Political Science Research and Methods (2026), 1–8
doi:10.1017/psrm.2026.10115

RESEARCH NOTE

Measuring racial attitudes pre-treatment in survey experiments

Leah Christiani[1], Nathan Kelly[2] and Jana Morgan[2]

[1]Political Science, CUNY Hunter College, New York, NY, USA and [2]Political Science, Rutgers University, New Brunswick, NJ, USA
Corresponding author: Nathan Kelly; Email: nathan.j.kelly@gmail.com

(Received 10 April 2025 UTC; revised 25 November 2025 UTC; accepted 2 January 2026 UTC)

Abstract
Experimental researchers often need to assess the moderating effect of racial attitudes. However, it is unclear whether the racial attitude battery deployed may bias treatment effects. Less subtle batteries could bias treatment effects toward zero if measuring racial attitudes primes these attitudes for subjects assigned to all conditions, including the control condition. We assess whether bias is introduced across four conditions, pre-treatment: not measuring racial attitudes, using stereotype endorsement measures, using the American National Election Study racial resentment battery, or using the simplified Cooperative Election Study racial resentment battery. We do not uncover widespread instances of treatment bias. However, we do identify bias when the stereotype endorsement measures are used pre-treatment. We conclude with recommendations for researchers.

Keywords: American politics; class and ethnicity; experimental research; race

1. Introduction

Scholars studying a variety of political attitudes in an experimental setting are often interested in assessing how pre-existing psychological attributes and attitudes moderate treatment effects. Racial attitudes, in particular, are often theorized to be moderating variables when studying how policy attitudes are affected by framing or other kinds of information provision.

The typical strategy for assessing moderating (or conditioning) effects of background characteristics is to measure the moderator prior to treatment and then assess the moderation effects. But there are potential problems with this approach when racial attitudes are the moderator. Specifically, measuring racial attitudes prior to treatment could bias treatment effects by priming racial considerations that enhance or undermine the effect of treatment conditions and/or contaminate control conditions designed to be devoid of racialized content.

While studies have demonstrated that the placement of racial attitude moderators pre-treatment is unlikely to bias effects, it is still unclear whether the particular battery used to measure racial attitudes matters. In this note, we report the results of an experiment designed to test the viability of several pre-treatment racial attitude batteries as moderators of treatment effects. We first summarize what we know from research on the placement of moderators, and then extend this discussion to the particular measure of racial attitudes. Finally, we present the results of our study, which show that low-cost

© The Author(s), 2026. Published by Cambridge University Press on behalf of EPS Academic Ltd. This is an Open Access article, distributed under the terms of the Creative Commons Attribution licence (http://creativecommons.org/licenses/by/4.0), which permits unrestricted re-use, distribution and reproduction, provided the original article is properly cited.

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

options for the pre-treatment measurement of racial attitudes as moderators are available and can be effectively deployed in experimental settings.

2. What we (mostly) know: where to place moderators

Researchers have given considerable attention to evaluating whether the placement of moderators can bias treatment effects. The options for where to place moderators are: well before the treatment (i.e., on an earlier panel), on the same survey but before the treatment, or on the same survey but after the treatment. While there is a general consensus that measuring moderators well before treatment, such as in an earlier panel, is ideal, this is often not feasible given resource constraints and attrition concerns. When it comes to moderator placement on a single survey, two main issues have risen to the fore: post-treatment bias and priming effects.

Studies have demonstrated that moderating on variables that are measured post-treatment can bias treatment effects (Acharya *et al.*, 2016; Montgomery *et al.*, 2018). This is especially likely for moderators that are potentially influenced by the treatment itself. On the other hand, measuring moderators pre-treatment runs the risk of priming effects. Priming can occur when moderators bring particular concerns to mind for all respondents in a way that goes on to influence how they respond to the treatment (Mendelberg, 2008; Klar *et al.*, 2020).

