Sex, Stereotypes, and Security: A Study of the Effects of Terrorist Threat on Assessments of Female Leadership

Mirya R. Holman\textsuperscript{a}, Jennifer L. Merolla\textsuperscript{b} & Elizabeth J. Zechmeister\textsuperscript{c}

\textsuperscript{a} Florida Atlantic University, Boca Raton, Florida, USA
\textsuperscript{b} Claremont Graduate University, Claremont, California, USA
\textsuperscript{c} Vanderbilt University, Nashville, Tennessee, USA

Available online: 14 Jul 2011

To cite this article: Mirya R. Holman, Jennifer L. Merolla & Elizabeth J. 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, 173-192

To link to this article: http://dx.doi.org/10.1080/1554477X.2011.589283

Please scroll down for article

Full terms and conditions of use: http://www.tandfonline.com/page/terms-and-conditions

This article may be used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan, sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.
Sex, Stereotypes, and Security: A Study of the Effects of Terrorist Threat on Assessments of Female Leadership

MIRYA R. HOLMAN
Florida Atlantic University, Boca Raton, Florida, USA

JENNIFER L. MEROLLA
Claremont Graduate University, Claremont, California, USA

ELIZABETH J. ZECHMEISTER
Vanderbilt University, Nashville, Tennessee, USA

This project examines the extent to which terrorist threat shapes affect assessments of female political leadership and whether these assessments are affected by the characteristics of the evaluated leader. Using experimental and survey data, we find that conditions of terrorist threat increase the expression of certain gender-trait stereotypes and negatively affect evaluations of female Democratic leaders. Our analysis further suggests that Republican women and/or those with strong national security experience may be able to overcome the otherwise negative implications of security threats and gender-trait stereotypes.

KEYWORDS threat, terrorism, gender, sex, stereotypes, leadership

International terrorist plots and attacks have increased over the last few decades (Glasser 2005), and media reports often make salient the possibility of another, more devastating attack than 9/11. How does a political climate characterized by terrorist threat affect evaluations of female leaders? While considerable research focuses on evaluations of women in politics, we know little about the influence of conditions of terrorist threat compared to better times. Moreover, while some work suggests women are disadvantaged by
security threats, we know little about whether this result is universal across female leaders. We address the following questions. First, what effect does a condition of terrorist threat have on assessments of female leaders, both in the abstract and with respect to particular individuals? Second, can certain characteristics of the politician—including experience and political party—alter the effects of conditions of threat on evaluations of female leaders?

We argue that conditions of terrorist threat will increase the expression of certain gender-trait stereotypes and directly and indirectly affect evaluations of actual politicians. We further argue that characteristics which counter gender stereotypes can alleviate the negative effect that conditions of threat have on evaluations of female political figures. To this end, we examine two female leaders who differ in terms of their partisanship and their experience with national security: Hillary Clinton and Condoleezza Rice. We test our expectations using two types of data: data from an original experiment conducted in 2005 and data from the 2002 and 2004 American National Election Studies (ANES).

Analyses of the experimental and survey data collectively support our argument regarding the relationships among terrorist threat, gender-trait stereotypes, and evaluations of leaders. Some of our findings are consistent with those found in existing research. We also make an additional contribution by showing with the experimental data that these relationships differ when the political leader is varied. Specifically, conditions of terrorist threat (and potentially security threats more generally) do not have the theorized negative effects on female leaders with characteristics (party and background) that counter gender stereotypes.

The findings have important implications for understanding how elevated concerns about terrorist threat influence the expression of trait stereotypes and evaluations of female leaders. When the threat of terrorism increases in salience, individuals perceive certain women as less capable leaders; therefore, it may be difficult for those women to make gains in political office. At the same time, certain characteristics, including experience and political party, may override traditional stereotypes, allowing Republican women and/or women with foreign policy experience to govern with greater support in times of security crisis.

FACTORS THAT INFLUENCE EVALUATIONS OF FEMALE POLITICAL LEADERS

While scholars have documented many factors that make it more difficult for women to win elected office,¹ one part of the uphill battle that female political figures face is the presence of belief- and trait-based gender stereotypes.² With respect to belief stereotypes, for example, women are generally seen as more liberal than they actually are (e.g., Koch 2000;
2002; Matland and King 2002; McDermott 1997; 1998; but see Hayes 2005 for an analysis of when party trumps gender in trait attribution). Women are also seen as better able to handle certain issues, such as education, civil rights, and poverty, but less able to handle duties perceived as traditionally “male,” such as the military, foreign policy, and crime (e.g., Burrell 1994; Gordon 2001; Huddy and Terkildsen 1993a; 1993b; Lawless 2004; Matland 1994; Rosenwasser and Dean 1989; Rosenwasser et al. 1987). With respect to trait stereotypes, women are generally seen as more compassionate, trustworthy, willing to compromise, and more empathetic, while men are typically viewed as stronger leaders, more assertive, more active, and more self-confident (Burrell 1994; Gordon 2001; Huddy and Terkildsen 1993a; 1993b; Leeper 1991; Matland and King 2002; Rosenwasser and Dean 1989; Wilson 2004; but see Gordon 2001 for no differences on decisiveness). It is commonly held that these trait and belief stereotypes present an important challenge to women running for political office, especially higher levels of office (Fox and Oxley 2003; Huddy and Terkildsen 1993a), since masculine traits and issue competencies are more highly valued in the political arena.3

