Routledge
Taylor & Francis Group

Journal of Women, Politics & Policy

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/wwap20

Gender Stereotyping Questions Accurately Measure Beliefs About the Traits and Issue Strengths of Women and Men in Politics

Mirya R. Holman

To cite this article: Mirya R. Holman (2023) Gender Stereotyping Questions Accurately Measure Beliefs About the Traits and Issue Strengths of Women and Men in Politics, Journal of Women, Politics & Policy, 44:1, 90-104, DOI: 10.1080/1554477X.2023.2162285

To link to this article: https://doi.org/10.1080/1554477X.2023.2162285

JOURNAL OF WOMEN, POLITICS & POLICY
2023, VOL. 44, NO. 1, 90–104
https://doi.org/10.1080/1554477X.2023.2162285

Routledge
Taylor & Francis Group

Gender Stereotyping Questions Accurately Measure Beliefs About the Traits and Issue Strengths of Women and Men in Politics

Mirya R. Holman

Department of Political Science, Tulane University, New Orleans, Louisiana, USA

ABSTRACT

The public often believes that men in political office are better at handling some issues or possess specific traits, compared to women. Do individuals reveal their true preferences on surveys that inquire about these political gender stereotypes? This article employs methods that allow researchers to examine true attitudes without pressuring individuals to explicitly reveal sensitive preferences. I use three experiments: a list experiment, a new group-count sensitive measure, and a question-wording experiment employed on the 2016 CES. I find little evidence of reluctance to share true attitudes about gender stereotypes across any of the measures. The results presented help confirm the importance of gender stereotypes in shaping political preferences in American politics today and undergird evidence in prior scholarship on stereotypes.

KEYWORDS

Political gender stereotypes; measurement; social desirability bias; surveys; experiments; gender roles

In a *Washington Post* op-ed discussing Hillary Clinton's use of her gender in her 2016 Presidential campaign, Ruth Marcus opined "The 2016 Clinton campaign has no doubt that gender benefits Clinton. The single-digit slice of voters who say they would not elect a woman are not Clinton supporters in any event, in the campaign's view, and are outweighed by the facts that women voters make up more than a majority of the electorate and tend to vote more heavily Democratic." Polls backed the consensus that voters no longer hold explicit bias against women running for political office, with most individuals reporting that they would vote for a woman for president. These views can be paired with declining endorsements of political gender stereotypes, or beliefs about the behavior and attributes of political candidates and leaders based on their gender. For example, as Figure 1 shows, the segment of the American public that report believing that "Most men are better suited emotionally for politics than are most women" has declined from just under 50% in 1975 to a low of 7% in 2002, although that number increased to 12–18% in the 2010s.

Yet, these trends may be misleading. Social desirability bias, or the motivation to disguise actions or beliefs because of a wish to conform to social norms (Berinsky 2004; Klar, Weber, and Krupnikov 2016), may promote particular individuals to underreport their beliefs about the capability of women to govern effectively. Some evidence suggests that true attitudes about women in politics are less positive than traditional polls would lead us to believe (Krupnikov, Piston, and Bauer 2016; Mo 2014; Streb et al. 2008). For example, question wording shapes voters willingness to support women in experimental contexts (Krupnikov, Piston, and Bauer 2016). At the same time, work during the 2016 election found very little underreporting due to social desirability bias on the question of support for a female president (Burden, Ono, and Yamada 2017; Claassen and Ryan 2016).

In this article, I focus on the possibility that voters downplay or hide their attitudes about *political gender stereotypes*, or beliefs about the strengths and weaknesses of political leaders based exclusively on their gender (Schneider and Bos 2019; Sweet-Cushman 2022). These views are distinct from, but

CONTACT Mirya R. Holman mholman@tulane.edu Department of Political Science, Tulane University, New Orleans, Louisiana, USA

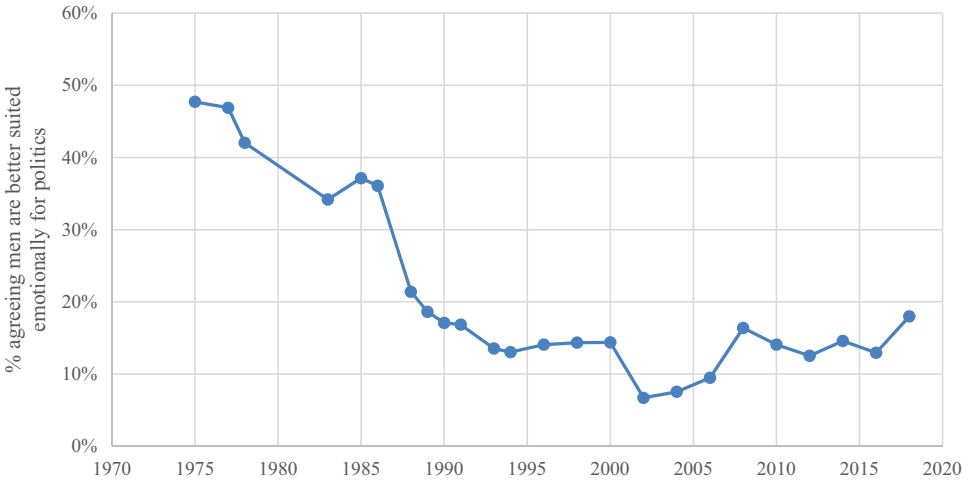

Figure 1. Share of the Population Who See Men as Better Suited Emotionally for Politics. Data from the General Social Survey.

deeply rooted in views of gender itself and nonpolitical gender stereotypes. Gender is constructed from social interactions and expectations about a social category; this is distinct from sex, which a category associated with biological markers. These expectations about gender then shape broader gendered attitudes, such as the idea that women are better suited for the home (Holman, Merolla, and Zechmeister 2022; McDermott 2016). Political gender stereotypes might include views that men are better at handling certain issues as political leaders, such as foreign policy, or are more likely to possess traits associated with success in politics, like assertiveness and strong leadership (Cassese and Holman 2018; Holman, Merolla, and Zechmeister 2016; Huddy and Terkildsen 1993).

The research on women in office and social desirability bias focuses on willingness to support a woman on the ballot or actual vote choice. While important, these attitudes only capture a small set of gendered attitudes about political actors. A substantial majority of voters might have been willing to vote for a woman for president, but still held a wide set of gender stereotypical beliefs that shaped their willingness to support Clinton in 2016 (Bernhard 2022; Bracic, Israel-Trummel, and Shortle 2019; Cassese 2020; Cassese and Barnes 2018; Cassese and Holman 2019; Conroy 2018; Conroy, Joesten Martin, and Nalder 2020). And, while whether people would vote for a woman for president is important, these questions are not directly applicable to 99.9% of women who run for and serve in office at the local, state, and national levels in the United States or the millions of women engaged in politics in other countries.

I evaluate whether social desirability is associated with the underreporting of gender stereotypes about political leaders using a combination of three experiments. Each experiment utilizes a common gender stereotype question wording, varied with an alternative approach designed to provide a more accurate measure of stereotypes. These alternatives include a list experiment, a question response experiment, and a new measure, piloted here, that provides an easily accessible group-based measure of social desirability bias and true attitudes. Across the studies, I show participants accurately answer questions about gender stereotypes, with very few differences between traditional methods of question formation (i.e., just asking respondents what they think) and sensitive methods (i.e., asking in a way that allows the respondent to hide their views). I conclude that researchers should feel comfortable using traditional means of assessing gender stereotypes.