Researchers have studied the placement of sensitive moderators, like racial attitudes, in particular. By and large, studies have found that measuring racial attitude and/or identity moderators pre- or post-treatment does not introduce substantial bias. Valentino *et al.* (2018) find no difference when symbolic racism, as a moderator, was placed pre-treatment versus post-treatment versus in an earlier wave. Similarly, Albertson and Jessee (2023) find no difference between pre- or post-treatment measurement of racial resentment. Sheagley and Clifford (2025) largely uncover no difference in pre- or post-treatment moderator placement, though they do note that measuring racial resentment pre-treatment, in the same survey, versus an earlier wave led to somewhat stronger treatment effects (though not statistically significantly stronger at a 95% confidence level). This finding is actually the opposite of the typical concern, that priming would wash treatment effects out (Sheagley and Clifford, 2025). Finally, across four separate experiments, d'Urso *et al.* (2025) uncover no effects of moderator placement on the measurement of the moderators, the measurement of experimental outcomes, or experimental outcomes based on treatment conditions. However, they do find some inconsistent effects for moderator placement on the dependent variables when placement is interacted with the treatment condition and the moderator itself. Even in these instances, though, d'Urso *et al.* (2025) note that the differences were minimal, and conclusions unlikely to change. Nevertheless, they advise researchers in most cases to follow Montgomery *et al.* (2018)'s recommendations to place moderators pre-treatment when possible, to limit or avoid any potential bias. Thus, while the bulk of the literature on the placement of sensitive items indicates that measuring moderators pre- or post-treatment has no or minimal effects on the experiment, the most conservative approach is still to measure the items pre-treatment.

3. Our question: which measure of racial attitudes to use pre-treatment?

While existing studies suggest that placing moderators pre-treatment is preferred, less is known about whether the particular *type* of moderator used to measure racial attitudes matters for avoiding bias. Though many studies have demonstrated that experimental results hold regardless of the particular racial attitude battery used (e.g., Valentino *et al.*, 2018; Stephens-Dougan, 2020), none to our knowledge have directly tested whether the potential for biasing treatment effects is greater for some batteries than others. Further, most of the evidence that tests placement of racial attitude moderators uses one common battery: the racial resentment scale (Albertson and Jessee, 2023; d'Urso *et al.*, 2025;

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

Sheagley and Clifford, 2025) or its close variant, symbolic racism (Valentino *et al.*, 2018).[1] This makes sense, given that racial resentment is the dominant measure of racial attitudes in American politics research, but it leaves uncertainty as to whether some racial attitude measures have more potential for bias (via priming effects) than others.

Given the emerging consensus in favor of pre-treatment measures of racial attitudes as moderators, it is important to evaluate whether any specific measures of racial attitudes are more or less likely to bias treatment effects downstream. The reason why the particular racial attitude battery may matter is because some are subtler than others—and the more subtle the measure, the less likely that priming effects will emerge. Thus, the goal is to choose a racial attitude battery that is *least* likely to prime racial considerations for the entire sample, which would run the risk of washing out any treatment effects (especially if the treatment is related to race).

But deciding how to measure racial attitudes is tough for both methodological and theoretical reasons. The ideal measure is direct enough to accurately capture the racial attitudes held by respondents, but subtle enough that respondents are not primed to think about race in a way that changes their response to treatment or control conditions. As a result, a plethora of measures for racial attitudes have emerged over the years.

One commonly used *direct* measure of racial attitudes asks respondents about the degree to which they associate different groups with characteristics that reflect certain negative stereotypes. By rating groups in terms of a particular stereotype, such as the proclivity of group members to be "hard-working" or "lazy," respondents' answers to these stereotype endorsement items are designed to directly capture racial animus (Huddy and Feldman, 2009).[2] A potential benefit to the stereotype-endorsement measures is that they can be asked about a variety of groups, including non-racial groups. As a result, it may be the case that the stereotype-endorsement measures could be made subtler if racial groups are embedded within other, non-racial social groups, such as age groups or professions. Thus, while this measure of racial attitudes is direct, embedding the racial groups within a larger pool might increase its subtlety.

