Shifting Standards: How Voters Evaluate the Qualifications of Female and Male Candidates

Nichole M. Bauer, Louisiana State University

Existing empirical research finds that female candidates have higher levels of qualifications for political office compared to male candidates. An untested assumption behind this finding is that female candidates must have stronger qualifications to overcome feminine stereotypes that characterize women as ill qualified for leadership positions. I test this assumption by drawing on psychology research to develop a theory that explains how a candidate's sex affects the way voters evaluate the qualifications of political candidates. Using innovative survey experiments, the results show that, across multiple experiments, voters hold female candidates, relative to male candidates, to more stringent qualification standards, and these higher standards limit the ability of female candidates to secure electoral support. These findings uncover a subtle but pernicious source of bias facing female candidates. The implications speak to how candidate sex affects voter decision-making and the ability of democratic institutions to select the best candidates for leadership.

I can say with confidence there has never been a man or a woman more qualified than Hillary Clinton to serve as President of the United States of America.
—President Barack Obama, July 2016, address to the Democratic National Convention

A critique Hillary Clinton faced during the 2016 general election is that she lacked the qualifications necessary to serve as president. Clinton's resume included winning election to the Senate, twice, and serving as Secretary of State. These are all qualities scholars use to classify a candidate as qualified, yet Clinton frequently had to persuade voters that she had the right qualifications for the office. There is certainly room to debate Clinton's performance while serving in political office, but Clinton exemplifies the qualifications scrutiny female candidates frequently face when running for political office (Ditonto, Hamilton, and Redlawsk 2014, Holman, Merolla, and Zechmeister 2016). Despite these criticisms, empirical research consistently finds that female candidates, including incumbents and challengers, have more political experience, stronger professional profiles, and higher levels of education relative to male candidates (Anzia and Berry 2011; Fulton 2012). An untested assumption behind these dynamics is that female candidates must meet higher qualification standards to overcome bias rooted in conventional feminine stereotypes that characterize women as ill suited for political office. I test this assumption.

A long-standing empirical conclusion is that female candidates win elections at rates equal to male candidates (Ekstrand and Eckert 1981; Seltzer, Newman, and Leighton 1997). Observational surveys regularly show that voters rate the qualifications of female candidates positively—often even more positively than a male candidate's qualifications (see, e.g., Brooks 2013; Dolan 2014; Hayes and Lawless 2016). It is easy to conclude from these findings that female candidates do not face gender bias. This conclusion rests on two implicit and untested assumptions. First, scholars assume that rating a female candidate as "very qualified" or "very experienced" directly translates into electoral support (Brooks 2013; Hayes and Lawless 2016). Second, previous scholarship assumes that a female candidate rated as highly qualified looks just like a male candidate also rated as highly qualified (Dolan 2014). The pervasive underrepresentation of women suggests that positive ratings do not necessarily lead to increased vote support. Women hold

---

000

an average of 30% of state legislative seats, 9 out of 50 gubernatorial seats, and just under 25% of seats in Congress (CAWP 2018). Current approaches to explaining women's underrepresentation measure differences in voter trait ascriptions (Bauer 2015), perceived issue competencies (Ditonto 2017; Huddy and Terkildsen 1993), and the likelihood of electoral support (Dolan 2014). These approaches do not address whether voters hold female candidates to higher, or different, standards than male candidates.

This project has several critical findings. First, I show that gender bias affects how voters form impressions of candidate qualifications. This gender bias does not necessarily manifest in empirical comparisons of election outcomes because these entail comparisons in which the female candidate is, on average, of higher quality than her male opponent (Milyo and Schlosberg 2000; Pearson and McGhee 2013; Seltzer et al. 1997). Second, this research illustrates that male candidates, regardless of their qualifications, consistently have an electoral advantage. Third, I show that female candidates face particularly steep gatekeeping barriers in primary elections, as copartisan voters will rate a female candidate less positively than an equally qualified male candidate. Holding female candidates to higher standards can discourage women from entering the electoral arena.

DIFFERENCES IN FEMALE AND MALE CANDIDATE QUALIFICATIONS

Female candidates win elections at rates equal to male candidates (Seltzer et al. 1997), but this parity comes from the higher quality of female candidates and not from a lack of gender bias on the part of voters (Fulton 2012, 2014; Pearson and McGhee 2013). Female House and Senate candidates, as challengers and incumbents, relative to male candidates have a stronger set of objective qualifications as measured through legislative productivity, professional profiles, and academic accomplishments (Anzia and Berry 2011; Lazarus and Steigerwalt 2018; Volden, Wiseman, and Wittmer 2013). As challengers, female candidates have more experience serving in lower political offices (Maestas et al. 2006; Palmer and Simon 2001; Pearson and McGhee 2013). Existing research identifies several sources of the gendered qualification gap including institutional gatekeepers and socialization forces.

Local party networks are more likely to ask men rather than women to run for political office (Carroll 1994; Sanbonmatsu 2006), and these dynamics mean female candidates need to be exceptionally well qualified to get on the radar of local party leaders (Milyo and Schlosberg 2000). Illustrating this behavior is Fulton (2012), whose research finds that even political elites hold female candidates to a higher qualification standard compared to male candidates. There are also differences in recruit-

ment practices across parties. The Democratic Party is more likely to recruit and field female candidates compared to the Republican Party (Crowder-Meyer and Lauderdale 2014). Candidate recruitment patterns can create high entry barriers for female candidates, especially Republican women, but these insights do not clarify whether voters hold female candidates to higher qualification standards.

Women in the pool of potential political candidates often downgrade their own qualifications for political office, even if these potential contenders have the same or better qualifications as male candidates (Lawless and Fox 2010). Consequently, female candidates spend more time in lower levels of office establishing their records before ascending to a higher level of office (Fulton 2006; Maestas et al. 2006). Potential female candidates also express aversion to electoral competition, conflict, and the power-related goals involved in pursuing political office (Kanthak and Woon 2015; Schneider et al. 2016). The anticipation of bias among voters leads female candidates to believe they "must be better than their male counterparts to be elected" (Anzia and Berry 2011, 481).

Research, both observational and experimental, asking voters to evaluate candidate qualifications along dimensions such as experience or knowledge finds no differences across candidate sex (Brooks 2013; Ekstrand and Eckert 1981; Hayes and Lawless 2016). There are two reasons to be skeptical of these conclusions. First, social desirability pressures may affect how voters respond to subjective questions about the traits of female candidates (Burden, Ono, and Yamada 2017; Claassen and Ryan 2016; Krupnikov, Piston, and Bauer 2016; Streb et al. 2008). Second, this research implicitly assumes that an absence of differences across candidate sex indicates that voters evaluate candidates through a gender-neutral process (Brooks 2013; Dolan 2014). For example, Hayes and Lawless (2016) find that voters rate female and male candidates as equally qualified for political office. The authors assume that a "qualified" female candidate looks exactly the same as a "qualified" male candidate and that high qualification ratings will lead to electoral support. If this were the case, then female candidates would win elections at higher rates than male candidates. Voters hold distinctly masculine expectations for political candidates—indicating that voter decision-making may not be gender neutral.

