American Political Science Review (2026) 120, 2, 707–724
doi:10.1017/S0003055425101056 © The Author(s), 2025. Published by Cambridge University Press on behalf of American Political Science Association. 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.

Race, Democracy, and Public Support for War

MICHAEL TOMZ *Stanford University, United States*
JESSICA L. P. WEEKS *University of Wisconsin–Madison, United States*

Studies have found that voters in democratic countries are far more reluctant to use military force against democracies than against nondemocracies. This pattern may help explain why democracies almost never wage war against other democracies. In an important contribution, Rathbun, Parker, and Pomeroy (2025) propose that the apparent democratic peace in public opinion is an artifact of failing to account for race. Rather than democracy itself influencing support for war, they argue, the term "democracy" cues assumptions about the adversary's racial composition, and those racialized assumptions are the true drivers of support for war. We reevaluate RPP's evidence, concluding that their data do not support their predictions. In fact, their novel experiments provide powerful evidence that democracy affects support for war, independent of race. Our findings contribute to major debates about both regime type and race in international relations, as well as the design and interpretation of survey experiments.

INTRODUCTION

One of the most influential findings in international relations is the "democratic peace," the discovery that democracies rarely, if ever, wage wars against other democracies. Although scholars have debated why this phenomenon occurs, one possible explanation involves public opinion. Perhaps democracies have remained at peace partially because their voters are more reluctant to use military force against democracies than against nondemocracies. This idea is supported by evidence from numerous survey experiments across a wide variety of countries (e.g., Bassan-Nygate et al. 2024; Chaudoin, Gaines, and Livny 2021; Johns and Davies 2012; Tomz and Weeks 2013).

In a recent article, Rathbun, Parker, and Pomeroy (2025) challenge this growing consensus by arguing that the democratic peace in public opinion is an artifact of failing to account for racism. Past studies, they contend, have mistakenly concluded that democracy influences public support for war. "Rather than regime type per se doing the causal work, the term 'democracy' inadvertently primes the presumption that target countries are predominantly white. This implicit racialization, in turn, explains the reluctance of the American public to support aggression against fellow democracies…." (RPP, 621). Based on original experiments and analyses of English-language texts, they conclude that controlling for race "erases the effect of democracy" (RPP, 625) and that "the democratic peace in public opinion is, largely, an ethnocentric and racialized peace" (RPP, 621).

Establishing whether race explains the democratic peace in public opinion has profound implications for the study and practice of international relations. In recent years, scholars have turned increasing attention to the role of race in IR. Studies have shown how race and racial bias influence the structure of the international system, the content of international law, norms of good governance, patterns of foreign aid and migration, attitudes about war, the behavior of international organizations, and elite perceptions of foreign countries (e.g., Baker 2015; Buzas 2013; 2021; Carson, Min, and Van Nuys 2024; Freeman 2023; Freeman, Kim, and Lake 2022; Green-Riley and Leber 2023; Grovogui 1996; Mills 1997; Oksamytna and von Billerbeck 2024; Rosenberg 2019; Vitalis 2018; Zvobgo and Loken 2020; but cf. Chu and Lee 2024). RPP contribute to this growing literature by suggesting another way race could matter: by driving the alleged effect of democracy on peace in widely cited survey experiments.

Knowing whether race drives the apparent democratic peace in public opinion is vitally important for at least three reasons. First, disentangling the effects of democracy versus race would shed light on foundational theories about the causes of war. Based on previous experiments, researchers proposed public opinion as a pathway through which shared democracy dampens interstate conflict. If, however, previous findings were an artifact of overlooking race, scholars would need to turn elsewhere to explain peace among democracies.

Second, studying the relative importance of democracy versus race could enhance our understanding of outcomes beyond war. Countless theories of international relations involve regime type. Scholars have,

Michael Tomz, William Bennett Munro Professor in Political Science, Department of Political Science, Stanford University, United States; Senior Fellow, Stanford Institute for Economic Policy Research, Stanford University, United States, tomz@stanford.edu.
Corresponding author: Jessica L. P. Weeks, Professor of Political Science and H. Douglas Weaver Chair of Diplomacy and International Relations, Department of Political Science, University of Wisconsin–Madison, United States, jweeks@wisc.edu.

Handling editor: Scott de Marchi.

Received: October 09, 2024; revised: March 12, 2025; accepted: March 12, 2025. First published online: October 08, 2025.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

therefore, used survey experiments to estimate the effects of democracy on reputations for resolve (Renshon, Yarhi-Milo, and Kertzer 2023), public support for free trade (Carnegie and Gaikwad 2022; Chen, Pevehouse, and Powers 2023), and public attitudes about regulating foreign direct investment (Chilton, Milner, and Tingley 2020), among other themes. If race explains the democratic peace in public opinion, that might cast doubt on previous findings about the effects of democracy in other spheres of international relations.

Finally, if racialized attitudes are "hidden in plain sight" (Henderson 2013) in experiments involving democracy, this would not only upend how scholars interpret past studies, but also transform the design of future experiments. Assuming RPP are correct, experiments that vary a country's regime type would need to provide information about—or perhaps even randomize—the country's racial composition, as well. In summary, understanding the role of race in the democratic peace could have wide-ranging implications for theories about war, reputations, and international economic relations, as well as for the proper design and interpretation of experiments in IR.

The claims advanced by RPP therefore deserve careful scrutiny. We begin by distilling RPP's argument into three key claims: that a country's regime type affects public perceptions about the racial composition of its population; that the racial composition of a country's population affects public support for striking the country militarily; and that these racialized beliefs explain most if not all of the effect of regime type in survey experiments about the democratic peace. Refuting *any one* of these three claims would invalidate RPP's conclusion that "the democratic peace in public opinion owes, in large part, to racialized assumptions about democracy" (RPP, 621).

In addition to advancing three key claims about the population as a whole, RPP make an ancillary prediction about the behavior of certain subgroups: "these dynamics should be most pronounced for individuals scoring highest in ethnocentrism" (RPP, 622). Showing that the dynamics hold even more strongly among ethnocentric segments of society would, they reason, suggest that ethnocentrism "underpins" and "drives" the racialization. Failing to find differences by ethnocentrism would not, by itself, invalidate the main conclusion that "racialization explains much of the democratic peace in public opinion" (RPP, 630), but it would cast doubt on the ancillary claim that "ethnocentrism disproportionately explains the racialization of democracy" (RPP, 627).

Although RPP's theory is plausible, we conclude that their data do not support any of its predictions. First, describing a country as democratic does not cause ordinary people to spontaneously assume that the country has a majority-white population. Our reanalysis of data collected by RPP and others shows that survey respondents almost never make unprompted connections between a country's regime type and its racial makeup, and text analyses show no special association between references to democracy and words conveying whiteness. Second, RPP's survey experiments provide powerful, causally identified evidence *against* their prediction that assumptions about race shape support for war. Respondents were not significantly more supportive of attacking a country with a majority nonwhite population than a country with a majority white population. Third, consistent with these null results, we find no evidence that racialized beliefs explain the democratic peace or that accounting for race "erases" the effect of democracy. Rather, voters shrink from attacking democracies even when race is explicitly taken into account. Finally, RPP's data do not show that ethnocentrism causes respondents to behave in the racialized ways predicted by their theory.

These findings have both substantive and methodological implications. Substantively, our findings support the longstanding view that democracy plays an important causal role in international relations—here, by shaping the willingness of U.S. voters to attack an adversary. Although future surveys in the United States and other countries might suggest that democracy affects public support for war by changing perceptions of race, our reexamination of RPP's data shows that accounting for race does little to change widely accepted conclusions about how regime type influences public attitudes toward war. Thus, RPP's data strengthen, rather than undermine, existing insights about how the public thinks about international relations. Methodologically, our analyses clarify how future studies could document the role of race in international relations, whether through survey experiments fielded on the mass public or textual analyses of everyday language.

THE POTENTIAL ROLE OF RACE IN EXPERIMENTAL RESEARCH ABOUT THE DEMOCRATIC PEACE

Early survey experiments about the democratic peace (e.g., Johns and Davies 2012; Tomz and Weeks 2013) presented subjects with a hypothetical scenario involving a foreign country on the brink of developing nuclear weapons. The experiments manipulated information about the foreign country's regime type: respondents read that the country was either a democracy or a nondemocracy.[1] Participants were then asked whether they supported military strikes against the foreign country, and, in some cases, follow-up questions designed to measure causal mechanisms (Tomz and Weeks 2013).

These "now-canonical" studies (RPP, 622), fielded in the United States and the United Kingdom, found that voters were significantly less willing to attack democracies than autocracies, and that regime type shaped

[1] In Johns and Davies (2012), respondents also received randomized information about religion (Islamic versus Christian) and the expected civilian death toll. In Tomz and Weeks (2013), respondents also received randomized information about alliances, trade, and military power.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

perceptions of the threat the country posed and the morality of attacking. Subsequent studies employing variations on this design across a wide range of countries have consistently found a pacifying effect of democracy (e.g., Bassan-Nygate et al. 2024; Chaudoin, Gaines, and Livny 2021; Dafoe, Zhang, and Caughey 2018; Kim, Kim, and Kwak 2024; Lacina and Lee 2013).

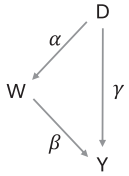
RPP contend, however, that the effect of democracy on public attitudes arises not because the target's regime type *per se* influences preferences about war, but because describing a hypothetical country as a democracy triggers assumptions about the racial composition of the target country, and those racialized assumptions are the true drivers of support for war. RPP claim that these racialized patterns are evident in the population as a whole and especially pronounced among individuals scoring high in ethnocentrism.

Figure 1 illustrates RPP's argument. In this causal diagram, the letter D represents the key treatment variable: whether the experiment randomly characterizes the potential adversary as a democracy $(D = 1)$ or not a democracy $(D = 0)$. The letter W represents the respondent's belief that the population of the potential adversary is "mostly white" $(W = 1)$ or mostly nonwhite $(W = 0)$, and the letter Y represents the dependent variable, the respondent's level of support for going to war against the country described in the experiment.