Another common measure of racial animus is the racial resentment scale. Racial resentment conceives of contemporary racism as a combination of anti-Black bias and proindividualist attitudes—a form of "new racism" that Kinder and Sanders (1996) argue emerged in response to the decline of biological racism (Schuman *et al.*, 1998) in the wake of the Civil Rights Movement. The common four-item battery asks respondents to report their level of agreement with four statements that purport to explain racial inequality in contemporary American society: (1) The Irish, Italians, Jews, and many other minorities overcame prejudice and worked their way up. Blacks should do the same without any special favors. (2) Generations of slavery and discrimination have created conditions that make it difficult for blacks to work their way out of the lower class. (3) It is really a matter of not trying hard enough; if blacks would only try harder, they could be just as well off as whites. (4) Over the past few years, blacks have gotten less than they deserve.

Some have critiqued the racial resentment scale, arguing that it measures conservatism rather than racial attitudes (Feldman and Huddy, 2005; Carmines *et al.*, 2011). However, recent work has demonstrated that racial resentment is a powerful measure of racial attitudes (Stephens-Dougan, 2020), even after accounting for ideological commitments (Enders, 2021). It is included on most large-scale national surveys, like the American National Election Study (ANES). The Cooperative

[1]Though note that d'Urso *et al.* (2025) use several: symbolic racism, racial resentment, and Muslim American resentment. However, these are not on the same survey and are not tested against one another.

[2]The exact phrasing used in the American National Election Study (ANES), which we adopt for this study, is: "Now, we have some questions about different groups in our society. We are going to show you a seven-point scale on which characteristics of the people in a group can be rated. In the first statement, a score of 1 means that you think that almost all of the people in that group tend to be 'hard-working.' A score of 7 means you think that 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."

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

Election Study (CES) uses a shortened version of the racial resentment scale, including only the first two items listed above. This reduced battery could make the racial resentment scale, which is already intended to be subtler than more direct measures of racial animus, even less likely to prime racial considerations.

When it comes to experimental research on race, being able to moderate by racial attitudes is extremely important, but it is also important that we limit the priming of racial considerations in ways that might bias treatment effects or contaminate the control. On the one hand, as the preceding discussion suggests, racial resentment, as a measure of "new racism," aims to be subtler than more direct measures, including stereotype endorsement. While stereotype endorsement measures are more direct, stereotype batteries that include items not only about racial groups but also non-racial ones could reduce potential priming effects. Thus, to better understand whether different measurement approaches might bias treatment effects, this study evaluates the stereotype endorsement approach against two versions of the racial resentment scale, all measured pre-treatment.

4. Experimental design

Our core aim is to determine whether different types of racial attitude measures are more or less likely to bias treatment effects when asked pre-treatment. Note that we are measuring racial attitudes early in the survey, then providing additional questions to buffer before the treatment. So, while we are including racial attitudes measures as well as an experiment on the same survey, there is space between the two. We are not testing moderator placement or advocating measuring racial attitudes *immediately* prior to the experiment.

We fielded a pre-registered survey experiment with a 4×2 design. Respondents were randomized into one of four measurement categories: (1) no racial attitudes measured, (2) stereotype endorsement items, (3) four-item ANES racial resentment scale, and (4) two-item CES racial resentment scale. Later in the survey, respondents were again randomized into one of two conditions: (1) experimental control and (2) a race-related experimental treatment. We use this 4×2 design to test whether the way that we measure racial attitudes pre-treatment has effects on an experiment that comes later in the survey. For the stereotype-endorsement measure, we asked about how hardworking or lazy the following groups are (randomized): Old people, Young people, White people, Black people, School teachers, and Police officers.

The survey was in the field with Lucid Marketplace from May 11 to 12, 2021.[3] Respondents first consented to the survey and answered a few basic demographic questions. Then, they were randomized into one of the four measurement conditions. Afterward, respondents were again asked some demographic questions to take their minds away from race—before being randomized into either the experimental treatment or control. Finally, all respondents answered the key dependent variables, a manipulation check, and were debriefed before exiting the survey.

The subsequent survey experiment tested whether mentioning racial inequality changes support for race-based public policies. Respondents in both conditions read news articles about the COVID-19 pandemic. The control article, titled "COVID-19 Reducing Life Expectancy in U.S.: Vaccines Aim to Help," reported on the pandemic and vaccine roll-out without any mention of race. The treatment article, titled "COVID-19 Reducing Life Expectancy in U.S.—Especially for Blacks: Vaccine Inequities Don't Help," reported on racial disparities in COVID-19 deaths and vaccine access. We present the full articles in the appendix.