Perceptions of female candidates and leaders

A wide variety of causes contribute to women's underrepresentation in politics, including a low level of political candidates because women lack interest running for office (Schneider et al. 2016; Shames et al.

2020) and that political parties and elites are less likely to recruit women (Crowder-Meyer 2013; Sanbonmatsu 2006). There is also the possibility that voters hold gendered and sexist attitudes, which reduce their willingness to support women running for office. Sanbonmatsu (2002, 20) finds that "all else equal, some voters would rather vote for and be represented by a man, and some voters would rather vote for and be represented by a woman."

Researchers' ability to evaluate how voters' attitudes shape their voting behavior may be limited by social desirability bias, or motivated actions aimed at conforming to social norms (Berinsky 2004). Survey respondents sometimes provide socially acceptable answers, where they "consciously misreport attitudes in order to avoid embarrassment" (Zaller and Feldman 1992, 610). Social desirability can shape responses ranging from voter turnout (Presser 1990) to support for racial integration (Krysan 1998).

While issues around race and social desirability bias have received the lion's share of scholarly attention, gendered attitudes are also subject to these inclinations. For example, men respondents offer different answers on gender equality questions to women and men interviewers (Huddy et al. 1997; Kane and Macaulay 1993). Implicit measures of gender stereotypes also reveal higher levels of stereotype endorsement than explicit measures (Chen et al. 2022; Mo 2014). Similarly, mechanisms designed to probe true responses indicate lower levels of support for women in public office than public opinion polls would indicate (Krupnikov, Piston, and Bauer 2016; Setzler 2019). And Streb et al. (2008) find that when they use a method developed for uncovering sensitive attitudes, the rate of support for a female president is much lower than indicated in public opinion polls.

But political gender *stereotypes* might be different. Sanbonmatsu (2002, 31) argues that gender stereotyping is not seen as unacceptable as racial stereotyping "perhaps because of biological differences between the sexes and a sexual division of labor that appears to be natural." Gender trait stereotypes are based on gender roles, where men and women are expected to have the traits associated with gendered socialization patterns (Eagly and Karau 2002; Schneider and Bos 2019). These include that women are more compassionate, have better interpersonal skills, and are less aggressive, while men are seen as stronger leaders, more assertive, and less capable at interpersonal relationships (Huddy and Terkildsen 1993). Gender issue stereotypes, where women are seen as more capable at handling policies relating to children and women and men are seen as capable of handling national defense and issues of security also emerge out of beliefs about appropriate gender roles, as well as accurate reflections of women's preferences in the general population and their behavior in political office (Atkinson 2020; Beauregard, Holman, and Sheppard 2022; Ondercin 2017; Shames et al. 2020).

It is also possible that social desirability biases are simply no longer in operation. In 2016, Burden, Ono, and Yamada (2017) found the share of those "hiding" their true attitudes had diminished to half of what Streb et al. (2008) uncovered. And studies run during the 2016 election found little evidence of voters hiding their opposition to Clinton on surveys (Bracic, Israel-Trummel, and Shortle 2019; Claassen and Ryan 2016); if anything, voters may have hid their views about Trump to a limited degree (Klar, Weber, and Krupnikov 2016).

Overall, research on social desirability bias suggests that voters might hold stereotypes about women in politics but may also be unwilling to express those attitudes in traditional public opinion surveys. Sensitive methods are better at measuring true attitudes, but these inquires have not examined gender stereotypes. Thus, there is clear room in the scholarship for evaluations of the general levels of gender stereotypes through more accurate research methods that take social desirability bias into account. Alternatively, the broad acceptance of stereotypes may mean that people are willing to express gender stereotypes openly. The later would be of comfort to gender and politics scholars, who routinely use a standard set of gender stereotype questions in their research.

The nature and measure of gender stereotypes

In many elections, particularly low information environments, "voters look for bargains when engaging in political decision-making" (Fulton and Ondercin 2015, 2). These information "deals" often

include heuristics like candidate gender, race, or incumbency status that replace more complete information about candidate issue competencies or traits as shortcuts toward a more informed evaluation of current leaders (Bauer 2019; Crowder-Meyer, Gadarian, and Trounstine 2015; Trounstine 2010). Gender stereotypes provide one of these easy shortcuts that allow voters to make decisions in information-scant environments. These stereotypes can take many forms, including that women are more liberal than men (i.e., Fulton and Ondercin 2015; Roberts and Utych 2022), that they are less viable candidates (Bateson 2020), and that women or men are better suited for particular offices like school boards (Bernhard and Anzia 2022). I focus on trait, issue, and emotions stereotypes in this article, given their widespread use in the scholarship.

Trait stereotypes are based on the idea that men are more likely to possess some traits like being more assertive or stronger leaders, while women are more kind or compassionate (Eagly and Karau 2002). These views are rooted in gender roles, where girls and boys are taught, through social expectations and role models, to behave in ways that are consistent with their gender (Bos et al. 2022; Schneider, Bos, and DiFilippo 2022). These trait stereotypes are commonly used in academic work on gender and politics (Aaldering and Van Der Pas 2020; Bauer 2020a; Cassese and Holman 2018) and cover a broad set of characteristics from aggressive and assertive to compassionate to honest (Barnes and Beaulieu 2019; Barnes, Beaulieu, and Saxton 2020).

Issue stereotypes represent the downstream consequences of trait stereotypes. Here, individuals make assumptions about men's and women's ability to handle specific issues, including that men are better able to handle issues like foreign policy and women are more capable of handling issues like education (Cassese and Holman 2018; Holman et al. 2019). These issue stereotypes emerge from several sources, including the aforementioned gender role expectations, gender segregation in occupations (Barnes and Holman 2020), gaps in policy preferences between men and women in the general population (Holman 2014; Beauregard, Holman, and Sheppard 2022; Ondercin 2017, 2018), and the behavior and expertise of women in political office (Atkinson 2020; Atkinson and Windett 2019; Shim 2021). Because women in the population and women in office regularly demonstrate their preferences for policies consistent with feminine issue gender stereotypes (Greene et al. 2022; Lizotte 2020) and because men dominate political roles associated with masculine issue gender stereotypes (Barnes and O'Brien 2018; Holman, Merolla, and Zechmeister 2022; Ortbals and Poloni-Staudinger 2018), these attitudes are reinforced by what individuals see in the world.

Emotional suitability is the social expectation that the emotions that women are socialized to express more are not the emotions that are acceptable in the political arena (Bauer and Carpinella 2018; Boussalis and Coan 2021; Boussalis et al. 2021; Gabriel and Masch 2017; Masch 2020; Meeks 2012). Social expectations include that women are more emotional than men, although research shows that this is not, in fact, the case in the general population (Plant et al. 2000) and that women express more happiness and sadness (Boussalis et al. 2021). These pass on to expectations about women and men in political office, such that men are believed to be better emotionally suited for politics than are women, although Figure 1 demonstrates that the share of the population endorsing this belief has fallen over time.

Data and methods

A wide variety of techniques exist for measuring responses to sensitive questions, including randomization, list experiments, and endorsement experiments. While randomization provides researchers with clarity on mechanisms and processes, it does not protect the identities of the participants. I turn to three experiments to evaluate true attitudes without the restriction of social desirability bias: a list experiment, a new measure of randomizing a counting process of group comparisons, and an answer-option experiment.