This article fills several gaps in the existing work. First, extant scholarship has yet to identify why female candidates win elections by smaller margins than male candidates given their exceptionally high qualifications (Pearson and McGhee 2013). Using theories from psychology research, I identify why voters hold female candidates to different, and frequently higher, qualification standards than male candidates. Second, I conduct a series of experiments that directly measure the qualification standards voters have for candidates, which is difficult

to do with observational analyses. Third, I uncover evidence that primary elections pose the greatest challenges for female candidates.

A SHIFTING-STANDARDS PERSPECTIVE ON EVALUATIONS OF CANDIDATE QUALIFICATIONS

I draw on a theory of shifting standards, developed in psychology research, to delineate why voters hold women and men to different standards and to identify the contents of these divergent standards. Voters will use either gender-typicality or role-typicality standards to evaluate female candidates. Voters will use gender-typicality standards to compare female candidates to the stereotypes held of "typical women." Voters will use role-typicality standards to compare female candidates to the stereotypes held of a "typical leader." I argue that voters will use gender-typicality standards and role-typicality standards to form different types of evaluations of female candidates.

Gender-typicality standards use a "within-group standard of judgment" for which a woman or a man is compared to stereotypes of a "typical woman" or a "typical man" (Biernat and Kobrynowicz 1997). Voters will use gender-typicality standards to form belief assessments about the abilities of an individual candidate, such as assessing whether a female candidate is very experienced or very knowledgeable. Belief assessments do not evoke an explicit comparison between two individuals, and this lack of an explicit comparison leads individuals to evaluate a single individual by comparison to a stereotypic gender group (Biernat and Kobrynowicz 1997; Foschi 1992). Individuals need a comparative anchor when forming impressions of other individuals because such comparative anchors ease the task of decision-making (Campbell, Lewis, and Hunt 1958; Postman and Miller 1945). Gender stereotypes provide individuals with an easy and accessible comparative metric they can use to evaluate the qualities of an individual woman or a man (Biernat and Manis 1994).

The key to gender-typicality standards is that these involve isolated assessments of a single candidate and do not necessarily involve comparisons between two candidates. For example, voters will use gender typicality when asked to rate whether a candidate is very experienced, somewhat experienced, somewhat inexperienced, or very inexperienced, a common question asked in survey and experimental research. This type of question does not tell people how many years of serving in political office constitutes a "very experienced" rating. Voters will develop beliefs about the experience of a candidate by comparing a female candidate's level of experience to a typical woman and a male candidate's level of experience to a typical man. Gender-typicality standards affect how citizens respond to questions asked by researchers, but gender typicality also affects how voters form impressions about candidates outside the research setting (Funk 1999; McGraw 2003).

Stereotypes about women do not comport with the masculine expectations voters hold for political leaders (Eagly and Karau 2002). As such, feminine stereotypes characterize the "typical woman" as not very qualified for political office. This incongruence between stereotypes about women and political leadership creates an initially low qualification bar for female candidates. A female candidate does not need much political experience for voters to consider her more qualified than a typical woman. Conversely, being male is congruent with serving in political leadership roles (Koenig et al. 2011). Calling to mind a typical man leads voters to think about broader stereotypes of men as leaders, and this leads voters to see the "typical man" as having a relatively high level of political experience. For male candidates, gender typicality creates a high qualification bar. Gender typicality means that voters hold female candidates to an initially low qualification bar, and voters hold male candidates to an initially high qualification bar. The gender-typicality hypothesis describes these dynamics:

> **Gender-Typicality Hypothesis.** All else equal, female candidates will receive more positive belief assessments relative to male candidates.

Female candidates will appear to receive more positive evaluations than equally qualified male candidates, but this does not necessarily mean that voters will evaluate female candidates as being more qualified relative to male candidates. Rather, these evaluations indicate that the female candidate has qualifications that outpace those of a typical woman.

Individuals will shift from gender-typicality to role-typicality standards when forming overall evaluations about how well a candidate can fill leadership roles (Biernat and Kobrynowicz 1997; Eagly and Karau 2002; Foschi 1992). Overall evaluations differ from belief assessments because voters must think about the ability of a candidate to fill the specific role of being a political leader, thereby invoking an implicit comparison between a specific candidate and a "typical political leader." There are a variety of evaluations voters form about candidates that evoke role typicality. Asking voters to indicate their likelihood of voting for a specific candidate, without explicitly referencing an opponent, will activate role-typicality standards. Voters form these decisions by envisioning how well a candidate fits into the mold of a political leader, and this evokes a comparison to masculine stereotypes. The high level of congruity between being male and being a leader makes it easier for voters to see male candidates as fitting into a political leadership role. Questions that ask voters to choose between a

woman and a man will also activate role-typicality standards because voters make these choices on the basis of which candidate best fits the masculine role-typicality standard. Under role-typicality standards, voters are not comparing a woman to a man but comparing a woman or a man to a "typical leader."

Masculine role-typicality standards create a high qualification bar for female candidates because such standards evoke a masculine standard of comparison, and there is a high level of incongruence between being female and being a leader (Conroy 2015; Holman, Merolla, and Zechmeister 2011, 2016; Huddy and Terkildsen 1993). In contrast, masculine role-typicality standards create a low qualification bar for male candidates. In this masculine context, role-typicality standards will "bias the evaluations so that a man's performance at a masculine task will be assessed as better than the same performance by a woman" (Foschi 1992, 185). Implicit comparisons between a specific candidate and "typical political leader" evoke masculine role typicality. The key to role-typicality standards is the implicit comparison between a specific candidate and a "typical leader"; thus, role-typicality standards can negatively affect female candidates even in races featuring two female candidates. The role-typicality hypothesis outlines these differences:

> **Role-Typicality Hypothesis.** All else equal, female candidates will receive less positive overall evaluations relative to male candidates when asking voters to consider the ability of a candidate to fill a political leadership role.

Role-typicality standards disadvantage female candidates but provide male candidates with a baseline advantage when voters form overall evaluations of candidates including vote choice decisions.

The shifting standards approach addresses two empirical puzzles in the literature. First, election polls and political science scholarship find that voters rate female candidates as highly qualified for political office, but these evaluations do not lead to increased vote support (see, e.g., Hayes and Lawless 2016). These positive ratings stem from gender-typicality standards. Second, female candidates win elections at rates equal to male candidates (Seltzer et al. 1997), but female candidates only get electoral parity by being more qualified than male candidates (Fulton 2012; Pearson and McGhee 2013). Masculine role-typicality standards mean female candidates must achieve higher qualifications to simply reach parity at the polls.