[3]We used this opt-in survey with a goal of targeting white, non-Hispanic Americans age 18 and over, with age, gender, region, education, and income matching national distributions. Respondents were given two attention check items. The results reported below include all respondents, but dropping those who fail both attention checks produces substantively similar conclusions.

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

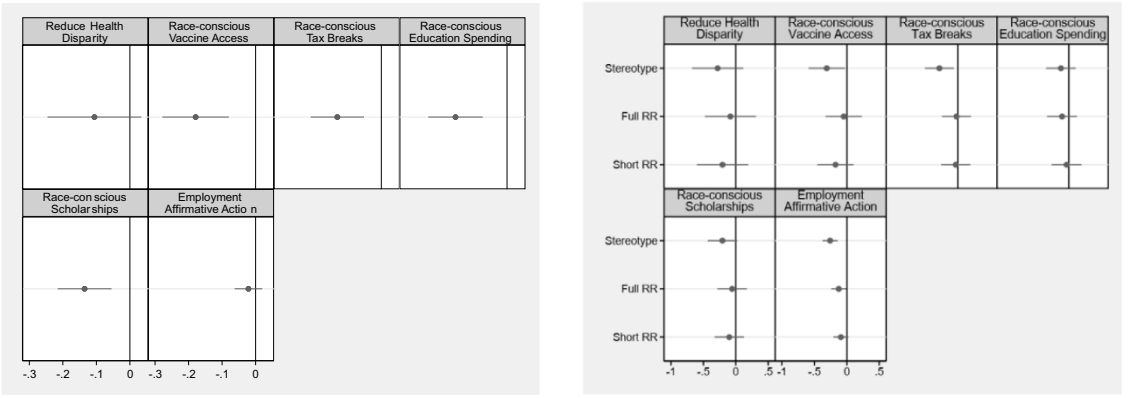


Figure 1. Average treatment effects across two experiments. (a) Treatment effects across multiple out. (b) Difference in treatment effects by racial comes attitude measurement strategy.
Notes: All outcomes coded such that higher values indicate more *opposition* to the programs listed. For Figure 1(b), only the coefficients on the interaction terms between the treatment and the measure of racial attitudes are plotted.

5. Results

In the first portion of our analysis, we seek to establish that our substantive treatment has statistically significant effects. In Figure 1a, we plot the treatment effect on all of the outcomes measured in our experiment. The results reported here show that our treatment was successful in making respondents more supportive of the egalitarian policies (note that higher values on these outcomes mean more opposition to the policies).

Next, we assess whether the particular pre-treatment measure of racial attitudes (or the absence of any such measure) had any influence on the treatment effects produced. Figure 1b reports how treatment effects vary across measurement groups for the six outcomes. Specifically, we report the interaction between the measurement group assignment and substantive treatment, which allows us to compare the difference between the treatment effect in each measurement group to the control measurement group in which racial attitudes were not measured pre-treatment. A negative value here indicates that the treatment effect is biased in a negative direction within the measurement group, while a positive value indicates positive bias. We use a significance level of $p < .10$ to avoid dismissing bias in treatment effects due to insufficiently precise estimates.

The results show minimal differences in treatment effects across the measurement groups. Of the 18 treatment effects estimated within pre-treatment racial attitudes measurement groups, only three differ significantly from the treatment effect estimated when racial attitudes were not measured pre-treatment. In most cases, the amount of bias introduced by measuring racial attitudes prior to treatment is very small and not significantly different from zero, using a p-value of .10.

Measuring racial resentment pre-treatment does not produce treatment-effect bias for any of these outcomes. Whether one uses a scaled-back two-item index or a more traditional four-item index, no statistically significant treatment-effect bias is introduced. There is one outcome, attitudes toward affirmative action in employment, where utilizing the four-item racial resentment battery prior to treatment introduces bias that approaches statistical significance, but the size of the bias is quite small and is not, strictly speaking, statistically significant even using a relaxed $p < .10$ threshold. The two-item racial resentment scale shows no evidence of introducing bias on this outcome measure or any other.