List experiment methods and data: The unobtrusive approach of a list experiment has been used extensively in political science to evaluate social desirability effects in a variety of areas, including in

attitudes on race (Kuklinski et al. 1997), religion (Benson, Merolla, and Geer 2011), and gender (Claassen and Ryan 2016; Streb et al. 2008). My list experiment followed existing methods (Rosenthal, Imai, and Shapiro 2016) and proceeded as follows: I randomly assigned individuals into control, treatment, and traditional survey conditions. The first (*control*) group receives this prompt: "The following are three things that Americans may agree with. Of this list, HOW MANY do you agree with? There is no need to indicate which items, just HOW MANY."[1]

> Presidential campaigns are too costly.
> The US should have stronger immigration laws.
> Congress is doing a really good job.

I provided the *treatment* group with one of the following as a fourth item on the list:

> Most men are better suited emotionally for politics than are most women. (Treatment I)
> Most men in elected office are better at handling foreign affairs than are most women. (Treatment II)

The list experiment works by taking the average number of items in the Control group and subtracting it from the average number of items reported in a Treatment group, which gives the percent of people in each treatment group that agree with the sensitive item. I asked a final group (*traditional method*) the questions through traditional survey questions. The responses to this question supply levels of agreement when social desirability bias is in effect. I use differences of proportions tests to evaluate whether the percent of people who agree with the statement differs statistically from the portion of people who agree with the statement when queried through traditional survey mechanisms.

The sample was recruited through Amazon.com's Mechanical Turk (MTurk) program, which provides a platform for individuals and businesses to hire workers to complete tasks, often online. I restricted participants to 18+ years old and US citizens through both the MTurk recruitment and screener questions in the survey. After receiving consent from the participants, I randomly assigned the participants into experimental treatments. Participants then answered a series of questions on demographic information. The experiment concluded with a short memo detailing the purpose of the study and the provision of a code to receive payment from MTurk. The survey was administered to 750 participants in December 2015. The respondents are balanced across treatments in their age, income, partisanship, and race.

Table 1 presents the mean number of items that respondents agree with in the control and test groups, as well as the calculated percent that "agree" from these comparisons and from the group asked the question through traditional questions. Table 1 also presents the percent agreeing from the Traditional Survey group, where the participants provided directly what their level of agreement was with statement. The question's format was consistent with the GSS: I presented participants with a series of statements, with the options of agree, disagree, and not sure. The percentage presented here are those who indicated agree out of the entire group.

Overall, the same proportion of individuals agree with the statement for the emotions between the traditional and sensitive methods. In the traditional survey group in my data, this percentage is firmly in the middle of national polls, with 23% of respondents agreeing with the statement using both the

Table 1. Share Selecting "Men" on Stereotype Questions.

	Percent selecting men (sensitive method)	Percent selecting men (traditional method)	Difference
Control Condition (number of items with no sensitive items)			
Most men are better suited emotionally for politics than are most women	23%	23%	0%
Most men in elected office are better at handling foreign affairs than are most women	24%	21%	3%

N = 743. *Difference* column = % traditional - % sensitive. * $p < 0.05$ in a comparison of the difference between sensitive and traditional methods, using difference of proportions tests.

sensitive and traditional methods.[2] Similarly, on the question about foreign policy, 21% of respondents agreed with the statement, "Most men in elected office are better at handling foreign affairs than are most women" using the traditional method and 24% agreed using the sensitive method.

Study 2: Pick a side

Common measures of gender stereotypes in politics often ask respondents to choose between whether men or women are "better at" a particular task or whether a particular trait applies more to men or women. For example, scholars studying gender traits ask about a broad set of attributes, but often ask specifically about whether men or women are better described as *strong leaders* or *compassionate* or if women or men are better at handling *foreign affairs* or *education* (Bauer 2015a, 2020b; Cassese and Holman 2018; Holman, Merolla, and Zechmeister 2022; Schneider 2014). To understand if asking respondents these sorts of questions elicit responses driven by social desirability bias, I created a new form of a list experiment. In the experiment, I provided respondents with this prompt: "We are interested in who you think is better at handling different types of policies in the United States. For each policy, you will be given a series of pairs – please just count the number of groups on SIDE A that you believe are better at handling this issue. There is no need to indicate which groups, just HOW MANY."[3] I randomly assigned the participants to two conditions. In the *first* condition, respondents are provided a table like Table 2.

In the *second* condition, respondents are provided with an identical prompt, but the position of "Men in Politics" and "Women in politics" is swapped as in Table 3.

I asked all respondents to tell me "How many groups from side A?" This allows for an aggregate evaluation of how many people will select "Men in politics" by evaluating the differences between the two conditions. Respondents are asked about feminine (*education*), masculine (*foreign policy*), and neutral (*the fiscal deficit*) issues. Gendered trait stereotypes are measured by asking respondents "We are interested in who you think is *better described with the following phrases* in the United States" with the same set-up. Respondents evaluated *compassionate*, *strong leaders*, and *knowledgeable*.

After the experimental treatment, all respondents also directly answered the questions, allowing for a direct comparison between the respondents' traditional and sensitive answers (Bullock, Imai, and Shapiro 2011). Respondents also answered a series of questions that measured their egalitarian and benevolent and hostile sexist attitudes (Swim et al. 1995). The sample was recruited through the LUCID Theorem program, which provides a platform for individuals and businesses to hire survey respondents, aggregated from a wide set of sources. LUCID provides demographics on their samples and uses a quota system to benchmark their samples against US national demographics. The survey was run in April 2017 and taken by 987 participants. The respondents are balanced across treatments in their age, income, partisanship, and race.

Table 2. Treatment 1.

Side A	Side B
Men in politics	Women in politics
Republicans	Democrats
Congress	The President
Business leaders	Government employees
Liberals	Conservatives

Table 3. Treatment 2.

Side A	Side B
Women in politics	Men in politics
Republicans	Democrats
Congress	The President
Business leaders	Government employees
Liberals	Conservatives

Table 4. Pick a Side Patterns.

Item	Percent selecting men (sensitive method)	Percent selecting men (traditional method)	Difference
Better at:			
Foreign policy	59%	50%	9%
Education	24%^	28%	−4%
Deficit policy	43%	50%	−7%
Better described as . . .			
Compassion	20%^	24%	−4%
Strong leader	47%	53%	−6%
Knowledgeable	56%	46%	10%

N = 987. Difference column: % traditional method - % sensitive methods. ^ $p < 0.05$ in a comparison of the difference between number of items in the "men's" side and "women's" side; * $p < 0.05$ in a comparison of the difference between sensitive and traditional methods, using difference of proportions tests (no differences are significant).

Do people pick sides?

As with the traditional list experiment, I find little evidence that people are hiding their true preferences on gender stereotype questions. Broadly, respondents viewed women as better at handling education and as compassionate via both the sensitive and traditional methods of evaluation and men as better at foreign policy, stronger leaders, and more knowledgeable (see Table 4).

Study 3: Question Wording

List experiments provide an opportunity to understand social desirability biases, but they are also unlikely to appear on traditional questionnaires as measures of gender stereotypes. I next test whether varying the traditional response options available on gender stereotype questions provides evidence of social desirability bias. Here, the primary difference is whether respondents are given the "no difference" option when asked about gender stereotypes, a common approach in public opinion research (Brown and Pope 2021). Here I explore the differences between the responses based on answer options.

