black respondents are given a test by a seemingly racially "neutral" (or at least not clearly identifiable) interlocutor. In any case, we find clear support for our overall expectation of higher performance on the test when blacks were interviewed by blacks.

Controlling for respondent gender and education. The foregoing analysis does not take into account other respondent characteristics that could account for some of the race differences in performance on the political knowledge test or, conceivably, the differences in performance associated with the race of the interviewers. One threat to the validity of the findings is that they could be produced by respondent characteristics such as gender and education.

Table 5 reports the results of OLS regressing the number of correct answers (as the dependent variable) onto perceived race of the interviewer, respondent's gender, and respondent's educational attainment. We should expect to find that men and persons with higher education are more knowledgeable about politics. This is indeed what we find, both for blacks and for whites.

Even with the effects of gender and education taken into account, however, among black respondents we still find a substantial and statistically significant effect of respondent-perceived race of the interviewer on the number of correct answers to the political knowledge test. Compared to the number of correct answers that they provide to black interviewers, black respondents provide an average of .436 fewer correct answers to white interviewers and .698 fewer correct answers to interviewers with "ambiguous" race. At the same time, for white respondents we find no significant differences in the number of correct answers associated with the perceived race of the interviewer.

Controlling for interviewer-respondent rapport. Another threat to the validity of our inference that

TABLE 5 OLS Regression Analysis of Effects of Perceived Race of Interviewer and Respondent's Gender and Education on Number of Correct Answers to Political Knowledge Questions

	Black Respondents		White Respondents	
	b	T-Ratio	b	T-Ratio
Constant	2.116**	(6.105)	1.910**	(4.052)
Race of Interviewer				
Perceived by Respondent ^a				
White	-.436*	(-2.098)	-.068	(-.283)
Don't Know, Other	-.698*	(-2.456)	-.180	(-.719)
Respondent Gender				
Male	.810**	(4.069)	.839**	(4.294)
Respondent Education ^b				
College Graduate	1.874**	(4.998)	2.444**	(5.354)
Some College	1.031**	(3.026)	1.950**	(4.227)
High School	.539	(1.548)	.764	(1.661)
Adjusted R ²	.245		.272	
SEE	1.33	1.41		
Base N	208	220		

* $p \leq .05$, 2-tailed test.

** $p \leq .01$, 2-tailed test.

^aReference (omitted) category: Black.

^bOmitted (omitted) category: Less than high school education.

race-of-interviewer affects the political knowledge test performance is that something else in the relationship between respondents and interviewers is responsible for the observed patterns of responses. Conceivably, black interviewers establish greater rapport with black respondents than do white interviewers. This greater rapport might reduce the level of anxiety that respondents feel during the interview. If so, the better test performance of the black respondents interviewed by black interviewers could be due to the rapport between respondents and interviewers. However, if we can establish that the differences in test performance associated with perceived race of the interviewer hold up even after we take into account the rapport between interviewers and respondents, we would have even greater confidence in our interpretation.

At the very end of each interview the interviewers were asked to evaluate how cooperative the respondent had been as well as how much interest he or she had shown in the survey.⁷ Since large majorities of the respondents were judged to be "very cooperative" and "very interested," we

dichotomized each of the initial four-point scales into "very" and "not very" (cooperative, interested).

We find no statistically significant difference among white respondents by race-of-interviewer in the percentage of respondents who are perceived as very cooperative or very interested in the survey (the data are not shown in a table). On average, 78 percent of the white respondents were judged to be "very cooperative" and 61 percent to be "very interested." But these percentages did not vary significantly with the race of the interviewer.

Among black respondents, too, we find no significant difference by race-of-interviewer in the percent who were perceived by the interviewers as very cooperative or very interested in the survey. On average, 75 percent of the black respondents were judged "very cooperative" and 61 percent "very interested."

