

Separate but Unequal: Ethnocentrism and Racialization Explain the “Democratic” Peace in Public Opinion

BRIAN C. RATHBUN *University of Southern California, United States*

CHRISTOPHER SEBASTIAN PARKER *University of California, Santa Barbara, United States*

CALEB POMEROY *Stanford University, United States*

Why are democratic publics reluctant to use force against fellow democracies? We hypothesize that the democratic peace in public opinion owes, in large part, to racialized assumptions about democracy. 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, most notably among respondents higher in ethnocentrism who disproportionately drive the democratic peace treatment effect. Two original survey experiments, a large-scale word embedding analysis of English texts, and reanalyses of published studies support this expectation. Our results suggest that the democratic peace in public opinion is, largely, an ethnocentric and racialized peace. The findings hold implications for the role of racism and racialization in foreign policy opinion research generally.

INTRODUCTION

The democratic peace, the finding that democracies almost never fight other democracies, is one of the most studied and contested phenomena in international relations. Most research focuses on the elite and state levels, but recent work purports to find strong microfoundational support at the level of public opinion. Democratic publics, notably in the United States and the United Kingdom, are less likely to support the use of force against democracies relative to nondemocracies (Johns and Davies 2012; Tomz and Weeks 2013). Public constraints on the use of force, along with other liberal institutions and practices, might contribute to a “separate peace” (Kahl 1998) among democracies in an otherwise anarchic and violent international system (Goldgeier and McFaul 1992).

Why are democratic publics reluctant to use force against fellow democracies? In contrast to conventional wisdom, we show that racialization of democracy drives much of the “democratic” peace in public opinion. Integrating insights from critical theory, as well as social and political psychology, we argue that ethnocentrism—

a felt sense of cultural and moral superiority of one’s group, often predicated on race (Bizumic, Monaghan, and Priest 2021; Kinder and Kam 2010)—underpins this racialization. Rather than regime type per se doing the causal work, the term “democracy” unwittingly primes presumptions of whiteness; respondents assume that nondemocracies are non-white. This implicit racialization, which varies across individuals, explains the reluctance of the U.S. public to support aggression against fellow democracies.

Although democracy as a form of government has no inherent racial content, survey research shows that western respondents presume that democracies are white (Dafoe, Zhang, and Caughey 2018). Critical theorists link this presumption to a pervasive sense of western civilizational superiority based in part on beliefs in a unique white capacity for enlightened democratic rule (e.g., Bowden 2021, 164; Henderson 2013). This “civilizationalism” shows a striking resemblance to the political psychological concept of ethnocentrism (Bizumic, Monaghan, and Priest 2021), as well as the social psychological finding that culture transmits implicit racial associations (Nosek, Greenwald, and Banaji 2005). Together, members of the public display what one might consider a “folk Eurocentrism” in their discrimination between democracies and nondemocracies, an ethnocultural chauvinism likely systemic in nature but with significant individual-level variation.

To identify the role of race in the public’s preference for democracies, we field two original survey experiments, reanalyze previous surveys, and conduct a large-scale word embedding analysis of quotidian English texts. Judging the effect of race on foreign policy opinion involves more than the mere addition of a racial manipulation to an existing survey experiment.

Brian C. Rathbun , Professor of International Relations, Department of Political Science and International Relations, University of Southern California, United States, brathbun@usc.edu.

Christopher Sebastian Parker, Professor of Political Science, Department of Political Science, University of California, Santa Barbara, United States, csp@ucsb.edu.

Corresponding author: Caleb Pomeroy , Stanton Nuclear Security Fellow, Center for International Security and Cooperation, Stanford University, United States, calebpom@stanford.edu.

Received: February 21, 2023; revised: November 02, 2023; accepted: April 15, 2024. First published online: May 27, 2024.

Instead, 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. We use this design to field an adaptation of the now-canonical Tomz and Weeks (2013) counterproliferation experiment in two surveys of the U.S. public.

In our first experiment, one third of respondents receive information about the regime type (democracy or nondemocracy) of a hypothetical country developing nuclear weapons, while two thirds receive information about regime type *and* the racial characteristics of the target's population (either white or non-white). We broadly expect that ethnocentrism explains the treatment effect of democracy, and this design allows us to examine whether, consistent with that premise, race eliminates the democratic peace effect. 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; racialization implies that the provision of white information provides no additional information beyond democracy. By contrast, provision of non-white information should erode the democratic peace effect, because this information contrasts with racialized assumptions about regime type that flow from a sense of cultural superiority. These dynamics should be most pronounced for individuals scoring highest in ethnocentrism.

Far from liberal or cosmopolitan amity, the results indicate that the preference shown for democracies is entirely driven by respondents who score higher in ethnocentrism. For these ethnocultural chauvinists, the effect of democracy is entirely eliminated when the hypothetical state is non-white, suggesting an implicit racialization of "democracy" even in the absence of explicit racial information. We also find a countervailing tendency among respondents low in ethnocentrism. These more culturally tolerant respondents exhibit no preference for democracies (relative to nondemocracies) and are slightly more likely to show democratic peace tendencies toward non-white countries. Without direct measurement of ethnocentrism, these subtle and countervailing forces would mask what is, in fact, a highly racialized phenomenon.

In a replication of our first survey with a larger but less representative sample, we field a more robust ethnocentrism battery, and omit the white mediator arm to address possible concerns about statistical power. In line with our first survey, we find that respondents higher in ethnocentrism are far more likely to discriminate between democracies and nondemocracies, and hypothetical non-white countries again face a substantially reduced democratic peace effect, a racial penalty of 45.1%. However, in contrast to the first survey, this racialization occurs at even lower thresholds of ethnocentrism. Here, racialization of the democratic peace is more widespread than our first survey suggests, which was possibly limited by statistical power.

Our experimental findings indicate that democracy carries associations of whiteness unless information is

given to the contrary, which leads more ethnocentric individuals to favor democracies. This suggests that race was hiding in plain sight in earlier surveys, to use Henderson's (2013) phrase. Indeed, in a reanalysis of Tomz and Weeks (2013) and Johns and Davies (2012), we find that cruder (but more racially explicit) measures of ethnocentrism moderate the treatment effect of democracy. The democratic peace in public opinion seems to be more of an ethnocentric peace.

We supplement our survey evidence with a large-scale analysis of everyday English texts. Recent work shows that word embeddings, which quantify associations between words, provide strong leverage on the analysis of racial bias and prejudice in speech (e.g., Caliskan, Bryson, and Narayanan 2017). We first show that democracy terms are positively associated with peace terms in the word embeddings, which suggests that quotidian English texts exhibit the same relationship between democracy and peace found in survey experimental research. Importantly, however, averaging the democracy terms with non-white/nonwestern terms eliminates 40% of the similarity between democracy and peace. By contrast, averaging the democracy terms with white/western terms leaves the similarity between democracy and peace entirely intact. This analysis provides external validity: the democratic peace in public opinion is implicitly racialized in the contemporary English language writ large, far beyond stylized experimental scenarios.

At the level of public opinion, then, the "separate peace" is a separate but unequal peace, one driven by discrimination on the basis of race implicitly invoked by "democracy." The theoretical implications are at least twofold. First, our results challenge the supposedly liberal foundations of democratic publics' pacifism toward democracies (see also Bell 2014; Hor 2024; Vucetic 2011a). Many scholars argue that the public plays an important role in explaining the democratic peace at the elite level (Doyle 2005; Tomz and Weeks 2013), and recent experimental work shows that elites in democracies are responsive to the public's views and desires (Chu and Recchia 2022; Tomz, Weeks, and Yarhi-Milo 2020). While the relationship between the public and elites is beyond our article's scope, we offer preliminary evidence that elites display the same racialization of the democratic peace in U.K. parliamentary speeches. This finding dovetails with research on the role of racial stereotypes at the elite level (Mercer 2023) and further suggests that our results are not a uniquely American phenomenon.

Second, beyond the democratic peace, the results point to subtle ways in which racism and racialization influence foreign policy opinion more generally. Rather than simply adding race as a "variable" to existing analyses, our results align with the critical theoretic insight that seemingly race-neutral constructs in fact carry race along with them, a lesson broader than any particular finding about the democratic peace (see, e.g., Vitalis 2000; Anievas, Manchanda, and Shilliam 2015; Sabaratnam 2020; Howell and Richter-Montpetit 2020). This provides a basis to rethink determinants of public opinion toward international security, such as

issues of deterrence, preventive war, alliances, and nuclear proliferation. For example, extensive research documents the American public's willingness to support the use of nuclear weapons (e.g., Press, Sagan, and Valentino 2013). Yet, this foreign policy opinion research has yet to grapple with critiques by scholars of race surrounding the racialization of nuclear weapons (Intondi 2020). This oversight is striking given that most mainstream foreign policy opinion research examines the opinions of majority white countries' publics toward potential conflict with non-white countries and peoples. If race is lurking in the concept of "democracy," one of the most predictive features of public opinion toward the use of force, then race is likely waiting to be noticed across foreign policy opinion research. Critical theoretic insights, notably by scholars of race, alert positivist social science work to this pressing possibility.

IS THE DEMOCRATIC PEACE COLORBLIND? A RACIALIZED PEACE

Noting the contestation over the mechanisms that link joint democracy to peace (Bueno de Mesquita et al. 1999; Dafoe 2011; Gartzke 2007; Rosato 2003), as well as the importance of the mass public in many theories of the democratic peace, Tomz and Weeks (2013) experimentally validated the deference of the U.S. and British mass publics when engaging with democracies as opposed to nondemocracies. In a hypothetical scenario of a country developing nuclear weapons, respondents were more likely to support the use of preventive strikes against a nondemocracy than a democracy. Respondents perceived nondemocracies as more dangerous and striking them to be less immoral, justifying violent action for self-defense. Johns and Davies (2012), using a similar scenario in a survey fielded at roughly the same time, also found a pacifying effect of democracy, as well as a greater willingness to use force against an Islamic as opposed to a Christian target.

