Hidden Racial Prejudice?: Impact of Social Desirability

Pressures on Endorsement of Racial Stereotypes

April 14, 2015

Word Count: 6258

Abstract

Far fewer white Americans report anti-black views in oral face-to-face survey interviews today than did sixty years ago. Is it because what whites believe about blacks has changed? Or is it because the social norms against expressing anti-black views have grown stronger? To estimate the extent to which social desirability pressures have led whites to hide anti-black views in oral face-to-face interviews, we compared oral reports to confidential reports obtained using two different methods: Audio Computer Assisted Self Interviewing (ACASI), and online self-complete surveys. Both, answering via ACASI, and online led to small increases in percent reporting negative views of blacks. Confidential reports, however, exhibited no greater convergent validity than oral reports. Both the findings are partly explained by the fact that a large majority of white respondents interviewed orally were interviewed by white interviewers. Results bring further nuance to our understanding of perceived social norms about expression of negative stereotypes of blacks in America.

Proportion of white Americans reporting prejudicial views about blacks has declined sharply over the last six decades (Bobo and Dawson 2009; Jaynes and Williams 1989; Kluegel 1990; Schuman 1997). For instance, between 1942 and 1972, the percent of whites who said that African Americans "should have as good a chance as white people to get any job" rose from 45% to 97% (p. 104, Schuman 1997). Similarly large majorities of white Americans in recent years have said that they oppose anti-miscegenation laws, oppose residential segregation, and favor equal treatment of blacks and whites (Bobo 2001; Schuman 1997).

One way to interpret the declines in reports of prejudicial views is as evidence of rising racial liberalism. However, many scholars view the declines far more pessimistically: as artifacts of declining acceptability of expressing racist attitudes socially; racist beliefs haven't declined but people are increasingly unwilling to admit them during face-to-face conversations with survey interviewers (Corstange 2009; Dovidio and Fazio 1992; Feagin and Sikes 1995). Such concerns have led some scholars to question the wisdom of asking explicit racism questions in face-to-face oral interviews. For example, Corstange (2009) asks, "whether or not it is worth all the trouble to administer such surveys in the first place" (p. 46).

In this paper we explore whether such concerns are merited. We test whether administering explicit racism measures in ways that substantially reduce social desirability pressures cause large increases in proportion of people expressing prejudicial views. A variety of techniques - mostly aimed at preserving the confidentiality of the response - have been used to reduce social desirability pressures, including the item count technique (Holbrook and Krosnick 2010; Kuklinski et al. 1997; Miller 1984; Miller et al. 1986)[2], the randomized response technique (Warner 1965), and the bogus pipeline technique (Jones and Sigall 1971; Sigall and Page 1971). In the research reported here, we relied on two other techniques that allow respondents to respond confidentially: Audio Computer Assisted Self-Interviewing (ACASI), and online self-complete surveys. Both techniques yield small increases in percentage of whites reporting negative views about blacks. We also find that oral measures exhibit no lower convergent validity than confidential measures. Part of the explanation for both findings lies in the following twin facts: a) a vast majority of the oral interviews of white respondents were conducted by white interviewers, and b) most whites who hold aversive beliefs

---

[2]The technique is better known as the *list experiment* in political science (Kuklinski et al. 1997).

about blacks do not feel pressured to hide those beliefs in front of white interviewers.

We begin by discussing the relevant theory and evidence. We follow it with description of the research design, data, and our findings. We end by discussing implications of our findings for the study of racism, the current norms for expressing prejudicial views against blacks, and our understanding of prevalence of prejudicial beliefs in society.

Social Desirability Pressures On Racial Measures

In social interactions, people are thought to strive to present themselves in a favorable light (Goffman 1959). As a result, people are thought to refrain from reporting potentially embarrassing opinions and behaviors, a tendency referred to as impression management social desirability bias (Edwards 1957; Paulhus 1984, 1986, 1989, 1991). Such censoring depends on how embarrassing the person thinks the opinion is, the extent to which responses can be observed by others, expectations about the type and extent of observer's reactions to the responses, and the extent to which judgment of the observer matters to the person.

If responses are confidential, all other considerations become moot; ACASI and anonymous self-complete online surveys exploit this particular insight. By the same token, if the respondent doesn't think that the responses are embarrassing, subsequent considerations become moot. Predictions of social desirability pressures on racial measures are predicated on the notion that people who hold aversive beliefs about other racial groups think that these beliefs are prejudiced, and out of the ordinary. Given few, if any, prejudiced people are liable to see themselves as prejudiced, it is likely that many of those who hold aversive beliefs think such beliefs are unprejudiced, and that most people hold similar views. (Such kind of projection, also known as the 'false consensus effect', has been shown to be common (Ross 1977).)

However, it is also likely that those who hold aversive beliefs about other racial groups are aware that reporting these beliefs will not go down well with certain kinds of people, such as those who (they think) have vested interests in not seeing the truth, or, at any rate, discussing it openly. In a survey interview, respondent's expectations about whether the interviewer is that kind of a person is likely based on two sets of cues: preceding interaction with the interviewer, and readily observed characteristics of the interviewer that the respondent can use to conjecture interviewer's

reaction. Interaction with trained interviewers likely leads most respondents to expect a tepid or even agreeable response to whatever they say. On the other hand, it is likely that people would hesitate in saying something harsh about the racial group to which the interviewer belongs to, especially if the interviewer is of a different race than them (Hatchett and Schuman 1975).

Lastly, given the respondents' relation with an interviewer is typically short-lived and superficial, it is unlikely that interviewer's judgments matter especially to the respondent. Hence, it seems unlikely that respondents would go through the trouble of adjusting their reports to suit what they think the interviewer wants to hear. In all, we expect the magnitude of social desirability bias in racial measures in oral face-to-face interviews to be small.[3] We also expect the bias to be somewhat larger when people are reporting these attitudes to interviewers of a different race than interviewers of the same race.

There is a good deal of evidence that is consistent with such predictions. In a study conducted over the telephone during the 1989 Virginia gubernatorial race between Douglas Wilder, an African-American, and Marshall Coleman, a Caucasian, Finkel et al. (1991) found that a majority of respondents (52%) said they supported Wilder over Coleman when interviewed by African-Americans, but only 44% of those interviewed by whites reporting supporting Wilder. (For a yet more circumscribed theory, and even smaller estimates of the effect size, see Hopkins (2009).) Elsewhere, studies that manipulated privacy of response have found that the social desirability pressures on racial measures are neither consistent across measures (Krysan 1998) nor always substantial (Krysan 1998; Krysan and Couper 2003), and that it is the blatant measures of racism that are more impervious to social context (Krysan 1998).