All three instances of treatment-effect bias come when measuring racial attitudes as stereotype endorsement. This means that for half of the dependent variables, measuring stereotypes pre-treatment biased the treatment effects. Specifically, when respondents answered the stereotype

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

endorsement items pre-treatment, the causal effect of exposure to the substantive treatment with racial content (versus those who encountered no racial content in the substantive control condition) was enhanced relative to those whose racial attitudes were not measured pre-treatment. In other words, the gap in support for egalitarian policies between those who encountered the substantive treatment about racial inequality versus those in the control whose news article did not mention race was larger for those who answered stereotype endorsement questions pre-treatment compared to those who answered no pre-treatment racial attitudes questions. The main takeaway here is that this pattern of bias sends an important warning against measuring stereotype endorsements within a single survey prior to treatment.

6. Discussion and conclusions

This paper set out to provide scholars interested in measuring racial attitudes in survey experiments with guidance about how to do so. We provide new evidence about the appropriateness of measuring racial attitudes prior to treatment in the same survey. Specifically, using the racial resentment scale pre-treatment did not bias downstream treatment effects, but using stereotype endorsement measures did, in some cases. As such, our findings suggest that measuring racial attitudes before the treatment is possible, but that it is better to do so with the racial resentment scale than stereotype endorsement measures.

As scholars consider whether and how to measure racial attitudes as moderators in experiments, it is essential to examine what may be driving the bias induced by measuring stereotypes prior to treatment. In our study, measuring stereotype endorsement pre-treatment often biased substantive treatment effects. Specifically, when exposed to the substantive treatment containing information about racial inequality, respondents exposed to the stereotype endorsement measures became *more* supportive of race-conscious policies than those who saw no racial attitude measures. This indicates that the respondents may have been primed by the stereotype endorsement measures in a way that shaped their response to the treatment as compared to the control.

It is important to think through exactly what might be driving this result. The patterns here suggest that the more direct nature of the stereotype endorsement items makes them prone to introducing bias; whereas the racial resentment batteries, which are designed to be less direct and subtler, do not. Even embedding the stereotype items about racial groups alongside parallel items about non-racial groups does not sufficiently obscure the direct racial content of the stereotype endorsement approach. Mechanistically, the results we observe could appear if one or both of the following scenarios occur: (1) the scores on the post-treatment outcomes of those exposed to the substantive treatment condition decrease (i.e., become less opposed to racial-egalitarian policy) when stereotype endorsement is measured pre-treatment relative to not measuring racial attitudes pre-treatment, and/or (2) the scores on the post-treatment outcomes of those exposed to the substantive control condition increase (i.e., become more opposed to racial-egalitarian policy) when stereotype endorsement is measured pre-treatment relative to not measuring racial attitudes pre-treatment.

Why might these patterns emerge? The first scenario above (in which those who were both exposed to a stereotype endorsement battery pre-treatment and treated with exposure to information about racial inequality become more supportive of racially egalitarian policy) might be driven by social desirability. Answering the stereotype endorsement battery may have reminded respondents that negative judgments about racial groups are normatively unacceptable, which then pushes them to respond in a more egalitarian way to the message highlighting disparities between racial groups in the treatment. The second scenario above (in which the control condition is contaminated by measuring stereotype endorsement pre-treatment) might occur if measuring stereotype endorsement brings racial considerations to mind and reduces support for racial egalitarian policy, generally, for those who are not treated with exposure to information about racial inequality. Race is brought to mind by the stereotype endorsement battery, so support for potentially racialized policies declines.

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

Further research would be needed to fully test the mechanism by which measuring stereotype endorsement pre-treatment introduces bias in treatment effects. If we are correct, it is essential for experimental researchers to think carefully about the outcomes they are assessing. In contexts where priming racial considerations could shift attitudes on outcomes, which is often the case when studying how the racialization of policy domains shapes egalitarian attitudes, then measuring stereotype endorsement prior to treatment in the same study may be especially risky. Additionally, our study was not designed to assess whether the measure of racial attitudes utilized biases any conditioning effects in which people with different degrees of racial animus respond differently to treatments with racial content.