However, the experimental basis of the democratic peace in public opinion has recently faced criticism on the grounds of "information equivalence." As Dafoe, Zhang, and Caughey (2018, 400–1) explain, "Manipulating whether a country is described as 'a democracy' or 'not a democracy'...is likely to affect subjects' beliefs about such background features as the country's geographic location or *demographic composition*. If it does, then any differences between experimental groups cannot be reliably attributed to the effects of the beliefs of interest." In other words, in the minds of those surveyed, the term "democracy" is potentially indexed to a number of factors associated with democratic countries that might drive the effect of interest, rather than democratic regime type per se.

One of the most likely elements of "demographic composition" is race. We understand race not as an immutable biological characteristic but rather as the "product of a complex fusion of factors including societal values, skin color, cultural traits, physical attributes, diet, region of ancestry, *institutional power*

relationships and education" (Sen and Wasow 2016, 506). In the terms of Sen and Wasow (2016), perhaps democracy is part of race's "bundle of sticks." In their replication of Tomz and Weeks (2013), Dafoe, Zhang, and Caughey (2018) asked respondents to report beliefs about the characteristics of the hypothetical country described in the experiment, characteristics that were not explicitly mentioned in the vignette. Subjects assigned to a hypothetical democracy were more likely to infer that the country is wealthy, Christian, *majority white*, a military ally of the United States, and a significant trading partner, despite the fact that the vignette did not explicitly mention these traits (Dafoe, Zhang, and Caughey, 2018, 400). And, although Tomz and Weeks (2013) and Johns and Davies (2012) fix or manipulate a number of strategic factors associated with the hypothetical scenario, no previous study, to the best of our knowledge, fixes the racial characteristics of the country in question. This raises the possibility that what appears to be a "democratic" peace effect might instead be one driven by implicit racial inferences about the country's population; perhaps members of the public read "democracy" and infer "white" despite the vignette's lack of explicit information about race.

Recent contributions to critical IR theory, remarkably given their strikingly different epistemological foundations, offer a criticism that parallels this concern about information equivalence: the concept of democracy is racialized. Racialization is the "extension of racial meaning to a previously racially unclassified relationship, social practice or group" (Omi and Winant 2014, 64). As Sabaratnam (2020, 10) explains, "A key point, yet one readily forgotten, is that 'the West' is on any plausible reading a racialised category indexed to 'Whiteness.'" While the concept has often been applied to the racialization of particular social groups as "white" or "non-white" through historical processes of subjugation and domination (Maass 2023), racialization also applies to the institutions that emerge from such social practices—institutions like democracy. If American identity is amenable to implicit racialization (Devos and Banaji 2005), it is not a stretch to argue that assumptions about democracies are as well.

In a way that methodologically minded criticisms of "information equivalence" do not, critical theorists offer an explanation for this racialization: the self-perceived cultural superiority of white democratic societies. "Though recently established, white men's countries sought legitimacy through locating themselves in the long tradition of Anglo-Saxon race history that dated back to the mythic glories of Hengist and Horsa. They shared an English-speaking culture and newly ascendant democratic politics, priding themselves, as Anglo-Saxons, on a distinctive capacity, indeed a genius, for self-government," write Lake and Reynolds (2008, 6). Importantly, much of this sense of civilizational superiority, particularly for Anglo-Saxon countries, rested on beliefs in a unique white capacity for enlightened *democratic* rule (Bowden 2021, 164). In a careful study of the origins of the Anglo-American "special relationship," often attributed to a shared

commitment to democracy, Vucetic (2011b, 413) finds instead that “Anglo-American elites indeed understood each other as fellow democrats, but their antecedent ontology was always race, not regime type. To the extent that democracy was recognized as a cause of Anglo-American peace and cooperation, this frame was typically deployed in the context of a racial theory of history... [I]t can therefore be said that what caused peace...was *Anglo-Saxon* democracy, not Anglo-Saxon *democracy*.”

The implication of these critiques is that democracies receive better treatment from democracies because they are thought to be equivalently culturally superior, and this cultural superiority is implicated by race. The racial connotations of democracy are crucial, because it has previously been theorized that democracies favor other democracies based on a sense of common identity that rests on the common institutions themselves, without reference to race (Hayes 2013). In other words, the democratic peace in public opinion might be *ethnocentric*, a broader concept from political psychology with clear parallels to “civilizationalist” critiques of liberal IR theory. Ethnocentrism as a psychological attribute brings these elite-level critiques down to the level of the mass public.

TIES THAT BLIND: AN ETHNOCENTRIC PEACE

Ethnocentrism broadly refers to preferential divisions of the world into in-groups and out-groups (Kinder and Kam 2010). Bizumic and Duckitt (2012) and Bizumic, Monaghan, and Priest (2021) report numerous facets of this group self-centeredness: devotion to the group, a preference for group cohesion, a preference of one's group over others, a sense of superiority of one's group over another, a desire for group purity, and a willingness to exploit those outside the group. Importantly, Bizumic and Duckitt (2012, 893) note in their review of existing conceptualizations of ethnocentrism that “superiority appears to be the most widely emphasized facet.”

This superiority is primarily cultural and moral, rather than innate and biological. Ethnocentrism is a broader construct than racism: individuals can feel culturally superior to those of the same race. But the two concepts overlap significantly. Associations between a sense of cultural superiority and race have more in common with “symbolic” than “old-fashioned,” biological racism (Kinder and Sears 1981). In the U.S. context, this symbolic racism is more about the perceived violation of cherished American values than the belief that Black people are “naturally” inferior. Therefore, we conceive of ethnocentrism as a sense of cultural superiority of particular ethnic groups, which might (but need not) be defined racially. According to Bizumic and Duckitt (2012, 897), ethnic groups are “the real-world groups that their members perceive as having a unique system of shared meanings (e.g., distinct customs, mores, norms, language, or dialect), the perception of common

historical past and future, and usually belief in a common origin.”

The concept of ethnocentrism is particularly well-suited for our purposes given its emphasis on culture and broad applicability to different social contexts. Culture often drives unconscious biases, including stereotypes (Banaji and Greenwald 2016). The presumption that democracies are white is likely such an implicit bias. Ethnocentrism, therefore, is a good way for us to index cultural assumptions. More ethnocentric individuals are likely to exhibit racialized assumptions if those assumptions reflect the dominant culture. However, ethnocentrism is a concept that transcends parts of the west historically cleaved along domestic racial lines. For instance, perceptions that immigrants pose a threat to “Dutch culture” are the primary source of hostility to newcomers in the Netherlands (Sniderman, Hagendoorn, and Prior 2004).

Nevertheless, we expect that a sense of cultural superiority in the United States, as well as in other western countries, likely contains a racial element. If assumptions about race defined eligibility for first-class citizenship well into the twentieth century in the United States (Smith 1997), the supposed “beacon of democracy,” it only makes sense that those assumptions inform perceptions of democracy itself. As Mills (1997) argues, non-whites were never intended as parties to the western social contract. This paved the way for discrimination to take root in societies that identified as democracies (Hanchard 2018).

A clear parallel exists between this conceptualization of ethnocentrism and “civilizationist” distinctions found in critiques of liberal IR theory, of which democracy is an integral part. Western societies define themselves against others by their degree of advancement, notably material, technological, and moral maturity. This concept of advancement explicitly and implicitly ties to race (Persaud and Walker 2001; Vucetic 2011a). As Henderson (2013, 72) explains historically, “Uniquely among the races, whites were assumed to possess civilization while nonwhites were assumed to occupy a lower stage of development characterized as either barbarism or savagery...The lesser races were assumed to be not only biologically inferior to whites, but in a state of almost perpetual conflict.” Hobson (2015, 83) argues that “after 1945 it transmogrified into a much more subliminal form... [A]ll the old manifest Eurocentric-institutional (and scientific racist) tropes, civilisation, barbarism, savagery and imperialism, were whitewashed but appeared in terms...that dare not speak their name, such as ‘tradition versus modernity’ or ‘core versus periphery.’” This Eurocentrism is a special case of ethnocentrism, one specific to the superiority of European culture, values, and so on (Blaut 2000, 4). Racism and Eurocentrism are often intertwined. Civilizationalism draws a stark line between a civilized, white west and an uncivilized, non-white other that had profound historical implications in practices such as colonialism and humanitarian intervention (Vitalis 2000).