However, a few studies have found that the social desirability bias is sizable. In a study in which undergraduates were randomly assigned to answer questions orally, or under a 'bogus pipeline', a procedure designed to convey to the respondent that a mechanism exists for the researcher to know the respondent's 'true' attitude, participants said various derogatory attributes were far more characteristic of "Negroes" in the bogus pipeline condition than in the oral interview condition (Sigall and Page 1971). But, as Sigall and Page (1971) note, it is possible that respondents favor "negative" errors under bogus pipeline condition —"it may be better to admit being a bigot and

---

[3]We use the word 'small' to characterize the numerical magnitude on an absolute scale, and not to convey that the effect is inconsequential.

have it shown that you are fairly liberal than to claim tolerance while being revealed as a bigot" (Jones and Sigall 1971). Hence, the estimates of social desirability bias may be upwardly biased. It is also likely that undergraduates, who were the subjects used in the study, were especially prone to social desirability pressures at a time when racial issues were being discussed forcefully on college campuses. Elsewhere, questions have been raised about validity of results obtained from some unobtrusive techniques. For instance, Sears (2004) notes that figures such as "98% white Southerners expressed covert anger about affirmative action in the early 1990s" (Kuklinski et al. 1997) raise questions about the validity of the measure. In all, some research does indeed purport to show that extensive social desirability pressures exist on survey reports of racial beliefs and attitudes. However, weaknesses in research design, use of non-representative samples, and the era in which some of these studies were conducted, raise questions about the validity and generalizability of the results. In all, we expect the magnitude of the social desirability bias to be small.

Hitherto, we have focused on the impact of impression management social desirability bias on 'descriptive validity' of racial measures. Next, we discuss its impact on their correlational validity. Impact on correlational validity depends on the structure of measurement error. For instance, if everyone hesitated to a similar degree in reporting aversive beliefs, correlation with 'true measure' would remain unaffected; an intercept shift leaves correlations unchanged. On the other hand, if the people who hold the most aversive beliefs about blacks were the ones who were most likely to censor their comments, correlations with the 'true measure', and thus, criteria, would be lower. However, it seems unlikely that such a scenario would come to pass given those who hold the most aversive beliefs often seem most eager to publicly declaim such beliefs.

In all, we posit a small downward bias in white respondents' reports of racial prejudice. We expect the bias to be greater when the interviewer is of a different race than the respondent than when the interviewer is of the same race as the respondent. We also posit that the correlational validity of oral racial measures will be as high as that of confidential measures.

The Present Investigation

In this paper, we assess whether social desirability pressures distort measurement of racial stereotypes in oral face-to-face interviews. Racial stereotype measures, which date back to at least

Katz and Braly (1933), have proven to be valuable in the study of racial attitudes, especially so because they are thought to be less confounded with political ideology than other measures (p. 429, Huddy and Feldman 2009). Studies suggest that racial stereotypes are consequential: they predicted support for segregationist policies such as anti-miscegenation laws, opposition to general government assistance to blacks, and opposition to opportunity-enhancing programs such as enterprise zones (Gilens 1998; Hurwitz and Peffley 1997; Peffley et al. 1997; Sniderman and Piazza 1993). Furthermore, whites who endorsed negative stereotypes of blacks were substantially less likely to support welfare, and substantially more likely to support police search for blacks than similarly described whites (Peffley et al. 1997).

To explore the degree to which racial stereotype measures are compromised by social desirability pressures, we conducted two studies. In the first study, we assessed whether rising social pressures to hide racism have led to a decline in proportion of white Americans endorsing negative stereotypes of blacks in oral face-to-face interviews between 1992 and 2004. In the second study, we exploited data from two different sources: a within-subjects experiment in which respondents reported racial stereotypes using ACASI in one interview and orally face-to-face in another interview, and an online panel study conducted over a national sample which fielded two waves at about the same time as the first study. We used the two sets of data to estimate social desirability bias in oral reports of racial stereotypes in three different ways - a) a within-subjects analysis of differences across modes (and across waves), b) a difference-in-difference analysis to account for any confounding changes across waves, and c) a between-subjects comparison of responses from the oral interview and the contemporaneously fielded online self-complete survey. We pair these analyses with various robustness checks. We follow the analyses of descriptive validity with analyses of correlational validity. We conclude with an analysis of impact of race of the interviewer.

Trends in Oral face-to-Face Reports of Stereotypes

Concern about social desirability is predicated on the notion that people know of, and heed to the social norm that it is undesirable to express prejudicial judgments (Corstange 2009; Kuklinski et al. 1997). Based on this reasoning, we should expect reports of prejudicial stereotypes in oral interviews to be quite rare in recent times. If instead, expressions of prejudicial judgments have

been common, it would challenge the notion that a strong norm against expressing prejudicial views of blacks pervades society.

To assess how common reports of anti-black racial stereotypes have been in oral face-to-face interviews, we used data from face-to-face oral interviews from the 1992, 1996, 2000, and 2004 American National Election Studies (ANES) Time Series Surveys. (See Appendix A for details about sampling and measures.) In particular, we analyzed trait ratings of whites and blacks on hardwork and intelligence by self-identified white non-Hispanic respondents.

Results

In 1992, 58.87% percent of the white respondents rated whites as more hardworking than blacks. The percent of white respondents rating whites as more hardworking than blacks dropped to 52.15% in 1996, then rose to 59.10% in 2000, and then dropped again, to 48.55%, in 2004 (see Table 1). The same seesaw pattern appeared in ratings of intelligence. In 1992, 52.05% of whites rated whites as more intelligent than blacks, 45.11% did so in 1996, 49.13% in 2000, and 40.71% in 2004.

(Insert Table 1 here)

Thus, even as of 2004, a large proportion of white Americans were willing to tell an interviewer in an oral face-to-face interview that they viewed blacks as less hardworking and less intelligent than whites. This challenges the notion that a strong norm against endorsing anti-black stereotypes in social settings pervades America.

Comparing Oral Reports to Anonymous Reports

Next, we used ACASI and online self-complete surveys to explore the extent to which percentages reported in Study 1 are downwardly biased due to social desirability pressures in oral face-to-face interviews.

In the 2008 ANES (See Appendix A for details about sampling), respondents were asked the same questions examined in Study 1 via ACASI during the pre-election interview, and orally during the post-election interview. The ACASI administration was conducted as follows: the interviewers gave the laptop that they had been using to the respondent, the respondent put on the headphones, and the computer played an audio recording of the questions and answer choices; the text of the question along with the response options was displayed on the screen. Respondents answered

7

privately by typing on the laptop. In the oral interview, respondents were given a card showing them the rating scale and asked to respond orally as to where they placed each racial group.

A great deal of research shows that people are more likely to admit to engaging in potentially embarrassing behaviors and suffering from sensitive conditions when interviewed via ACASI than when interviewed orally (e.g. Metzger et al. 2000; Newman et al. 2002; Tourangeau and Smith 1996; Turner et al. 1998). For example, when interviewed using ACASI, more people admitted to using illegal drugs, engaging in risky sexual behaviors, being HIV positive, and having a sexually transmitted disease, than when asked orally (Des Jarlais et al. 1999; Gribble et al. 2000; Villarroel et al. 2008). So we start by exploring whether moving to ACASI induced more white respondents to offer negative assessments of blacks vis-à-vis whites. We do so by comparing the percent of whites rating whites above blacks across the two modes. From relative assessments of blacks and whites, we next move to absolute assessments of blacks and whites. We compare changes in the percent of whites rating whites and blacks above and below the midpoint in the oral and confidential interviews. Lastly, we compare difference in difference ratings of whites and blacks in the ACASI and the oral interviews.