In Figure 1, the total effect of democracy on support for war can be decomposed into two paths: an indirect racialized path and a direct nonracialized path. Along the indirect path proposed by RPP, democracy changes beliefs about race $(D \rightarrow W)$, which in turn affect support for war $(W \rightarrow Y)$. The parameters α and β quantify each step in this causal chain. α measures the effect of democracy on the perception that the country is predominantly white, and β measures how perceptions of whiteness, in turn, affect support for war. The product $\alpha \times \beta$ represents the effect of democracy that operates through this indirect pathway by influencing the mediating variable W.

Figure 1 also includes a direct path, $D \rightarrow Y$, encompassing any effects of democracy that do not operate by changing beliefs about the racial composition of the target. This direct path represents the aggregation of various nonracial mechanisms discussed in previous observational and experimental work on the democratic peace (e.g., Maoz and Russett 1993; Johns and Davies 2012; Tomz and Weeks 2013). The parameter γ summarizes the importance of this path; it quantifies how D affects Y through all other processes that do not involve the mediator W. The total effect of democracy on support for war in this diagram is the sum of the racialized and nonracialized paths, $\alpha \times \beta + \gamma$.

FIGURE 1. Causal Mechanisms in RPP's Theory



The argument advanced by RPP entails three key implications for the general population. First, the term "democracy" must cause people to assume that the target country is predominantly white $(\alpha > 0)$. Second, respondents must be less willing to use military force against a country whose population is primarily white than against a country whose population is primarily nonwhite $(\beta < 0)$. Third, if the democratic peace is predominantly if not entirely a racialized peace, then most or all of the effect of democracy should flow through the racialized path $\alpha \times \beta$, rather than the nonracialized path γ.[2] These three predictions are logically sound and theoretically plausible, in light of past research documenting numerous ways in which racial bias affects perceptions of foreign countries and support for policies directed at them.

Importantly, *all three* predictions must hold to sustain RPP's conclusion that "racialization explains the pacifying effect of democracy in previous public opinion work on the democratic peace" (RPP, 630). If information about democracy did not shape perceptions of whiteness, race could not explain the democratic peace, even if information about race affected support for war (for a proof, see Appendix Proposition 1). Likewise, if information about race did not shape support for war, race could not explain the democratic peace, even if democracy generated assumptions of whiteness (for a proof, see Appendix Proposition 2). Finally, if most of the effects of democracy flowed through nonracialized pathways, one could not argue that the democratic peace was primarily a racialized peace.

RPP also make an ancillary prediction: "these dynamics should be most pronounced for individuals scoring highest in ethnocentrism" (RPP, 622). Applying the notation from Figure 1, ethnocentrism should intensify the effect of regime type on racial beliefs (α), the effect of racial beliefs on support for military force (β), and the importance of the racialized pathway ($\alpha \times \beta$) relative to the nonracialized pathway (γ). Testing whether ethnocentrism moderates these links in the causal graph "allows for better confirmation of the ethnocentric and culturally superior nature of any preference shown to democracies based on racial presumptions" (RPP, 626). If racialized patterns were not significantly stronger among respondents with ethnocentric worldviews, this would not necessarily negate the main conclusion that "racialization of democracy drives much of the 'democratic' peace in public opinion" (RPP, 621), but it would call into question the idea that ethnocentrism "underpins" and "drives" the racialization. In the remainder of this article, we revisit RPP's data to reevaluate each of their claims.

[2] Both paths are hypothesized to be negative, but RPP's theory predicts that the racialized path is more powerful, such that $\alpha \times \beta < \gamma$.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

DOES DEMOCRACY TRIGGER PERCEPTIONS OF WHITENESS?

RPP's first key claim is that democracy conveys impressions of whiteness ($\alpha > 0$ in Figure 1). They offer two pieces of evidence to support this idea: previous research by Dafoe, Zhang, and Caughey (2018) and an original analysis of word embeddings. We reexamine both pieces of evidence, concluding that neither supports the claim that democracy conveys impressions of whiteness.

Revisiting Evidence from the Dafoe, Zhang, and Caughey Survey

RPP's first piece of evidence that democracy suggests whiteness relies on a survey by Dafoe, Zhang, and Caughey (2018). Dafoe, Zhang, and Caughey replicated the canonical democratic peace experiment, with a twist: after presenting vignettes in which the country was randomized to be a democracy or a nondemocracy, they inserted a series of "placebo" questions asking respondents to provide their "best guess of what the country was like 10 years ago" along nine dimensions not explicitly mentioned in the vignette.

One placebo question prompted respondents to guess the racial composition of the country: "How likely do you think it is that the majority of the country's population was white (Caucasian)?" The democracy treatment produced a statistically significant increase in beliefs that the country was majority white (Dafoe, Zhang, and Caughey 2018, 409). RPP cite this finding as indicating that "survey research shows that western respondents presume that democracies are white" (RPP, 621).

However, explicitly asking respondents to guess the country's racial makeup may cause respondents to make conjectures about race that would not have occurred to them unprompted. A better test would involve analyzing what respondents said about democracies and nondemocracies when they had not been primed to think about race. RPP proposed exactly this test: "if we asked members of the public to describe democracies versus nondemocracies, we would expect terms like white or nonwhite to, respectively, coappear, despite the fact that we did not explicitly ask respondents about racial characteristics" (RPP, 631).

Although RPP did not carry out the proposed test, data from Dafoe, Zhang, and Caughey allowed us to implement it. Dafoe, Zhang, and Caughey included an open-ended question: "Think about the scenario you read. Write down what you think the country in the scenario is like. Write down at least five things that come to your mind." Dafoe, Zhang, and Caughey randomized the placement of this open-ended question, with some respondents encountering it before they were prompted to think about race. *Unprompted* respondents answered the open-ended question before placebo questions about race and other country characteristics, and before manipulation checks that suggested an association between regime and race by listing white countries as exemplars of democracies and nonwhite countries as exemplars of nondemocracies.[3] *Partially prompted* respondents encountered the open-ended question before the placebo questions but potentially after manipulation checks that could have primed race indirectly. We tested whether these two sets of respondents made spontaneous associations between regime type and race.[4]

Two researchers independently hand-coded the open-ended answers, without knowing which regime type treatment the respondent had received.[5] Coders began by judging whether each entry was relevant to the invitation to describe the country in the vignette.[6] We based all analyses on relevant entries, to ensure that non-mentions of race reflected real attempts at describing the country.

Coders next checked for references to race, using two approaches. First, they looked for any mention of skin color (white, nonwhite, Black, light-skinned, etc.), with instructions to code such mentions as referring to either white or nonwhite populations. As Table 1 shows, coders found *no* mentions of skin color among the hundreds of entries they coded. This conclusion held not only among unprompted respondents, but even among partially prompted respondents who potentially answered the open-ended question after manipulation checks that gave mostly white countries as examples of

TABLE 1. Explicit References to Skin Color in Open-Ended Responses

	Unprompted		Partially prompted	
	Nondem	Dem	Nondem	Dem
No mentions	147 100.0%	150 100.0%	235 100.0%	213 100.0%
"White" mentions	0 0.0%	0 0.0%	0 0.0%	0 0.0%
"Nonwhite" mentions	0 0.0%	0 0.0%	0 0.0%	0 0.0%
Total	147 100.0%	150 100.0%	235 100.0%	213 100.0%

[3] The manipulation checks asked, "How likely do you think it is that the country has the following types of government?" and gave examples of each government type. The "fully democratic" examples included three white countries (Canada, Germany, and the United Kingdom) and only one nonwhite country (Japan). The "fully nondemocratic" examples were China, North Korea, Iran, Saudi Arabia, and Vietnam.

[4] Our analysis excluded "Embedded Natural Experiment" (ENE) vignettes, which were designed not to convey information about factors potentially correlated with democracy. Including ENEs would have stacked the deck against finding that democracy conveys whiteness.

[5] See SM for coding procedures.

[6] Answers were judged as "relevant" if they offered at least one word or phrase that could describe a country. Compliance was high: 91% of both unprompted (297 out of 326) and partially prompted (448 out of 490) participants provided a relevant answer to the open-ended question. Among cases coded as irrelevant, all but two were instances in which the respondent had left the question blank.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

TABLE 2. Explicit References to Skin Color Plus Geographic and Ethnic Adjectives in Open-Ended Responses

	Unprompted		Partially prompted	
	Nondem	Dem	Nondem	Dem
No mentions	143 97.3%	141 94.0%	228 97.0%	203 95.3%
"White" mentions	0 0.0%	3 2.0%	0 0.0%	3 1.4%
"Nonwhite" mentions	4 2.7%	6 4.0%	7 3.0%	7 3.3%
Total	147 100.0%	150 100.0%	235 100.0%	213 100.0%

democracies and nonwhite countries as examples of nondemocracies.

Second, we broadened the coding procedures to include not only references to skin color, but also geographic and ethnic adjectives such as "Iranian," "Middle Eastern," "Eastern European," or "Western." A downside of this approach is that it classifies every use of such adjectives as indicating that a subject was thinking about race, even though race might not have occurred to a respondent. Nonetheless, Table 2 shows that extremely few respondents (fewer than 1 out of every 20, on average) volunteered terms that could be interpreted as either racial or ethnic descriptors, and their mentions bore little relation to regime type. In summary, our analysis of open-ended responses from Dafoe, Zhang, and Caughey showed that democracy did not prompt spontaneous impressions of whiteness.

Rather than conducting the test they proposed, RPP cited Dafoe, Zhang, and Caughey's finding that when people were *explicitly asked to guess* about race, they guessed "white" more often if the country was a democracy. Even this problematic test provides little support for RPP's theory, however. Using data from Dafoe, Zhang, and Caughey, we calculated the percentage of respondents in each regime type condition who guessed that it was "likely" or "very likely" that the country was majority white.[7] Approximately 16% made this guess when the country was a democracy, compared to 6% when the country was not a democracy. Although the percentage guessing white was higher in the democracy condition, the fact that only 16 out of 100 guessed white when presented with a democracy contradicts the claim that "survey research shows that western respondents presume democracies are white." (RPP, 621)[8]

[7] Again, we excluded ENE vignettes.

[8] Moreover, Dafoe, Zhang, and Caughey found that the estimated effect of democracy on support for war became slightly *larger* after using their ENE experimental design, which they designed to prevent people from associating democracy with race and other variables. Thus, evidence from Dafoe, Zhang, and Caughey *directly contradicts* the conclusion that the democratic peace in public opinion is actually a racialized peace.