Because the two dichotomous variables are highly correlated with one another (Pearson's $r = .54$), we combined them to form a three-point "rapport" scale which takes on the value of 2 if the respondent was judged by the interviewer to be *both* "very cooperative" and "very interested" in the survey, 1 if the respondent was *either* very cooperative or very interested, and 0 if respondent is *neither* very cooperative nor very interested. We then created dummy variables, Hirapport which takes the value of 1 if the combined rapport score was 2, and the value

⁷"How cooperative was the respondent? Very cooperative, somewhat cooperative, not very cooperative, not at all cooperative?" "How interested was the respondent in the survey? Very interested, somewhat interested, not very interested, not at all interested?"

TABLE 6 OLS Regression Analysis of Effects of Respondent-Interviewer Rapport and the “Test Threat” Introduction

	Black Respondents		White Respondents	
	b	T-Ratio	b	T-Ratio
Constant	1.707**	(3.658)	1.468**	(2.712)
Race of Interviewer				
Perceived by Respondent ^a				
White	-.462*	(-2.186)	-.037	(-.158)
Don't Know, Other	-.761*	(-2.530)	-.166	(-.673)
Respondent Gender				
Male	.763**	(3.776)	.788**	(4.074)
Respondent Education ^b				
College Graduate	1.671**	(4.303)	2.199**	(4.850)
Some College	.792*	(2.211)	1.729**	(3.760)
High School	.328	(.902)	.611	(1.334)
Rapport ^c				
High	.466	(1.940)	.915**	(3.757)
Medium	.625*	(2.029)	.664*	(2.254)
Test Threat ^d				
“This is a test”	.165	(.874)	-.015	(-.076)
Adjusted R ²	.249		.269	
SEE	1.33	1.38		
Base N	204	217		

* $p \leq .05$, 2-tailed test.

** $p \leq .01$, 2-tailed test.

^aReference (omitted) category: Black.

^bOmitted (omitted) category: Less than high school education.

^cHigh = 1 if respondents were judged by the interviewer to be *both* “very cooperative” and “very interested” in the survey, and 0 if not. Medium = 1 if respondent is *either* very cooperative *or* very interested, and 0 if not. The omitted category: those who were neither very cooperative nor very interested.

^dDummy variable: 1 if respondents received the “this is sort of a test” introduction to the political knowledge questions, and 0 if respondents received the “this is not a test” introduction.

of 0 if not; and Mdrapport which takes the value of 1 if the combined rapport score was 1, and the value of 0 if not.⁸

When we introduce the terms Hirapport and Mdrapport into the previous regression equations (Table 6), we find not surprisingly that respondents who have high or medium level of rapport with the interviewers are likely to have offered more correct answers to the political knowledge questions. Those respondents were probably more

motivated to perform the survey tasks. Of course this relationship can also be reciprocal: interviewers were more likely to judge respondents as cooperative or interested if they took the survey tasks more seriously.

Even with both education and rapport taken into account, however, the race of interviewer effects remain among black respondents (and still do not appear among white respondents). Black respondents gave fewer correct answers to white interviewers or to those whose race was ambiguous (from the respondent's perspective) than they did to black interviewers.

⁸In the regression results reported in Table 6, this approach produced marginally better results than simply putting in the three-point rapport score. Both approaches are preferable to entering the two highly collinear “very cooperative” and “very interested” variables as separate variables in the equation. Also, combining the two extreme values (the “very cooperative” and “very interested” dummy) performed better than combining the two four-point scales.

Test anxiety. Standard introductions to questions on political knowledge are designed to reduce the potential threat or stigma associated with giving “wrong” answers. The 1985 NES pilot study (Zaller 1986) employed an introduction to such a battery of questions that explicitly

states “this is not a test of any kind.” We modeled one variant of the introduction on this approach. That variant was given to a randomly selected half of the respondents who took the political knowledge test. We also hypothesized, however, that the effect of any “stereotype threat” would be intensified if the respondents were told explicitly that the political knowledge questions were a “kind of test.” That variant of the introduction was given to the other randomly selected half of the respondents to the political knowledge questions.