EXPERIMENTAL EVIDENCE OF SHIFTING STANDARDS

I use an experimental method to test my predictions. An experimental approach is valuable for two reasons. First, the method directly traces how citizens think about candidate qual-

ifications. Observational comparisons do not measure the actual process of voter decision-making, only the outcome. Second, an experiment is the only method that allows me to track differences across candidate sex when holding qualifications constant because there is no "real-world" scenario that pits an equally qualified woman against an equally qualified man. In practice, female candidates, on average, have better qualifications than their male opponents. Holding qualifications constant across candidate sex is crucial because I can directly test whether voters downgrade the qualifications of female candidates. I rely on two main experiments with smaller extension studies that use different designs to test the two predictions.

Candidate productivity experimental design

The candidate productivity experiment presented participants with information about a female or a male incumbent's accomplishments in the legislature. The qualification information about the female and male candidates is the same across the two conditions. I only manipulated candidate sex. The stimulus stated that the candidate had just finished a first term in the House of Representatives, served on multiple committees, and sponsored several pieces of legislation during the last legislative term. I provide information about legislative productivity for several reasons. First, research measuring legislative productivity consistently finds that female lawmakers sponsor and cosponsor more bills compared to male lawmakers, bring home more federal dollars to their districts, and are more likely to broker legislative compromises (Barnes, Branton, and Cassese 2017; Branton et al. 2018; Lazarus and Steigerwalt 2018; Pearson and McGhee 2013).[1] Second, this experiment tests how voters use productivity information. Third, positioning both candidates as incumbents in competitive races reflects empirical research showing that female lawmakers face high-quality challengers (Milyo and Schlosberg 2000).

I manipulated candidate sex with names, Carol or Chris Hartley, and female and male photos.[2] I excluded partisan labels in the first study to ensure that partisan affinity does not affect the results, but study 2 incorporates partisanship into the experimental design. The only factor varied is candidate sex, and each condition featured either a female or a male candidate. This is a strategic choice because including a single candidate allows me to test the expectation that vote choice decisions can activate a shift to masculine role-typicality standards because these overall evaluations evoke an implicit comparison between a female candidate and masculinity.

---

1. See app. 1 for the full set of stimuli.
2. There were no significant differences in candidate age ($p = .5976$), education ($p = .7024$), and attractiveness ratings ($p = .8709$). See app. 3 for full pretest results.

The experimental sample comes from Amazon's Mechanical Turk (MTurk; $N = 197$), and the data collection occurred in June 2016. MTurk is an online recruitment platform where participants complete small tasks for a nominal reward. The results produced by MTurk samples frequently mirror those conducted with nationally representative samples (Berinsky, Huber, and Lenz 2012). MTurk samples are particularly useful for research on perceptions of women in politics because these samples are less likely to misreport preferences for female candidates due to social desirability pressures (Krupnikov et al. 2016). Appendix 2 (apps. 1–8 are available online) includes the sample characteristics along with comparisons to other national internet-based surveys.

The first outcome I measured asked participants to rate the legislative skills of each candidate.[3] The legislative skills included working hard; being organized, determined, principled, able to stand their ground, compromising, consensus building, and willing to engage in bipartisan behavior; sharing credit with others; public speaking; being able to manage multiple priorities; and working well with others. Participants indicated whether the individual possessed each skill. I combined the individual legislative skill ratings into a single scale of *legislative skills*.[4] The legislative skills scale ranges from 0 to 12, where each value indicates the number of skills a candidate possesses. I combine the skills into a single measure because voters want candidates who can display excellence on as many skills as possible. Support for gender typicality will emerge if the female candidate receives more positive ratings than the male candidate.

There are several benefits to this measurement approach. First, conventional trait-based measures often ask participants to indicate whether a specific trait describes a candidate very well or not very well (see, e.g., Bauer 2015; Huddy and Terkildsen 1993). These types of questions often find that both female and male candidates receive positive ratings on being experienced, knowledgeable, or qualified (see, e.g., Hayes and Lawless 2016). These studies do not tell us what constitutes a qualified female or a qualified male candidate. A very experienced female candidate may have six or seven skills, while a very experienced male candidate may have only one or two specific skills. Second, I consider skills rather than traits. A skill indicates an individual's ability to do something well, whereas a trait is a distinguishing characteristic of one's personal nature. Skills can be learned, taught, and refined, whereas traits are innate qualities of an individual that remain relatively stable through an individual's life. I ask participants to

think about skills as observable actions and behaviors rather than the innate qualities of an individual candidate.

I also asked participants about the likelihood of vote support and electoral viability.[5] Both vote choice and viability questions evoke a comparison of each candidate to a masculine role-typicality standard. I ask about viability because expressing electoral support for a female candidate in an experimental setting can lead to overreports of support for female contenders (Burden et al. 2017; Krupnikov et al. 2016; Streb et al. 2008). Asking about how likely it is that others will support a candidate can alleviate social desirability pressures (Claassen and Ryan 2016).

Candidate productivity results

The gender-typicality hypothesis predicts that the female candidate will receive more positive evaluations than the male candidate on the legislative skills ratings.[6] Each bar of figure 1 displays the legislative skill assessments across candidate sex.[7] Participants rated the female candidate as having 4.58 skills (SD $= 2.88$) and the male candidate as having only 3.48 skills (SD $= 2.33, p = .0036$).[8] The result is in line with the gender-typicality prediction. Female candidates receive more positive ratings because voters compare the female candidate's skills to those of a typical woman and not directly to a male candidate.[9]

The female candidate receives more positive ratings than the male candidate, and I argue that this is due to the low bar set by gender-typicality standards. An alternative explanation is that participants are not comparing the female candidate to a "typical woman" but to a female politician stereotype. Voters stereotype female politicians as being well educated, competitive, ambitious, and confident (Schneider and Bos 2014)—qualities that do not fit into stereotypes about women but fit into stereotypes about leaders. The "female politician" standard creates a high bar for female candidates because many of the qualities associated with this stereotype reflect the same qualities that voters use to stereotype political leaders. If participants used a female politician standard then the evaluation of the female candidate with the legislative skills measure would not differ from the evaluation of the male candidate. On

---

3. See app. 3 for information about the skills pretest.

4. Cronbach's $\alpha = 0.8622$, indicating that this scale has a high level of internal consistency.

5. See app. 4 for the full question wording.

6. Ninety percent of participants correctly identified the correct candidate name. Participant demographics (sex, education, region, etc.) do not predict group assignment ($\chi^2(9) = 6.25, p = .7150$).

7. All the reported p-values are the result of two-tailed tests.

8. Factor analysis shows that all the skills, except two, load onto a single factor. Excluding these skills in the measure does not alter the results. I also conducted analyses that separate the skills into feminine and masculine items in app. 5.