Revisiting Evidence from RPP's Analysis of Word Embeddings

As a second way of testing whether democracy is associated with whiteness, RPP analyzed correlations among words in a large corpus of English texts. If democracy is associated with whiteness, they argue, references to democracy should coincide with references to whiteness "in the English language writ large" (RPP, 630). Such a pattern could suggest that democracy and whiteness are related in the minds of the authors who produced the texts, and by extension, among potential survey respondents, because "textual data … reflect the same sorts of biases and implicit associations that we would typically measure via survey methods" (RPP, 631).

To measure the correspondence between democracy and whiteness in everyday language, RPP relied on word embeddings. Word embedding models convert individual words into *N*-dimensional numerical vectors based on the contexts in which the words are used (e.g., Rodriguez and Spirling 2022). Words used in similar contexts have vectors close to each other, while words with fewer shared contexts have more distant vectors. RPP used word vectors from the GloVe model, which was trained on a corpus of contemporary English language online texts, including newswires and Wikipedia (Pennington, Socher, and Manning 2014).

RPP first defined lists of terms for democracy, white, and nonwhite (RPP, 631):

- Democracy terms: democracy, democratic, democratically, elect, elections, elected.
- White terms: white, western, caucasian, european.
- Nonwhite terms: non-white, non-western, non-caucasian, non-european.

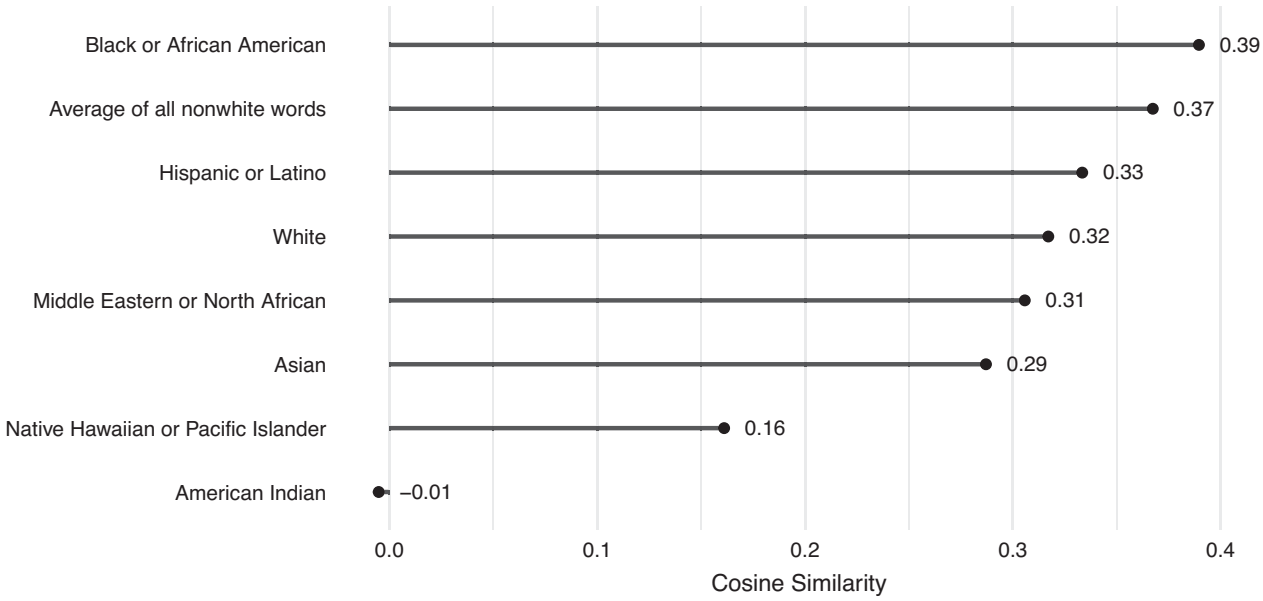
RPP then tested for an association between democracy and whiteness by measuring the cosine similarity between the democracy vector and each of the racial vectors. Cosine similarity, defined as the cosine of the angle between two vectors, is a common measure of similarity in text analysis. The closer the value is to 1, the greater the match between the two vectors. RPP found that the cosine similarity between democracy terms and white terms was 0.389, whereas the cosine similarity between democracy terms and nonwhite terms was –0.003. RPP interpreted this pattern as suggesting "that democracy implicitly associates with whiteness in this corpus of English language texts" (RPP, 631).

This finding is extremely fragile, however. RPP did not cite sources to justify their lists of racial terms or report whether their conclusions were robust to using other words to describe racial groups. As we show below, replacing their list of racial terms with a more authoritative list reverses their conclusion.

For a more authoritative set of racial terms, we turned to the U.S. government's race and ethnicity reporting categories. The U.S. government defines seven main racial categories: White, Asian, Black or African American, Hispanic or Latino, Middle Eastern or North African, Native Hawaiian or Pacific Islander,

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

FIGURE 2. Cosine Similarity of Official U.S. Race/Ethnicity Categories with Democracy Words



and American Indian or Alaska Native.[9] The U.S. definitions also include six examples for each racial category. For instance, the official definition of "Black or African American" lists African American, Jamaican, Haitian, Nigerian, Ethiopian, and Somali as examples. We constructed lists of racial terms based on these official definitions and examples, with minor adjustments to permit analysis using GloVe (see SM). We then replicated RPP's procedure by calculating the cosine similarity between their list of democracy terms and each list of U.S. government race words. Figure 2 plots the results, with cosine similarities on the X-axis.

Contrary to RPP's theory, democracy was not more closely associated with whiteness than with other racial categories. In fact, the democracy terms were closer in vector space to Black and Hispanic words than white words, and the cosine similarity between democracy and the average of all nonwhite terms was larger than the cosine similarity between democracy and the average of all white terms.

In summary, we revisited both Dafoe, Zhang, and Caughey's survey data and RPP's word embeddings analysis to probe the first step in RPP's causal pathway: that democracy primes assumptions of whiteness. Our analyses provided little support for this idea, either among respondents in democratic peace experiments or in everyday English usage. Given that *each* link in the causal chain is essential to RPP's broader claim that race explains the democratic peace in public opinion, these non-findings contradict their core conclusion.

Although RPP themselves proposed the two empirical approaches we used (analysis of open-ended responses and word embeddings), it is possible that neither was sensitive enough to detect subtle associations between democracy and whiteness. Perhaps respondents in the Dafoe, Zhang, and Caughey survey were thinking about race, but were unaware of their own racialized thinking or consciously avoided mentioning race in open-ended responses. Similarly, perhaps word embeddings failed to detect subtle, potentially unconscious, and therefore unwritten connections between democracy and whiteness. Future research could test for racialized assumptions about regime type using tools designed to measure implicit and/or unconscious biases, such as the Implicit Association Test (IAT) (Greenwald, McGhee, and Schwartz 1998), the Affect Misattribution Procedure (AMP) (Payne et al. 2005), and others (for reviews, see Gawronski and De Houwer 2014 and Stark, Krosnick, and Scott 2025). For now, though, there is little if any evidence that respondents associate democracy with whiteness.

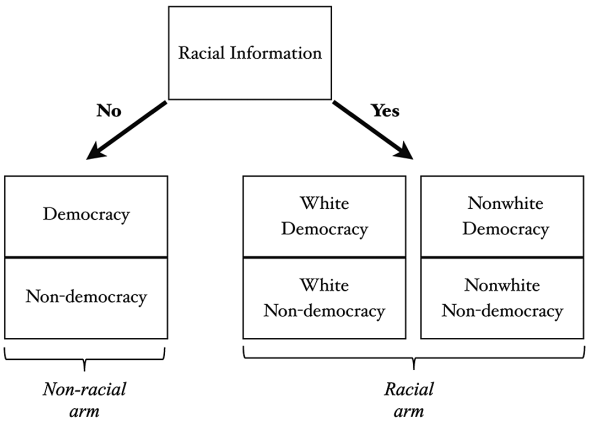
DOES RACE AFFECT SUPPORT FOR WAR?

RPP's theory requires not only that "the term 'democracy' inadvertently primes the presumption that target countries are predominantly white," but also that "this implicit racialization, in turn, explains the reluctance of the American public to support aggression among fellow democracies" (RPP, 621). Survey respondents must be less likely to support war against a country they perceive as predominantly white than against one they believe is predominantly nonwhite. In terms of Figure 1, the data must show that β is negative. If experiments failed to reject the null hypothesis that

[9] https://www.federalregister.gov/documents/2024/03/29/2024-06469/revisions-to-ombs-statistical-policy-directive-no-15-standards-for-maintaining-collecting-and. See page 22191, "Categories and Definitions," for the formal definitions.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

FIGURE 3. RPP's Experimental Design (Survey 1)



Note: Reproduced from RPP (625), with labels to identify the nonracial and racial arms

$\beta = 0$, this alone would undermine RPP's conclusions, whether or not we observed that the term "democracy" primed respondents to think about the target's race.

Fortunately, RPP designed and administered highly innovative experiments to shed light on their key predictions, including the hypothesis that $\beta < 0$. In their main experiment (Survey 1: Qualtrics Sample), some respondents were assigned to the nonracial arm (left side of Figure 3), which randomized information about the target country's political regime but provided no information about its racial composition. Others were assigned to the racial arm (right side of Figure 3), which randomized information not only about regime but also about race, resulting in four types of targets: white democracies, white nondemocracies, nonwhite democracies, and nonwhite nondemocracies. Using data from the racial arm of Survey 1, one can estimate not only the effect of regime but also the effect of race. Notably, this approach can detect the effects of racial bias, even if respondents remain completely unaware of it.

Before running this experiment, RPP preregistered the testable implications of their theory.[10] Their first hypothesis was "H1: Main effect. As previous research has found, we hypothesize that subjects randomly told that a target state is a democracy will be less inclined to support the use of force against that target state, relative to subjects told that the target state is a nondemocracy." Their second hypothesis was "H2: Main effect: Subjects randomly told that a target state has a majority white population will be less inclined to support the use of force against that target state, relative to subjects told that the target state has a majority nonwhite population" (RPP Preregistration, 1–2). RPP said they would "use linear regression to test H1 and H2, estimating the main effect of the race treatment (white versus nonwhite) and regime type (democracy or nondemocracy) on each DV," with the key dependent variable being support for military strikes against the target country (RPP Preregistration, 13).