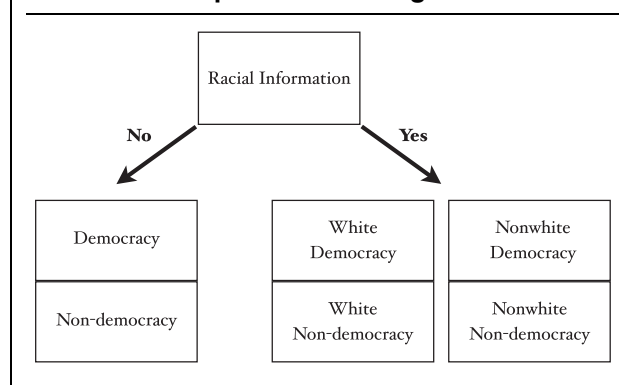
Importantly, a sense of cultural superiority lessens inhibitions against violence. Scholars of foreign policy opinion increasingly document a strong relationship between ethnocentrism and hawkishness, the willingness to support the use of military force against out-groups. Ethnocentrism maps onto militant postures toward the world in general (Hurwitz and Peffley 1987), and much recent work has focused on “terrorism” in particular (Hansen and Dinesen 2023; Kam and Kinder 2007). For example, individuals higher in ethnocentrism displayed stronger support for the United States’ “War on Terror.” What is unclear is whether this support for force against nefarious out-groups extends to regime type, namely preferences for democracies relative to nondemocracies, and what it is about democracies that garners ethnocentric support. Cultural superiority is a likely answer, and the sense of being better than others could lead to a greater willingness to use violence. Ethnocentrism can translate into perceptions that adversaries are more dangerous (or “barbaric”) and possess less moral worth. These perceptions parallel the two faces of racism described by Freeman, Kim, and Lake (2022) and also emerge empirically as the two most prominent mechanisms found by Tomz and Weeks (2013).¹

All of this raises the possibility that past survey experimental work on the democratic peace might be unwittingly tapping into ethnocentrism. When “democracy” is invoked in prior instruments, the term might have less to do with checks and balances, a free press, and so on. Rather, democracy becomes a signifier for “ethnologically similar to us.” Nondemocracy implies ethnocultural inferiority and increases the willingness to use force. Perhaps a set of assumptions about the cultural superiority of the west, associated with race, exists in the minds of western publics which form the primary experimental microfoundations for the democratic peace to date. The remainder of the article empirically examines this possibility.

RACE ERASES THE EFFECT OF DEMOCRACY: EVIDENCE FROM TWO ORIGINAL SURVEY EXPERIMENTS

To assess the effect of race potentially misattributed to democracy in past survey experimental work, we use a randomized mediator design described by Acharya, Blackwell, and Sen (2018). The intuition of the design is straightforward: if racial inferences drive the effect of “democracy” in past survey work—that is, if democracy conveys “whiteness” that decreases support for force—then the treatment effect of democracy should not change when respondents are told that a hypothetical country is also white; *whiteness is already associated with democracy*. By contrast, when subjects are told that a democratic country has a non-white population, then this racial manipulation should deflate and

FIGURE 1. Experimental Design



possibly even eliminate the democratic peace in public opinion. All democracies would not be created equal, and whiteness drove the original effect rather than democratic regime type per se.

Figure 1 displays our design. The natural-mediator arm (the left side of the figure) assigns regime type information without explicit racial information about the hypothetical country, identical to past democratic peace experiments. This arm provides a “natural” baseline to estimate implicit judgments about race conveyed by democracy in the absence of explicit racial information (and serves as the typical average treatment effect [ATE] reported in past work on the democratic peace).

By contrast, the manipulated-mediator arm (the right side of Figure 1) includes the same democracy versus nondemocracy assignment *in addition to* explicit racial information about the hypothetical country. This arm allows us to estimate the same effect of democracy (relative to nondemocracy) while fixing the country’s race to either white *or* non-white (also known as “average controlled direct effects” [ACDEs]). That is, these regime type assignments in the presence of explicit racial information allow us to reestimate the same democratic peace effect for white countries and non-white countries separately. If whiteness is associated with the democratic peace, then the democratic peace effect should be similar with and without explicit information that the country is white, whereas the provision of non-white information should substantially deflate the democratic peace effect. The difference between the effect of democracy with no racial information and the effects of democracy with explicit white *or* non-white information quantifies the amount of democracy’s effect eliminated by white and non-white racial information. Acharya, Blackwell, and Sen (2018) refer to these differences as “eliminated effects.” Large and significant eliminated effects indicate that race plays a role in the mechanism that links democracy to peace.

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

¹ On racialized threat perception, see also Búzás (2013).

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. Conversely, we expect that a predominantly non-white population will significantly eliminate the effect of democracy. This would indicate that respondents presume that democracies are white and, if told otherwise, no longer display the same reluctance to use force against the country.

Furthermore, we expect that the effect of democracy eliminated by non-white information will be even larger for respondents higher in self-reported ethnocentrism, measured in multiple ways, as explained below. 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. Moderation of this type allows for better confirmation of the ethnocentric and culturally superior nature of any preference shown to democracies based on racial presumptions.²

We field this design through a replication of the now-canonical Tomz and Weeks (2013) counterproliferation instrument. Our fundamental interest is to test whether the causal mechanism underlying existing studies of the democratic peace in public opinion is properly specified. Therefore, we utilize the same instrument as the most influential study in this area. Theoretical changes to the scenario or dependent variable would undermine our efforts.

All subjects were first presented with the following scenario and fixed characteristics associated with a hypothetical country:

A country is developing nuclear weapons and will have its first nuclear bomb within six months. The country could then use its missiles to launch nuclear attacks against any country in the world.

The country's motives remain unclear, but if it builds nuclear weapons, it will have the power to blackmail or destroy other countries. The country has refused all requests to stop its nuclear weapons program.

Here are some things to know about the strategic situation.

- The country is **not a military ally** of the U.S.
- The country has **low levels of trade** with the U.S.
- The country's military forces are **half as strong** as American military forces in the region.

The above material retains all of the substantive features from Tomz and Weeks (2013) with minor adjustments for parsimony. Note that we fix rather than randomize alliance and trade status, given that we are interested in the democracy treatment in

particular. The next screen presented our experimental interventions:

Here are some other basic characteristics to know about the country:

- The country [is a **democracy** and shows every sign that it will remain a democracy/is **not a democracy** and shows no sign of becoming a democracy].
- [The country's population is predominantly **white**/The country's population is predominantly **non-white**/No racial information].
- The country is predominantly **Christian**.

As described above, subjects were assigned to either a democracy *or* nondemocracy bullet, as well as one of the following racial assignments: no racial information (and therefore no bullet point) *or* a predominantly white bullet point *or* a predominantly non-white bullet point. Because the mention of race could induce information equivalence problems of its own, we re-analyzed data from Dafoe, Zhang, and Caughey (2018).³ Their study measured respondents' posttreatment judgments about multiple characteristics of the country in question beyond regime type. We found that religion posed the greatest risk to inference. Therefore, we fix religion to Christian for external validity reasons. There are more plausible real-world candidates for non-white Christian countries than white non-Christian countries.

Following treatment, subjects responded to the same primary dependent variable from Tomz and Weeks (2013):

By attacking the country's nuclear development sites now, the United States could prevent the country from making any nuclear weapons. Would you favor or oppose using the U.S. military to attack the country's nuclear development sites?

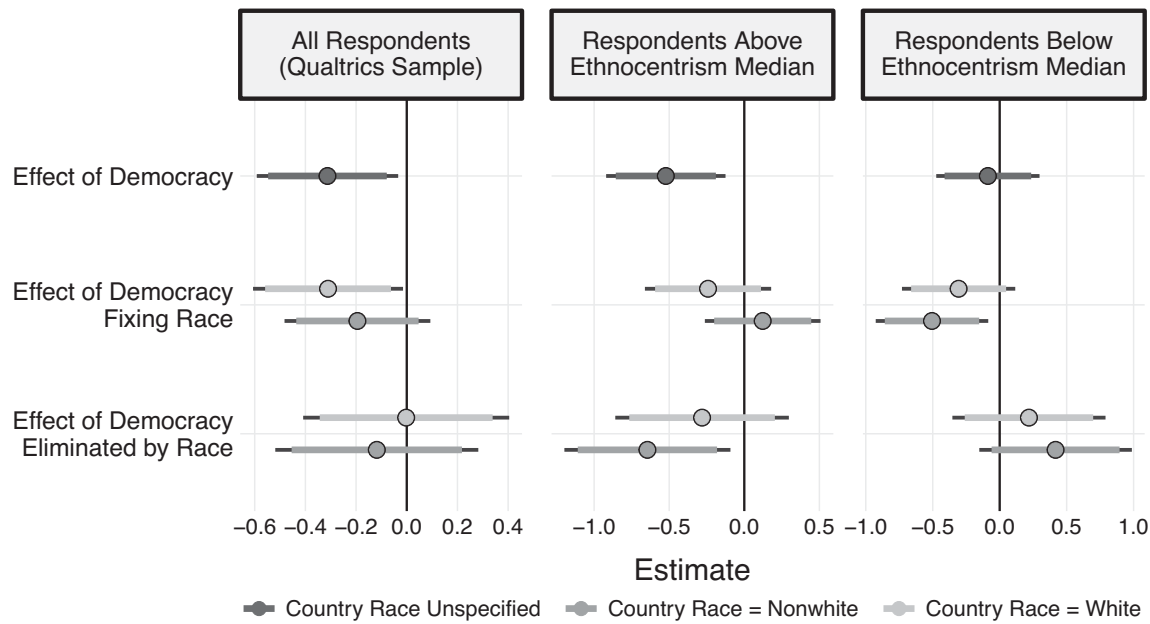
Responses were gathered on a five-point scale from "oppose strongly" (=1) to "favor strongly" (=5), where lower values in the democracy condition indicate a negative, pacifying effect of democracy on support for strikes. The key finding from Tomz and Weeks (2013) is that respondents were *less likely to support strikes against a fellow democracy* (ATE = -11.4, 95% CI [-17.0, -5.9]).⁴

As found in Tomz and Weeks (2013), we expect that respondents will be less likely to support strikes against a hypothetical democracy (relative to nondemocracy). However, we expect that ethnocentric respondents will primarily drive this treatment effect. Further, beyond the typical ATE of regime type, as described above, we expect that the provision of non-white racial information will eliminate this democratic peace effect. By contrast, the provision of white racial information will not eliminate the democratic peace in public opinion to any appreciable degree.