We find no relationship between whether the threatening or the nonthreatening introduction was used and the number of correct answers offered to the political knowledge questions (Table 6). The coefficients for the test threat variable are small and not statistically significant for both white and black respondents. Nor are there any statistically significant interaction effects between the threat condition and other variables in the equation (not shown in the table). One reason for this result could be that even telling the respondent that “this is *not* a test” could heighten anxiety for some respondents. Because of the small sample size, we were not able to include a third variant—one in which the interviewers did not use the word “test” at all. So we do not regard our result here as definitive. Further experiments with question wording and order are warranted.

Conclusions

We find that black respondents to a battery of questions about political knowledge in a telephone survey get fewer answers right when interviewed by a white interviewer than when interviewed by a black interviewer. These results are consistent with expectations based on the theory of “stereotype threat” that has been developed and applied to account for performance on standardized achievement and intelligence tests as well as athletic performance.

The observed differences in performance on the political knowledge questions cannot be accounted for by differences in the educational background or gender of the individual respondents. Nor can the higher scores achieved by black respondents who are interviewed by blacks be accounted for by greater rapport between respondents and interviewers. Among both black and white respondents, the level of rapport does not differ significantly between those who were interviewed by black interviewers and those who were interviewed by interviewers from other racial groups.

Although we can rule out the respondents’ education and gender, as well as respondent-interviewer rapport, as plausible rival explanations of the differences in the

political knowledge test scores, we cannot say for sure that “stereotype anxiety” accounts for the differences. But the results are interesting in part because they suggest that research that heretofore has focused on standardized tests or on experiments with relatively low *N*s, can be extended to “tests” of factual information in telephone surveys and to much larger samples in which it is possible to control explicitly for a variety of potential explanatory factors.

Furthermore, these results suggest another way to think about the race-of-interviewer effects in social surveys. To the extent that minority group respondents regard their answers to survey questions as tests—whether the questions address factual issues or issues of attitudes or beliefs—they may be susceptible to anxiety produced by their role as subjects to a process in which they are at risk of being judged as giving “right” or “wrong” answers. In typical accounts of race-of-interviewer effects, respondents are said to mask their true feelings in an effort to please the interviewer or to appear to hold socially desirable attitudes. In some accounts, the respondents may exaggerate their conformity with socially approved norms and may even modify their subsequent behavior to fit the norms.

However, in the present study, in which the task set before the respondents is to tell the interviewers what they know—and in which the interviewers (the survey researchers)—have an external standard for determining whether the answers are correct or incorrect, it is not possible for respondents (on average) to provide correct answers to questions to which they do not know the answers. Most respondents cannot make themselves look more knowledgeable than they actually are (except for those who may guess some correct answers). But it is possible for the respondents to appear to be uninformed or ill-informed. And respondents who belong to racial minorities may experience added anxiety when they risk being uninformed in the presence of a member of the dominant racial group.

The respondents were, of course, assured confidentiality and told also that they did not have to answer every question. They were not under the type of pressure to perform that they might experience if they had been given a large standardized test of knowledge or achievement. Also, as cooperative respondents committed to the task, few of the respondents who did not know the answers tried to *avoid* giving wrong answers by refusing to answer the questions. Instead, they mostly just gave more wrong answers. But black respondents were more likely to give wrong answers—to questions to which in some cases they probably knew the answers—when the questioner was from a different racial group than their own. And

this, we suggest, looks like the consequences of stereotype anxiety.

In future research, we plan to expand the types of tests and to experiment with a variety of test conditions to see whether we can replicate the results. One consideration is that the respondents in this survey may have been “race primed” (Steele and Aronson 1998) because the survey asked the respondents to identify their race before it asked the political knowledge questions.⁹ This may have increased the race sensitivity of the respondents to stereotype threat. This does not mean that the effects that we have observed are wrong, but it suggests the need to examine the effects of question order and content on the results.