We follow these analyses with three robustness checks. First, we investigate whether curtailing analyses to respondents who offered substantive response at both times biases the results; if missingness is correlated with potential outcomes, estimates are liable to be biased. Second, we discuss how election of a black president between the two measurement waves may have affected our results, and present a difference-in-difference analysis that speaks to the concern. Third, we assess whether another way of reducing social desirability pressures by providing confidentiality —online self-complete surveys —gives us very different estimates of social desirability bias on oral measures. After these robustness tests, we estimate the impact of providing confidentiality on the convergent and concurrent validity of the measures. Lastly, we investigate whether our findings are driven by vast overlap between interviewer and respondent race.

Results

The results presented below are based on data from white non-Hispanic respondents who gave a substantive response to the stereotype questions on both the oral and ACASI interview. Analyses were done with weights on and while accounting for clustering due to the complex survey design.

Descriptive Validity

The percent of whites rating whites more favorably than blacks on hardwork increased by about 4 percentage points in the confidential interview, from 45.10% in the oral interview to 49.49% in the ACASI administration ($p < .01$; see Table 2). A similar pattern appeared in ratings of intelligence - the percent of whites rating whites more favorably than blacks rose by about 5 percentage points in the confidential interview, from 39.80% in the oral interview to 44.54% in the ACASI administration ($p < .001$). [4]

(Insert Table 2 here)

These estimates, however, may be biased for two reasons. For one, the estimates are based on data from non-Hispanic whites who gave a substantive response on both the oral and the ACASI interview and list-wise delete missing data. For two, the estimates of the differences across modes may be confounded by other between wave changes, foremost among them, election of a black president. Here below we discuss how we address the two concerns:

Attrition and Non-Response: One natural concern about ignoring missing data is about selection bias —the difference in negative stereotype reports between the oral and the ACASI interview may be different among those who did not answer the stereotype questions on at least one of the interviews, compared to those who offered a substantive response at both time points. Given data are missing, we cannot directly assess this claim. Hence, we estimate the impact of missing data on our estimates under the assumption that missing data are ignorable conditional on covariates. We find the bias to be small, about .1%. (See Appendix C for details of how much data were missing, and the method we used to estimate impact of missing data.)

Accounting for Mr. Obama's Election: Given that a black president was elected between the ACASI and oral administrations, differences between responses obtained under the two modes reflect not just mode differences, but also the impact (if any) of the election. As we discuss below, we expect the election to have had only a minimal impact on the stereotypes of blacks. We also expect any impact to be positive, with people less likely to hold negative stereotypes of blacks after the election than before it. We start by discussing the reasons behind our expectations, and then

---

[4]Alternate conceptualizations of the dependent variable - a continuous difference score - . See Appendix B.

go on and present Difference-in-Difference analyses that corroborate our expectations.

Stereotype measures deal with people's beliefs about traits common to different racial groups. These beliefs have been shown to dramatically shape a person's cognition —information that is inconsistent is often ignored, or, worse yet, processed in ways that confirms prior beliefs (Jost and Hamilton p. 210, 2005; Hamilton and Sherman p. 31, 1994). Such biases in processing attenuate impact of any discomformatory information; this shouldn't be surprising —stereotypes are often a result of ignoring plentiful conflicting data. Hence we expect stereotypes of blacks to change only minimally due to the election.

Next, we must posit the direction of change. There are at least two reasons to think that election of a black president would, if anything, reduce negative stereotypes of blacks —(1) People saw Obama as a black exemplar and updated their beliefs about the racial group he represents (Goldman 2012), (2) People rallied around the new (Black) president and updated their beliefs about the racial group to which he belonged, or updated their beliefs after seeing that a majority of country thought differently, or both. Research confirms such predictions —there was indeed a minor reduction in negative stereotypes of blacks reported by whites between the period two months before the election till the election, and the period starting immediately after the election till three months after (Goldman 2012). If people came to hold more positive stereotypes of blacks, the ACASI-oral difference in negative stereotype reports of blacks is actually an upper bound of the potential increase in negative stereotype reports between oral and ACASI; the true percentage of people hiding negative stereotypes in the oral interview is likely to be lower.

To estimate the bias in ACASI-oral differences in negative black stereotypes due to the election of a black president, we used a Difference-in-Difference design. The intuition behind the design is as follows: ANES estimate is a sum of effect of mode and Obama's election. To estimate the effect of mode alone, we need to subtract out the effect of Obama's election. Effect of Obama's election can be estimated using a panel in which mode remains the same across waves. (To guard against any complex interactions between social desirability pressures and effect of Obama's election, the mode of interview for both waves should preferably be a confidential interview.) This estimate can then simply be subtracted from changes across waves (modes) in the ANES.

For data, we turn to the 2008 National Annenberg Election Study (NAES) Online Panel, which included stereotype questions in three waves. (See Appendix A for details about sampling and

measures used.) In particular, we focus on data from two waves that were fielded roughly at the same time as the ANES waves: Wave 4 (August 29 to November 4, 2008) and Wave 5 (November 5, 2008 to January 31, 2009). Respondents were asked to rate their own ethnic group, and one other ethnic group on hardwork and intelligence on a semantic 101 point scale, ranging from "extremely lazy/unintelligent" to "extremely hardworking/intelligent." As before, we recoded these ratings as dummy variables indicating whether or not the respondent had rated whites higher than blacks. And once again, we only analyze data from non-Hispanic white respondents who answered both Wave 4 and 5.

Between Waves 4 and 5, percentage of white respondents rating whites as more hardworking than blacks decreased by 3.44%, from 50.61% to 47.17% (see Table 3). During the same time, percentage of white respondents rating whites as more intelligent than blacks decreased by 1.83%, from 46.88% to 45.05%. Hence, as expected, the difference in difference estimates were sharply lower than the naive estimates; the difference in difference estimate for ratings of hardwork was .96%, and 2.91% for ratings of intelligence. These results suggest that the results reported in Piston (2010) are upwardly biased.

(Insert Table 3 here)

Comparison with Online Self-Complete Surveys

Turning to ACASI for a few questions in a face-to-face interview may not be as effective in reducing social desirability biases as a self-complete interview with no interviewer present. To respond of this conjecture, we turn once again to data from the 2008 NAES Online Panel Study. In particular, we compare data from Wave 5 of the NAES (unmonitored self-complete interviews) with data from (contemporaneous) post-election oral face-to-face NES survey to estimate impact of social desirability pressures on racial stereotypes using an alternate technique to reduce social desirability bias. Using this method, we find estimate of social desirability pressures on ratings of hardwork to be 2.07%, and on ratings of intelligence to be 7.08%. In all, evidence suggests that social desirability pressures led only a few respondents to hide prejudicial views of blacks in oral face-to-face interviews.

Concurrent and Convergent Validity

Until now we have focused on descriptive validity of oral reports of racial stereotypes. Now we turn our attention to assessing their concurrent and convergent validity. For reasons we discuss above, even if oral measures may not describe absolute levels of negative racial stereotyping well, they may still prove useful as correlates. To assess concurrent and convergent validity, we use some well established correlates of negative racial stereotypes that are unlikely to be affected by social desirability pressures.