Data was collected through a module on the 2016 Congressional Elections Study (formerly, the Congressional Cooperative Elections Study), a national, cooperative study where universities and researchers buy into the study and are provided space on both a pre-election and post-election questionnaire. Researchers' independent questions are asked of a subsample of the overall survey population, while a set of core questions are asked of all participants. The CES's sample of 60,000 respondents is drawn from YouGov panels, advertisements, and other survey draws.

Respondents were randomly assigned to either answer a battery of questions about gender stereotypes where they were provided the answer options of "Men" or "Women" (the *no middle option* response) or the answer options of "Men," "No difference," or "Women" (the *middle option* response). Again, I asked about the broad emotions stereotype, as well as issue stereotypes (about *foreign affairs* and *education*) and traits (including *assertive* and *compassionate*). Results for these comparisons are presented in Table 5 and 6.

Table 5. Difference Between Providing a Middle Response or Not on Gender Stereotype Questions.

	Select men, No middle option	Select men, Middle option	Difference
More emotionally suited for political office (Emotions)	53%	19%	33%*
Better at handling foreign affairs (Foreign Affairs)	63%	27%	36%*
Better at handling education (Education)	19%	5%	14%*
More assertive (Assertive)	74%	31%	43%*
More compassionate (Compassionate)	8%	4%	4%*

N = 504. Difference column: % selecting men with no middle option - % selecting men with a middle option. * $p < 0.05$ in a comparison of the difference between answer options, using difference of proportions tests.

Table 6. Sexism and Stereotypes.

	Foreign Policy	Education	Assertive	Compassionate
Hostile Sexism (No middle)	0.05	0.06*	0.03	0.06*
	(0.03)	(0.03)	(0.03)	(0.02)
Hostile Sexism (Middle)	0.07*	0.01	0.05	0.04*
	(0.03)	(0.01)	(0.03)	(0.01)
Benevolent Sexism (No middle)	0.05*	0.04	−0.03	0.02
	(0.02)	(0.02)	(0.02)	(0.01)
Benevolent Sexism (Middle)	0.08*	0.01	0.04	0.01
	(0.02)	(0.01)	(0.02)	(0.01)

Standard errors in parentheses. Each row is a separate logistical regression model with controls for gender, race, education, partisanship, and ideology. * $p<0.05$.

Consider the groups of people who might select a "no difference" option when presented with a common gender stereotype: one possibility is that people do not have an opinion on whether men or women are better described by a trait or better at handling a policy (the *don't knows*). Another possibility is that some respondents might truly think that there are no differences between the groups (the *no difference*). Finally, some people use the no difference option to "hide" their true opinions: they do have an opinion, but just do not want to provide it. I use two methods to identify the groups one and two from each other and from the third option: examining respondent characteristics and whether levels of sexism shape these answers.

To differentiate between the *don't knows* and the *no differences*, I first attempt to identify whether characteristics of the respondents differentially predict answers on preferences for men on the questions with and without the middle option. Specifically, I look at whether education is associated with any difference in preference, with the assumption that, as with other political attitudes, the "don't know" response declines with information and knowledge (Gooch and Vavreck 2019). Thus, if education is more strongly associated with attitudes when there *is* a middle option (or smaller standard errors), this would suggest that people who really do not know are selecting the middle option.

Alternatively, to differentiate between options one and two versus option three, I look for reliable and consistent patterns in how respondent characteristics or sexist attitudes shape response to those questions with or without a middle option (Brown and Pope 2021). For example, if women's responses on the questions with no middle option are significantly different than men's responses but not when there is a middle option, this would suggest underlying patterns in social desirability. However, if people are providing true responses to questions, regardless of the answer options, I will not see systematic responses.

Respondent characteristics and responses: I start with a series of logistic regression models that estimate the likelihood that any person selects "men" on the gender stereotype questions for both the no middle point and middle point treatments and gender, race, education, income, ideology, and partisanship as control variables (see Table 5 and Figure 2). I present a coefficient plot of the outcomes of the two options across the gender stereotype questions. Education is not consistently associated with one option or the other and I find little evidence that respondent characteristics differentially are associated with responses to the no middle or middle options. The patterns in the data suggest that individuals who select the middle option are doing so because they do not believe that there is a difference between men and women on that dimension.

Next, I examine whether endorsements of hostile and benevolent sexism are systematically associated with responses on these questions by the answer responses. Both forms of sexism are associated with views of gender more broadly in society and women in politics specifically (Beauregard and Sheppard 2020; Cassese and Barnes 2018), so it would not be unsurprising if these views were correlated generally with gender stereotypes. Despite social desirability, a substantial share of people still report holding sexist views (Cassese and Barnes 2018; Gothreau, Arceneaux, and Friesen 2022;

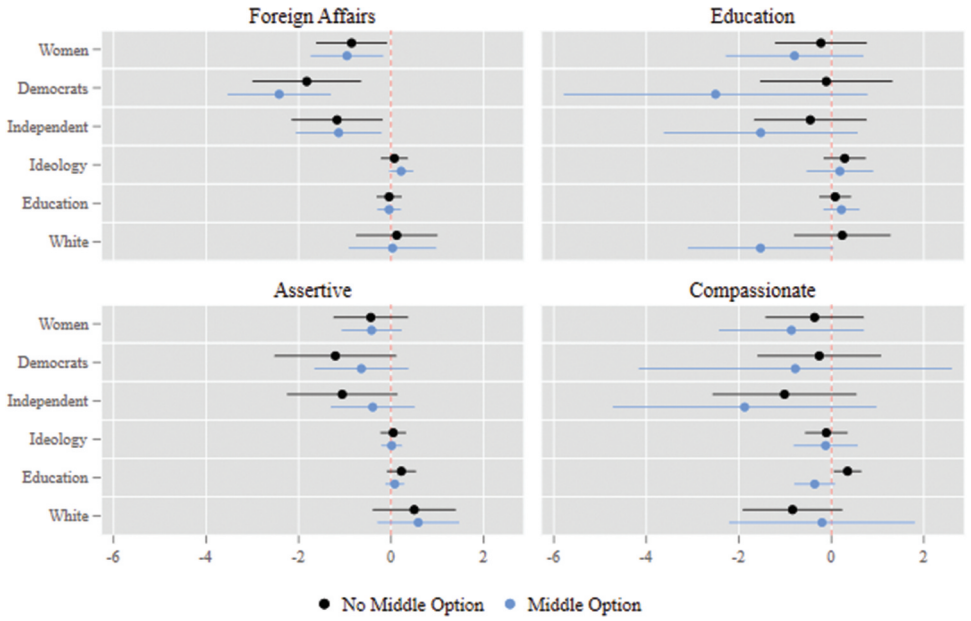

Figure 2. Respondent Characteristics and Selecting Men on Stereotype Questions. Coefficient plots derived from four separate logistical regression models. Survey weights applied.

Benegal and Holman 2021) and sexism shapes political preferences and voting behavior (Beauregard, Holman, and Sheppard 2022; Cassese and Holman 2019).