² These expectations, as well as the design and analysis plan, were preregistered with the Open Science Foundation prior to data collection. See <https://osf.io/x4wvf/>. Replication code and data are available in Rathbun, Parker, and Pomeroy (2024).

³ Dataverse Appendix B5 presents this analysis.

⁴ See, e.g., Table 1 in Tomz and Weeks (2013, 854).

FIGURE 2. Qualtrics Sample—Ethnocentrism Drives the “Democratic” Peace Effect

Note: Estimates and confidence intervals come from linear regressions with robust standard errors at 95% (thin line) and 90% (thick line) levels. Supplementary Tables A4–A6 present these results numerically.

Survey 1: Qualtrics Sample

Our first survey sample was recruited by Qualtrics and fielded on April 7–13, 2022 ($N = 1,626$ U.S. adults). The survey sample included quotas for race, education, and gender to increase representativeness relative to the U.S. population.⁵ All respondents completed the survey through Qualtrics’s online platform.

In addition to the instrumentation described above, the survey included self-report measures of ethnocentrism to assess whether ethnocentrism disproportionately explains the racialization of democracy.⁶ To measure ethnocentrism, we used three items from the commonly used Neuliep and McCroskey (1997) scale:

- Most other cultures are backward compared with my culture.
- My culture should be the role model for other cultures.
- I am not really interested in the customs and values of other countries.

Given that our argument hinges on cultural superiority, this scale is attractive. For example, another popular

scale pioneered by Kinder and Kam (2010) relies on generalized hostility toward specific ethnic groups, which is too specific in its target and too culture-specific for our purposes (Bizumic, Monaghan, and Priest, 2021, 13). We aim to measure a general sense of cultural superiority. Moreover, because our experimental treatment is explicitly racial—indeed it must be in that we hypothesize that democracy is operating as a euphemistic stand-in for race—we wanted to account for possible social desirability concerns (Huddy and Feldman 2009). If, as we hypothesize, ethnocentrism moderates the deflation of democracy’s effect for non-white countries, this would suggest that the sense of cultural superiority is racial to some significant degree.⁷ We use factor analysis to reduce the responses to single-dimensional factor scores, where higher values indicate greater ethnocentrism.⁸ For ease of presentation, we split respondents into two bins: high ethnocentrism (above the ethnocentrism median) versus low ethnocentrism (at or below the median).

If the democratic peace is racialized, then ethnocentrism should drive the pacifying effect of democracy. Figure 2 displays the results estimated with linear regression, including the full sample (left panel) and the sample broken down by subjects above or below the ethnocentrism median (middle and right panels,

⁵ Supplementary Section A2.1.1 presents the sample characteristics.

⁶ Dataverse Appendix B1.1 presents the relevant instrumentation. As described in our preregistration files, our first survey included a mix of ethnocentrism items from different scales, but the combined factor structures of these items were very poor. Therefore, here we focus only on the most coherent ethnocentrism items, which all come from Neuliep and McCroskey (1997).

⁷ We randomized the order of the experiment and the questions that measure ethnocentrism and racial resentment to avoid potential order effects.

⁸ Cronbach’s $\alpha = 0.68$, SS loadings = 1.32.

respectively).⁹ The first row displays the well-known ATE of democracy (relative to nondemocracy) for subjects assigned no racial information. As found in Tomz and Weeks (2013) and replicated multiple times since, the treatment effect is substantively and statistically noteworthy in the full sample (coef = -0.31 , 95% CI [$-0.592, -0.034$]). However, the middle and right panels suggest that respondents above the ethnocentrism median drive the treatment effect of democracy (coef = -0.52 , 95% CI [$-0.920, -0.125$]). Subjects below the ethnocentrism median do not discriminate between democracies and nondemocracies at all (coef = -0.09 , 95% CI [$-0.474, 0.298$]). This result is noteworthy, because ethnocentrism should play no role in the democratic peace in theory.

The second row of Figure 2 reports this same treatment effect of democracy (relative to nondemocracy) for white and non-white countries (i.e., the ACDEs for respondents randomly assigned racial information). In the full sample, the democratic peace effect for white countries is substantively and statistically identical to the democratic peace effect without racial information (coef = -0.31 , 95% CI [$-0.607, -0.015$]). Adding white information does nothing to alter democracy's effect. By contrast, the effect of democracy for non-white countries deflates to a level statistically indistinguishable from zero, such that non-white countries enjoy no democratic peace effect (coef = -0.19 , 95% CI [$-0.482, 0.093$]). Respondents assigned to non-white countries are equally likely to strike democracies and nondemocracies in the full sample. However, we do not overinterpret these ACDEs, because eliminated effects provide the most direct test of our argument.

The third and final row of Figure 2 reports the eliminated effects—that is, the difference between the ATE and ACDE quantities—which assess whether race plays a role in the causal mechanism linking democracy to peace. The full sample results suggest that racial information does not significantly eliminate the effect of democracy, and therefore race plays no role in the causal mechanism linking democracy to peace. However, the middle and right panels illuminate this null finding. For respondents above the ethnocentrism median, the provision of non-white information *entirely* eliminates the democratic peace effect (coef = -0.65 , 95% CI [$-1.198, -0.092$]), strong evidence for the racialization of democracy among the exact subpopulation that drives the average effect of democracy. This is not the case when the target is identified as white. Without measurement of ethnocentrism, this result would be missed at the average level: respondents below the ethnocentrism median respond in the opposite direction, slightly more likely to display democratic peace tendencies when the countries are non-white (coef = 0.42 , 95% CI [$-0.152, 0.987$]). This suggests that individuals low in ethnocentrism are perhaps expressing anti-racist views, which work against the overall effect in the sample and mask this subtle,

racialized process. In sum, individuals higher in ethnocentrism drive the treatment effect of democracy, and the provision of non-white information entirely eliminates the democratic peace effect within this subpopulation.

Furthermore, while the above results follow past research that bins respondents by ethnocentrism level (e.g., Hainmueller and Hopkins 2015; Hansen and Dinesen 2023), we note that a non-white population eliminates the democratic peace effect for respondents at or above the bottom 43% of the ethnocentrism scale. That is, far from a handful of ethnocentric respondents driving our results, approximately 57% of respondents display no democratic peace tendencies toward non-white countries.

Finally, to assess whether ethnocentrism uniquely drives the democratic peace effect in public opinion, Supplementary Section A2.1.3 examines a host of individual differences that could plausibly drive hawkishness toward out-groups. We find that other variables do not moderate democracy's treatment effect. Further, Dataverse Appendix B3 shows that we find no differences by respondent race, but our survey was not designed with sufficient statistical power to detect those differences. Still, we note that in theory racialized assumptions can affect the views of even non-white Americans.

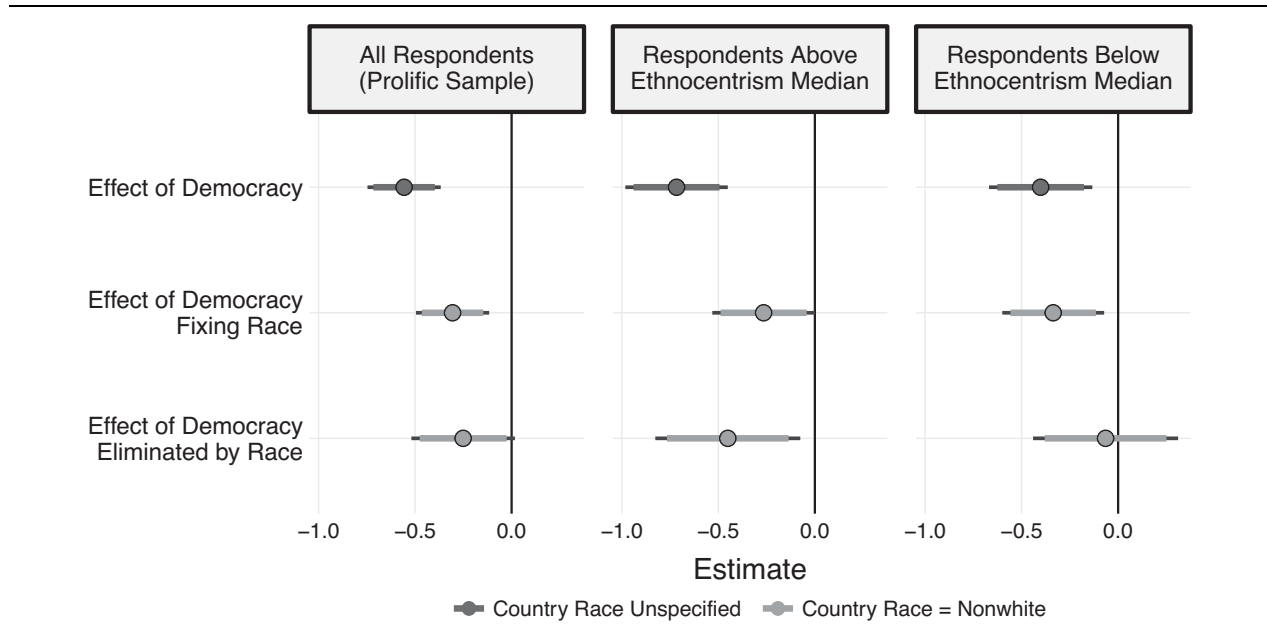
Survey 2: Prolific Sample

Our first survey indicates that non-white countries incur a racial penalty and that ethnocentrism drives this effect. However, we were concerned about two potential issues with our first survey. First, given that randomized mediator designs can face statistical power constraints (Acharya, Blackwell, and Sen, 2018, 363), our first survey might be underpowered. For example, it is difficult to determine whether the absence of eliminated effects in the full sample is a substantively important result or simply because we based our power calculations on Tomz and Weeks (2013), which was not explicitly designed to detect this quantity. Second, although we used well-established items from Neuliep and McCroskey (1997) to measure ethnocentrism in the first survey, we wanted to validate the above results with a more robust set of ethnocentrism items from a different inventory.