A great deal of research shows that anti-black beliefs among whites vary by region, level of education, and certain dispositional characteristics. The more educated typically hold more egalitarian views on race than the less educated (Greeley and Sheatsley 1971; Case et al. 1989; Wagner and Zick 1995; Heerwig and McCabe 2009). While whites living in the South typically hold less egalitarian views than those living elsewhere (Kuklinski et al. 1997; Black 1987). And those with an authoritarian disposition hold negative stereotypes of other ethnic groups more often than those who aren't authoritarian (Cribbs and Austin 2011; Whitley Jr 1999; Feldman and Stenner 1997). We use these well corroborated correlates to test the concurrent validity of stereotype measures obtained using the two modes.

If oral measures are less valid than confidential measures, regressing oral measures on well known correlates ought to yield smaller coefficients. We assessed whether there was a greater divergence between Southern whites and whites from elsewhere in the US on ACASI measures than on oral measures by regressing difference in stereotype ratings of whites and blacks obtained orally and confidentially on an indicator variable for South. Coefficients for the oral measures were as large as they were for the confidential measures ($b = .02; p < .01$; see Table 4).

Next, to assess whether oral stereotype measures were more weakly correlated with education than confidential measures, we regressed difference in rating of whites and blacks on stereotype measures on education. Oral stereotype ratings were as strongly related to education as confidential ratings confidentially obtained ratings (Hardwork: $b_{\text{ACASI}} = -.035, b_{\text{ORAL}} = -.04$, Diff.$_{\text{ORAL-ACASI}}$ = -.01, $p_{\text{ORAL-ACASI}} = .44$; Intelligence: $b_{\text{ACASI}} = -.05, b_{\text{ORAL}} = -.06$, Diff.$_{\text{ORAL-ACASI}} = -.01$, $p_{\text{ORAL-ACASI}} = .46$).

(Insert Table 4 here)

12

Our last piece of evidence for concurrent validity comes from relationship between oral and confidential stereotypes to authoritarianism. Regressing difference in white and black stereotype ratings of hardwork obtained using the oral interview on authoritarianism yielded a coefficient that was no different from the coefficient we got when we regressed the confidential stereotype ratings on authoritarianism (Hardwork: $b_{\text{ACASI}} = .10, b_{\text{ORAL}} = .10$, Diff.$_{\text{ORAL-ACASI}} = .001$, $p_{\text{ORAL-ACASI}} = .63$). Lastly when we regressed the intelligence ratings obtained orally and confidentially, the coefficient for oral stereotype ratings was slightly higher than coefficient for confidential ratings of intelligence (Hardwork: $b_{\text{ACASI}} = .11, b_{\text{ORAL}} = .09$, Diff.$_{\text{ORAL-ACASI}} = .02$, $p_{\text{ORAL-ACASI}} = .08$).

To assess the congruent validity of oral measures, we checked how the stereotype measures obtained using the two different modes related to a measure of racism gotten using a different procedure shown to be largely immune to social desirability pressures —the Affect Misattribution Procedure or AMP (see Payne et al. 2005). (See Appendix A for description of the measure.) Regressing oral reports of hardwork on AMP yielded a coefficient that was as large as one obtained when we regressed confidential reports of stereotype ($b_{\text{ORAL}} = .10, p < .001; b_{\text{ACASI}} = .10, p < .001$; Diff.$_{\text{ORAL-ACASI}} = .00$, $n.s.$). On stereotypes about intelligence, regression coefficient on AMP was slightly stronger for oral reports than for confidential reports ($b_{\text{ORAL}} = .11, p < .001; b_{\text{ACASI}} = .09, p < .001$; Diff.$_{\text{ORAL-ACASI}} = .02$, $p = .08$). In all, these results strongly suggest that oral ratings are no less valid than confidential ratings.

Race of the Interviewer

Lastly, we turn our attention to one potential explanation for our findings. In particular, we investigate whether the reason why we find only small amounts of hidden negative stereotypes of blacks is the correspondence between respondent and interviewer race in a vast majority of the oral interviews.

As we note above, whites may hesitate less in reporting anti-black stereotypes to a white interviewer than to an interviewer of another race. Add to this the fact that over 80% of white respondents have been interviewed by white interviewers in the National Election Studies and General Social Surveys for which interviewer race data are available; in the 2008 NES, the figure was 84.7%. To test this possibility, we tallied how ratings changed among respondents interviewed by

interviewers of the same race as that of the respondent, vis-à-vis those interviewed by interviewers of a different race.

Of the respondents interviewed by white interviewers (n = 893), and who rated whites higher than blacks on hardwork via ACASI, 68.92% persisted with their ratings in the oral interview. The parallel figure for those interviewed by non-white interviewers (n = 139) was 59.83%. On ratings of intelligence, among those interviewed by white interviewers, 65.90% of those rating whites higher than blacks via ACASI persisted with their rating in the oral interview. The corresponding figure for those interviewed by non-white interviewers was 52.66%. So there is some suggestive evidence that there was more hesitation to report negative stereotypes of blacks when interviewed by non-white interviewers.

Since interviewers were not randomly assigned to respondents, comparisons between raw estimates of hidden prejudice during conversations with white and non-white interviewers are potentially biased due to confounding differences between the two groups. To account for these differences, we control for some covariates that are expected to be related to the extent of hidden negative stereotypes in the oral interview. Social desirability concerns apply only to those who have something to hide. More generally, it is likely that those who think blacks are less intelligent than whites have more to hide in the oral interview than those who believe that blacks are as intelligent as whites. It is possible that race of interviewer was correlated with true level of negative stereotypes of blacks (as measured via ACASI). So we control for trait ratings obtained under ACASI. It is also possible that race of interviewer was correlated with how susceptible individuals were to social desirability pressures. One correlate of susceptibility to social desirability pressures around race is education (Krysan 1998; Hatchett and Schuman 1975). So we controlled for respondent level of education.[5]

We estimated a generalized hierarchical model modeling predicting whether or not a respondent endorsed a negative stereotype of blacks in the oral interview as a function of race of the interviewer (coded: 1= non-Hispanic White, 0 = all else), education (coded in four categories - High school or less, Some College, College, More than College), age (in years; rescaled to range from 0 and 1),

---

[5]Controlling for a variety of other observed characteristics of the respondent, including income, whether the respondent lives in an urban or a rural area, census region, etc. leaves the coefficient for race of interviewer virtually unchanged.

and stereotype reports obtained using ACASI. We included random effects for primary sampling unit and each interviewer, and fixed effects for strata. Results suggest that being interviewed by a non-Hispanic white interviewer increases the probability of reporting blacks as less intelligent than whites in the oral interview (b= .84, $p < .01$), and very likely marginally increases the probability of reporting blacks as less hardworking (b = .39, $p = .11$) (see Table 4).