Consistent patterns in significance for one answer form but not the other could suggest broader patterns of social desirability. Again, I do not find evidence of systematic patterns (see Table 3). In three of the four models (hostile sexism, middle option, and both benevolent sexism models), sexism is associated with a higher likelihood of saying men are better at handling foreign policy, although the coefficient for the hostile sexism, no middle is in the same direction and just outside statistical significance. Hostile sexism is associated with seeing men as more compassionate across both answer options. In one model (hostile sexism, no middle), sexism is associated with seeing men as better at education. Broadly, these results do not suggest patterns in the relationship between sexism and gender stereotypes by the answer option.

Researchers regularly make decisions about the format of questions and answers when obtaining information about political gender stereotypes from target populations. The research presented here suggests that one answer option does not lead to satisficing on the part of respondents. It is not that answer options do not matter as the CES data demonstrates how the choice to include a middle option can dramatically shift the answers on the question. Researchers should weigh this choice carefully as they craft their research design.

Conclusion

In this article, I have provided three tests to evaluate the probability that individuals in the general population are hiding their true attitudes about gender stereotypes. I find very little evidence that social desirability is driving these results. Instead, these experiments would suggest that researchers should be comfortable using gender stereotype questions as the general population is reporting accurate beliefs and views across questions that favor men and women, on issues, traits, and emotional competency for political office. One pattern that stands out is that people are least likely to lie about stereotypes where women are advantaged, such as the compassionate trait; this suggests

that the public is willing to be truthful about women's advantages, but perhaps less so about their disadvantages.

The limits to the research on gender stereotypes also apply here, specifically in geographic focus and a general ignoring of intersectional identities. Research on political gender stereotypes is stymied by a focus almost exclusively on the United States, but new research shows that stereotypes shape voter behavior across country contexts (Devroe 2021; Holman, Merolla, and Zechmeister 2022; López-Cariboni and Reyes-Housholder 2022). And some research suggests that there may be more social desirability biases around these attitudes in other countries (Chen et al. 2022). Work on stereotypes has also almost universally focused on gender without a consideration of how other identities might shape these gendered expectations. For example, we know much less about how people stereotype women of color in political office, although recent work is starting to fill in these gaps (Cargile 2015; Gonzales and Bauer 2020, 2022; Runderkamp et al. 2022). Given that racial attitudes are also often hidden by social desirability biases, understanding whether people are more reluctant to express true views about women of color could be a fruitful avenue for future research. Additional work on gender and class also offers an opportunity to consider how the overlapping identities of women in office might shape stereotypes and people's willingness to express them (Barnes, Beall, and Holman 2021; Kim and Kweon 2022); for example, individuals may be particularly likely to agree with statements about men's or women's strengths in a particular policy area if first cued with information about a leader's experience in those areas.

As research on political gender stereotypes advances in political science and related fields, researchers increasingly note the importance of context in shaping the application of stereotypes (Bauer 2015a; Holman, Merolla, and Zechmeister 2022; Rohrbach 2022). One possibility is that the setting of research matters and social desirability may increase in settings where the dual processing model applies and voters' stereotypes become political meaningful (Bauer 2020b; Kunda 1999; Kunda and Spencer 2003). Researchers may want to consider whether panel data can be utilized (i.e., Holman, Merolla, and Zechmeister 2022) that allows for the assessment of the stereotype in settings separate from the context of candidate evaluation to decrease the possibility of underreporting on gender stereotypes. Still, experimental work rarely finds that experimental treatments affect the reported level of gender stereotypes in meaningful ways, even if the stereotype questions are asked after the treatment (e.g., Holman, Merolla, and Zechmeister 2011).

A possibility exists that there are heterogeneities in social desirability within groups that are disguised by the broad patterns – in essence, everyone is lying but in different directions. This might be particularly true given the uneven distribution of gender stereotypes and views about women in the general population (Bauer 2013, 2015b; Beauregard and Sheppard 2020; Gothreau, Warren, and Schneider 2022; Holman, Merolla, and Zechmeister 2022). As a final test, I examine this question and find little evidence: women and men, and Democrats and Republicans all are equally likely to review true beliefs about gender stereotypes, even as these groups reveal very different base attitudes on the stereotype measures. Taken together, this research should give those studying gender stereotypes confidence in the general measures of both issue and trait stereotypes commonly used in the scholarship.

Notes

1. I include one item with chronically low approval (Congress is doing a really good job) to avoid issues of a ceiling effect, or that individuals may be reluctant to provide their true answer if they agree with *all* the items in the treatment condition (Oliveros 2016). I randomized the order of all items.
2. Traditional polls generally find between 7 and 28% of people agree that "Most men are better suited emotionally for politics than are most women."
3. I randomized the order of groups within the table, as well as the order in which respondents saw the issue or trait questions, and the specific issues and traits within each of those question blocks.

Acknowledgments

Thanks to Erin Cassese, Angie Bos, Kris Kanthack, Caitlin Andrews Lee, Bethany Albertson, and Christina Xydias for their thoughts on project and the Gender and Political Psychology Writing Group for comments. Funds for the research provided by the Newcomb Institute and the Lavin Bernick Grants at Tulane University (approximate cost: $14,000 for data collection, including CCES module purchase and research assistance). Research approved by the Tulane Institutional Review Board (818773-1).

Disclosure statement

No potential conflict of interest was reported by the author(s).

Notes on contributor

Mirya R. Holman is an associate professor of political science at Tulane University. She is an expert on gender and politics, local politics, and political behavior. She's published widely in these areas, including her book, *Women in Politics in the American City* and her co-edited volume *Good Reasons to Run*. She's authored more than 60 articles published in leading outlets like *American Political Science Review, Journal of Politics, Perspectives on Politics, Political Behavior, Political Psychology*, and *Political Analysis*. She writes and distributes a weekly mentoring newsletter to academics called Mirya Holman's Aggressive Winning Scholars (MHAWS).

ORCID

Mirya R. Holman http://orcid.org/0000-0001-6648-4122

References

Aaldering, Loes, and Daphne Joanna Van Der Pas. 2020. "Political Leadership in the Media: Gender Bias in Leader Stereotypes During Campaign and Routine Times." *British Journal of Political Science* 50 (3):911–31. doi:10.1017/S0007123417000795.

Atkinson, Mary Layton. 2020. "Gender and Policy Agendas in the Post-War House." *Policy Studies Journal* 48 (1):133–56. doi:10.1111/psj.12237.

Atkinson, Mary Layton, and Jason Windett. 2019. "Gender Stereotypes and the Policy Priorities of Women in Congress." *Political Behavior* 41 (3):769–89. doi:10.1007/s11109-018-9471-7.

Barnes, Tiffany D., Victoria Beall, and Mirya R. Holman. 2021. "Pink Collar Representation and Policy Outcomes in U.S. States." *Legislative Studies Quarterly* 46 (1):119–54. doi:10.1111/lsq.12286.

Barnes, Tiffany D., and Emily Beaulieu. 2019. "Women Politicians, Institutions, and Perceptions of Corruption." *Comparative political studies* 52 (1):134–67. doi:10.1177/0010414018774355.

Barnes, Tiffany D., Emily Beaulieu, and Gregory Saxton. 2020. "Sex and Corruption: How Sexism Shapes Voters' Responses to Scandal." *Politics, Groups and Identities* 8 (1):103–21. doi:10.1080/21565503.2018.1441725.