To address these issues, we fielded a replication of our first survey on September 27 to October 10, 2022 with respondents recruited on Prolific ($N = 2,659$ U.S. adults).¹⁰ Our second survey differed from the first in only two respects. First, to ensure that we have sufficient statistical power to detect non-white eliminated effects, here we do not include the white country randomization. Our first study already suggested that the provision of white information does not change democracy's overall treatment effect. Second, we used the 12-item "superiority" subscale from Bizumic et al. (2009), because this subscale best meshes with the

⁹ Supplementary Section A2.1.2 shows the key results are substantively unchanged when adding individual covariates to the models.

¹⁰ Supplementary Section A2.2.1 presents the sample characteristics.

FIGURE 3. Prolific Sample—A Predominantly Non-White Population Deflates the “Democratic” Peace Effect

Note: Estimates and confidence intervals come from linear regressions with robust standard errors at 95% (thin line) and 90% (thick line) levels. Supplementary Tables A11–A13 present these results numerically.

theoretical literature that motivates our expectations. Respondents agree or disagree with statements like “The world would be a much better place if all other cultures modeled themselves on my culture.” The items we use do not mention race explicitly, but the scale correlates highly with Kinder and Kam’s (2010) ethnocentrism scale, which asks respondents to assess racial in-groups and out-groups on positive and negative traits.¹¹ Responses were gathered on five-point scales from “strongly disagree” to “strongly agree.” We again use factor analysis to reduce the responses to unidimensional factor scores (Cronbach’s $\alpha = 0.93$, SS loadings = 6.19), where higher values indicate greater ethnocentrism, and we split respondents at the ethnocentrism median to ease presentation.

Figure 3 (left panel) displays the results of our second survey estimated with linear regression.¹² The first row shows that we again replicate the well-known ATE of democracy relative to nondemocracy for subjects assigned no racial information (coef = -0.56 , 95% CI $[-0.747, -0.367]$). The only notable difference from our first survey is that respondents below the ethnocentrism median show a significant democratic peace effect (coef = -0.40 , 95% CI $[-0.669, -0.134]$). But this effect is significantly smaller in magnitude than the ATE of democracy for respondents above the ethnocentrism median (coef = -0.72 , 95% CI $[-0.982, -0.451]$).

Therefore, we again find evidence that ethnocentrism disproportionately explains the treatment effect of democracy.

The second row of Figure 3 reports this same treatment effect for respondents explicitly told that the country’s population is predominately non-white (the ACDE). Fixing the country’s population to non-white substantially deflates democracy’s ATE in the full sample (coef = -0.31 , 95% CI $[-0.495, -0.116]$). As such, non-white countries *do* enjoy a democratic peace effect, but this effect is 45.1% smaller than countries without race specified. Respondents higher in ethnocentrism primarily drive this deflation of democracy’s effect when the hypothetical country is non-white. However, it is notable that, even among respondents low in ethnocentrism, the estimate for democracy’s effect deflates by 16.2% for non-white countries (albeit not to the extent that we can conclude a significant deflation).

The final row of Figure 3 shows that, at least at the $\alpha = 0.10$ level, race plays a role in the mechanism that links joint democracy to pacifism in the full sample (coef = -0.25 , 95% CI $[-0.519, 0.017]$). The middle and right panels confirm again that respondents higher in ethnocentrism drive this effect of democracy eliminated by race. Furthermore, although we median split ethnocentrism for presentation, it is worth noting that non-white information significantly eliminates the effect of democracy for a full 76% of this sample; only respondents in the bottom quarter of ethnocentrism hesitate to discriminate on the basis of race. This provides strong evidence that these results do not reduce to

¹¹ Dataverse Appendix B1.2 presents the relevant instrumentation.

¹² Supplementary Section A2.2.2 shows that the results are substantively unchanged when adding individual covariates to the models.

a handful of ethnoculturally chauvinistic apples. Racialization of the democratic peace in public opinion is likely more widespread than our first survey revealed, which was possibly limited by statistical power.

Finally, we note that our Prolific survey included posttreatment measures of threat perception and moral inferences based on Tomz and Weeks (2013), with the expectation that subjects would view non-white countries as more threatening and believe that striking them is less immoral, thus easing the psychological brakes on doing those countries harm. Supplementary Section A2.2.3 and Dataverse Appendix B4 show that the beneficial effect of democracy on threat perception and morality substantially declines for non-white countries. These results align with the two faces of racism outlined by Freeman, Kim, and Lake (2022): one face that approaches non-white countries in a paternalistic and legalistic manner, centered in part on moral deficiency, and a second face that views non-white countries as hostile and threatening. This tendency to divide the world into “virtuous in-groups and nefarious out-groups” is central to the concept of ethnocentrism (Bizumic, Monaghan, and Priest 2021; Kam and Kinder 2007), further suggesting that ethnocentrism explains the democratic peace in public opinion. Importantly, these results imply that racialization likely influences other important variables in foreign policy opinion research, beyond our focus on support for preventive strikes.

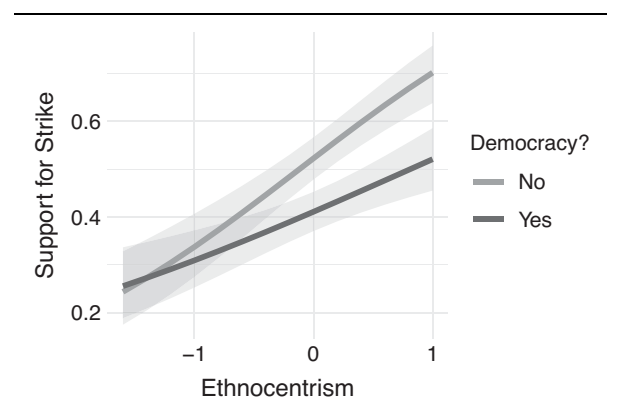
Taken together, our results across two surveys suggest that racialization explains much of the democratic peace in public opinion. Respondents higher in ethnocentrism drive the overall treatment effect of democracy, and the provision of non-white information eliminates the democratic peace effect within this subpopulation. This implies that the term “democracy” (relative to “nondemocracy”) conveys whiteness, and respondents with ethnoculturally superior beliefs support the use of force when race is specified otherwise.

HIDING IN PLAIN SIGHT: ETHNOCENTRISM IN PREVIOUS FINDINGS

Our original surveys indicate that racial presumptions about democracy help to explain the democratic peace in public opinion, particularly for respondents higher in ethnocentrism. Do these results shed light on past findings? Here, we reanalyze Tomz and Weeks (2013). If democracy implies whiteness, then ethnocentrism will be a major force behind the preference shown to democracies.

Tomz and Weeks (2013) measured ethnocentrism with items that center on anti-immigration attitudes. These items are a useful proxy for ethnocentrism, because recent reviews suggest that hostility to immigration is a function of perceived cultural threat (Hainmueller and Hopkins 2014). As described in Supplementary Section A1.2, we reduce the items to unidimensional factors scores, where higher values indicate greater ethnocentrism.

FIGURE 4. Tomz and Weeks (2013) Reanalysis: Ethnocentrism Moderates the Democratic Peace Effect



Note: Moderation effect estimated with linear regression. Column 2 of Supplementary Table A2 presents these results numerically.

If ethnocentrism explains the democratic peace effect in past work, then ethnocentrism should moderate the treatment effect of democracy. Figure 4 displays the relationship between support for a strike and the interaction between ethnocentrism and regime type assignment. Although ethnocentrism correlates positively with aggressive responses in general, individuals higher in ethnocentrism are significantly less likely to support the use of force against democracies in particular (coef = -0.32 , $p = 0.032$). At lower levels of ethnocentrism, respondents do not differentiate at all between democracies and nondemocracies. At higher levels of ethnocentrism, there is a substantively large, 18.1 point difference in the likelihood of endorsing the use of force. Further, splitting the sample at the ethnocentrism median reveals that the ATE for subjects below the median is statistically indistinguishable from zero (coef = -0.19 , $p = 0.28$). Respondents above the median entirely drive democracy’s effect (interaction coef = -0.53 , $p = 0.033$).

Supplementary Section A1 reports the same substantive results for the longitudinal data presented in Tomz and Weeks (2013) and for the survey of the British public conducted by Johns and Davies (2012). The latter also gives initial indication that our argument travels beyond the U.S. public.

PUTTING THE DEMOCRATIC PEACE INTO CONTEXT: RACIAL ASSOCIATIONS IN THE ENGLISH LANGUAGE

Our survey experiments and reanalyses suggest that racialization explains the pacifying effect of democracy in previous public opinion work on the democratic peace. This final empirical section moves beyond the survey experimental world to assess whether these same associations exist in the English language writ large, a demonstration of external validity. This analysis allows

us to validate the above results with an entirely different method—word embeddings—and generalize the results far beyond a handful of survey samples.