(Insert Table 5 here)

More concretely, change to a non-white interviewer reduces the percent of respondents willing to report blacks are less intelligent than whites in the oral interview by 14%. Corresponding figure for hardwork is 8.5%. One lesson is immediately clear —negative stereotypes of blacks on hardwork are likely thought off by respondents to be more socially acceptable, than negative stereotypes of intelligence. This is also backed by the fact that larger percentage of white respondents endorses negative stereotypes of blacks on hardwork than on intelligence. In all, it seems that being interviewed by interviewers of the same race partly explains why the declines in percent reporting negative stereotypes between the ACASI and oral interview did not show an even sharper decline.

Some of this greater hesitation when reporting stereotypes to non-white or Hispanic interviewers also appears to harm the validity of oral reports. Among white respondents interviewed by non-white or Hispanic interviewers, regressing oral and ACASI reports on criteria and other measures of anti-black prejudice (AMP), the coefficients for oral reports are uniformly lower than ACASI. For instance, regressing difference in ACASI-based intelligence ratings of whites and blacks on AMP yields a coefficient, .09 ($p < .05$), substantially larger than .05 ($p < .05$), obtained when we regress difference in oral intelligence reports on AMP. Expectedly, among white respondents interviewed by non-Hispanic white interviewers, coefficients for oral reports are larger than ACASI reports; difference in ACASI intelligence reports yields a coefficient of .09 ($p < .001$), while difference in oral reports a coefficient of .12 ($p < .001$).

The same pattern holds for ratings of hardwork. Among white respondents interviewed by non-white interviewers, regressing difference in hardwork ratings obtained using ACASI on AMP yields a coefficient of .09 ($p < .05$), while regressing the difference measure based on oral reports on AMP yields a coefficient that is again more than 40% lower, .05 ($p < .1$). As before, among white respondents interviewed by non-Hispanic white interviewers, difference in ACASI based hardwork

ratings yields a coefficient of .10 ($p < .001$), and difference in oral reports a coefficient of .12 ($p < .001$).[6]

Discussion

Lots of whites feel comfortable expressing negative stereotypes of blacks in oral face-to-face interviews with white interviewers. This finding is at odds with conventional wisdom, but not the evidence (see for instance, pg. 41, Sniderman and Piazza 1993). The picture the results paint about race in the US is one that is all together more troubling than commonly apprehended. The traditional story of widespread social desirability pressures about expressing negative stereotypes of blacks speaks of strong social norms against such practice. If such norms were indeed universal, they would be an advance[7] from the situation where lots of whites hold aversive stereotypes about blacks, and feel, by and large, comfortable in expressing those stereotypes openly.

One possible interpretation of our results is that many more whites harbor anti-black views secretly and that ACASI failed to adequately eliminate social desirability pressures. However, one should hesitate before reaching this conclusion, because a large literature of studies on ACASI has produced results suggesting that this method does in fact lead respondents to admit more socially embarrassing facts about themselves than they would admit in oral interviews (e.g. Metzger et al. 2000; Newman et al. 2002; Tourangeau and Smith 1996; Turner et al. 1998,?). Nonetheless, future studies of representative national samples might employ alternative techniques, such as the bogus pipeline (Sigall and Page 1971) (though getting human subjects approval for this 'deceptive' technique is liable to prove hard), the bonafide pipeline (Fazio et al. 1995), unobtrusive monitoring (Crosby et al. 1980), the item count technique (Holbrook and Krosnick 2010; Kuklinski et al. 1997) to see whether substantial social desirability-driven bias on racial stereotypes exists.

Another possible interpretation of the results is that racial stereotypes are not notably distorted by social desirability pressures during in-person interviews. As Study 1 demonstrates, plenty of white respondents in past ANES surveys were willing to report explicitly that they viewed whites more favorably than blacks in terms of hard work and intelligence.[8] Perhaps this occurs because

---

[6]Similar patterns hold for other criteria. These results can be requested from the author.

[7]Schuman (1997) make the same argument.

[8]The large percentage of whites willing to express prejudicial stereotypes also stands at odd with claims of rampant

high-level professional interviewers may do exactly what they are hired to do: to establish rapport with and trust from the respondent. Under those circumstances, respondents may feel that they can answer sensitive questions honestly. Therefore, perhaps oral self-reports of racial stereotypes in such national surveys can be trusted. Indeed, this conclusion is supported by other studies that used methods other than ACASI and found relatively small increases in reports of anti-black racial attitudes (e.g. Krysan 1998; Krysan and Couper 2003).

The results in this paper lend themselves to two specific recommendations for measurements of racial stereotypes in oral interviews - 1) Assign interviewers of the same race as the respondent, 2) Because ACASI adds a small procedural complexity to the CAPI questionnaire programming and administration processes, and because the ANES has administered the stereotypes measures orally in many past surveys, the current results suggest that this extra complexity is not worthwhile. Because ACASI might even have produced a reduction in predictive validity without increasing the expression of anti-black stereotypes notably, oral administration of these questions in future surveys seems desirable.

Our findings also address a bolder assertion made about in-person surveys: "The fact that people may misrepresent themselves about sensitive topics on attitude surveys should force us to ask whether or not it is worth all the trouble to administer such surveys in the first place. How much damage do these misreports do to the data we collect and the inferences we try to draw from them? The answer is, unsurprisingly, "a great deal," if we proceed naively as if the data we have are not measured with bias" (Corstange 2009).

In contrast, Sears (2004) said: "Social desirability biases can be shown in experiments, and no doubt do occur to some extent in surveys. Neither tells us whether they occur in sufficient magnitude to threaten the main findings of survey research" (pg. 296).

Our results answer Sears' question and challenge Corstange's (2009) assertion. At least in this instance, it appears that white respondents did not intentionally misrepresent themselves when reporting their perceptions of blacks' and whites' traits when speaking to (well-trained) white non-Hispanic interviewers.

---

self-deception social desirability bias —effort to create a socially desirable impression in one's own eyes (Paulhus 1984, 1986).

References

Black, E. (1987). *Politics and Society in the South*. Harvard University Press.

Bobo, L. (2001). Racial attitudes and relations at the close of the twentieth century. *America becoming: Racial trends and their consequences 1*, 264–301.

Bobo, L. D. and M. C. Dawson (2009). A change has come. *Du Bois Review: Social Science Research on Race 6*(01), 1–14.

Case, C. E., A. M. Greeley, and S. Fuchs (1989). Social determinants of racial prejudice. *Sociological Perspectives*, 469–483.

Corstange, D. (2009). Sensitive questions, truthful answers? modeling the list experiment with listit. *Political Analysis 17*(1), 45–63.

Cribbs, S. E. and D. M. Austin (2011). Enduring pictures in our heads: The continuance of authoritarianism and racial stereotyping. *Journal of Black Studies 42*(3), 334–359.

Crosby, F., S. Bromley, and L. Saxe (1980). Recent unobtrusive studies of black and white discrimination and prejudice: A literature review. *Psychological Bulletin 87*(3), 546.

Des Jarlais, D. C., D. Paone, J. Milliken, C. F. Turner, H. Miller, J. Gribble, Q. Shi, H. Hagan, and S. R. Friedman (1999). Audio-computer interviewing to measure risk behaviour for hiv among injecting drug users: a quasi-randomised trial. *The Lancet 353*(9165), 1657–1661.