Barnes, Tiffany D., and Mirya R. Holman. 2020. "Essential Work is Gender Segregated: This Shapes the Gendered Representation of Essential Workers in Political Office." *Social science quarterly* 101 (5):1827–33. doi:10.1111/ssqu.12850.

Barnes, Tiffany D., and Diana Z. O'Brien. 2018. "Defending the Realm: The Appointment of Female Defense Ministers Worldwide." *American Journal of Political Science* 62 (2):355–68. doi:10.1111/ajps.12337.

Bateson, Regina. 2020. "Strategic Discrimination." *Perspectives on Politics* 18 (4):1068–87. doi:10.1017/S153759272000242X.

Bauer, Nichole M. 2013. "Rethinking Stereotype Reliance." *Politics & the Life Sciences* 32 (1):22–42. doi:10.2990/32_1_22.

Bauer, Nichole M. 2015a. "Emotional, Sensitive, and Unfit for Office? Gender Stereotype Activation and Support Female Candidates." *Political Psychology* 36 (6):691–708. doi:10.1111/pops.12186.

Bauer, Nichole M. 2015b. "Who Stereotypes Female Candidates? Identifying Individual Differences in Feminine Stereotype Reliance." *Politics, Groups, and Identities* 3 (1):94–110. doi:10.1080/21565503.2014.992794.

Bauer, Nichole M. 2019. "The Effects of Partisan Trespassing Strategies Across Candidate Sex." *Political Behavior* 41 (4):897–915. doi:10.1007/s11109-018-9475-3.

Bauer, Nichole M. 2020a. "Running Local: Gender Stereotyping and Female Candidates in Local Elections." *Urban Affairs Review* 56 (1):96–123. doi:10.1177/1078087418770807.

Bauer, Nichole M. 2020b. "Shifting Standards: How Voters Evaluate the Qualifications of Female and Male Candidates." *The Journal of Politics* 82 (1):1–12. doi:10.1086/705817.
Bauer, Nichole M., and Colleen M. Carpinella. 2018. "Visual Information and Candidate Evaluations: The Influence of Feminine and Masculine Images on Support for Female Candidates." *Political Research Quarterly* 71 (2):395–407. doi:10.1177/1065912917738579.
Beauregard, Katrine, Mirya R. Holman, and Jill Sheppard. 2022. "Sexism and Attitudes Toward Policy Spending in Australia and the United States." *Frontiers in Political Science*. https://doi.org/10.3389/fpos.2022.892111
Beauregard, Katrine, and Jill Sheppard. 2020. "Antiwomen but Proquota: Disaggregating Sexism and Support for Gender Quota Policies." *Political Psychology* 42 (2):219–37. doi:10.1111/pops.12696.
Benegal, Salil, and Mirya R. Holman. 2021. "Understanding the Importance of Sexism in Shaping Climate Denial and Policy Opposition." *Climatic Change* 167 (3–4):1–19. doi:10.1007/s10584-021-03193-y.
Benson, Brett V., Jennifer L. Merolla, and John G. Geer. 2011. "Two Steps Forward, One Step Back? Bias in the 2008 Presidential Election." *Electoral studies* 30 (4):607–20. doi:10.1016/j.electstud.2011.05.004.
Berinsky, Adam J. 2004. "Can We Talk? Self-Presentation and the Survey Response." *Political Psychology* 25 (4):643–59. doi:10.1111/j.1467-9221.2004.00391.x.
Bernhard, Rachel. 2022. "Wearing the Pants(suit)? Gendered Leadership Styles, Partisanship, and Candidate Evaluation in the 2016 U.S. Election." *Politics & Gender* 18 (2):513–45. doi:10.1017/S1743923X20000665.
Bernhard, Rachel, and Sarah F. Anzia. 2022. "Gender Stereotyping and the Electoral Success of Women Candidates: New Evidence from Local Elections in the United States." *British Journal of Political Science* 52 (4):1544–63. doi:10.1017/S0007123421000570.
Bos, Angela L., Jill S. Greenlee, Mirya R. Holman, Zoe M. Oxley, and J. Celeste Lay. 2022. "This One's for the Boys: How Gendered Political Socialization Limits Girls' Political Ambition and Interest." *The American Political Science Review* 116 (2):484–501. doi:10.1017/S0003055421001027.
Boussalis, Constantine, and Travis Coan. 2021. "Facing the Electorate: Computational Approaches to the Study of Nonverbal Communication and Voter Impression Formation." *Political Communication* 38 (1–2):75–97. doi:10.1080/10584609.2020.1784327.
Boussalis, Constantine, Travis G. Coan, Mirya R. Holman, and Stefan Müller. 2021. "Gender, Candidate Emotional Expression, and Voter Reactions During Televised Debates." *The American Political Science Review* 115 (4):1242–57. doi:10.1017/S0003055421000666.
Bracic, Ana, Mackenzie Israel-Trummel, and Allyson F. Shortle. 2019. "Is Sexism for White People? Gender Stereotypes, Race, and the 2016 Presidential Election." *Political Behavior* 41 (2):281–307. doi:10.1007/s11109-018-9446-8.
Brown, Adam R., and Jeremy C. Pope. 2021. "Mechanical Turk and the 'Don't Know' Option." *PS: Political Science & Politics* 54 (3):416–20. doi:10.1017/S1049096520001651.
Bullock, Will, Kosuke Imai, and Jacob N. Shapiro. 2011. "Statistical Analysis of Endorsement Experiments: Measuring Support for Militant Groups in Pakistan." *Political Analysis* 19 (4):363–84. doi:10.1093/pan/mpr031.
Burden, Barry C., Yoshikuni Ono, and Masahiro Yamada. 2017. "Reassessing Public Support for a Female President." *The Journal of Politics* 79 (3):1073–78. doi:10.1086/691799.
Cargile, Ivy. 2015. "Latina Issues: An Analysis of the Policy Issue Competencies of Latina Candidates." In *Distinct Identities: Minority Women in U.S. Politics*, eds. Sarah A. Gershon and Nadia E. Brown. New York: Routledge, 134–50.
Cassese, Erin C. 2020. "Straying from the Flock? A Look at How Americans' Gender and Religious Identities Cross-Pressure Partisanship:." *Political Research Quarterly* 73 (1):169–83. doi:10.1177/1065912919889681.
Cassese, Erin C., and Tiffany D. Barnes. 2018. "Reconciling Sexism and Women's Support for Republican Candidates." *Political Behavior* 41 (3):677–700. doi:10.1007/s11109-018-9468-2.
Cassese, Erin C., and Mirya R. Holman. 2018. "Party and Gender Stereotypes in Campaign Attacks." *Political Behavior* 40 (3):785–807. doi:10.1007/s11109-017-9423-7.
Cassese, Erin C., and Mirya R. Holman. 2019. "Playing the Woman Card: Ambivalent Sexism in the 2016 Presidential Race." *Political Psychology* 40 (1):55–74. doi:10.1111/pops.12492.
Chen, Philip, Melanee Thomas, Allison Harrell, and Tania Gosselin. 2022. "Implicit and Explicit Gender Stereotyping in Canadian Politics." Working Paper.
Claassen, Ryan L., and John Barry Ryan. 2016. "Social Desirability, Hidden Biases, and Support for Hillary Clinton." *PS: Political Science & Politics* 49 (4):730–35. doi:10.1017/S1049096516001414.
Conroy, Meredith. 2018. "Strength, Stamina, and Sexism in the 2016 Presidential Race." *Politics & Gender* 14 (1):116–21. doi:10.1017/S1743923X17000642.
Conroy, Meredith, Danielle Joesten Martin, and Kim L. Nalder. 2020. "Gender, Sex, and the Role of Stereotypes in Evaluations of Hillary Clinton and the 2016 Presidential Candidates." *Journal of Women, Politics & Policy* 41 (2):194–218. doi:10.1080/1554477X.2020.1731280.
Crowder-Meyer, Melody. 2013. "Gendered Recruitment Without Trying: How Local Party Recruiters Affect Women's Representation." *Politics & Gender* 9 (4):390–413. doi:10.1017/S1743923X13000391.