[10] https://osf.io/x4wvf/.

TABLE 3. Effect of Democracy and Race on Support for Military Strikes

Country Is a Democracy	–0.25*
	(0.11)
Country Race Is White	–0.15
	(0.11)
Intercept	3.93**
	(0.09)
Number of observations	1087

Note: Estimates from linear regression of support for military strikes on indicators for whether the country is a democracy (1 = democracy, 0 = nondemocracy), and whether the race of the country is specified as white (1 = white, 0 = nonwhite). Robust standard errors in parentheses. Data from racial arm of RPP Survey 1. ** $p < 0.01$, * $p < 0.05$.

In Table 3, we carried out RPP's preregistration by analyzing the racial arm of the experiment, regressing support for military strikes on indicator variables for democratic institutions (1 if democracy, 0 if nondemocracy) and white population (1 if predominantly white, 0 if predominantly nonwhite). Consistent with H1, the estimated coefficient on democracy was negative and statistically significant ($t = 2.41$, $p = 0.02$). Thus, RPP's data show that U.S. voters were less willing to use military force against a democracy than against an autocracy, *even* when taking race into account. However, contrary to H2, the estimated coefficient on white was statistically indistinguishable from zero ($t = 1.43$, $p = 0.15$). Thus, based on RPP Survey 1, we cannot reject the null hypothesis that $\beta = 0$.

The published article barely mentions this null finding. RPP refer to it only in the concluding paragraph, which directs readers to their Dataverse Appendix B2. There, they report "a notable absence of an independent effect of race on support for strikes" but do not point out that this null finding contradicts both their preregistration and their theory.

In sum, RPP's first experiment tests a crucial preregistered implication of their theory: support for military strikes should be lower among respondents told that the target was majority white, than among respondents told that the target was majority nonwhite, i.e., β should be negative. The data do not support this prediction. Without evidence of this essential link in the causal chain, we cannot conclude that the democratic peace is a racialized peace.

HOW MUCH DOES DEMOCRACY MATTER AFTER EXPERIMENTALLY CONTROLLING RACE?

Comparing the Two Arms of the Experiment

RPP's third key claim is that democracy affects support for war mainly through the racialized path $\alpha \times \beta$, rather

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

than the nonracialized path γ. RPP said they would test the strength of these two paths by comparing the effect of democracy in the experimental arm that withheld racial information about the target country (nonracial arm) to the effect of democracy in the experimental arm that specified the race of the target (racial arm). The difference between these two values quantifies how much of the original effect of democracy was eliminated by providing racial information. According to RPP, "large and significant eliminated effects indicate that race plays a role in the mechanism that links democracy to peace" (RPP, 625).[11]

The intuition is that, by specifying the race of the target instead of letting respondents make unprompted assumptions about race, one can knock out the racial mechanism in Figure 1 (Acharya, Blackwell, and Sen 2018). Mathematically, assigning the race of the target eliminates any correlation between D and W, thereby setting $\alpha = 0$ and zeroing-out the indirect $\alpha \times \beta$ pathway. Assigning race should, therefore, eliminate the portion of the democracy effect that was operating via race. Any remaining effect of democracy would represent the "controlled direct effect," that is, the direct effect of democracy (γ) after controlling for the race of the target.

Hence, RPP preregistered "H3: Controlled direct effect: Because regime type might convey information about race (more specifically respondents might believe democracies to be white), we will find a reduction in the main effect of democracy when we take into account the indirect effect of racial characteristics of the target population" (RPP Preregistration, 2). They committed to test this hypothesis with data from Survey 1 "by comparing the treatment effect of democracy in the arm of the survey that only receives information about regime type against the treatment effect of democracy in the arm of the survey that also manipulates the racial characteristics of the target country" (RPP Preregistration, 13). Thus, RPP said they would compare the average effect of democracy among all respondents assigned to the nonracial arm of Figure 3 to the average effect of democracy among all respondents assigned to the racial arm of Figure 3.[12]

RPP did not report the results of this test. To carry it out, we regressed support for military strikes on whether the country was a democracy (1 if yes, 0 if no), whether the respondent was assigned to the racial arm (1 if yes, 0 if no), and the interaction of the two. The coefficient on democracy represents the effect of democracy in the nonracial arm, and the coefficient on the interaction quantifies whether/how the effect of democracy changed as a result of specifying the country's race.

TABLE 4. Effect of Democracy, by Whether Race Was Specified

Country Is a Democracy	–0.31*
	(0.14)
Country Race Is Specified	–0.20
	(0.13)
Country Is a Democracy x Country Race Is Specified	0.06
	(0.18)
Intercept	4.05**
	(0.10)
Number of observations	1626

Note: Estimates from linear regression of support for military strikes on indicators for whether the country is a democracy (1 = democracy, 0 = nondemocracy), whether the race of the country is specified (1 = yes, 0 = no), and their interaction. Robust standard errors in parentheses. Data from RPP Survey 1. ** $p < 0.01$, * $p < 0.05$.

As in previous survey experiments about the democratic peace, the coefficient on democracy should be negative, meaning that citizens assigned to the nonracial arm are less willing to strike a democracy than a nondemocracy. H3 would be supported if, in addition, we found a positive and significant coefficient on the interaction, indicating that the effect of democracy was partially eliminated by specifying race.

Using RPP's data, though, we found no support for H3. Table 4 shows that the estimated coefficient on the interaction was not only substantively small (0.06) but also statistically insignificant ($t = 0.34$, $p = 0.74$). Thus, we cannot reject the null hypothesis that democracy was equally consequential in the racial and nonracial arms of the experiment.

Table 4 also allows us to calculate two other quantities that RPP mention. The "average controlled direct effect," defined as the effect of democracy after fixing race as white for half of the racial arm and nonwhite for the other half of the racial arm, is the coefficient on democracy plus the coefficient on the interaction, resulting in $-0.31 + 0.060 = -0.25$ ($t = 2.41$, $p = 0.02$). The "eliminated effect," defined as the effect of democracy eliminated by providing racial information, is the negative of the coefficient on the interaction, –0.06 ($t = 0.34$, $p = 0.74$). These quantities further clarify that controlling for race did not change the effect of democracy.

Splitting the Racial Arm of the Experiment

The previous section showed that the effect of democracy was the same in the nonracial and racial arms of the experiment. Rather than reporting this preregistered test, RPP split the racial arm into white and nonwhite treatment groups. First, they compared the effect of democracy when the target was characterized as white to the effect of democracy in the nonracial arm.

[11] This method can detect the effects of racial biases—in this case, the combination of an association between regime type, and a bias against attacking majority-white countries—even when respondents are unaware that they are engaged in racialized thinking.

[12] More specifically, this comparison pits the *average effect* of democracy (given by the nonracial arm) against the *average controlled direct effect* of democracy (given by the racial arm). The racial arm gives an average with approximately equal weight to white and nonwhite targets, because RPP assigned half of the respondents in the racial arm to consider a white target and the other half to consider a nonwhite target. For formal definitions of conditional and average controlled direct effects, see Appendix.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

TABLE 5. Effect of Democracy, by Whether Race Was Specified as White or Nonwhite

	1	2	3
Country Is a Democracy	−0.31*	−0.31*	−0.56**
	(0.14)	(0.14)	(0.10)
Country Race Is White	−0.24		
	(0.15)		
Country Is a Democracy x Country Race Is White	0.00		
	(0.21)		
Country Race Is Nonwhite		−0.15	−0.09
		(0.14)	(0.10)
Country Is a Democracy x Country Race Is Nonwhite		0.12	0.25
		(0.20)	(0.14)
Intercept	4.05**	4.05**	3.85**
	(0.10)	(0.10)	(0.07)
Number of observations	1086	1079	2659

Note: Estimates from linear regression of support for military strikes on indicators for the country's regime type, the country's race, and their interaction. In model 1, "Country Race Is White" is coded 1 if race was specified as white, and 0 if race was not specified. In models 2 and 3, "Country Race Is Nonwhite" is coded 1 if race was specified as nonwhite, and 0 if race was not specified. Robust standard errors in parentheses. Models 1 and 2 are based on RPP Survey 1. Model 3 is based on RPP Survey 2. ** $p < 0.01$, * $p < 0.05$.

Second, they compared the effect of democracy when the target was characterized as nonwhite to the effect of democracy in the nonracial arm. We reproduce these comparisons, which again fail to show that race explains the effect of democracy.

We first used data from Survey 1 to test whether the effect of democracy weakened when the country was characterized as white, compared to when race was not mentioned (Table 5 model 1). The coefficient on the interaction between democracy and white was substantively and statistically indistinguishable from zero ($t = 0.01$, $p = 0.99$), meaning the effect of democracy did not budge when the experiment characterized the country as white instead of leaving race unspecified.

Although this null finding contradicts the preregistered prediction that controlling for race should diminish the effect of democracy, the published article characterizes the null finding as if it were an anticipated and logical implication of RPP's theory. According to the article, "If racialization explains the democratic peace in public opinion, then the effect size of democracy with and without white information will be similar—whiteness will not eliminate the effect of democracy. In the absence of explicit racial information, respondents are likely to assume that a democratic country is majority white compared to a nondemocratic country, such that any racial information about whiteness in the context of democracy is superfluous" (RPP, 625–6).

This unregistered hypothesis is logically flawed, because it overlooks how information about whiteness would affect beliefs about *non*democracies. According to the published article, mentioning that a democracy is predominantly white would constitute "superfluous" information, because when respondents read about a democracy, they automatically assume a white population. Thus, "the provision of white information provides no additional information beyond democracy" (RPP, 622). But following this same logic, a vignette describing a nondemocracy as predominantly white should be highly informative, rather than superfluous; it should reverse the "nonwhite" assumption respondents allegedly would have made if the vignette had described a nondemocracy without mentioning race. Thus, under RPP's theory, characterizing *both* types of regimes as white should *eliminate any perceived racial difference* between democracies and nondemocracies. If the democratic peace in public opinion remains the same after removing this racial difference, then beliefs about race cannot explain why the distinction between democracies and nondemocracies affects public support for war.