References

- Anderson, Barbara A., and Brian D. Silver. 1986. “Measurement and Mismeasurement of the Validity of the Self-Reported Vote.” *American Journal of Political Science* 31(4):771–85.
- Anderson, Barbara A., Brian D. Silver, and Paul R. Abramson. 1988a. “The Effects of Race of the Interviewer on Measures of Electoral Participation by Blacks in SRC National Election Studies.” *Public Opinion Quarterly* 52(1):53–83.
- Anderson, Barbara A., Brian D. Silver, and Paul R. Abramson. 1988b. “The Effects of the Race of the Interviewer on Race-Related Attitudes of Black Respondents in SRC/CPS National Election Studies.” *Public Opinion Quarterly* 52(3):289–324.
- Aronson, Joshua, Lustina, M.J., Good, C., Keough, K., Steele, C.M., and Brown, J. 1999. “When White Men Can’t Do Math: Necessary and Sufficient Factors in Stereotype Threat.” *Journal of Experimental Social Psychology* 35(1):29–46.
- Aronson, Joshua, Claude M. Steele, M.F. Salinas, and M.J. Lustina. 8th ed. “The Effects of Stereotype Threat on the Standardized Test Performance of College Students.” In *Readings About the Social Animal*, ed. E. Aronson. New York: Freeman.
- Asher, Herbert. 1998. *Polling and the Public*, 4th ed. Washington, DC: CQ Press.
- Blascovich, Jim, Wendy B. Mendes, Sarah B. Hunter, Brian Lickel, and Neneh Kowai-Bells. 2001. “Perceiver Threat in Social Interactions with Stigmatized Others.” *Journal of Personality and Social Psychology* 80(2):253–67.
- Croizet, Jean-Claude, and Theresa Claire. 1998. “Extending the Concept of Stereotype Threat to Social Class: The Intellectual Underperformance of Students from Low Socioeconomic Backgrounds.” *Personality and Social Psychology Bulletin* 24(6):588–94.
- Danso, Henry A., and Victoria M. Esses. 2000. “Black Experimenters and the Intellectual Test Performance of White Participants: The Tables Are Turned.” *Journal of Experimental Social Psychology* 37(2):158–65.
- Davis, Darren W. 1997. “The Direction of Race of Interviewer Effects Among African-Americans: Donning the Black Mask.” *American Journal of Political Science* 41(1):309–22.
- Delli Carpini, Michael X., and Scott Keeter. 1993. “Measuring Political Knowledge: Putting First Things First.” *American Journal of Political Science* 37(4):1179–1206.
- Delli Carpini, Michael X., and Scott Keeter. 1996. *What Americans Know About Politics and Why It Matters*. New Haven: Yale University Press.
- Finkel, Steve E., Thomas M. Guterbock, and Marian J. Borg. 1999. “Race-of-Interviewer Effects in a Preelection Poll: Virginia 1989.” *Public Opinion Quarterly* 55(3):313–30.
- Goffman, Erving. 1963. *Behavior in Public Places: Notes on the Social Organization of Gatherings*. New York: Free Press of Glencoe.
- Goffman, Erving. 1973. *The Presentation of Self in Everyday Life*. Woodstock, NY: Overlook Press.
- Hembroff, Larry A., and Brian D. Silver. 2001. *State of the State Survey: SOSS 20, Methodological Report*. Institute for Public Policy and Social Research, Michigan State University.
- Iyengar, Shanto. 1986. “Whither Political Information.” Report to the National Election Studies Board of Overseers, Center for Political Studies, University of Michigan.
- Jackson, James S., Shirley Hatchett, and Patricia Gurin. 1990. *Hope and Independence: Blacks’ Response to Electoral and Party Politics*. New York: Russell Sage.
- Jennings, M. Kent. 1996. “Political Knowledge Over Time and Across Generations,” *Public Opinion Quarterly* 60(2):228–52.
- Kinder, Donald R., and David O. Sears. 1981. “Prejudice and Politics: Symbolic Racism Versus Racial Threats to the Good Life.” *Journal of Personality and Social Psychology* 40(3):414–31.
- Marx, D., Brown, J., and C.M. Steele. 1999. “Allport’s Legacy and the Situational Press of Stereotypes.” *Journal of Social Issues* 55(3):491–502.
- McConahay, John B., and J.C. Hough. 1976. “Symbolic Racism.” *Journal of Social Issues* 32(2):23–45.
- Mondak, Jeffery J. 2001. “Developing Valid Knowledge Scales.” *American Journal of Political Science* 45(1):224–38.
- Oyserman, Daphna, and Janet K. Swim. 2001. “Stigma: An Insider’s View.” *Journal of Social Issues* 57(1):1–14.
- Quinn, Diane M., and Steven J. Spencer. 2001. “The Interference of Stereotype Threat with Women’s Generation of Mathematical Problem-Solving Strategies.” *Journal of Social Issues* 57(1):55–71.
- Schuman, Howard, and Jean M. Converse. 1971. “The Effects of Black and White Interviewers on Black Responses in 1968.” *Public Opinion Quarterly* 35(1):44–68.
- Sears, David O. 1988. “Symbolic Racism.” In *Eliminating Racism: Profiles in Controversy*, ed. P.A. Katz and D.A. Taylor. New York: Plenum, 53–84.
- Silver, Brian D., Barbara A. Anderson, and Paul R. Abramson. 1986. “Who Overreports Voting?” *American Political Science Review* 80(2):613–24.