9. Candidate sex could be used as a low-information heuristic, but this supports gender typicality because voters are comparing a female candidate to stereotypes of a typical woman.
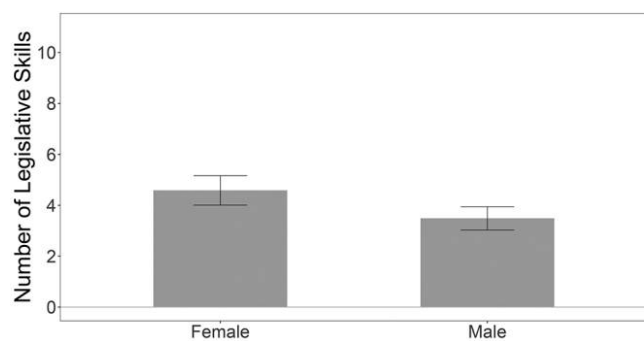
Figure 1. Candidate productivity study, legislative skills assessment (95% confidence intervals included). Each bar displays the number of positive skills ratings received by each candidate. Higher values indicate more positive legislative skill assessments.

these legislative skill ratings, a female politician standard creates a higher bar than gender-typicality standards.[10]

A masculine role-typicality standard means the female candidate will not have an electoral advantage over the male candidate. The use of role typicality is based on the implicit comparison between a specific candidate and the masculine expectations voters hold for those serving in leadership roles. The left side of figure 2 displays the percentage of participants who indicated they were either very likely or somewhat likely to vote for each candidate. At first glance, the results appear to counter the role-typicality prediction. Approximately 15% of participants supported the female candidate, while only about 9% of participants supported the male candidate ($p = .2173$). Another way to think about this finding is that female candidates must have at least one additional skill than a male candidate to secure a six-point electoral advantage.[11]

The right side of figure 2 displays the electoral viability results. On this metric, there is no female candidate advantage. Participants rated the female and male candidates as equally likely to win the election ($p = .2687$). The lack of differences across candidate sex supports role-typicality standards. The initially positive skills rating the female candidate received does not mean that voters think a female candidate is better than a male candidate at filling a masculine role. The female candidate's rating on the viability outcome suggests that even if voters are willing to support her she is likely to face bias that hinders her electoral success.

10. Social desirability could lead to more positive female candidate ratings (Burden et al. 2017; Claassen and Ryan 2016; Krupnikov et al. 2016; Streb et al. 2008). This effect, however, is consistent with gender-typicality expectations. The low bar of feminine stereotypes makes it easier for high self-monitors to proffer the socially acceptable response.

11. Appendix 5 includes the full results broken down by participant sex and participant party.

Candidate productivity robustness checks

The productivity information in the stimulus, while resembling the legislative productivity of first-term incumbents, did not explicitly frame either candidate as having many legislative accomplishments. I conducted an extension of the productivity experiment positioning both candidates as higher-status incumbents. Social psychology suggests that having specific individuating information can reduce negative patterns of stereotyping (Fiske and Neuberg 1990), and the productivity extension study tests this premise. The full text of the extension stimulus is included in appendix 6. I use the same measures from the main productivity study. In the productivity extension, the female candidate has 8.90 skills (SD = 2.52), and the male candidate has 9.06 skills (SD = 2.04), but this difference is not statistically significant ($p = .6241$). The lack of differences across candidate sex suggests that voters shift from gender-typicality to role-typicality standards to rate highly successful female incumbents. Previous research finds that female incumbents do not face the same disadvantages as female challengers and that voters are more likely to see high-status female incumbents similarly to male incumbents (Fridkin and Kenney 2009).

In the extension study, there are no statistically significant differences in the levels of vote support, with 82% indicating positive support for the female candidate, and 87% supporting the male candidate ($p = .3171$). Wider gaps emerge in perceived candidate viability, with 53% of participants rating the female candidate as electorally viable and 57% of participants rating the male candidate as electorally viable, and these values differ ($p = .0338$). Putting together the results from the extension study with the results from study 1 shows that voters are more likely to use gender-typicality standards to rate the legislative skills of female incumbents of a more junior status. This is important because female incumbents frequently face
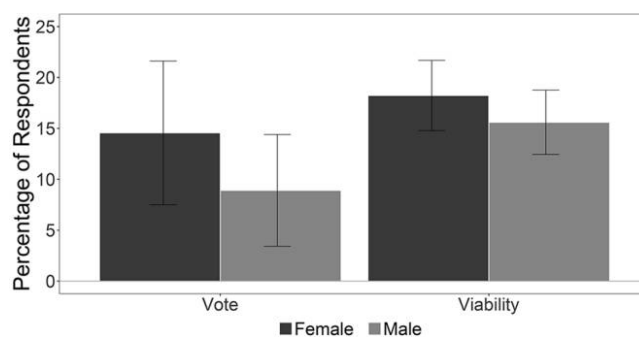

Figure 2. Perceptions of vote support and electoral viability (95% confidence intervals included). Percentage of participants indicating they were very or somewhat likely to vote for each candidate, and the percentage of participants indicating the candidate was very or somewhat likely to win the election.

high-quality reelection challengers during their first or second reelection bids (Milyo and Schlosberg 2000).

Competitive context experimental design

The first experiment shows that female candidates initially receive positive ratings on qualifications for holding political office because of gender-typicality standards, but these positive ratings do not increase the female candidate's electoral support because of role-typicality standards. I build on these results with the competitive context experiment to test how standards shift in the context of a primary election between two candidates of the same political party. Presenting participants with two candidates bolsters the external validity of the study, as most female candidates run against male opponents (Palmer and Simon 2005). In this study, I focus on how voters use role-typicality standards when asked to implicitly consider how well a candidate can fill the political leader role and when asked to explicitly choose a woman or a man.

Each condition always included two candidates. The sex of the manipulated candidate was either female or male, and the sex of the second candidate was always male. I manipulated candidate sex with the same names and photos from the productivity study: Carol or Chris Hartley.[12] Hartley's opponent was always Tom Larson.[13] Hartley and Larson always belonged to the same political party. I matched participants into conditions based on shared partisanship. This means Democratic participants received information about two Democratic candidates running in the Democratic primary for an open House seat, while Republican participants received information about two Republican candidates. I sorted Independents into partisan conditions based on the party they leaned most closely toward.[14] With this design, I can be sure that any negative effects in the female candidate condition come from candidate sex and not from partisan inferences. Table 1 outlines the full set of conditions.