For further intuition about the logical flaw in RPP's pivot to claiming, contra their preregistration, that fixing race as white should not diminish the effect of democracy, recall two important definitions. First, the *total effect* of democracy on support for war when no racial information is given is the sum of the racialized and nonracialized paths, represented by $\alpha \times \beta + \gamma$ in Figure 1. Second, the *controlled direct effect* of democracy, i.e., the effect of democracy when racial information is given (by fixing the race of the target regardless of the target's regime type) is γ. Suppose these two quantities are equal, i.e., the effect of democracy remains the same after fixing the race of the target. In that case, the total effect equals the controlled direct effect, i.e., $\alpha \times \beta + \gamma = \gamma$. Cancelling γ from both sides gives $\alpha \times \beta = 0$, which can only be true if regime type has no effect on beliefs about the race of the target $(\alpha = 0)$ and/or the race of the target has no effect on support for the use of force after controlling for democracy $(\beta = 0)$. Either way, the indirect racial pathway would be zero.

For simplicity, the previous paragraph implicitly assumed no interaction between the treatment (democracy) and the mediator (race). An interaction could arise, for example, if the effect of race on support for war depended on whether the target was a democracy or not, such that $\beta|D = 1$ differed from $\beta|D = 0$. Other interactions between democracy and race are possible, as well. In the Appendix we relax the no-interaction assumption. Even then, Appendix Proposition 3 proves that, given the assumptions in RPP, if the total effect of democracy remains the same after setting the race of the country equal to white, then D has no effect on W (i.e., $\alpha = 0$) and/or W has no effect on Y conditional on nondemocracy (i.e., $\beta = 0|D = 0$). In summary, if the democratic peace in public opinion is similar with and without stipulating that the target is white, as in Table 5 model 1, this would substantially undermine RPP's claim that the democratic peace is a racialized peace.

Next, we consider the second of RPP's comparisons: whether the effect of democracy weakens when the country is characterized as nonwhite, compared to not mentioning anything about the country's race. As RPP note, their theory implies that "provision of nonwhite information should erode the democratic peace effect" (RPP, 622). We tested this prediction with Survey 1 by

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

regressing support for military strikes on democracy, the provision of nonwhite information, and the interaction of the two. The regression estimates in Table 5 model 2 show, however, that the coefficient on the interaction term was statistically insignificant ($t = 0.58$, $p = 0.56$). Thus, contrary to RPP's theory, characterizing the country as nonwhite did not significantly change the effect of democracy.

RPP wondered whether their first survey might have been underpowered, and therefore fielded a second survey with a larger sample. In Survey 2, some respondents received the nonracial arm; others received the portion of the racial arm describing the target country as nonwhite. Because Survey 2 omitted the white condition, it cannot shed additional light on the null effect in Table 5 model 1, but it can help us reexamine the null effect in Table 5 model 2.

Table 5 model 3 uses data from Survey 2 to compare the effect of democracy when the country was described as nonwhite to the effect of democracy when race was not mentioned. Although the coefficient on the interaction term in model 3 carried the expected sign, it remained just shy of statistical significance at conventional levels ($t = 1.84$, $p = 0.07$). Thus, even after leveraging the larger sample, we cannot confidently conclude that the effect of democracy is smaller when race is fixed at nonwhite than when race is not fixed at all.

Interpreting the Evidence

In every column of Tables 4 and 5, the interaction between democracy and racial information was statistically indistinguishable from zero, undermining RPP's theory. Suppose, however, that we took the estimated interaction coefficients at face value. Even then, the data would not indicate that race "erases" or "eliminates" the effect of democracy. In Table 4, the effect of democracy on support for strikes was $-0.31 + 0.06 = -0.25$ in the racial arm versus -0.31 in the nonracial arm. Hence, controlling for race would reduce the estimated effect of democracy by only 19%. Likewise, most of the original effect of democracy would remain even if one accepted the interaction coefficients in Table 5. Thus, the data provide no basis for concluding that race "erases" the effect of democracy.

Finally, even if the interaction terms in Tables 4 and 5 had been positive, statistically significant, and substantial in magnitude, this would have provided only suggestive evidence that race mediates—"explains"—the effect of democracy on support for war, as implied by RPP's theory (Figure 1). A positive and significant interaction could instead mean that race *moderates* the effect of democracy, i.e., that democracy has larger effects for some values of the race treatment than for others, a pattern known as causal interaction.[13] As a moderator, race would not "explain" the effect of democracy, but instead would amplify or reduce its effect.

FIGURE 4. The Mediator–Moderator Distinction

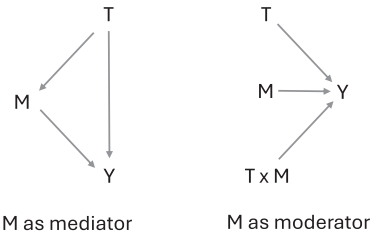


Figure 4 illustrates the conceptual difference (Baron and Kenny 1986). On the left panel, M mediates the relationship between the treatment T and the outcome Y, i.e., T causes M, which in turn causes Y.[14] On the right, M moderates the relationship between T and Y, either amplifying or dampening the effect of T on Y.[15] Importantly, the moderation graph contains no arrow from T to M, as required by RPP's theory.

These two possibilities are not mutually exclusive. A single variable M could both mediate and moderate the effect of the treatment on the outcome. Unfortunately, even if RPP's experiment had found a significantly positive interaction (i.e., that racial information significantly reduced the effect of democracy), we could not tell whether the interaction was due to mediation, moderation, or both.

RPP do not acknowledge this ambiguity. They write, "we use a randomized mediator design described by Acharya, Blackwell, and Sen (2018), which allows us to identify whether the effect of regime type on pacifism flows via racialized assumptions about the hypothetical target country" (RPP, 622). However, Acharya, Blackwell, and Sen (2018, 358) emphasize that their method "cannot separately identify the indirect effect and the causal interaction," where indirect effect refers to mediation and causal interaction refers to moderation. At best, their method can "recover a combination of the indirect and interaction effects induced by the manipulation" (358). Without additional assumptions, "disentangling the relative contribution of the indirect and interaction effects in contributing to the eliminated effect is impossible" (367).

In summary, RPP's data do not support their prediction that specifying race reduces the effect of democracy, and even if it had, one could not tell whether the pattern was evidence of mediation, moderation, or both. Recall, however, that we failed to find evidence that $\alpha > 0$ and $\beta < 0$, both necessary for the mediation

[13] For example, the target's regime type could be more important to respondents when the target's race is unknown than when the target's race is known, or when the target's race is nonwhite than when it is white.

[14] The graph on the left also allows the possibility that T could affect Y through mechanisms other than M.

[15] The graph on the right also allows the possibility that M could have an independent effect on Y, above and beyond its moderating effect on the relationship between T and Y.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

hypothesis. Thus, even if specifying race *had* reduced the effect of democracy, the evidence would suggest that race was acting as a moderator rather than a mediator: altering the magnitude of the democratic peace in public opinion, rather than explaining why such a peace exists.

ETHNOCENTRISM AND THE DEMOCRATIC PEACE

Above, we reevaluated RPP's three main claims: that democracy conveys information about whiteness; that voters are less willing to strike countries with white populations; and that accounting for race diminishes and even eliminates the democratic peace effect in public opinion. We now investigate RPP's contention that their main claims hold even more strongly among the most ethnocentric individuals.

Does Ethnocentrism Moderate the Effect of D on W?

RPP assert that "more ethnocentric individuals are likely to exhibit racialized assumptions" about the target country, including "the presumption that democracies are white" (RPP, 624). Unfortunately, the article provides no direct evidence for this hypothesis, and—to our knowledge—existing data do not permit a direct test. Although the surveys fielded by RPP measured ethnocentrism, they did not measure beliefs about the race of the target. And although the survey by Dafoe, Zhang, and Caughey (2018) elicited beliefs about race, it did not contain measures of ethnocentrism. Hence, we have no evidence that people scoring higher on ethnocentrism are more likely to presume that democracies are white.

Does Ethnocentrism Moderate the Effect of W on Y?

RPP also argue that ethnocentrism should moderate the effect of race (W) on support for war (Y). They write, "If the racialization of democracy is part of prominent cultural assumptions in the United States or even the west more broadly, even if implicitly, then more ethnocentric individuals who defend that culture will be more likely to make such presumptions *and act on them* through support for the use of force" (RPP, 626, emphasis added). Thus, "Subjects higher in self-reported ethnocentrism will be even more likely to support the use of force against hypothetical nonwhite target states" (RPP Preregistration, 2). They committed to "test H4 using linear regression by interacting the regime type and race treatments with self-reported ethnocentrism scores" (RPP Preregistration, 13).

Neither the article nor RPP's supplementary materials report whether RPP's experiments supported H4. To find out, we conducted their proposed test by analyzing the racial arm of Survey 1, which measured ethnocentrism while also randomizing the race of the target country (white or nonwhite).[16] Table 6 regresses support for military strikes on the race of the target country (1 if predominantly white, 0 if predominantly nonwhite), RPP's measure of ethnocentrism (1 if high, 0 if low), and the interaction of the two.

TABLE 6. Does Ethnocentrism Moderate the Effect of Race on Support for War?

Country Is a Democracy	–0.41**
	(0.15)
Country Race Is White	–0.15
	(0.15)
Ethnocentrism Above Median	0.15
	(0.18)
Country Is a Democracy x Ethnocentrism Above Median	0.35
	(0.21)
Country Race Is White x Ethnocentrism Above Median	0.01
	(0.21)
Intercept	3.85**
	(0.13)
Number of observations	1087

Note: Estimates from linear regression of support for military strikes on indicators for whether the country is a democracy (1 = democracy, 0 = nondemocracy), whether the race of the country is specified as white (1 = white, 0 = nonwhite), whether the respondent is above the ethnocentrism median (1 = yes, 0 = no), and interactions between ethnocentrism and the democracy and race treatments. Robust standard errors in parentheses. Data from racial arm of RPP Survey 1. ** $p < 0.01$, * $p < 0.05$.

H4 predicted that the interaction between white population and ethnocentric respondents should be negative, meaning that ethnocentric respondents would be especially averse to using force against white targets. However, the estimated coefficient on "Country Race Is White × Ethnocentrism Above Median" in Table 6 is incorrectly signed and statistically insignificant ($t = 0.05$, $p = 0.96$). We thus found no support for RPP's prediction.

Does Ethnocentrism Moderate the Eliminated Effect?