⁹As noted earlier, the respondents’ self-identified race was used to determine which respondents would be asked the political knowledge questions—all self-identified African Americans and a random one-third of self-identified whites.

- Spencer, S.J., C.M. Steele, and D.M. Quinn. 1999. "Stereotype Threat and Women's Math Performance." *Journal of Experimental Social Psychology* 35(1):4–28.
- Stangor, Charles, Christine Carr, and Lisa Kiang. 1998. "Activating Stereotypes Undermines Task Performance Expectations." *Journal of Personality and Social Psychology* 75(5):1191–97.
- Steele, Claude M. 1997. "A Threat in the Air: How Stereotypes Shape Intellectual Identity and Performance." *American Psychologist* 52(6):613–29.
- Steele, Claude M. 1999. "Thin Ice: 'Stereotype Threat' and Black College Students." *The Atlantic Monthly* 284(2):44–54.
- Steele, Claude M., and Joshua Aronson. 1995. "Stereotype Threat and the Intellectual Test Performance of African-Americans." *Journal of Personality and Social Psychology* 69(5):97–111.
- Steele, Claude M., and Joshua Aronson. 1998. "How Stereotypes Influence the Test Performance of Academically Successful African Americans." In *The Black-White Test Score Gap*, ed. Christopher Jencks and Meredith Phillips. Washington, DC: Brookings Institution, 401–27.
- Stone, Jeff, Christian I. Lynch, Mike Sjomeling, and John M. Darley. 1999. "Stereotype Threat Effects on Black and White Athletic Performance." *Journal of Personality and Social Psychology* 77(6):1213–27.
- Walsh, Margaret, Crystal Hickey, and Jim Duffy. 1999. "Influence of Item Content and Stereotype Situation on Gender Difference in Mathematical Problem Solving." *Sex Role* 41(3–4):219–40.
- Zaller, John. 1986. "Analysis of Information Items on the 1985 NES Pilot Study" Report to the National Election Studies Board of Overseers, Center for Political Studies, University of Michigan.
- Zeller, Richard A., and Edward G. Carmines. 1980. *Measurement in the Social Sciences: The Link Between Theory and Data*. New York: Cambridge University Press.