Participants read a newspaper article about a primary election for an open seat to the US House of Representatives featuring Hartley pitted against Larson. I use a newspaper article because this resembles how most citizens learn about elections (West 2005), and research shows that media coverage of candidate competency differs for female candidates compared to male candidates (Bligh et al. 2012). The article always mentioned Hartley first and Larson second and stated that both candidates had prior experience serving in the state legislature. I held the information about candidate quality con-

stant. I rely on prior political experience as a component of candidate qualifications because this is one of the most common objective measures of qualifications used in observational research (Pearson and McGhee 2013).[15]

The experimental sample comes from Survey Sampling International (SSI; $N = 226$), and the data collection occurred in November 2016. SSI is a market-based research company that recruits adult participants to complete studies online. Robustness tests of SSI find that these samples include populations often excluded from online survey platforms (Berinsky, Margolis, and Sances 2014). Appendix 2 compares the SSI sample to the MTurk sample in the candidate productivity experiment. The sample overrepresents women at 57% but resembles other survey populations on demographics including age, race, political interest, and party identification.

I measured the legislative skills of each candidate, but I use a measure that asks participants to rate the number of behaviors candidates should possess rather than asking participants to form belief assessments. The experiment presented participants with a list of 10 legislative behaviors.[16] Participants indicated how many, but not which, of these behaviors candidates should have to be an effective legislator. This measure evokes an implicit comparison to role typicality because the question primes the masculine leadership role by asking about the skills of an ideal legislator. The legislative skills measure in the productivity experiment, study 1, asked voters to form beliefs about the candidate's perceived abilities, activating a gender-typicality standard, while the skills measure in this experiment asks voters how many skills a qualified candidate should possess, thereby activating a shift to role-typicality standards.

This measure is valuable for several reasons. First, asking participants to indicate how many skills a candidate possesses, rather than which specific skills, can alleviate social desirability pressures to overreport a candidate's skills (Biernat and

Table 1. Competitive Context Experimental Design

Candidate Party	Candidate Sex	Opposing Candidate	N
Democratic	Carol Hartley	Tom Larson	75
Democratic	Chris Hartley	Tom Larson	73
Republican	Carol Hartley	Tom Larson	40
Republican	Chris Hartley	Tom Larson	38

Note. Study conducted via Survey Sampling International, November 2016. This experiment matched participants into conditions based on shared partisanship. $N = 226$.

---

12. No significant differences occur in the male photos across age ($p = .6941$), education ($p = .8767$), or attractiveness ($p = .2618$).

13. See app. 3 for information about the stimuli pretest.

14. Twenty percent initially identified as Independent.

15. See app. 1 for the full set of stimuli.

16. Appendix 4 lists the full set of activities.

Kobrynowicz 1997). Participants may feel pressure to indicate that a female candidate has many skills if asked to evaluate these skills individually, but asking participants to indicate the total number of skills a female candidate has can alleviate this pressure to overreport. Second, this question asks participants about expectations that can pinpoint whether voters expect more from female relative to male candidates during campaigns. Asking about expectations for candidates offers insight into the qualification bar female and male candidates need to reach to win elections.

I also asked participants to choose between Hartley and Larson on vote choice as well as primary general election viability questions. I coded these variables as 1 if participants selected Hartley as the likely winner. These questions explicitly ask participants to think about which candidate fits the "typical political leader" role. On these questions, masculine role-typicality standards should disadvantage the female candidate but advantage the male candidate.

Competitive context results

Figure 3 documents the difference in the number of tasks Hartley had to complete relative to Larson, the opposing candidate, when Hartley is a woman and when Hartley is a man.[17] When Hartley is a woman, participants indicated she must demonstrate ability on 1.09 tasks more than Larson. As a female candidate, Hartley must demonstrate ability on an average of 5.99 tasks (SD = 3.26) while Larson only must demonstrate ability on 4.90 (SD = 3.59) tasks, and this difference is statistically significant (p = .0242). In the male candidate condition, there is no statistically significant difference in the number of legislative skills Hartley must complete (M = 6.11, SD = 3.09) relative to Larson (M = 5.57, SD = 3.63, p = .1900). Comparing the female Hartley's task number to the task number for the male Hartley shows no significant differences (p = .7826). The key here is that Larson must demonstrate fewer skills when Hartley is a woman.

On the vote choice outcome, there are no significant differences in the percentage of participants who indicated vote support for the female Hartley (68%) relative to the male Hartley (61%, p = .3272). This null effect initially suggests that role-typicality standards do not undermine female candidates. Pairing these results with the skills assessment outcomes, however, shows that the female candidates must be better than the male candidate just to achieve parity.
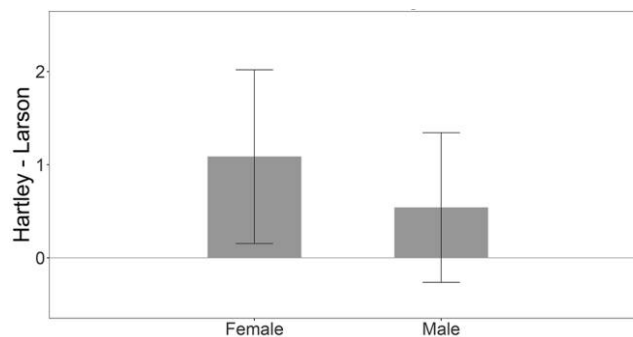

Figure 3. Competitive context: sex differences in legislative skills (95% confidence intervals included). Each bar displays the difference in the number of skills Hartley had to demonstrate minimum ability on relative to the male opponent, Larson.

Figure 4 displays the difference in Hartley's electoral viability relative to Larson, the male opponent. In the female condition, 49% of participants selected Hartley as the likely winner of the primary while 51% selected Larson as the primary winner (p = .7810). When Hartley is a male candidate, 58% of participants Hartley as the likely primary winner while 42% selected Larson as the likely winner, and this difference is marginally significant (p = .0861). The key difference between these two conditions is that the female Hartley is at a −2.61 disadvantage while the male Hartley has a much larger 17-point advantage. These results offer evidence of higher role-typicality standards that disadvantage female candidates.

The results are similar on general election viability (fig. 4B). When Hartley is a woman, she is at an electoral disadvantage, as only 48% selected her as the likely winner while 52% selected Larson as the likely winner, although this 4% difference in support is not statistically significant (p = .6401). Participants favored the male Hartley, with 53% selecting him as more likely to win a general election compared to Larson, with only 47% of participants selecting him as the general election winner (p = .5696). Again, the male Hartley has a larger advantage over Larson on viability, while the female Hartley is at a disadvantage. These patterns suggest the use of role-typicality standards. As a woman and as a man, Hartley always has the same set of qualifications, and these qualifications are sufficient for participants to see the male Hartley as electorally viable but not sufficient for the female Hartley.

Competitive context robustness check

I used a partisan extension study of the competitive context experiment that allows me to conduct comparisons within partisan groups, as there may be differences across partisanship in how voters evaluate qualifications (Bauer 2018a).[18] I focus here

---

17. Ninety-one percent of participants identified the level of office the candidates previously held. Participant demographics (sex, education, etc.) do not predict group assignment ($\chi^2(9)$ = 13.42, p = .1447). All the reported p-values are the result of two-tailed tests.