Our results are based on a specific treatment and measurement of specific types of racial attitudes and outcomes. The finding that some racial attitudes can be measured pre-treatment without biasing treatment effects, while others cannot, suggests that future research could more broadly assess the effects of different approaches to pre-treatment racial attitudes measures. For instance, future work could evaluate whether the presence or degree of bias introduced by different pre-treatment measures of racial attitudes varies according to the substantive focus of the treatment itself or by the outcomes of interest. Here the treatment focused on racial inequalities in vaccine access, and our outcome measures considered support for a series of different egalitarian policies. Though we found similar patterns across a range of outcome measures, studies that varied the substance of the treatment or considered other dependent variables would further extend our understanding of the nature of bias that stereotype endorsement items produce. Finally, future work could also provide evidence about *precisely* how far in advance of the treatment racial attitudes that do bias treatment effects in our analysis must be measured in order to avoid the bias uncovered here.

Supplementary material. The supplementary material for this article can be found at https://doi.org/10.1017/psrm.2026.10115.

Funding statement. Funding for this project was provided by the Russell Sage Foundation. The views presented are the authors' alone and do not represent the views of the Russell Sage Foundation.

Competing interests. The authors declare no competing interests.

References

Acharya A, Blackwell M and Sen M (2016) Explaining causal findings without bias: Detecting and assessing direct effects. *American Political Science Review* **110**(3), 512–529.

Albertson B and Jessee S (2023) Moderator placement in survey experiments: Racial resentment and the "welfare" versus "assistance to the poor" question wording experiment. *Journal of Experimental Political Science* **10**(3), 448–454.

Carmines EG, Sniderman PM and Easter BC (2011) On the meaning, measurement, and implications of racial resentment. *The Annals of the American Academy of Political and Social Science* **634**(1), 98–116.

d'Urso AS, Bonilla T and Bogdanowicz G (2025) Where to place sensitive questions? Experiments on survey response order and measures of discriminatory attitudes. *Political Science Research and Methods*, 1–22.

Enders AM (2021) A matter of principle? On the relationship between racial resentment and ideology. *Political Behavior* **43**(2), 561–584.

Feldman S and Huddy L (2005) Racial resentment and white opposition to race-conscious programs: Principles or prejudice? *American Journal of Political Science* **49**(1), 168–183.

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

Kinder DR and Sanders LM (1996) *Divided by Color: Racial Politics and Democratic Ideals*. Chicago, IL: University of Chicago Press.

Klar S, Leeper T and Robison J (2020) Studying identities with experiments: Weighing the risk of posttreatment bias against priming effects. *Journal of Experimental Political Science* **7**(1), 56–60.

Mendelberg T (2008) Racial priming revived. *Perspectives on Politics* **6**(1), 109–123.

Montgomery JM, Nyhan B and Torres M (2018) How conditioning on posttreatment variables can ruin your experiment and what to do about it. *American Journal of Political Science* **62**(3), 760–775.

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

Schuman H, Steeh C, Bobo LD and Krysan M (1998) *Racial Attitudes in America: Trends and Interpretations*. Cambridge, MA: Harvard University Press.

Sheagley G and Clifford S (2025) No evidence that measuring moderators alters treatment effects. *American Journal of Political Science* **69**(1), 49–63.

Stephens-Dougan L (2020) *Race to the Bottom: How Racial Appeals Work in American Politics*. Chicago: University of Chicago Press.

Valentino NA, Neuner FG and Vandenbroek LM (2018) The changing norms of racial political rhetoric and the end of racial priming. *The Journal of Politics* **80**(3), 757–771.

https://doi.org/10.1017/psrm.2026.10115 Published online by Cambridge University Press

Cite this article: Christiani L, Kelly N and Morgan J (2026) Measuring racial attitudes pre-treatment in survey experiments. *Political Science Research and Methods*, 1–8. https://doi.org/10.1017/psrm.2026.10115