Dovidio, J. F. and R. H. Fazio (1992). New technologies for the direct and indirect assessment of attitudes. *Questions about questions: Inquiries into the cognitive bases of surveys 21*, 204–237.

Edwards, A. L. (1957). *The social desirability variable in personality assessment and research*. Dryden Press.

Fazio, R. H., J. R. Jackson, B. C. Dunton, and C. J. Williams (1995). Variability in automatic activation as an unobtrusive measure of racial attitudes: a bona fide pipeline? *Journal of personality and social psychology 69*(6), 1013.

Feagin, J. R. and M. P. Sikes (1995). *Living with racism: The black middle-class experience*. Beacon Press.

Feldman, S. and K. Stenner (1997). Perceived threat and authoritarianism. *Political Psychology 18*(4), 741–770.

Finkel, S. E., T. M. Guterbock, and M. J. Borg (1991). Race-of-interviewer effects in a preelection poll virginia 1989. *Public Opinion Quarterly 55*(3), 313–330.

Gilens, M. (1998). Racial attitudes and race-neutral social policies: White opposition to welfare and the politics of racial inequality. *Perception and prejudice: Race and politics in the United States*, 171–201.

Goffman, E. (1959). The presentation of self in everyday life. *New York: Doubleday Anchor*.

Goldman, S. K. (2012). Effects of the 2008 obama presidential campaign on white racial prejudice. *Public opinion quarterly 76*(4), 663–687.

Greeley, A. M. and P. B. Sheatsley (1971). Attitudes toward racial integration. *Scientific American*.

Gribble, J. N., H. G. Miller, P. C. Cooley, J. A. Catania, L. Pollack, and C. F. Turner (2000). The impact of t-acasi interviewing on reported drug use among men who have sex with men. *Substance use & misuse 35*(6-8), 869–890.

Hamilton, D. L. and J. W. Sherman (1994). Stereotypes. In *The Handbook of social cognition*, Volume 2, pp. 1–68. Erlbaum Mahwah, NJ.

Hatchett, S. and H. Schuman (1975). White respondents and race-of-interviewer effects. *The Public Opinion Quarterly 39*(4), 523–528.

Heerwig, J. A. and B. J. McCabe (2009). Education and social desirability bias: The case of a black presidential candidate*. *Social Science Quarterly 90*(3), 674–686.

Holbrook, A. L. and J. A. Krosnick (2010). Social desirability bias in voter turnout reports tests using the item count technique. *Public Opinion Quarterly 74*(1), 37–67.

Hopkins, D. J. (2009). No more wilder effect, never a whitman effect: When and why polls mislead about black and female candidates. *The Journal of Politics 71*(03), 769–781.

Huddy, L. and S. Feldman (2009). On assessing the political effects of racial prejudice. *Annual Review of Political Science 12*, 423–447.

Hurwitz, J. and M. Peffley (1997). Public perceptions of race and crime: The role of racial stereotypes. *American journal of political science*, 375–401.

Jaynes, G. D. and R. M. Williams (1989). *A common destiny: Blacks and American society*. National Academies Press.

Jones, E. E. and H. Sigall (1971). The bogus pipeline: a new paradigm for measuring affect and attitude. *Psychological Bulletin 76*(5), 349.

Jost, J. T. and D. L. Hamilton (2005). Stereotypes in our culture. *On the nature of prejudice*, 208–225.

Katz, D. and K. Braly (1933). Racial stereotypes of one hundred college students. *The Journal of Abnormal and Social Psychology 28*(3), 280.

Kluegel, J. R. (1990). Trends in whites' explanations of the black-white gap in socioeconomic status, 1977-1989. *American Sociological Review*, 512–525.

Krysan, M. (1998). Privacy and the expression of white racial attitudes: A comparison across three contexts. *Public Opinion Quarterly*, 506–544.

Krysan, M. and M. P. Couper (2003). Race in the live and the virtual interview: Racial deference, social desirability, and activation effects in attitude surveys. *Social psychology quarterly*, 364–383.

Kuklinski, J. H., M. D. Cobb, and M. Gilens (1997). Racial attitudes and the" new south". *Journal of Politics*, 323–349.

Metzger, D. S., B. Koblin, C. Turner, H. Navaline, F. Valenti, S. Holte, M. Gross, A. Sheon, H. Miller, P. Cooley, et al. (2000). Randomized controlled trial of audio computer-assisted self-interviewing: utility and acceptability in longitudinal studies. *American Journal of Epidemiology 152*(2), 99–106.

Miller, J. D. (1984). *A new survey technique for studying deviant behavior.* Ph. D. thesis, George Washington University.

Miller, J. D., I. Cisin, and A. Harrell (1986). A new technique for surveying deviant behavior: Item-count estimates of marijuana, cocaine, and heroin. In *Paper presented the Annual Meeting of the American Association for Public Opinion Research, St. Petersburg, Florida.*

Newman, J. C., D. C. Des Jarlais, C. F. Turner, J. Gribble, P. Cooley, and D. Paone (2002). The differential effects of face-to-face and computer interview modes. *American Journal of Public Health 92*(2), 294–297.

Paulhus, D. L. (1984). Two-component models of socially desirable responding. *Journal of personality and social psychology 46*(3), 598.

Paulhus, D. L. (1986). Self-deception and impression management in test responses. In *Personality assessment via questionnaires*, pp. 143–165. Springer.

Paulhus, D. L. (1989). Socially desirable responding: Some new solutions to old problems. In *Personality Psychology*, pp. 201–209. Springer.

Paulhus, D. L. (1991). Measurement and control of response bias. In *Measures of personality and social psychological attitudes. Measures of social psychological attitudes.* Academic Press.

Payne, B. K., C. M. Cheng, O. Govorun, and B. D. Stewart (2005). An inkblot for attitudes: affect misattribution as implicit measurement. *Journal of personality and social psychology 89*(3), 277.

Peffley, M., J. Hurwitz, and P. M. Sniderman (1997). Racial stereotypes and whites' political views of blacks in the context of welfare and crime. *American Journal of Political Science*, 30–60.

Piston, S. (2010). How explicit racial prejudice hurt obama in the 2008 election. *Political Behavior 32*(4), 431–451.

Ross, L. (1977). The false consensus effect: An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology*, 279–301.

Schuman, H. (1997). *Racial attitudes in America: Trends and interpretations.* Harvard University Press.

Sears, D. O. (2004). A perspective on implicit prejudice from survey research. *Psychological Inquiry 15*(4), 293–297.

Sigall, H. and R. Page (1971). Current stereotypes: A little fading, a little faking. *Journal of Personality and Social Psychology 18*(2), 247.

Sniderman, P. M. and T. L. Piazza (1993). *The scar of race.* Cambridge Univ Press.

Tourangeau, R. and T. W. Smith (1996). Asking sensitive questions the impact of data collection mode, question format, and question context. *Public opinion quarterly 60*(2), 275–304.

Turner, C. F., B. H. Forsyth, J. M. O'Reilly, P. C. Cooley, T. K. Smith, S. M. Rogers, and H. G. Miller (1998). Automated self-interviewing and the survey measurement of sensitive behaviors. *Computer assisted survey information collection*, 455–74.