RPP also predicted that controlling for the race of the target would reduce/eliminate the effect of democracy, and this pattern would be especially pronounced for respondents with ethnocentric worldviews. Hence, they preregistered an addendum to H4, that "The eliminated effect of democracy when taking into account the race of the target will be greater for those who score higher in ethnocentrism" (RPP Preregistration, 2). RPP committed to test this prediction with "linear regression by interacting the regime type and race treatments with self-reported ethnocentrism scores" (RPP Preregistration, 13).

According to this prediction, the coefficients on the triple interactions of regime, race, and ethnocentrism

[16] We could not conduct this test with Survey 2, which did not randomize the race of the target.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

in Surveys 1 and 2 should be positive. Moreover, if race plays a role in the mechanism connecting democracy to peace, RPP emphasize that the eliminated effects should be "large and significant" (RPP, 625). However, the evidence does not meet these standards. Figures 2 and 3 of RPP show eliminated effects for respondents above and below the ethnocentrism median but do not report whether the difference between the two groups was statistically significant. When we conducted formal tests, the difference between groups (i.e., the coefficient on the triple interaction) was significant for only one of their three comparisons.[17]

Suppose, however, that the differences in eliminated effects (i.e., the triple interactions between democracy, race, and ethnocentrism) had been distinguishable from zero. Even then, we would not have sufficient evidence to conclude that democracy was affecting support for war via racialized channels or that ethnocentrism was driving the process.

First, as noted earlier, eliminated effects could arise either because race is a mediator explaining the democratic peace, as hypothesized by RPP, or because race is a moderator conditioning the effect of democracy. Moderation seems far more likely, given that the data contradict so many other testable implications of the racial mediation theory. If eliminated effects were greater among people with ethnocentric worldviews, this could be because race moderates the effect of democracy and does so more strongly among ethnocentric respondents.

Second, even if race were a plausible mediator, we would need additional assumptions before interpreting differences by ethnocentrism as evidence for that mediator (Fu and Slough 2025). For example, we would need to assume that ethnocentrism affected *only* the racial mechanism, without moderating any of the nonracial explanations for the democratic peace that have been proposed by previous scholars. To the extent that ethnocentrism could affect nonracial mechanisms, differences by ethnocentrism would not constitute proof of a racial mechanism.[18]

Finally, to conclude that ethnocentrism was "driving" racialization, we would need to rule out confounders. RPP characterize ethnocentrism as a "major force" (RPP, 630) that "uniquely drives the democratic peace effect in public opinion" (RPP, 628) by causing people to react to democracy in a racialized way. However, it would be difficult to establish that ethnocentrism is exerting this force, given that ethnocentrism is not randomly assigned (Bansak 2021). Before making a causal claim about ethnocentrism, one would need to measure and control for many potential confounders *and* the interaction between each of those confounders and democracy, race, and democracy × race.[19] Choosing an appropriate set of covariates and, thus, interaction terms, is a daunting task. But without doing so, one could not claim that ethnocentrism "drives" the democratic peace, even if the eliminated effects in surveys 1 and 2 had been robust and statistically significant.

Robustness of Conclusions to Alternative Measures of Ethnocentrism

Finally, we consider whether alternative measures of ethnocentrism would have produced different conclusions. RPP conceptualized ethnocentrism as "a sense of cultural superiority of particular ethnic groups, which might (but need not) be defined racially" (RPP, 624). Survey 1 contained seven measures of ethnocentrism, but the published article used only three: "Most other cultures are backward compared with my culture"; "My culture should be the role model for other cultures"; and "I am not really interested in the customs and values of other countries." RPP combined these three items into a single score, which they used to distinguish respondents above and below the ethnocentrism median.

RPP planned to compare "various measures of ethnocentrism from different scales to ensure that findings are not scale-dependent (that is, that any heterogeneous treatment effects amongst ethnocentrics are robust to different measures of ethnocentrism)" (RPP Preregistration, 6). Although RPP did not report these robustness tests, carrying them out is important because the measures RPP selected are potentially controversial. All three of their selected items focused on "culture" or "other countries," but scholars warn that such measures could inadvertently capture other attitudes, such as nationalism, beliefs about foreign nationals (Bizumic, Monaghan, and Priest 2021, 13), or democratic political culture. If the goal is to measure ethnocentrism, scholars instead recommend asking about "ethnic groups."[20]

Fortunately, two of the items RPP discarded asked about ethnic groups: "The world would be a much better place if all other ethnic groups modeled themselves on my ethnic group" and "In general, I prefer

[17] In Survey 1, the difference in eliminated effect between the high and low ethnocentrism groups was significant when race was fixed as nonwhite, but not when race was fixed as white ($t = 1.21$, $p = 0.23$). In Survey 2, which fixed race as nonwhite, the difference was not significant ($t = 1.43$, $p = 0.15$). See Supplementary Materials, Tables S1–S2.

[18] Recent research has also shown heterogeneous treatment effects (such as differences by ethnocentrism) can arise even in the absence of mediation or moderation (Fu and Slough 2025), due to nonlinear transformation of a latent outcome to the observed outcome. This critique applies to RPP, who analyze a discrete realization (a five-point Likert scale) of a latent variable (preference for war). We thank an anonymous reviewer for making this point.

[19] In some analyses, RPP control for covariates that could predict both ethnocentrism and support for war, but they do not include the full set of *interaction* terms required by Bansak (2021).

[20] Moreover, questions about culture may tap preferences for democratic political culture. Had RPP found that respondents scoring high on "culture" questions were more sensitive to democracy, this could have been evidence of the moderating effect of democratic norms anticipated by existing theories (Bassan-Nygate et al. 2024), rather than of the importance of ethnocentrism *per se*.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

doing things with people from my own ethnic group than with people from other ethnic groups."[21] Using RPP's procedures, we combined these two items into a single score, split the sample at the median, and conducted the preregistered tests of H4.

The analyses, shown in the SM, contradict H4. First, the alternative measure of ethnocentrism did not moderate the effect of race on support for war. The effect of race was null not only for respondents below the ethnocentrism median, but also for highly ethnocentric individuals, and there was no significant difference between the two groups.[22] Second, the alternative measure of ethnocentrism did not moderate the eliminated effect; the interaction between race, regime, and ethnocentrism was insignificant, whether the country's race was fixed as white or nonwhite.[23] In sum, alternative measures of ethnocentrism, including measures more suited to testing RPPs theory, do not show any support for their claims.[24]

CONCLUSION

The study of race is becoming increasingly central to the field of international relations. Scholars have begun to document the effects of race on a wide range of issues, including military security, economic policies, and international law and organization.

RPP contribute to this rapidly growing literature by revisiting one of the most influential findings in international relations: the democratic peace. Previous survey experiments found that voters in democratic countries are far more reluctant to wage war against democracies than against otherwise similar nondemocracies. Scholars interpreted these results as evidence that voter preferences could help explain the zone of peace among democratic nations.

However, previous research failed to consider the role of race in the democratic peace. RPP advance the debate by proposing and testing an innovative theory that puts race, rather than regime type, at the heart of the causal story. They argue that "the democratic peace in public opinion owes, in large part, to racialized assumptions about democracy" (RPP, 621). In their theory, when respondents read that a country is democratic, they spontaneously assume that the country is predominantly white. "This implicit racialization," rather than regime type itself, "explains the reluctance of the American public to support aggression against fellow democracies" (RPP, 621).

The stakes in this debate are high. If correct, RPP's argument would cast doubt on widely cited research about the connection between public opinion and the democratic peace. Their logic could also undermine existing findings about the effect of democracy on many other aspects of international relations, including reputations for resolve and support for trade, investment, and foreign assistance. Finally, RPP's claims could change the trajectory of research in international relations by affecting how scholars design future survey experiments, text analyses, and observational studies.

Given these stakes, we asked whether the available data were sufficient to conclude, as RPP suggest, that "the democratic peace in public opinion is, largely, an ethnocentric and racialized peace" (RPP, 621). We identified several testable implications of their theory: that the term democracy triggers spontaneous assumptions about race; that assumptions about race, in turn, explain public preferences about war; that taking beliefs about race into account "erases" the effect of democracy; and that these patterns are especially evident among respondents with ethnocentric worldviews. We used data from RPP and others to test each of these predictions.

Notwithstanding the innovative and plausible nature of RPP's theory, we did not find support for any of its predictions. Table 7 summarizes our findings. We uncovered no evidence that the term democracy triggered unprompted assumptions about race, either in survey experiments or in analysis of English texts. Further, using RPP's experiments, we found no significant difference in public support for attacking countries that were mostly white versus mostly nonwhite. Moreover, the effect of democracy remained powerful even after controlling for race, contradicting the claim that race erases the effect of democracy. Finally, the patterns predicted by RPP were not more evident among respondents who scored highly on ethnocentrism than among respondents with lower levels of ethnocentrism.

Our findings in no way suggest that race is unimportant in international relations. A growing body of research highlights the substantial but underappreciated effects of race in many domains, and we suspect that future studies will further establish race as a central theme in IR scholarship. However, existing evidence does not demonstrate that the democratic peace in public opinion is, in reality, an ethnocentric and racialized peace. On the contrary, the weight of evidence continues to suggest that democracy has a powerful and pacifying effect on public opinion, even after taking race into account. Thus, our findings should strengthen confidence in past survey experiments about the relationship between democracy and support for war, and by extension, our understanding of the link between regime type and many important outcomes in international relations.

[21] The final two measures discarded by RPP did not mention ethnic groups.

[22] The coefficient on the interaction had the wrong sign and was not significant, $p = 0.65$. See Supplementary Table S3.

[23] The p-values on the interactions were $p = 0.67$ and $p = 0.24$, respectively. See Supplementary Tables S4–S5.

[24] We could not conduct a similar robustness test for Survey 2, because the ethnocentrism questions in that survey never mentioned ethnic groups. The items in Survey 2 "come from the 12-item superiority subscale from Bizumic et al. (2009)." However, RPP modified Bizumic's original items by deleting all references to ethnicity (RPP Dataverse appendix, 5).