Crowder-Meyer, Melody, Shana Kushner Gadarian, and Jessica Trounstine. 2015. "Electoral Institutions, Gender Stereotypes, and Women's Local Representation." *Politics, Groups, and Identities* 3 (2):318–34. doi:10.1080/21565503.2015.1031803.

Devroe, Robin. 2021. "Stereotypes, Who to Blame? Exploring Individual-Level Determinants of Flemish Voters' Political Gender Stereotypes." *Political studies* 69 (4):791–813. doi:10.1177/0032321720924808.

Eagly, Alice H., and Steven J. Karau. 2002. "Role Congruity Theory of Prejudice Toward Female Leaders." *Psychological Review* 109 (3):573–98. doi:10.1037/0033-295X.109.3.573.

Fulton, Sarah A, and Heather L Ondercin. 2015. "Candidate Sex and the Costs of Voting: New Implications for Turnout." In *MPSA Annual Meeting*. Chicago.

Gabriel, Oscar W., and Lena Masch. 2017. "Displays of Emotion and Citizen Support for Merkel and Gysi: How Emotional Contagion Affects Evaluations of Leadership." *Politics and the Life Sciences* 36 (2):80–103. doi:10.1017/pls.2017.15.

Gonzales, Sylvia, and Nichole M. Bauer. 2020. "Using Gender and Partisan Stereotypes to Evaluate Female Candidates." In *Politicking While Female the Political Lives of Women*, ed. Nichole M. Bauer. Baton Rouge, LA: Louisiana State University Press, 93–114.

Gonzales, Sylvia, and Nichole M. Bauer. 2022. "Strong and Caring? Is There is a Unique 'Women of Color' Political Stereotype Category?" *Politics, Groups and Identities*Advance online publication. *Politics, Groups, and Identities*, 1–18. doi: 10.1080/21565503.2022.2144389.

Gooch, Andrew, and Lynn Vavreck. 2019. "How Face-To-Face Interviews and Cognitive Skill Affect Item Non-Response: A Randomized Experiment Assigning Mode of Interview." *Political Science Research and Methods* 7 (1):143–62. doi:10.1017/psrm.2016.20.

Gothreau, Claire, Kevin Arceneaux, and Amanda Friesen. 2022. "Hostile, Benevolent, Implicit: How Different Shades of Sexism Impact Gendered Policy Attitudes." *Frontiers in Political Science* 4:817309. doi:10.3389/fpos.2022.817309.

Gothreau, Claire M., Clarisse Warren, and Stephen Schneider. 2022. "Looking the Other Way: How Ideology Influences Perceptions of Sexual Harassment." *European Journal of Politics and Gender* 5 (2):211–31. doi:10.1332/251510821X16445951244136.

Greene, Steven, Melissa Deckman, Laurel Elder, and Mary-Kate Lizotte. 2022. "Do Moms Demand Action on Guns? Parenthood and Gun Policy Attitudes." *Journal of Elections, Public Opinion and Parties* 32 (3):655–73. doi:10.1080/17457289.2020.1862130.

Holman, Mirya R. 2014. "Sex and the City: Female Leaders and Spending on Social Welfare Programs in U.S. Municipalities." *Journal of Urban Affairs* 36 (4):701–15. doi:10.1111/juaf.12066.

Holman, Mirya R., Jennifer L. Merolla, and Elizabeth Zechmeister. 2011. "Sex, Stereotypes, and Security: A Study of the Effects of Terrorist Threat on Assessments of Female Leadership." *Journal of Women, Politics & Policy* 32 (3):173–92. doi:10.1080/1554477X.2011.589283.

Holman, Mirya R., Jennifer L. Merolla, and Elizabeth J. Zechmeister. 2016. "Terrorist Threat, Male Stereotypes, and Candidate Evaluations." *Political Research Quarterly* 69 (1):134–47. doi:10.1177/1065912915624018.

Holman, Mirya R., Jennifer L. Merolla, and Elizabeth J. Zechmeister. 2022. "The Curious Case of Theresa May and the Public That Did Not Rally: Gendered Reactions to Terrorist Attacks Can Cause Slumps Not Bumps." *The American Political Science Review* 116 (1):249–64. doi:10.1017/S0003055421000861.

Holman, Mirya R., Jennifer L. Merolla, Elizabeth J. Zechmeister, and Ding Wang. 2019. "Terrorism, Gender, and the 2016 U.S. Presidential Election." *Electoral studies* 61:102033. doi:10.1016/j.electstud.2019.03.009.

Huddy, Leonie, Joshua Billig, John Bracciodieta, Lois Hoeffler, Patrick J. Moynihan, and Patricia Pugliani. 1997. "The Effect of Interviewer Gender on the Survey Response." *Political Behavior* 19 (3):197–220. doi:10.1023/A:1024882714254.

Huddy, Leonie, and Nayda Terkildsen. 1993. "Gender Stereotypes and the Perception of Male and Female Candidates." *American Journal of Political Science* 37 (1):119–47. doi:10.2307/2111526.

Kane, Emily W., and Laura J. Macaulay. 1993. "Interviewer Gender and Gender Attitudes." *Public Opinion Quarterly* 57 (1):1–28. doi:10.1086/269352.

Kim, Hyun-Jeong, and Yesola Kweon. 2022. "Double Penalty? Examining the Class Effects of Gender Representation." Working Paper.

Klar, Samara, Christopher R. Weber, and Yanna Krupnikov. 2016. "Social Desirability Bias in the 2016 Presidential Election." *The Forum* 14 (4):433–43. doi:10.1515/for-2016-0037.

Krupnikov, Yanna, Spencer Piston, and Nichole M. Bauer. 2016. "Saving Face: Identifying Voter Responses to Black Candidates and Female Candidates." *Political Psychology* 37 (2):253–73. doi:10.1111/pops.12261.

Krysan, Maria. 1998. "Privacy and the Expression of White Racial Attitudes: A Comparison Across Three Contexts." *Public Opinion Quarterly* 62 (4):506–44. doi:10.1086/297859.

Kuklinski, James H., Paul M. Sniderman, Kathleen Knight, Thomas Piazza, Philip E. Tetlock, Gordon R. Lawrence, and Barbara Mellers. 1997. "Racial Prejudice and Attitudes Toward Affirmative Action." *American Journal of Political Science* 41 (2):402–19. doi:10.2307/2111770.

Kunda, Ziva. 1999. "Parallel Processing of Stereotypes and Behaviors." In *Dual-Process Theories in Social Psychology*, eds. Shelly Chaiken and Yaacov Trope. New York: Guilford Press, 314–22.