Over the past decade, word embeddings, or vector space models of text, have quickly become the gold standard for the analysis of prejudice in human language (e.g., Caliskan, Bryson, and Narayanan 2017; Garg et al. 2018; Kozlowski, Taddy, and Evans 2019). In contrast to traditional bag-of-words analyses, which rely on comparisons of simple word frequencies in a given document, word embeddings operationalize the notion that we can know a word by the company it keeps (Rodriguez and Spirling 2022). Embedding models use large inputs of digitized text to estimate the coordinate positions of each unique word relative to all other unique words in some N -dimensional geometric space. The estimated coordinates of each word reveal that words that share (lack) many linguistic contexts occupy similar (dissimilar) locations in this vector space, and these positions tell us something about the semantic similarity of the terms. For example, “if the distance between ‘immigrants’ and ‘hardworking’ is smaller for liberals than for conservatives, we learn something about their relative worldviews” (Rodriguez and Spirling 2022, 101).

Word embeddings provide particularly strong leverage on the analysis of prejudice. Although explicit bias certainly exists in English documents, embedding models are able to detect far more subtle instances. Inspired by the “general idea that text corpora capture semantics, including cultural stereotypes and empirical associations,” Caliskan, Bryson, and Narayanan (2017, 1) replicate findings from the implicit association paradigm, such as the finding that European-American names associate with “pleasant” terms and African-American names associate with “unpleasant” terms in the English language. Similarly, Kozlowski, Taddy, and Evans (2019) quantify gendered stereotypes associated with occupation, noting, for example, that the vector for “male” falls closer to “engineer,” whereas “female” falls closer to “nurse.” Garg et al. (2018) go so far as to quantify well-known gender and ethnic stereotypes across a century’s worth of the English language. As Garg et al. (2018, E3635) explain, “word embeddings... capture common stereotypes because these stereotypes are likely to be present, even if subtly, in the large corpora of training texts.” In other words, quotidian texts provide a window into culture.

This emerging work suggests that word embeddings quantify prejudice at massive linguistic scales. Do these same prejudices—notably racial and ethnocentric prejudices—associate with the democratic peace in the English language? We believe that word embeddings are useful for research on the mass public, because textual data (and the embeddings estimated from these data) reflect the same sorts of biases and implicit associations that we would typically measure via survey methods. That is, members of the public generate these texts, and these texts therefore provide useful information about implicit biases and prejudices. For example, if we asked members of the public to describe

democracies versus nondemocracies, we would expect terms like white or non-white to, respectively, coappear, despite the fact that we did not explicitly ask respondents about racial characteristics. Embeddings allow us to make similar average-level assessments on a societal scale.

Recall our argument’s intuition: if racialization explains the democratic peace in public opinion, then the relationship between democracy and peace should substantially decrease for non-white countries. In a word embeddings context, we can similarly assess whether a positive correlation exists between “democratic” terms and “peace” terms, followed by a comparison of that same correlation when the “democratic” terms are averaged with “white” or “non-white” terms. To make these assessments, we use well-established pretrained word embeddings from Stanford’s NLP group (Pennington, Socher, and Manning 2014), the most commonly used embeddings in past work on stereotypical associations in the English language. The embeddings were estimated using the global vectors for word representation (GloVe) model, and the underlying corpus derives from massive web crawls of digitized, contemporary English materials (like newswire texts and the entirety of Wikipedia). We use the vectors trained in two hundred dimensions, with four hundred thousand unique English terms.

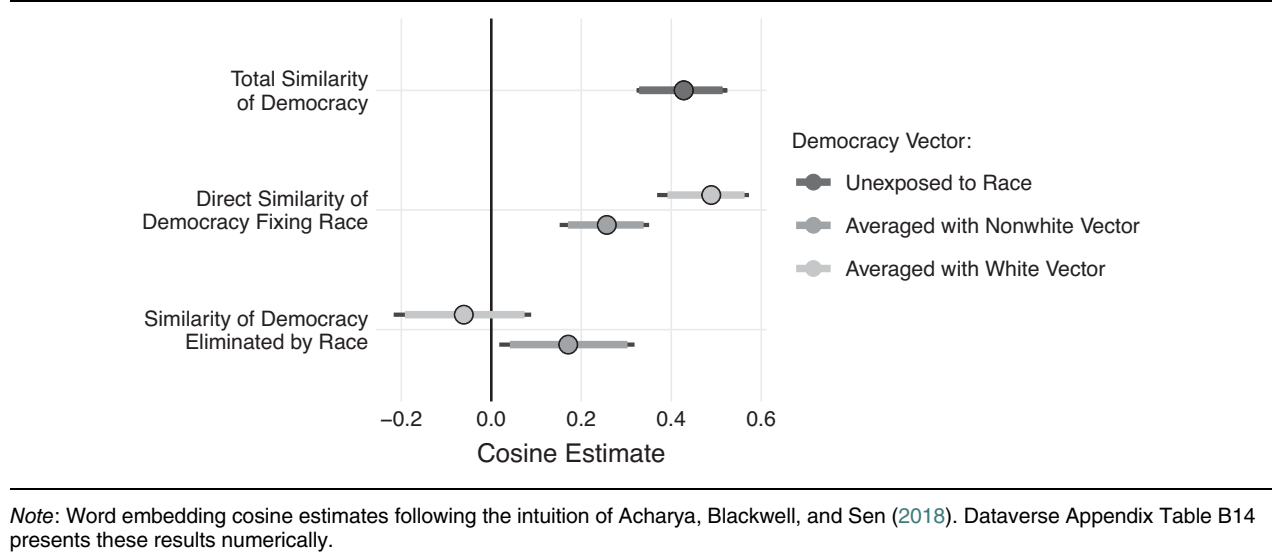
With these English language vectors in hand, we use the following terms to estimate our theoretical constructs of interest:

- *Democracy terms*: democracy, democratic, democratically, elect, elections, elected.
- *Peace terms*: peace, harmony, agreement, diplomacy.
- *White terms*: white, western, caucasian, european.
- *Non-white terms*: non-white, non-western, non-caucasian, non-european.

The average word coordinates for each set of dictionary terms defines the construct of interest, which is the most common method of combining vectors in embedding research on prejudice. For example, the average coordinates for the terms “white,” “western,” “caucasian,” and “european” provide a single averaged vector that represents the estimate of our white construct in the vector space. Mirroring the verbiage of our experiments, we use the prefix “non-” to form contrasts to the white terms.

Before estimating our primary quantities of interest, we note interesting descriptive results from these word embeddings. In line with our expectation that democracy is racialized, we find that the democracy terms show a much larger similarity to the white/western terms ($\cos \text{sim} = 0.389$) than non-white/nonwestern terms ($\cos \text{sim} = -0.003$), which suggests that democracy implicitly associates with whiteness in this corpus of English language texts. This association also parallels Dafoe, Zhang, and Caughey’s (2018) finding that respondents presume that democracies are predominately white.

Next, guided by the intuition of our experiments above, we seek to estimate two relationships in

FIGURE 5. Racialization of the Democratic Peace in the English Language


particular. First, akin to the ATE of democracy on pacifism, we calculate the cosine similarity between democracy and peace:

$$\text{Total similarity} = \text{cosine}(v_{\text{democracy}}, w_{\text{peace}}), \quad (1)$$

where $v_{\text{democracy}}$ and w_{peace} represent the average of dictionary term vectors for the democracy and peace constructs, respectively. Large, significant similarities would indicate that democracy correlates positively with peace in the English language.

Second, to assess whether the racialization of democracy helps to explain this association between democracy and peace, we estimate the same similarity between democracy and peace for “white democracies” versus “non-white democracies” separately (akin to the ACDEs above). To do so, we average the democracy terms with white versus non-white terms and simply recalculate the similarity between white democracies and peace, as well as non-white democracies and peace. More formally,

$$\begin{aligned} \text{Similarity fixing white} = \\ \text{cosine}[\text{mean}(v_{\text{democracy}}, u_{\text{white}}), w_{\text{peace}}], \end{aligned} \quad (2)$$

$$\begin{aligned} \text{Similarity fixing nonwhite} = \\ \text{cosine}[\text{mean}(v_{\text{democracy}}, u_{\text{nonwhite}}), w_{\text{peace}}], \end{aligned} \quad (3)$$

where u_{white} and u_{nonwhite} represent the average of dictionary term vectors for the white and non-white constructs, respectively. According to our expectations, democracy should display the same similarity to peace with or without white terms. By contrast, averaging the democracy terms with non-white terms will likely deflate the similarity between democracy and peace, such that non-white terms eliminate the positive

association between democracy and peace in the English language.

If past survey work on the democratic peace generalizes to the English language writ large, then democracy terms should correlate positively and significantly with the peace terms (akin to the ATE of democracy on pacifism in the above surveys). Figure 5 displays the results of our cosine similarity estimates, with 95% and 90% confidence intervals calculated using one thousand dictionary permutations. The first row displays the total cosine similarity between democracy and peace, the standard cosine similarity estimate reported in almost all social science applications of word embeddings. It is positive and substantively noteworthy (cos sim = 0.43, 95% CI [0.323, 0.524]). This similarity indicates that we can successfully replicate associations between democracy and peace at the mass public level in a corpus that essentially represents the entire contemporary English language, a notable finding in its own right.

Furthermore, if racialization underwrites the democratic peace in public opinion, then whiteness should do little to alter the similarity between democracy and peace, whereas non-whiteness should substantially impinge on the relationship between democracy and peace. The second row of Figure 5 shows that the similarity between democracy and peace substantially declines when we average the democracy terms with the non-white terms (cos sim = 0.26, 95% CI [0.152, 0.351]). That is, just as we found in our original experiments above, the pacifying effect of democracy substantially declines in the presence of non-white information. By contrast, averaging the democracy terms with white terms yields no statistical change in the similarity between democracy and peace (cos sim = 0.49, 95% CI [0.369, 0.573]). Similar to our experimental results, whiteness is baked into conceptions of democracy and therefore the democratic peace in public opinion.