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

TABLE 7. Summary of Predictions and Evidence

Empirical prediction	Analysis in RPP (2025)	Analysis in this Article	Supported?
Regime type affects beliefs about the race of the target.			
"If we asked members of the public to describe democracies versus nondemocracies, we would expect terms like white or nonwhite to, respectively, coappear, despite the fact that we did not explicitly ask respondents about racial characteristics" (RPP, 631).	None. Did not show how survey respondents describe democracies versus nondemocracies when not explicitly prompted to think about race.	Analyze the content of open-ended descriptions by survey respondents in Dafoe, Zhang, and Caughey (2018) who had not been prompted to think about race.	**No**. See Tables 1–2.
If democracy is associated with whiteness, references to democracy should coincide with references to whiteness "in the English language writ large" (RPP, 630).	Calculated cosine similarity between democracy terms and an unofficial, ad hoc list of racial terms (RPP, 631).	Using RPP's data, calculate cosine similarity between democracy terms and official U.S. government racial terms.	**No**. See Figure 2.
Racial beliefs, in turn, affect support for war.			
"Subjects randomly told that a target state that has a majority white population will be less inclined to support the use of force against that target state, relative to subjects told that the target state has a majority nonwhite population" (RPP Prereg., 1–2). Use regression to estimate the main effects of race and democracy (RPP Prereg., 13).	Reported that race had no independent effect on support for strikes (RPP, 634), but did not acknowledge that this null finding contradicts their theory and preregistration.	Using RPP's data, run the preregistered regression to estimate the main effects of race and democracy on support for strikes.	**No**. See Table 3.
Racialized beliefs explain most/all of the effect of regime type.			
"We will find a reduction in the main effect of democracy when we take into account the indirect effect of racial characteristics of the target population" (RPP Prereg., 2). Test "by comparing the treatment effect of democracy in the arm of the survey that only receives information about regime type against the treatment effect of democracy in the arm of the survey that also manipulates the racial characteristics of the target" (RPP Prereg., 13).	None. Did not report the outcome of this preregistered test; see below for their alternative tests.	Using RPP's data, run the preregistered regression comparing the effect of democracy in the racial versus nonracial arms of the experiment.	**No**. See Table 4.
The effect of democracy should be smaller when the target is specified as white than in the nonracial arm of the experiment. RPP did not preregister this prediction, but it is a logical implication of their theory (for a proof, see Appendix Proposition 3).	Reported that specifying race as white did not reduce the effect of democracy (RPP, 628). Characterized this null finding as anticipated, even though the null effect was not preregistered and contradicts their theory.	Using RPP's data, run a regression comparing the effect of democracy when the target was white to the effect in the nonracial arm of the experiment.	**No**. See Table 5 column 1.
The effect of democracy should be smaller when the target is specified as nonwhite than in the nonracial arm of the experiment. RPP did not preregister this prediction, but it is a logical implication of their theory.	Reported that specifying race as nonwhite did not significantly reduce the effect of democracy (RPP, 628–9), but did not acknowledge that this contradicts their theory.	Using RPP's data, run a regression comparing the effect of democracy when the target was nonwhite to the effect in the nonracial arm of the experiment.	**No**. See Table 5 cols. 2–3.

(*Continued*)

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

TABLE 7. (*Continued*)

Empirical prediction	Analysis in RPP (2025)	Analysis in this Article	Supported?
These patterns are strongest among people high in ethnocentrism.			
"More ethnocentric individuals are likely to exhibit racialized assumptions" about the target country, including "the presumption that democracies are white" (RPP, 624).	None. Did not show a direct test of this claim.	We know of no dataset that would allow a direct test of this claim.	**Not testable** with RPP data.
"Subjects higher in self-reported ethnocentrism will be even more likely to support the use of force against hypothetical nonwhite target states" (RPP Prereg., 2). Test "using linear regression by interacting the regime type and race treatments with self-reported ethnocentrism scores" (RPP Prereg., 13).	None. Did not report the outcome of this preregistered test.	Using RPP's data, run the preregistered regression to see if ethnocentrism moderates the effect of race on support for using force.	**No**. See Table 6.
"The eliminated effect of democracy when taking into account the race of the target will be greater for those who score higher in ethnocentrism" (RPP Prereg., 2). Test with "linear regression by interacting the regime type and race treatments with self-reported ethnocentrism scores" (RPP Prereg., 13).	Did not report whether the predicted interaction between regime type, race, and ethnocentrism was statistically significant.	Using RPP's data, run the preregistered regression to see if ethnocentrism significantly moderates the eliminated effect.	**No**. See discussion.
Examine "various measures of ethnocentrism from different scales to ensure that findings are not scale-dependent (that is, that any heterogeneous treatment effects amongst ethnocentrics are robust to different measures of ethnocentrism)" (RPP Prereg., 6).	None. Did not report the outcome of this preregistered test.	Using RPP's data, run the preregistered test by replacing RPP's chosen measure of ethnocentrism with one referring to ethnic groups, as recommended in the literature.	**No**. See discussion.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit http://doi.org/10.1017/S0003055425101056.

DATA AVAILABILITY STATEMENT

Research documentation and data that support the findings of this study are openly available at the American Political Science Review Dataverse: https://doi.org/10.7910/DVN/NHOUXJ.

ACKNOWLEDGEMENTS

We are grateful to Christopher Sebastian Parker, Caleb Pomeroy, Brian Rathbun, and Baobao Zhang for sharing data. We also thank Avidit Acharya, Genevieve Bates, Derek Chong, Jonathan Chu, Lauren Davenport, Justin Grimmer, Chris Manning, Michaela Mattes, Jon Pevehouse, Jonathan Renshon, Xunchao Zhang, and seminar participants at CISAC at Stanford University for extremely helpful feedback, and Zhenchao Hu and Sloane Ward for research assistance.

CONFLICT OF INTEREST

The authors declare no ethical issues or conflicts of interest in this research.

ETHICAL STANDARDS

The authors affirm this research did not involve human participants.

REFERENCES

Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. "Analyzing Causal Mechanisms in Survey Experiments." *Political Analysis* 26 (4): 357–78.
Baker, Andy. 2015. "Race, Paternalism, and Foreign Aid: Evidence from U.S. Public Opinion." *American Political Science Review* 109 (1): 93–109.
Bansak, Kirk. 2021. "Estimating Causal Moderation Effects with Randomized Treatments and Non-Randomized Moderators." *Journal of the Royal Statistical Society Series A* 184 (1): 65–86.
Baron, Reuben M., and David A. Kenny. 1986. "The Moderator–Mediator Variable Distinction in Social Psychological Research." *Journal of Personality and Social Psychology* 51 (6): 1173–82.
Bassan-Nygate, Lotem, Jonathan Renshon, Jessica L.P. Weeks, and Chagai M. Weiss. 2024. "The Generalizability of IR Experiments Beyond the U.S." *American Political Science Review*, 1–16. https://doi.org/10.1017/S0003055424001199.
Bizumic, Boris, Conal Monaghan, and Daniel Priest. 2021. "The Return of Ethnocentrism." *Political Psychology* 42 (S1): 29–73.
Bizumic, Boris, John Duckitt, Dragan Popadic, Vincent Dru, and Stephen Krauss. 2009. "A Cross-Cultural Investigation into a Reconceptualization of Ethnocentrism." *European Journal of Social Psychology* 39 (6): 871–99.
Búzás, Zoltán I. 2013. "The Color of Threat: Race, Threat Perception, and the Demise of the Anglo-Japanese Alliance (1902–1923)." *Security Studies* 22 (4): 573–606.
Búzás, Zoltán I. 2021. "Racism and Antiracism in the Liberal International Order." *International Organization* 75 (2): 440–63.
Carnegie, Allison, and Nikhar Gaikwad. 2022. "Public Opinion on Geopolitics and Trade: Theory and Evidence." *World Politics* 74 (2): 167–204.
Carson, Austin, Eric Min, and Maya Van Nuys. 2024. "Racial Tropes in the Foreign Policy Bureaucracy: A Computational Text Analysis." *International Organization* 78 (2): 189–223.
Chaudoin, Stephen, Brian J. Gaines, and Avital Livny. 2021. "Survey Design, Order Effects, and Causal Mediation Analysis." *Journal of Politics* 83 (4): 1851–56.
Chen, Frederick R., Jon C.W. Pevehouse, and Ryan M. Powers. 2023. "Great Expectations: The Democratic Advantage in Trade Attitudes." *World Politics* 75 (2): 316–52.
Chilton, Adam S, Helen V. Milner, and Dustin Tingley. 2020. "Reciprocity and Public Opposition to Foreign Direct Investment." *British Journal of Political Science* 50 (1): 129–53.
Chu, Jonathan A., and Carrie A. Lee. 2024. "Race, Religion, and American Support for Humanitarian Intervention." *Journal of Conflict Resolution* 68 (10): 2076–100.
Dafoe, Allan, Baobao Zhang, and Devin Caughey. 2018. "Information Equivalence in Survey Experiments." *Political Analysis* 26 (4): 399–416.
Freeman, Bianca. 2023. "Racial Hierarchy and Jurisdiction in Status of Forces Agreements." *Security Studies* 32 (4–5): 748–74.
Freeman, Bianca, D. G. Kim, and David A. Lake. 2022. "Race in International Relations: Beyond the 'Norm Against Noticing'." *Annual Review of Political Science* 25: 175–96.
Fu, Jiawei, and Tara Slough. 2025. "Heterogeneous Treatment Effects and Causal Mechanisms." Working Paper.
Gawronski, Bertram, and Jan De Houwer. 2014. "Implicit Measures in Social and Personality Psychology." In *Handbook of Research Methods in Social and Personality Psychology*, eds. Harry T. Reis and Charles M. Judd, 283–310. Cambridge: Cambridge University Press.
Green-Riley, Naima, and Andrew Leber. 2023. "Whose War Is It Anyway? Explaining the Black-White Gap in Support for the Use of Force Abroad." *Security Studies* 32 (4–5): 811–45.
Greenwald, Anthony G., Debbie E. McGhee, and Jordan L.K. Schwartz. 1998. "Measuring Individual Differences in Implicit Cognition: The Implicit Association Test." *Journal of Personality and Social Psychology* 74 (6): 1464–80.
Grovogui, Siba N'Zatioula. 1996. *Sovereigns, Quasi Sovereigns, and Africans: Race and Self-Determination in International Law*. Minneapolis: University of Minnesota Press.
Henderson, Errol A. 2013. "Hidden in Plain Sight: Racism in International Relations Theory." *Cambridge Review of International Affairs* 26 (1): 71–92.
Johns, Robert, and Graeme A. M. Davies. 2012. "Democratic Peace or Clash of Civilizations? Target States and Support for War in Britain and the United States." *Journal of Politics* 74 (4): 1038–52.
Kim, Gidong, Yu Bin Kim, and Dongjin Kwak. 2024. "Public Opinion, Rivalry, and the Democratic Peace: Experimental Evidence from South Korea." *International Studies Quarterly* 68 (2): sqae027.
Lacina, Bethany, and Charlotte Lee. 2013. "Culture Clash or Democratic Peace? Results of a Survey Experiment on the Effect of Religious Culture and Regime Type on Foreign Policy Opinion Formation." *Foreign Policy Analysis* 9 (2): 143–67.
Maoz, Zeev, and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946–1986." *American Political Science Review* 87 (3): 624–38.
Mills, Charles. 1997. *The Racial Contract*. Ithaca, NY: Cornell University Press.
Oksamytna, Kseniya, and Sarah von Billerbeck. 2024. "Race and International Organizations." *International Studies Quarterly* 68 (2): sqae010.
Payne, B. Keith, Clara Michelle Cheng, Olesya Govorun, and Brandon D. Stewart. 2005. "An Inkblot for Attitudes: Affect Misattribution as Implicit Measurement." *Journal of Personality and Social Psychology* 89 (3): 277–93.
Pennington, Jeffrey, Richard Socher, and Christopher D. Manning. 2014. "GloVe: Global Vectors for Word Representation." In *Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing (EMNLP)*, eds. Alessandro