Turner, C. F., L. Ku, S. M. Rogers, L. D. Lindberg, J. H. Pleck, and F. L. Sonenstein (1998). Adolescent sexual behavior, drug use, and violence: increased reporting with computer survey technology. *Science 280*(5365), 867–873.

Villarroel, M. A., C. F. Turner, S. M. Rogers, A. M. Roman, P. C. Cooley, A. B. Steinberg, E. Eggleston, and J. R. Chromy (2008). T-acasi reduces bias in std measurements: the national std and behavior measurement experiment. *Sexually transmitted diseases 35*(5), 499–506.

Wagner, U. and A. Zick (1995). The relation of formal education to ethnic prejudice: Its reliability, validity and explanation. *European Journal of Social Psychology 25*(1), 41–56.

Warner, S. L. (1965). Randomized response: A survey technique for eliminating evasive answer bias. *Journal of the American Statistical Association 60*(309), 63–69.

Whitley Jr, B. E. (1999). Right-wing authoritarianism, social dominance orientation, and prejudice. *Journal of Personality and Social Psychology 77*(1), 126.

Table 1: Percent of Non-Hispanic White Respondents Rating Whites More Favorably than Blacks.

	Hardworking	Intelligent
1992 ANES	60.02% [57.34%, 62.69%] ($n = 1294$)	50.49% [47.75%, 53.22%] ($n = 1287$)
1996 ANES	52.40% [48.37%, 56.44%] ($n = 592$)	45.59% [41.55%, 49.64%] ($n = 586$)
2000 ANES	57.83% [53.59%, 62.07%] ($n = 525$)	48.02% [43.72%, 52.31%] ($n = 523$)
2004 ANES	49.46% [45.92%, 53.00%] ($n = 768$)	40.06% [36.59%, 43.54%] ($n = 768$)

Note: The numbers in the square brackets are the 95% confidence intervals.

Table 2: Effect of Mode of Administration on Trait Ratings of Blacks and Whites by Non-Hispanic White Respondents.

ANES	Oral Face-to-Face	ACASI	Diff.
Hardworking (Whites > Blacks)	44.67% [41.64%, 47.71%] ($n = 1003$)	49.43% [46.35%, 52.52%] ($n = 1003$)	4.76%**
Intelligent (Whites > Blacks)	39.33% [36.35%, 42.32%] ($n = 1006$)	44.45% [41.39%, 47.50%] ($n = 1006$)	5.12%***
NAES	Wave 4	Wave 5	Diff.
Hardworking (Whites > Blacks)	50.61% ($n = 12869$)	47.17% ($n = 12869$)	3.44%***
Intelligent (Whites > Blacks)	45.05% ($n = 12871$)	46.88% ($n = 12871$)	1.83%***

Note —***$p < .001$, **$p < .01$, *$p < .05$, +$p < .1$

Table 3: Coefficients of Bivariate Regressions of Oral and ACASI Stereotype Measures on Criteria.

Criteria	Hardworking			Intelligent		
	Oral	ACASI	Oral-ACASI	Oral	ACASI	Oral-ACASI
South	.02**	.02**	.00	.02**	.02**	.00
Education	-.04**	-.03*	-.01	-.06***	$-.05$***	-.01
AMP	.09***	.08**	.01	.09***	.08**	.01
Authoritarianism	.10**	.10**	.00	.11**	.09**	.02[+]

Note —***$p < .001$, **$p < .01$, *$p < .05$, [+]$p < .1$

Table 4: Generalized Hierarchical Logistic Model coefficients predicting log odds of reporting negative stereotype of blacks in the oral interview among Non-Hispanic white respondents using race of interviewer, education, age, ratings under ACASI, with random effects for interviewer, and PSU, and fixed effects for strata.

Predictors	Hardworking	Intelligent
White Interviewer	.38	.84**
ACASI		
Hardwork (whites - Blacks)	6.37***	3.1***
Intelligence (whites - Blacks)	4.98***	7.95***
Education		
Some College	-.24	-.29+
College	-.33+	-1.06***
More than College	-.32*	-.64*
Age (Rescaled 0 to 1)	1.11***	2.34***
N	959	960
Log Likelihood	-776.7	-711.5

Note —$^{***}p < .001$, $^{**}p < .01$, $^{*}p < .05$, $^{+}p < .1$

Appendix A: Data and Measures

Data

ANES Time Series: In the American National Election Studies (ANES) Time Series Surveys conducted in 1992, 1996, 2000, and 2004, a representative sample of American adults was interviewed. In each of the surveys at least a large proportion of the interviews were conducted face-to-face in respondents' homes. ANES Time Series surveys have typically used multi-stage area probability sampling; though the exact sampling scheme and sampling frame have varied over the years, some parts of the sampling scheme have remained roughly constant —typically first the counties are sampled, followed by census tracts within counties, and lastly census block groups are sampled from within census tracts. Within each selected census block group, individual households are randomly selected, and finally one adult is randomly selected within each household. The Time Series studies consist of one pre-election interview and one post-election interview. The AAPOR Response Rate 1 was more than 60 percent for each of the years for which the data are included.

ANES 2008: The 2008 ANES Time Series Study survey interviewed a sample of American adults face-to-face in their homes. It selected the respondents using an area probability design. During the pre-election wave, interviewing began on September 5, 2008 and ended on November 3, 2008. The post-election interviewing was conducted between November 5 and December 21, 2008. The AAPOR Response Rate 3 was 64.3 percent. A total of 1,151 white Non-Hispanic respondents provided data. These respondents were interviewed by 111 non-randomly assigned interviewers. 975 of the respondents were interviewed by non-Hispanic white interviewers.

2008 NAES Online Panel: Knowledge Networks maintains a large Online panel, which it recruits using Random Digit Dialing, among other methods. It provides free Internet connections to those without Internet access, and provides other financial incentives to those with Internet access.

Measures

Stereotype Ratings (ANES Time Series and 2008): Respondents were shown a seven-point scale on which the group traits were to be rated. The interviewers explained to the respondents how to use the scale as follows, "a score of '1' means that you think almost all of the people in that group tend to be 'hard-working.' A score of '7' means that you think most people in the group are 'lazy.' A score of '4' means that you think that most people in the group are not closer to one end or

the other, and of course, you may choose any number in between." The interviewer then asked the respondent, "Where would you rate [blacks/whites] in general on this scale?"

The ACASI administration was conducted as follows: the interviewer gave a laptop that s/he had been using to the respondent, the respondent put on the headphones, and the computer played an audio recording of the questions and answer choices aloud to respondents. The question and the response options were also displayed on the screen. Respondents answered privately by typing questions on a computer keyboard.

Stereotype Ratings (2008 NAES Online Panel): Respondents were asked to rate their own racial group, and one other racial group on a 101 point semantic scale, going from "extremely hardworking/intelligent" to "extremely lazy/unintelligent". In particular, respondents were asked, "Where would you rate (whites/blacks/Hispanic Americans) in general on this scale?" Respondents could enter their responses by using the slider on the scale.

Race of the respondent: Respondents were asked: "What racial or ethnic group or groups best describes you? Black, Asian, Native-American, Hispanic or Latino, White, or Other?" Respondents were further asked if they were "of Spanish or Hispanic origin or descent?"