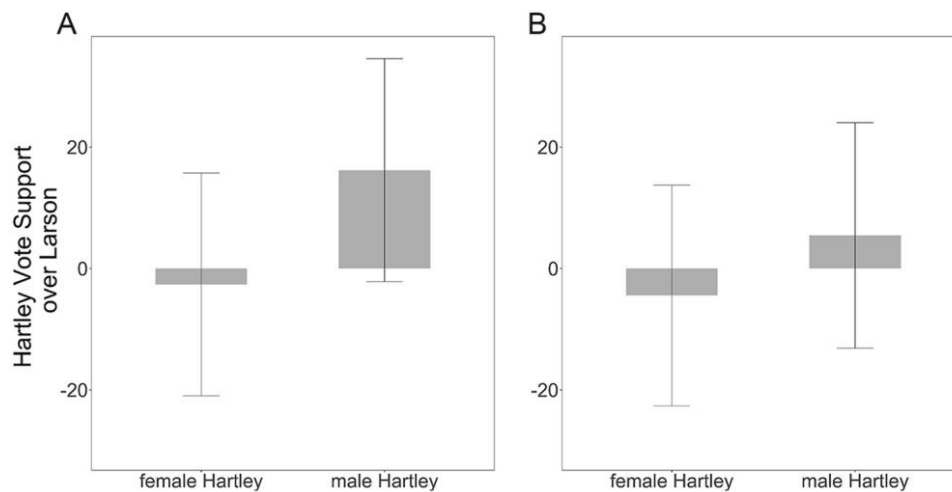18. See app. 7 for partisan comparisons in the main competitive context experiment.

Figure 4. Competitive context: differences in levels of perceived electoral viability (95% confidence intervals included). Each bar displays the difference in the percentage of participants who rated Hartley as the likely winner relative to Larson. *A*, Primary; *B*, general.

on the results using the electoral viability outcome.[19] Figure A2 in appendix 8 graphs the margin of victory on electoral viability across the partisan conditions. In the Democratic conditions, there are no statistically significant differences in the perceived viability of the female Hartley relative to the male Hartley. On the primary election question in the Republican condition, Hartley has a narrow five-point margin of victory over Larson, her male opponent, but when Hartley is male, he has a 66-point margin of victory over Larson, and these vote margin differences are statistically significant ($p = .0010$). At the general election level, the negative nine-point value indicates that Republican participants thought the Republican female Hartley was unlikely to win the election. When Hartley is male, the candidate has a 70-point advantage over Larson in the general election ($p = .0010$). The results of this partisan extension study suggest that Republican voters are more leery of supporting a female candidate compared to Democratic participants. Voters may hold Republican female candidates to a partisan standard that reflects both the masculine stereotypes ascribed to the Republican Party and the masculine role-typicality standards for political leaders (Winter 2010). Future work should explore in more depth the challenges Democratic and Republican female candidates face in securing electoral support from within their political party.

DISCUSSION AND CONCLUSION

A common finding in the literature is that when women run for political office they win elections at rates equal to male candidates (Seltzer et al. 1997). This equity at the ballot does not mean that voters evaluate candidates in a gender-neutral way. The experiments in this article show that male candidates face little pressure to prove their fitness for political leadership roles, while female candidates must display more evidence of their qualifications. These findings have implications for the candidate selection process, campaign strategy, role of bias in voter decision-making, and representation of women.

The productivity study found that the female candidate received a positive result on the initial skills evaluation, and this finding reflects other research showing that female candidates receive positive evaluations (Hayes and Lawless 2016). It is possible that female candidates receive positive belief assessments not because of gender-typicality standards but because voters may just feel more favorably toward female candidates. If participants simply feel more positively toward a specific woman, then the female candidate should also receive more favorable evaluations on vote choice and viability, which does not occur. Previous scholarship in political science and psychology finds that people rate women in masculine roles as cold and unlikable because of a social role violation (Bauer 2017; Bauer, Yong Harbridge, and Krupnikov 2017; Phelan, Moss-Rascusin, and Rudman 2008; Vraga 2017). This article does not test whether female candidates face a punishment for violating social role expectations, but that is an excellent question for subsequent scholarship.

This study only examined congressional races. The level of office will not adjust the gender-typicality and role-typicality framework, as long as voters associate masculinity with political leadership. Empirical evidence suggests that voters hold masculine standards for candidates even at lower levels of office (Bauer 2018b; Oliver and Conroy 2018). Role-typicality expectations may become more strongly masculine as the level of office increases, especially because evidence shows that female candidates face more bias as the level of office increases,

---

19. See app. 8 for the full partisan extension results.

particularly at the presidential level (Ono and Burden 2019). These qualification barriers are likely in other democracies where women face a representation disparity. Comparative research finds that men have an electoral advantage in other democracies (Murray, Krook, and Opello 2012), and women face barriers ascending to political leadership positions (Kerevel 2019; O'Brien 2015). Voters may hold strongly feminine expectations for an "ideal officeholder," such as a school board election, and in these cases the feminine role-typicality standard could give women an advantage.

The research in this article did not test how gender-typicality and role-typicality standards affect races with two female candidates. The recent surge of women running for political office increases the likelihood of more races with women running against other women. I expect that gender-typicality and role-typicality standards will still shape how voters think about female candidates, but the use of these standards may not automatically disadvantage a woman. A few studies examine all-female electoral contexts, and this literature finds that voters see a female incumbent similarly to a male candidate (Palmer and Simon 2005). Thus, the female incumbent squaring off against a female challenger may benefit from masculine role-typicality standards. It is also possible that voters will shift to entirely different standards in an all-female race, using partisan stereotypes, especially if both candidates are challengers.

The findings of the partisan extension study show that Republican female candidates face greater qualification barriers compared to Democratic female candidates. Republican female candidates are vastly underrepresented in political office. Indeed, a growing body of research shows that Republican female candidates face perceptual obstacles among voters based on the extent to which these candidates ideologically fit with their party (Bauer 2018c; Cassese and Holman 2018; Thomsen 2015). To win elections, especially Republican primaries, these candidates need to expend campaign resources extolling their qualifications but also need to persuade voters that they are the right ideological fit for the Republican Party. Republican male candidates do not have to overcome these perceptual obstacles. Partisan-typicality standards, which reflect masculinity for Republicans, can strongly disadvantage Republican women; partisan-typicality standards, however, may benefit Democratic women because feminine stereotypes about Democrats align with stereotypes about women (Winter 2010).

Scholarship on candidate qualifications operationalizes this concept through objective metrics such as a candidate's level of education, professional background, and previous political experience. But, the concept of candidate qualifications is not always well defined. The use of previous political experience is an inherently biased measure of qualifications for women because of women's underrepresentation in elected office. Sim-

ply put, women are more likely than men to lack previous political experience. Potential female candidates frequently come from backgrounds that differ from those of male candidates. Women interested in running for political office come from backgrounds in the nonprofit sector, social work, and community activism (Carroll and Sanbonmatsu 2013). It is not clear whether voters value these experiences.