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

Moschitti, Bo Pang, and Walter Daelemans, 1532–43. Doha, Qatar: Association for Computational Linguistics.
Rathbun, Brian C., Christopher Parker, and Caleb Pomeroy. 2025. "Separate but Unequal: Ethnocentrism and Racialization Explain the 'Democratic' Peace in Public Opinion." *American Political Science Review* 119 (2): 621–36.
Rodriguez, Pedro L., and Arthur Spirling. 2022. "Word Embeddings: What Works, What Doesn't, and How to Tell the Difference for Applied Research." *Journal of Politics* 84 (1): 101–15.
Renshon, Jonathan, Keren Yarhi-Milo, and Joshua D. Kertzer. 2023. "Democratic Reputations in Crises and War." *Journal of Politics* 85 (1): 1–18.
Stark, Tobias H., Jon A. Krosnick, and Amanda L. Scott. 2025. "Taking Stock: Explicit and Implicit Prejudice." In *The Cambridge Handbook of Implicit Bias and Racism*, eds. Jon A. Krosnick, Tobias H. Stark, and Amanda L. Scott, 1–24. Cambridge: Cambridge University Press.
Rosenberg, Andrew S. 2019. "Measuring Racial Bias in International Migration Flows." *International Studies Quarterly* 63 (4): 837–45.
Tomz, Michael, and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107 (4): 849–65.
Tomz, Michael, and Jessica L. P. Weeks. 2025. "Replication Data for: Race, Democracy, and Public Support for War." Harvard Dataverse. https://doi.org/10.7910/DVN/NHOUXJ.
Vitalis, Robert. 2018. *White World Order, Black Power Politics: The Birth of American International Relations*. Ithaca, NY: Cornell University Press.
Zvobgo, Kelebogile, and Meredith Loken. 2020. "Why Race Matters in International Relations." *Foreign Policy* 237: 11–3.

APPENDIX: FORMALIZATION OF RPP'S ARGUMENT

Setup: Let Y_i denote respondent i's support for going to war against a country, where $Y_i = 1$ if the respondent favors war and $Y_i = 0$ if they oppose war. Posit that Y_i depends on whether the country is described as a democracy or a nondemocracy, where $d = 1$ if democracy and $d = 0$ if nondemocracy.

Hypothesize that the effect of democracy is mediated by beliefs about whether the country is "mostly white" $(w = 1)$ or "mostly nonwhite" $(w = 0)$. Let $W_i(d)$ represent the "natural value" of this racial mediator for respondent i, where natural value refers to the assumptions i would spontaneously make about race if they were not told about the racial makeup of the country. Thus, $W_i(1)$ represents the assumptions i would naturally make about the racial makeup of a democracy, whereas $W_i(0)$ represents the assumptions i would naturally make about the racial makeup of a nondemocracy.

In this framework, the total effect on support for war of being told the country is a democracy, without explicit information about the racial composition of the country, is $TE_i = Y_i(1, W_i(1)) - Y_i(0, W_i(0))$. The first term represents the respondent's support for war when the country is a democracy, and the respondent makes whatever assumptions they would naturally make about the racial makeup of a democracy. The second term represents the respondent's support for war when the country is not a democracy, and the respondent makes whatever assumptions they would naturally make about the racial makeup of a nondemocracy.

We can decompose the total effect of democracy into two pathways: an indirect effect that operates via race, as hypothesized by RPP, and a direct effect that operates through mechanisms other than race, as suggested by previous literature on the democratic peace.

The indirect effect is defined as

$$\delta_i(d) = Y_i(d, W_i(1)) - Y_i(d, W_i(0)).$$

This expression isolates the indirect pathway by quantifying how support for war would change if the mediator moved from its nondemocratic value, $W_i(0)$, to its democratic value, $W_i(1)$, while holding the regime treatment constant at d. In the special case of no interaction between the treatment and the mediator, the indirect effect simplifies to $\delta_i = \delta_i(1) = \delta_i(0)$.

The direct effect is defined as

$$\zeta_i(d) = Y_i(1, W_i(d)) - Y_i(0, W_i(d)).$$

This expression quantifies how support for war would change if regime type moved from its nondemocratic value to its democratic value, while holding perceptions of the country's race constant at the level that would be realized if the country's regime were d. If there is no interaction between the treatment and the mediator, $\zeta_i = \zeta_i(1) = \zeta_i(0)$.

It is easy to show that indirect and direct effects sum up to the total effect, $TE_i = \delta_i(d) + \zeta_i(1-d)$. If there is no interaction between the treatment and the mediator, $TE_i = \delta_i + \zeta_i$.

Finally, the controlled direct effect of democracy after fixing the value of the mediator at w is

$$CDE_i(w) = Y_i(1, w) - Y_i(0, w).$$

This expression gives the effect of democracy on the outcome after controlling for the racial mechanism by holding the race of the country constant at either white $(w = 1)$ or nonwhite $(w = 0)$. The expression $ACDE_i = (CDE_i(1) + CDE_i(0))/2$ gives the average controlled direct effect, averaging over $w = 1$ and $w = 0$ by giving equal weight to each. If there is no interaction between the treatment and the mediator, $CDE_i = CDE_i(1) = CDE_i(0) = ACDE_i$.

Proposition 1: For the indirect pathway $\delta_i(d)$ to have a non-zero value, democracy must affect perceptions about race.

Proof: Suppose that democracy does not affect perceptions of race, i.e., $W_i(0) = W_i(1)$. Then we can substitute $W_i(1)$ for $W_i(0)$, resulting in $\delta_i(d) = Y_i(d, W_i(1)) - Y_i(d, W_i(1)) = 0$ for all d.

Proposition 2: For the indirect pathway $\delta_i(d)$ to have a non-zero value, the race of the target country must affect support for war after controlling for democracy.

Proof: Suppose that race had no effect on support for war after controlling for democracy, i.e., $Y_i(d, w') = Y_i(d, w)$ for all w, w' and for all d. Then $Y_i(d, W_i(1)) = Y_i(d, W_i(0))$ and therefore $\delta_i(d) = Y_i(d, W_i(1)) - Y_i(d, W_i(0)) = 0$ for all d.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press

Finally, we expose the logical flaw in RPP's claim that "If democracy is implicitly associated with whiteness, then the democratic peace effect should be similar with (and without) explicit information that the country is white" (RPP, 622).

The authors assume that people would naturally think a democracy was majority-white, such that telling people the democracy was majority-white would provide *no additional information* that was not already conveyed by democracy itself. They write, "in the absence of any explicit racial information, respondents will likely assume that a democratic country is majority white compared to a nondemocratic country, such that any racial information about whiteness in the context of democracy is *superfluous*" (RPP, 625–6). They add that their theory of "racialization implies that the provision of white information provides *no additional information* beyond democracy" (RPP, 622). We can express this assumption as $W_i(1) = 1$, meaning that people would naturally assume a democracy was white.

Given this assumption, what could we conclude if the effect of democracy were the same with and without explicit information that the country is white? The following proposition shows that this implies that the indirect pathway must be zero conditional on nondemocracy. As the proof of the proposition shows, this is because either democracy has no effect on perceptions of the race of the target, or the race of the target has no effect on support for the use of force conditional on nondemocracy, or both.

Proposition 3: Given the assumption in RPP that $W_i(1) = 1$, if the total effect of democracy remains the same after setting the race of the country equal to white, i.e., $TE_i = CDE_i(1)$, then the indirect pathway conditional on nondemocracy is zero, i.e., $\delta_i(0) = 0$.

Proof: If $TE_i = CDE_i(1)$, then $Y_i(1, W_i(1)) - Y_i(0, W_i(0)) = Y_i(1, w=1) - Y_i(0, w = 1)$. Applying RPP's assumption that $W_i(1) = 1$, we can write $Y_i(1, w = 1) - Y_i(0, W_i(0)) = Y_i(1, w = 1) - Y_i(0, w = 1)$, which simplifies to $Y_i(0, W_i(0)) = Y_i(0, w = 1)$.

For this to be true, either $W_i(0) = 1$ or w has no effect on the outcome conditional on $d = 0$.

But if $W_i(0) = 1$ and, as RPP assume, $W_i(1) = 1$, then $W_i(0) = W_i(1)$, meaning that democracy has no effect on perceptions of race. We could then write $\delta_i(d) = Y_i(d, W_i(1)) - Y_i(d, W_i(1)) = 0$ for both values of d.

Alternatively, if w has no effect on the outcome conditional on $d = 0$, then even if democracy affects assumptions about whiteness, $Y_i(0, w') = Y_i(0, w)$ for all w, w', which implies that $Y_i(0, W_i(1)) = Y_i(0, W_i(0))$, and thus again, $\delta_i(0) = Y_i(0, W_i(1)) - Y_i(0, W_i(0)) = 0$. QED.

https://doi.org/10.1017/S0003055425101056 Published online by Cambridge University Press