Authoritarianism (ANES 2008): Respondents were asked to choose between qualities emphasizing order, and those emphasizing freedom. They were asked to pick the quality they thought was more important for a child to have. They were asked to pick between independence and respect for elders, between obedience and self-reliance, between curiosity and good manners, and between being considerate and being well behaved. The 'authoritarian' choices, respect for elders, obedience, good manners, and well behaved, were coded as 1 while the rest were coded as 0. The scores from the four items were averaged.

Affect Misattribution Procedure (AMP) (ANES 2008): Respondents were shown an image of a black or a white face for a very brief duration, and then asked to rate a Chinese pictograph as pleasant or unpleasant (see Payne et al. 2005). There were a total of 48 items —half showed white faces before the rating task, and the other half showed black faces before the rating task. We scored ratings of pleasant as 1 and unpleasant as 0. We then summed up the scores for white and black faces separately, and subtracted the scores for black faces from the scores for white faces and rescaled the difference to lie between 0 and 1.

Appendix B: Alternate Ways of Coding the Dependent Variables

In the main body of the paper, we focus our attention on dichotomized difference ratings. While the categorization was theoretically motivated, there are other theoretical perspectives. We investigate here whether we get an alternate understanding if we were to code variables differently. In particular, we look at absolute categorical ratings of blacks and whites, and a continuous measure - the extent to which whites think whites are more hardworking (or intelligent) than blacks.

When we examine ratings of blacks and whites separately, again, at first, we see evidence consistent with the notion that confidentiality increased expression of anti-black sentiment. Among white respondents, confidentiality significantly increased the percent of respondents who rated blacks below the midpoint on the hardworking scale by about 10 percentage points, from 22.75% in the oral face-to-face condition to 32.65% in the ACASI condition ($p < .001$). A similar pattern appeared in whites' ratings of blacks' intelligence. Confidentiality significantly increased the proportion of whites who rated blacks below the midpoint on the intelligence rating scale by about 7 percentage points, from 14.37% to 21.21%. These results are consistent with the notion that confidentiality allowed more people to admit to holding derogatory views of blacks.

However, the same patterns appeared in whites' ratings of whites. Providing confidentiality increased the percentage of white respondents rating whites below the midpoint on the hardworking scale by about 5 percentage points, from 3.58% to 8.35% ($p < .001$). Likewise, on intelligence, confidentiality significantly increased the proportion of whites who rated whites below the midpoint by about 2 percentage points ($p < .05$). Thus, one effect of moving to oral condition was merely 'grade inflation': confidentiality increased derogatory ratings of *both* blacks and whites.

To assess changes in the extent to which whites favor whites over blacks, we computed the difference between respondent ratings of whites and blacks. We rescaled the difference to lie between 0 and 1, with more positive numbers indicating more derogatory ratings of blacks vis-à-vis whites. Providing confidentiality yields a small increase in the extent to which respondents rated whites higher than blacks on hardwork (Diff. = .006, $p = .11$). Similarly, moving to ACASI produced a small increase in the difference in ratings of intelligence (Diff. = .013, $p < .001$). To put the change in ratings across modes in context, the average change in intelligence ratings was less than one-seventh of a point on the 13 point difference scale, or about .1 standard deviation of ACASI ratings

of intelligence. Thus, once again, we find that providing confidentiality reveals little additional negative stereotyping. These small increases are partly explained by small differences in ratings under ACASI: even when provided confidentiality, only a small fraction (approximately 15%) of the respondents chose to rate blacks more than two points below whites.

We start by describing the sources and extent of missing data, follow it with a discussion of whether we think missingness is ignorable, and finally estimate the effects under the assumption that missing data are ignorable conditional on observables.

Attrition between pre and post-election interview claimed about 9% of the sample (n = 105). Setting aside unit non-response, 41 respondents did not give ratings of hardwork via ACASI and 33 respondents did not give ratings of intelligence when interviewed via ACASI. Corresponding figures for the oral interview were 17 and 18 respectively. So, somewhat counter to expectations, there was more missing data on ACASI than on the oral interview. Since expectations of strong social desirability driven distortion are far greater for those who chose to respond via ACASI, and not orally, than vice-versa, we focus on these respondents. Only 12 respondents (or about 1% of the remaining sample) provided ratings of intelligence only via ACASI, while only 7 respondents ($\sim$ .5% of the remaining sample) chose to not provide ratings of hardwork orally, while responding via ACASI.

So what explains attrition between waves —the cause of majority of the missing data? Given respondents were not aware that the stereotype questions would be asked orally in the post-election interview, attrition could not have been a consequence of reluctance to respond to stereotype questions orally. To test this intuition empirically, we checked if respondents were more likely to respond if interviewed by a white interviewer, controlling for education, age and gender, than when interviewed by a non-white interviewer, on the back of the theory that social desirability pressures would be greater on average for those interviewed by non-Hispanic whites. Race of interviewer did not predict whether or not the respondent gave the post-election interview ($p = .77$). The null result holds when we predict whether or not the respondent did not respond on the on the oral interview after responding on ACASI (a combination of unit and item non-response) ($p_{\text{hardwork}} = .61; p_{\text{intelligence}} = .81$). Regressions also reveal that respondents with more than a college degree were significantly less likely to respond to the post-election interview ($p < .001$). This may prove consequential as the educated potentially face more social desirability pressures on race than the less educated (Krysan 1998; Hatchett and Schuman 1975).

To investigate the issue further, we explored other potential correlates of non-response. There

is the possibility that those who held strong negative stereotypes of blacks were less likely to give the post-election interview, possibly as a result of their unhappiness at the election of a black president. This in itself may not bias our estimates should the negativity of stereotypes of blacks be uncorrelated with change between ACASI and the oral interview (or, be independent of potential outcomes). We checked if attrition was predicted by stereotype reports obtained under ACASI, controlling for age, gender, and education. It appears that ratings of hardwork were not correlated to attrition ($p = .17$), but ratings of intelligence were ($p = .04$). Regressing a combination of unit and item non-response in the oral interview weakens the predictive power of ACASI ratings a little ($p_{\text{hardwork}} = .23, p_{\text{intelligence}} = .08$).

In all there is some suggestive evidence of small systematic differences between those who responded to the stereotype questions on the oral interview and those who did not, among those who provided the answer in the ACASI interview. To address the issue, we predicted the propensity to not respond to the stereotype question on the oral interview using interviewer race, age, gender, education, and stereotype ratings obtained under ACASI. We then used the fitted values from the estimation to reweight our sample —dividing our original weights with the probability to respond on the post-treatment wave —to derive estimates of the causal effect under assumption of missing independent of potential outcomes when conditioned on the selected covariates.[9] Doing so nudges our estimates of the amount of net hidden negative stereotypes on both hardwork and intelligence by a thousandth decimal point (an additional .1% and .2% of respondents are estimated to have hidden negative stereotypes of blacks on hardwork, and intelligence respectively in the oral interview). In all, it appears that any bias due to missing data is small.

---

[9]We assume missing data on other covariates such as age, gender, education, and interviewer race to be missing at random.