The third and final row of Figure 5 simply takes the difference between the aforementioned quantities to estimate the amount of the similarity between democracy and peace eliminated by the presence of racial information, akin to the eliminated effect estimate in Acharya, Blackwell, and Sen (2018). The eliminated similarity shows that the decline in association between democracy and peace in the presence of non-white terms is substantively and statistically noteworthy: fixing democracy to non-white (i.e., averaging the democracy and non-white terms) eliminates 40% of the total similarity between democracy and peace. This estimate closely parallels the proportion eliminated in our well-powered experiment fielded on Prolific (45.1%). Indeed, Caliskan, Bryson, and Narayanan (2017) find that word embeddings yield effect sizes that parallel those found in implicit association tests, despite the methodological difference. By contrast, averaging the democracy terms with the white terms does nothing to statistically or substantively eliminate the similarity between democracy and peace, as we found in our first experiment.¹³

Finally, we note that in the context of this English language corpus, we largely expect the results in Figure 5 to capture the dyadic democratic peace phenomenon captured in our survey experiments. However, Supplementary Section A3.1 empirically verifies this expectation, showing that the results are substantively the same when we add “us” and “them” terms to the democracy dictionary, thus explicitly designating a dyadic democratic peace phenomenon. These results provide further evidence for the racialization of the democratic peace in public opinion.

Together, this analysis helps to validate the results found in our original experiments but generalizes these results far beyond two survey samples. The standard interpretation in the embeddings literature is that this underlying corpus represents the contemporary English language writ large. Even if non-white democracies enjoy a democratic peace effect in public opinion, in the English language, this effect is some 40% less pacific than the peace enjoyed by white democracies, which presents a serious challenge to the theoretical microfoundations traditionally thought to underlie the democratic peace.

CONCLUSION

We find that the “democratic” peace among the American public is based, at least in part, on the racial connotations of democracy. When we present respondents with hypothetical non-white countries, the pacifying effect of democracy declines, and respondents

scoring higher in ethnocentrism display marked racial bias in their willingness to use force. Reanalyses of past studies show that ethnocentrism disproportionately explains those results as well, and the phenomenon is not confined to the American mass public. Beyond our survey evidence, this same racial penalty pervades the English language, attesting to the external validity of our results.

Given that our findings focus on public opinion, our results alone preclude drawing firm conclusions about the democratic peace as a transhistorical process of relations among states, often evaluated through large-*N* analyses of dyadic relations. Given that democracy and race are subjective constructs, our argument resists the sort of “objective” coding of states as democratic or (non)white required for such analyses. Just as democracy implies race, race might imply democracy, precisely because of the racialization of democracy. Similarly, who counts as “white” is a subjective, time-varying process of social construction, not the result of objective physical characteristics. Thus, our argument cannot translate to a simple assignment of a racial variable to countries in a large-*N* dataset.

Nonetheless, we can assess whether elites exhibit the same racialization of democracy, with consequences for perceptions of peacefulness and aggressiveness. Democratic elites might resemble the public either because they are part of the same culture and therefore exposed to the same socialization processes or because they face strategic incentives to mimic the public’s positions given their accountability. The “normative” account of the democratic peace stresses the former, whereas the “institutional” account emphasizes the latter (Maoz and Russett 1993).

As a preliminary step toward assessment at the elite level, we conduct the same word embedding analysis presented above but instead utilize House of Commons speeches in the United Kingdom from 1945 to 2000.¹⁴ Supplementary Section A3.3 reports the same pattern of results at the British political elite level that we find in the non-elite embeddings. This supplementary analysis suggests that British elites implicitly racialize democracy, just like the public. This implicit racialization dovetails with recent international security research on the role of racial stereotypes among elite decision-makers (Mercer 2023). Beyond elites, these results provide further evidence that our argument extends beyond the United States.

Even if our results are confined to the mass public, though, we believe they have far-reaching implications for future work on racialization in foreign policy opinion. Consistent with the experimental democratic peace tradition, we primarily focus on support for preventive strikes. But, discussed above, we find that the same racialization of democracy affects perceptions of threat and immorality, two central variables in public opinion toward international security more generally (Kertzer et al. 2014). Foreign policy opinion studies of ethnocentrism and racialization to date examine the

¹³ In Supplementary Section A3.2, we use the `nss()` function in R’s `conText` package (Rodriguez, Spirling, and Stewart 2023) to examine the nearest terms to the racialized democracy vectors. Consistent with our argument that democracy is subtly and implicitly racialized, non-white terms are more prominent than white terms in the nearest neighbors to the non-white and white democracy vectors, respectively.

¹⁴ Supplementary Section A3.3 describes the estimation process.

important roles of anti-Muslim and anti-Asian attitudes (Kam and Kinder 2007; Kim 2022). However, our results suggest that racialization is lurking in security determinants traditionally considered race-neutral in mainstream international relations research.

For example, experimentally varying the possession of nuclear weapons in past work might prime more than strategic considerations (Herrmann, Tetlock, and Visser 1999). Among western publics, possession of nuclear weapons might activate connotations of whiteness and advancement, while non-white states seeking the bomb might be viewed as disproportionate threats (Intondi 2020). Similarly, past work often finds contradictory or null results when experimentally varying relative military capabilities (Herrmann, Tetlock, and Visser 1999; Tomz and Weeks 2013). One reason might be that western publics assume that militarily weak countries are more likely to be non-white. A basic premise of mainstream international relations theory is that weaker states are less threatening, but such racial inferences might inflate a sense of racialized threat. Further, on alliances, Hemmer and Katzenstein (2002) document racial underpinnings of the North Atlantic Treaty Organization. Perhaps experimental variation of alliance status in past work primes racialized bonds beyond an alliance's strict military utility.

Finally, we note that an understanding of race as an independent theoretical or experimental factor might be insufficient to reveal race's subtle effects. Mainstream international relations scholars often place "race" in explanatory competition with variables drawn from paradigms like realism and liberalism (Snyder 2023). Yet Dataverse Appendix B2 shows that we find scant evidence of a main effect of race in our experiments. The effect of race would have gone unnoticed without an experimental design built to draw out implicit assumptions. In the case of the democratic peace in public opinion, race was hiding behind associations with democracy. This finding suggests that mainstream IR scholars could gain from engagement with critical theoretic insights on the racialization of seemingly race-neutral "independent" variables.

SUPPLEMENTARY MATERIAL

To view supplementary material for this article, please visit <https://doi.org/10.1017/S0003055424000509>.

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/9JWDKK>.

ACKNOWLEDGMENTS

For feedback, the authors are grateful to Desh Girod, Kara Hooser, David Peterson, Johanna Rodehau-

Noack, participants at the University of Notre Dame's 2023 Race and International Relations Workshop, audiences at the University of Chicago and University of Southern California, and the journal's editors and reviewers. The authors also thank Jessica Weeks, Robert Johns, and Graeme Davies for providing their data and codebooks for reanalysis.

FUNDING STATEMENT

This research was supported by grants from the University of Southern California.

CONFLICT OF INTEREST

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

ETHICAL STANDARDS

The authors declare the human subjects research in this article was reviewed and approved by the University of Southern California and certificate numbers are provided in the Supplementary Material. The authors affirm that this article adheres to the principles concerning research with human participants laid out in APSA's Principles and Guidance on Human Subject Research (2020).

REFERENCES

- Acharya, Avidit, Matthew Blackwell, and Maya Sen. 2018. "Analyzing Causal Mechanisms in Survey Experiments." *Political Analysis* 26 (4): 357–78.
- Anievas, Alexander, Nivi Manchanda, and Robbie Shilliam. 2015. *Race and Racism in International Relations: Confronting the Global Colour Line*. New York: Routledge.
- Banaji, Mahzarin R., and Anthony G. Greenwald. 2016. *Blindspot: Hidden Biases of Good People*. New York: Bantam.
- Bell, Duncan. 2014. "Before the Democratic Peace: Racial Utopianism, Empire and the Abolition of War." *European Journal of International Relations* 20 (3): 647–70.
- Bizumic, Boris, and John Duckitt. 2012. "What Is and Is Not Ethnocentrism? A Conceptual Analysis and Political Implications." *Political Psychology* 33 (6): 887–909.
- 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.
- Bizumic, Boris, Conal Monaghan, and Daniel Priest. 2021. "The Return of Ethnocentrism." *Political Psychology* 42 (S1): 29–73.
- Blaut, James Morris. 2000. *Eight Eurocentric Historians*. New York: Guilford Press.
- Bowden, Brett. 2021. "Eurocentrism and Civilization." In *Routledge Handbook of Historical International Relations*, eds. Benjamin De Carvalho, Julia Costa Lopez and Halvard Leira, 162–70. New York: Routledge.
- Bueno de Mesquita, Bruce, James D. Morrow, Randolph M. Siverson, and Alastair Smith. 1999. "An Institutional Explanation of the Democratic Peace." *American Political Science Review* 93 (4): 791–807.