Kunda, Ziva, and Steven J. Spencer. 2003. "When Do Stereotypes Come to Mind and When Do They Color Judgment? A Goal-Based Theoretical Framework for Stereotype Activation and Application." *Psychological Bulletin* 129 (4):522–44. doi:10.1037/0033-2909.129.4.522.

Lizotte, Mary-Kate. 2020. *Gender Differences in Public Opinion: Values and Political Consequences*. Philadelphia, PA: Temple University Press.

López-Cariboni, Santiago, and Catherine Reyes-Housholder. 2022. "Parallel Conjoint Experiments for Measuring Gender Stereotypes and Analyzing Preferences." Working paper.

Masch, Lena. 2020. *Politicians' Expressions of Anger and Leadership Evaluations: Empirical Evidence from Germany*. Baden-Baden, Germany: Nomos Verlag.

McDermott, Monika L. 2016. *Masculinity, Femininity, and American Political Behavior*. New York: Oxford University Press.

Meeks, Lindsey. 2012. "Is She 'Man Enough'? Women Candidates, Executive Political Offices, and News Coverage." *The Journal of Communication* 62 (1):175–93. doi:10.1111/j.1460-2466.2011.01621.x.

Mo, Cecilia Hyunjung. 2014. "The Consequences of Explicit and Implicit Gender Attitudes and Candidate Quality in the Calculations of Voters." *Political Behavior* 37 (2):357–95. doi:10.1007/s11109-014-9274-4.

Oliveros, Virginia. 2016. "Making It Personal: Clientelism, Favors, and the Personalization of Public Administration in Argentina." *Comparative Politics* 48 (3):373–91. doi:10.5129/001041516818254437.

Ondercin, Heather L. 2017. "Who is Responsible for the Gender Gap? The Dynamics of Men's and Women's Democratic Macropartisanship, 1950–2012." *Political Research Quarterly* 70 (4):749–61. doi:10.1177/1065912917716336.

Ondercin, Heather L. 2018. "The Evolution of Women's (And Men's) Partisan Attachments." In *100 Years of the Nineteenth Amendment: An Appraisal of Women's Political Activism*, eds. Holly J. McCammon and Lee Ann Banaszak. New York: Oxford University Press, 46–68.

Ortbals, Candice D., and Lori M. Poloni-Staudinger. 2018. "Terrorism and the Public: Gender, Public Opinion, and Voting Behavior." In *Gender and Political Violence: Women Changing the Politics of Terrorism*, eds. Candice D. Ortbals and Lori M. Poloni-Staudinger. Cham: Springer International Publishing, 247–71.

Plant, E. Ashby, Janet Shibley Hyde, Dacher Keltner, and G. Devine. Patricia. 2000. "The Gender Stereotyping of Emotions." *Psychology of Women Quarterly* 24 (1):81–92. doi:10.1111/j.1471-6402.2000.tb01024.x.

Presser, Stanley. 1990. "Can Changes in Context Reduce Vote Overporting in Surveys?" *Public Opinion Quarterly* 54 (4):586–93. doi:10.1086/269229.

Roberts, Damon C., and Stephen Utych. 2022. "A Delicate Hand or Two-Fisted Aggression? How Gendered Language Influences Candidate Perceptions." *American Politics Research* 50 (3):353–65. doi:10.1177/1532673X211064884.

Rohrbach, Tobias. 2022. "Loud and Negative: Exploring Negativity in Voter Thoughts About Women and Men Politicians." *Politics and Governance* 10 (4). doi:10.17645/pag.v10i4.5752.

Rosenfeld, Bryn, Kosuke Imai, and Jacob N. Shapiro. 2016. "An Empirical Validation Study of Popular Survey Methodologies for Sensitive Questions." *American Journal of Political Science* 60 (3):783–802. doi:10.1111/ajps.12205.

Runderkamp, Zahra, Anne Louise Schotel, Daphne Joanna Van der Pas, and Lisa Mügge. 2022. "Space Invaders and Norm-Politicians: How the Media Represent the Intersectional Identities of Members of Parliament." *European Political Science Review* 14 (4):566–82. doi:10.1017/S1755773922000339.

Sanbonmatsu, Kira. 2002. "Gender Stereotypes and Vote Choice." *American Journal of Political Science* 46 (1):20–34. doi:10.2307/3088412.

Sanbonmatsu, Kira. 2006. "Do Parties Know That 'Women Win'? Party Leaders Beliefs About Women's Electoral Chances." *Politics & Gender* 2 (4):431–50. doi:10.1017/S1743923X06060132.

Schneider, Monica C. 2014. "The Effects of Gender-Bending on Candidate Evaluations." *Journal of Women, Politics & Policy* 35 (1):55–77. doi:10.1080/1554477X.2014.863697.

Schneider, Monica C., and Angela L. Bos. 2019. "The Application of Social Role Theory to the Study of Gender in Politics." *Political Psychology* 40 (S1):173–213. doi:10.1111/pops.12573.

Schneider, Monica C., Angela L. Bos, and Madeline DiFilippo. 2022. "Gender Role Violations and Voter Prejudice: The Agentic Penalty Faced by Women Politicians." *Journal of Women, Politics & Policy* 43 (2):117–33. doi:10.1080/1554477X.2021.1981095.

Schneider, Monica C, Mirya R. Holman, Amanda B Diekman, and Thomas McAndrew. 2016. "Power, Conflict, and Community: How Gendered Views of Political Power Influence Women's Political Ambition." *Political Psychology* 37 (4):515–31. doi:10.1111/pops.12268.

Setzler, Mark. 2019. "Measuring Bias Against Female Political Leadership." *Politics & Gender* 15 (4):695–721. doi:10.1017/S1743923X18000430.

Shames, Shauna L., Rachel Bernhard, Mirya R. Holman, and Dawn Langan Teele, eds. 2020. *Good Reasons to Run*. Philadelphia, PA: Temple University Press.

Shim, Jaemin. 2021. "Gender and Politics in Northeast Asia: Legislative Patterns and Substantive Representation in Korea and Taiwan." *Journal of Women, Politics & Policy* 42 (2):138–55. doi:10.1080/1554477X.2021.1888677.

Streb, Matthew J., Barbara Burrell, Brian Frederick, and Michael A. Genovese. 2008. "Social Desirability Effects and Support for a Female American President." *Public Opinion Quarterly* 72 (1):76–89. doi:10.1093/poq/nfm035.

Sweet-Cushman, Jennie. 2022. "Legislative Vs. Executive Political Offices: How Gender Stereotypes Can Disadvantage Women in Either Office." *Political Behavior* 44:411–34. doi:10.1007/s11109-021-09721-x.
Swim, Janet K., Kathryn J. Aikin, Wayne S. Hall, and Barbara A. Hunter. 1995. "Sexism and Racism: Old-Fashioned and Modern Prejudices." *Journal of Personality and Social Psychology* 68 (2):199–214. doi:10.1037/0022-3514.68.2.199.
Trounstine, Jessica. 2010. "Representation and Accountability in Cities." *Annual Review of Political Science* 13 (1):407–23. doi:10.1146/annurev.polisci.032808.150414.
Zaller, John, and Stanley Feldman. 1992. "A Simple Theory of the Survey Response: Answering Questions versus Revealing Preferences." *American Journal of Political Science* 36 (3):579–616. doi:10.2307/2111583.