A novel aspect of the experiments conducted in this article is the use of questions that ask about candidate skills in different ways. Asking individuals to evaluate whether a female or male candidate is very or not very qualified often leads to findings that voters rate female candidates as equally qualified to male candidates (Brooks 2013; Dolan 2014). But, this research does not offer insight into what it means for a candidate to be very or not very qualified. Questions that ask for participants to provide objective responses can offer a clearer indication of bias. For example, asking participants to indicate their impression of how much previous political experience a candidate has in number of years rather than whether a candidate is very experienced or very inexperienced asks participants to offer an objective rather than a subjective belief assessment about candidate skills. Future work should continue to develop alternative measures of candidate evaluations to gain a clearer understanding of the similarities and differences in how voters form impressions of female and male candidates.

A natural question arises from this research: Is it possible to train voters to evaluate a candidate's qualifications in an objective and gender-neutral manner? Developing strategies for disrupting gender bias can have ramifications for other contexts beyond politics. Individuals form comparisons, either implicit or explicit, between women and men in a variety of contexts shaped by norms of masculinity such as hiring decisions in the workplace, student evaluations of professors, or reporter decisions to seek out expert sources. These are all contexts long dominated by men and shaped by strong norms of masculinity—thus, these are all contexts with the potential to disadvantage women. Developing tools to teach citizens how to evaluate candidate qualifications, regardless of a person's sex, is critical to improving the social, political, and economic status of women.

REFERENCES

Anzia, Sarah F., and Christopher R. Berry. 2011. "The Jackie (and Jill) Robinson Effect: Why Do Congresswomen Outperform Congressmen?" *American Journal of Political Science* 55 (3): 478–93.

Barnes, Tiffany D., Regina P. Branton, and Erin C. Cassese. 2017. "A Re-examination of Women's Electoral Success in Open Seat Elections: The Conditioning Effect of Electoral Competition." *Journal of Women, Politics and Policy* 38 (3): 298–317.

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

Bauer, Nichole M. 2017. "The Effects of Counter-Stereotypic Gender Strategies on Candidate Evaluations." *Political Psychology* 38 (2): 279–95.

Bauer, Nichole M. 2018a. "The Effects of Partisan Trespassing Strategies across Candidate Sex." *Political Behavior* (forthcoming).

Bauer, Nichole M. 2018b. "Running Local: Gender Stereotyping and Female Candidates in Local Elections." *Urban Affairs Review* (forthcoming).

Bauer, Nichole M. 2018c. "Untangling the Relationship between Partisanship, Gender Stereotypes, and Support for Female Candidates." *Journal of Women, Politics and Policy* 39 (1): 1–25.

Bauer, Nichole M., Laurel Yong Harbridge, and Yanna Krupnikov. 2017. "Who Is Punished? Conditions Affecting Voter Evaluations of Legislators Who Do Not Compromise." *Political Behavior* 39 (2): 379–400.

Berinsky, Adam J., Gregory A. Huber, and Gabriel S. Lenz. 2012. "Evaluating Online Labor Markets for Experimental Research: Amazon.com's Mechanical Turk." *Political Analysis* 20 (3): 351–68.

Berinsky, Adam J., Michele F. Margolis, and Michael W. Sances. 2014. "Separating the Shirkers from the Workers? Making Sure Respondents Pay Attention on Self-Administered Surveys." *American Journal of Political Science* 58 (3): 739–53.

Biernat, Monica, and Diane Kobrynowicz. 1997. "Gender- and Race-Based Standards of Competence: Lower Minimum Standards but Higher Ability Standards for Devalued Groups." *Journal of Personality and Social Psychology* 72 (3): 554–57.

Biernat, Monica, and Melvin Manis. 1994. "Shifting Standards and Stereotype Based Judgments." *Journal of Personality and Social Psychology* 66 (1): 5–20.

Bligh, Michelle C., Michele M. Shlehofer, Bettina J. Casad, and Amber M. Gaffney. 2012. "Competent Enough, but Would You Vote for Her? Gender Stereotypes and Media Influences on Perceptions of Women Politicians." *Journal of Applied Social Psychology* 42 (3): 560–97.

Branton, Regina, Ashley English, Samantha Pettey, and Tiffany D. Barnes. 2018. "The Impact of Gender and Quality Opposition on the Relative Assessment of Candidate Competency." *Electoral Studies* 54 (August): 35–43.

Brooks, Deborah Jordan. 2013. *He Runs, She Runs*. Princeton, NJ: Princeton University Press.

Burden, Barry C., Yoshikuni Ono, and Masahiro Yamada. 2017. "Reassessing Public Support for a Female President." *Journal of Politics* 79 (3): 1–7.

Campbell, Donald T., Nan A. Lewis, and W. A. Hunt. 1958. "Context Effects with Judgmental Language That Is Absolute, Extensive, and Extra-experimentally Anchored." *Journal of Experimental Psychology* 55:220–28.

Carroll, Susan J. 1994. *Women as Candidates in American Politics*. Bloomington: Indiana University Press.

Carroll, Susan J., and Kira Sanbonmatsu. 2013. *More Women Can Run: Gender and Pathways to the State Legislatures*. New York: Oxford University Press.

Cassese, Erin C., and Mirya R. Holman. 2018. "Party and Gender Stereotypes in Campaign Attacks." *Political Behavior* 40 (3): 785–807.

CAWP (Center for American Women in Politics). 2018. "2018: Women Candidates for U.S. Congress and Statewide Elected Executive." CAWP, Rutgers University, Eagleton Institute of Politics.

Claassen, Ryan L., and John Barry Ryan. 2016. "Social Desirability, Hidden Biases, and Support for Hillary Clinton." *PS: Political Science and Politics* 49 (4): 730–35.

Conroy, Meredith. 2015. *Masculinity, Media, and the American Presidency*. New York: Palgrave Macmillan.

Crowder-Meyer, Melody, and Benjamin Lauderdale. 2014. "A Partisan Gap in the Supply of Female Potential Candidates in the United States." *Research and Politics* 1:1–7.

Ditonto, Tessa M. 2017. "A High Bar or a Double Standard? Gender, Competence, and Information in Political Campaigns." *Political Behavior* 39 (2): 301–25.

Ditonto, Tessa M., Allison J. Hamilton, and David P. Redlawsk. 2014. "Gender Stereotypes, Information Search, and Voting Behavior in Political Campaigns." *Political Behavior* 36 (2): 335–58.

Dolan, Kathleen. 2014. *When Does Gender Matter? Women Candidates and Gender Stereotypes in American Elections*. New York: Oxford University Press.

Eagly, Alice H., and Steve J. Karau. 2002. "Role Congruity Theory of Prejudice toward Female Leaders." *Psychological Review* 109 (3): 573–94.