- 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.
- Caliskan, Aylin, Joanna J. Bryson, and Arvind Narayanan. 2017. "Semantics Derived Automatically from Language Corpora Contain Human-Like Biases." *Science* 356 (6334): 183–6.
- Chu, Jonathan A., and Stefano Recchia. 2022. "Does Public Opinion Affect the Preferences of Foreign Policy Leaders? Experimental Evidence from the UK Parliament." *Journal of Politics* 84 (3): 1874–7.
- Dafoe, Allan. 2011. "Statistical Critiques of the Democratic Peace: Caveat Emptor." *American Journal of Political Science* 55 (2): 247–62.
- Dafoe, Allan, Baobao Zhang, and Devin Caughey. 2018. "Information Equivalence in Survey Experiments." *Political Analysis* 26 (4): 399–416.
- Devos, Thierry, and Mahzarin R. Banaji. 2005. "American = White?" *Journal of Personality and Social Psychology* 88 (3): 447–66.
- Doyle, Michael W. 2005. "Three Pillars of the Liberal Peace." *American Political Science Review* 99 (3): 463–6.
- 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.
- Garg, Nikhil, Londa Schiebinger, Dan Jurafsky, and James Zou. 2018. "Word Embeddings Quantify 100 Years of Gender and Ethnic Stereotypes." *Proceedings of the National Academy of Sciences* 115 (16): E3635–E3644.
- Gartzke, Erik. 2007. "The Capitalist Peace." *American Journal of Political Science* 51 (1): 166–91.
- Goldgeier, James M., and Michael McFaul. 1992. "A Tale of Two Worlds: Core and Periphery in the Post-Cold War Era." *International Organization* 46 (2): 467–91.
- Hainmueller, Jens, and Daniel J. Hopkins. 2014. "Public Attitudes toward Immigration." *Annual Review of Political Science* 17: 225–49.
- Hainmueller, Jens, and Daniel J. Hopkins. 2015. "The Hidden American Immigration Consensus: A Conjoint Analysis of Attitudes toward Immigrants." *American Journal of Political Science* 59 (3): 529–48.
- Hanchard, Michael. 2018. *The Spectre of Race: How Discrimination Haunts Western Democracy*. Princeton, NJ: Princeton University Press.
- Hansen, Christina Novak, and Peter Thisted Dinesen. 2023. "Terrorism Activates Ethnocentrism to Explain Greater Willingness to Sacrifice Civil Liberties: Evidence from Germany." *Political Science Research and Methods* 11 (2): 402–9.
- Hayes, Jarrod. 2013. *Constructing National Security: US Relations with India and China*. New York: Cambridge University Press.
- Hemmer, Christopher, and Peter J. Katzenstein. 2002. "Why Is There No NATO in Asia? Collective Identity, Regionalism, and the Origins of Multilateralism." *International Organization* 56 (3): 575–607.
- Henderson, Errol A. 2013. "Hidden in Plain Sight: Racism in International Relations Theory." *Cambridge Review of International Affairs* 26 (1): 71–92.
- Herrmann, Richard K., Philip E. Tetlock, and Penny S. Visser. 1999. "Mass Public Decisions to Go to War: A Cognitive-Interactionist Framework." *American Political Science Review* 93 (3): 553–73.
- Hobson, John M. 2015. "Re-Embedding the Global Colour Line within Post-1945 International Theory." In *Race and Racism in International Relations: Confronting the Global Colour Line*, eds. Alexander Anievas, Nivi Manchanda and Robbie Shilliam, 81–97. New York: Routledge.
- Hor, Amos J. Y. 2024. *Making the World Safe for White Democracy: Racialized Liberalism and the Post-War International Order*. Book Manuscript in Preparation.
- Howell, Alison, and Melanie Richter-Montpetit. 2020. "Is Securitization Theory Racist? Civilizationism, Methodological Whiteness, and Antiracist Thought in the Copenhagen School." *Security Dialogue* 51 (1): 3–22.
- Huddy, Leonie, and Stanley Feldman. 2009. "On Assessing the Political Effects of Racial Prejudice." *Annual Review of Political Science* 12: 423–47.
- Hurwitz, Jon, and Mark Peffley. 1987. "How are Foreign Policy Attitudes Structured? A Hierarchical Model." *American Political Science Review* 81 (4): 1099–120.
- Intondi, Vincent J. 2020. *African Americans Against the Bomb: Nuclear Weapons, Colonialism, and the Black Freedom Movement*. Stanford, CA: Stanford University Press.
- 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.
- Kahl, Colin H. 1998. "Constructing a Separate Peace: Constructivism, Collective Liberal Identity, and Democratic Peace." *Security Studies* 8 (2–3): 94–144.
- Kam, Cindy D., and Donald R. Kinder. 2007. "Terror and Ethnocentrism: Foundations of American Support for the War on Terrorism." *Journal of Politics* 69 (2): 320–38.
- Kertzer, Joshua D., Kathleen E. Powers, Brian C. Rathbun, and Ravi Iyer. 2014. "Moral Support: How Moral Values Shape Foreign Policy Attitudes." *Journal of Politics* 76 (3): 825–40.
- Kim, Daegyeong. 2022. "Anti-Asian Racism and the Racial Politics of US–China Great Power Rivalry." PhD diss. University of California, San Diego.
- Kinder, Donald R., and Cindy D. Kam. 2010. *Us Against Them: Ethnocentric Foundations of American Opinion*. Chicago, IL: University of Chicago Press.
- Kinder, Donald R., and David O. Sears. 1981. "Prejudice and Politics: Symbolic Racism versus Racial Threats to the Good Life." *Journal of Personality and Social Psychology* 40 (3): 414–31.
- Kozlowski, Austin C., Matt Taddy, and James A. Evans. 2019. "The Geometry of Culture: Analyzing the Meanings of Class through Word Embeddings." *American Sociological Review* 84 (5): 905–49.
- Lake, Marilyn, and Henry Reynolds. 2008. *Drawing the Global Colour Line: White Men's Countries and the Question of Racial Equality*. Cambridge: Cambridge University Press.
- Maass, Richard W. 2023. "Racialization and International Security." *International Security* 48 (2): 91–126.
- Maoz, Zeev, and Bruce Russett. 1993. "Normative and Structural Causes of Democratic Peace, 1946–1986." *American Political Science Review* 87 (3): 624–38.
- Mercer, Jonathan. 2023. "Racism, Stereotypes, and War." *International Security* 48 (2): 7–48.
- Mills, Charles. 1997. *The Racial Contract*. Ithaca, NY: Cornell University Press.
- Neuliep, James W., and James C. McCroskey. 1997. "The Development of a US and Generalized Ethnocentrism Scale." *Communication Research Reports* 14 (4): 385–98.
- Nosek, Brian A., Anthony G. Greenwald, and Mahzarin R. Banaji. 2005. "Understanding and Using the Implicit Association Test: Method Variables and Construct Validity." *Personality and Social Psychology Bulletin* 31 (2): 166–80.
- Omi, Michael, and Howard Winant. 2014. *Racial Formation in the United States*. New York: Routledge.
- 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)*, 1532–43. Cedarville, OH: Association for Computational Linguistics.
- Persaud, Randolph, and R. B. J. Walker. 2001. "Apertura: Race in International Relations." *Alternatives* 26 (4): 373–6.
- Press, Daryl G., Scott D. Sagan, and Benjamin A. Valentino. 2013. "Atomic Aversion: Experimental Evidence on Taboos, Traditions, and the Non-Use of Nuclear Weapons." *American Political Science Review* 107 (1): 188–206.
- Rathbun, Brian C., Christopher Sebastian Parker, and Caleb Pomeroy. 2024. "Replication Data for: Separate but Unequal: Ethnocentrism and Racialization Explain the 'Democratic' Peace in Public Opinion." Harvard Dataverse. Dataset. <https://doi.org/10.7910/DVN/9JWDKK>
- 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.
- Rodriguez, Pedro L., Arthur Spirling, and Brandon M. Stewart. 2023. "Embedding Regression: Models for Context-Specific Description and Inference." *American Political Science Review* 117 (4): 1255–74.
- Rosato, Sebastian. 2003. "The Flawed Logic of Democratic Peace Theory." *American Political Science Review* 97 (4): 585–602.
- Sabaratanam, Meera. 2020. "Is IR Theory White? Racialised Subject-Positioning in Three Canonical Texts." *Millennium* 49 (1): 3–31.

- Sen, Maya, and Omar Wasow. 2016. "Race as a Bundle of Sticks: Designs That Estimate Effects of Seemingly Immutable Characteristics." *Annual Review of Political Science* 19: 499–522.
- Smith, Rogers M. 1997. *Civic Ideals: Conflicting Visions of Citizenship in US History*. New Haven, CT: Yale University Press.
- Sniderman, Paul M., Louk Hagendoorn, and Markus Prior. 2004. "Predisposing Factors and Situational Triggers: Exclusionary Reactions to Immigrant Minorities." *American Political Science Review* 98 (1): 35–49.
- Snyder, Jack. 2023. "How Central Is Race to International Relations?" *Security Studies* 32 (4–5): 892–906.
- Tomz, Michael R., and Jessica L. P. Weeks. 2013. "Public Opinion and the Democratic Peace." *American Political Science Review* 107 (4): 849–65.
- Tomz, Michael R., Jessica L. P. Weeks, and Keren Yarhi-Milo. 2020. "Public Opinion and Decisions about Military Force in Democracies." *International Organization* 74 (1): 119–43.
- Vitalis, Robert. 2000. "The Graceful and Generous Liberal Gesture: Making Racism Invisible in American International Relations." *Millennium* 29 (2): 331–56.
- Vucetic, Srdjan. 2011a. *The Anglosphere: A Genealogy of a Racialized Identity in International Relations*. Stanford, CA: Stanford University Press.
- Vucetic, Srdjan. 2011b. "A Racialized Peace? How Britain and the US Made Their Relationship Special." *Foreign Policy Analysis* 7 (4): 403–21.