Ekstrand, Laurie E., and William A. Eckert. 1981. "The Impact of Candidate's Sex on Voter Choice." *Western Political Quarterly* 34 (1): 78–87.

Fiske, Susan T., and Steven L. Neuberg. 1990. "A Continuum of Impression Formation, from Category-Based to Individuating Processes: Influences of Information and Motivation on Attention and Interpretation." In M. P. Zanna, ed., *Advances in Experimental Social Psychology*. New York: Academic Press, 1–69.

Foschi, Martha. 1992. "Gender and Double Standards for Competence." In Cecilia L. Ridgeway, ed., *Gender, Interaction, and Inequality*. New York: Springer.

Fridkin, Kim L., and Patrick J. Kenney. 2009. "The Role of Gender Stereotypes in U.S. Senate Campaigns." *Politics and Gender* 5:301–29.

Fulton, Sarah A. 2006. "The Sense of a Woman: Gender Ambition, and the Decision to Run for Congress." *Political Research Quarterly* 59 (2): 235–48.

Fulton, Sarah A. 2012. "Running Backwards and in High Heels: The Gendered Quality Gap and Incumbent Electoral Success." *Political Research Quarterly* 65 (2): 303–14.

Fulton, Sarah A. 2014. "When Gender Matters: Macro-Dynamics and Micro-Mechanisms." *Political Behavior* 36:605–30.

Funk, Carolyn L. 1999. "Bringing the Candidate into Models of Candidate Evaluation." *Journal of Politics* 61 (3): 700–720.

Hayes, Danny, and Jennifer L. Lawless. 2016. *Women on the Run: Gender, Media, and Political Campaigns in a Polarized Era*. New York: Cambridge University Press.

Holman, Mirya R., Jennifer L. Merolla, and Elizabeth J. Zechmeister. 2011. "Sex, Stereotypes, and Security: An Experimental Study of the Effect of Crises on Assessments of Gender and Leadership." *Journal of Women, Politics and Policy* 32 (3): 173–92.

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.

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.

Kanthak, Kristin, and Jonathon Woon. 2015. "Women Don't Run: Election Aversion and Candidate Entry." *American Journal of Political Science* 59 (3): 595–612.

Kerevel, Yann. 2019. "Empowering Women? Gender Quotas and Women's Political Careers." *Journal of Politics* 81 (4): 1167–80.

Koenig, Anne M., Alice H. Eagly, Abigail A. Mitchell, and Tiina Ristikari. 2011. "Are Leader Stereotypes Masculine? A Meta-analysis of Three Research Paradigms." *Psychological Bulletin* 137 (4): 616–42.

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

Lawless, Jennifer L., and Richard L. Fox. 2010. *It Still Takes a Candidate: Why Women Don't Run for Office*. New York: Cambridge University Press.

Lazarus, Jeffrey, and Amy Steigerwalt. 2018. *Gendered Vulnerability: How Women Work Harder to Stay in Office*. Ann Arbor: University of Michigan Press.

Maestas, Cherie D., Sarah A. Fulton, L. Sandy Maisel, and Walter J. Stone. 2006. "When to Risk It? Institutions, Ambition, and the Decision to Run for the U.S. House." *American Political Science Review* 199 (2): 195–208.

McGraw, Kathleen M. 2003. "Political Impressions: Formation and Management." In David O. Sears, Leonie Huddy and Robert Jervis, eds., *Handbook of Political Psychology*. Oxford: Oxford University Press.

Milyo, Jeffrey, and Samantha Schlosberg. 2000. "Gender Bias and Selection Bias in House Elections." *Public Choice* 105 (1/2): 41–59.

Murray, Rainbow, Mona Lena Krook, and Katherine A. R. Opello. 2012. "Why Are Gender Quotas Adopted? Party Pragmatism and Parity in France." *Political Research Quarterly* 65 (3): 529–43.

O'Brien, Diana Z. 2015. "Rising to the Top: Gender, Political Performance, and Party Leadership in Parliamentary Democracies." *American Journal of Political Science* 59 (4): 1022–39.

Oliver, Sarah, and Meredith Conroy. 2018. "Tough Enough for the Job? How Masculinity Predicts Recruitment of City Council Members." *American Politics Research* 46 (6): 1094–122.

Ono, Yoshikuni, and Barry C. Burden. 2019. "The Contingent Effects of Candidate Sex on Voter Choice." *Political Behavior* 41 (3): 583–607.

Palmer, Barbara, and Dennis M. Simon. 2001. "The Political Glass Ceiling: Gender, Strategy, and Incumbency in U.S. House Elections, 1978–1998." *Women and Politics* 23 (1/2): 59–78.

Palmer, Barbara, and Dennis M. Simon. 2005. "When Women Run against Women: The Hidden Influence of Female Incumbents in Elections to the US House of Representatives." *Politics and Gender* 1 (1): 39–63.

Pearson, Kathryn, and Eric McGhee. 2013. "What It Takes to Win: Questioning 'Gender Neutral' Outcomes in U.S. House Elections." *Politics and Gender* 9:439–62.

Phelan, Julie E., Corinne A. Moss-Rascusin, and Laurie A. Rudman. 2008. "Competent Yet out in the Cold: Shifting Criteria for Hiring Reflect Backlash toward Agentic Women." *Psychology of Women Quarterly* 32:406–13.

Postman, L., and G. A. Miller. 1945. "Anchoring of Temporal Judgments." *American Journal of Psychology* 58:43–53.

Sanbonmatsu, Kira. 2006. *Where Women Run: Gender and Party in the American States*. Ann Arbor: University of Michigan Press.

Schneider, Monica C., and Angela L. Bos. 2014. "Measuring Stereotypes of Female Politicians." *Political Psychology* 35 (2): 245–66.

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.

Seltzer, Richard, Jody Newman, and Melissa Leighton. 1997. *Sex as a Political Variable: Women as Candidates and Voters in U.S. Elections*. Boulder, CO: Rienner.

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.

Thomsen, Danielle. 2015. "Why So Few (Republican) Women? Explaining the Partisan Imbalance in the US Congress." *Legislative Studies Quarterly* 50 (2): 295–323.

Volden, Craig, Alan E. Wiseman, and Dana E. Wittmer. 2013. "When Are Women More Effective Lawmakers than Men?" *American Journal of Political Science* 57:326–41.

Vraga, Emily K. 2017. "Which Candidates Can Be Mavericks? The Effects of Issue Disagreement and Gender on Candidate Evaluations." *Politics and Policy* 45 (1): 4–30.

West, Darrell. 2005. *Air Wars: Television Advertising in Election Campaigns, 1952–2004*, vol. 4. Washington, DC: CQ Press.

Winter, Nicholas J. G. 2010. "Masculine Republicans and Feminine Democrats: Gender and Americans' Explicit and Implicit Images of the Political Parties." *Political Behavior* 32 (4): 587–618.