While the literature has contributed greatly to our understanding of some factors that influence the expression of trait- and belief-based stereotypes,4 only recently has scholarship examined how the existence, or nonexistence, of a context of threat influences the expression of these stereotypes. Falk and Kenski (2006) use poll-based data from the Annenberg Center and find that respondents who consider terrorism, Iraq, or homeland security as the most important issue are less likely to view a woman as better at handling the problem. They further find significantly less support for a female president among those most concerned with the conflict in Iraq. Lawless (2004) uses survey data to determine that in a time of crisis (war) those individuals who prefer masculine traits in candidates are less willing to vote for a female presidential candidate.5 In the realm of experiment-based research, Little and colleagues (2007) show that, in contexts of war, people prefer masculine faces to feminine faces; in contexts of peace, participants prefer a more feminine face. Our work expands on these studies, by analyzing survey data and applying an experimental research approach (in which we can systematically isolate the effects of a terrorist threat compared to good times) to the study of evaluations of gender stereotypes and contemporary female leaders.

THEORY

We posit theoretical links among environmental factors (salience of threat), gender-trait stereotypes, and evaluations of female leaders. In this section, we argue that conditions of national security threat, specifically terrorist threat, can influence the expression of trait stereotypes. Next, we argue that
both stereotypes and contexts of threat can influence evaluations of specific female political leaders. Again, we assert that these relationships are relevant to times of security threat; in addition, they should be conditional upon the leader’s background (party affiliation and experience).

These (Stiletto) Boots Are Made for Leading?

The first part of our argument is that the expression of trait stereotypes is affected by conditions of crisis. Trait stereotypes refer to the personal qualities and characteristics that individuals ascribe to others based on characteristics such as gender (Huddy and Terkildsen 1993b). Thus, individuals may think a female has more empathy than a male just by virtue of her gender. Some studies investigate these stereotypes within the context of real or hypothetical elections, typically asking respondents to evaluate specific traits about hypothetical candidates. In the absence of an election, other studies use instead abstract measures designed to tap generic gender-trait stereotypes. Two common questions ask whether male candidates make better political leaders or whether men are better suited emotionally for politics (i.e., Alexander and Andersen 1993; Sanbonmatsu 2003). Our study is situated outside of an electoral context, so we focus on these abstract gender-trait stereotypes.

Theoretically, there is reason to expect that given certain environmental conditions individuals may express higher or lower levels of generic trait stereotypes, because different threat conditions may prime the value of certain factors differently. That is, the elements affecting evaluations of whether men make better political leaders and are better suited emotionally for politics may vary depending on the context. The priming literature suggests that raising the salience of certain issues can make related attitudes more available, or accessible, to individuals when answering other survey questions (e.g., Iyengar and Kinder 1987; Krosnick and Kinder 1990; Miller and Krosnick 2000). If a given information context makes certain considerations more salient to an individual (e.g., the need for strong leadership in a context of national security threat), then we would expect those attitudes to influence other political assessments made by the individual (e.g., whether men or women make better leaders). Thus, the nature of the effect of threats on the expression of gender-trait stereotypes might vary depending on the types of considerations that threat brings to mind. We focus on a threat to national security: terrorism.

We expect that national security threats will increase the expression of generic gender-trait stereotypes, such that people will be more inclined to think that men are better political leaders. To cope with the threat of terrorism, individuals look to strong leadership; in fact, empirical research shows that this need may be so great that individuals actually come to project strong leadership (as opposed to other traits such as empathy) onto certain
leaders in times of terror threat (Landau et al. 2004; Merolla, Ramos, and Zechmeister 2007; Merolla and Zechmeister 2009). How might this context fare for perceptions of whether women are capable political leaders? First, it is possible that individuals will look for leaders with more masculine characteristics in this context. A study by Little and colleagues (2007) provides some empirical evidence that masculine characteristics (in this case, facial images) are preferred in times of war compared to in times of peace. Second, from the literature on belief-based stereotypes, we know that male candidates are generally perceived as better able to handle issues such as the military, national defense, and terrorism, compared to female candidates (Alexander and Andersen 1993; Fox and Oxley 2003; Rosenwasser and Dean 1989; Sanbonmatsu 2003). Thus, in a context in which national security becomes salient—which should tend to make stereotypical masculine characteristics and issue competencies more highly valued—we expect that individuals will hold relatively more negative assessments of the general capabilities of female political figures. In particular, the context of a security crisis, in this case terrorist threat, ought to increase perceptions that men make better political leaders and are better suited emotionally for politics than are women.

Putting the Abstract into Practice

The second part of our argument is that conditions of crisis and trait stereotypes will influence evaluations of specific female political leaders. The link between trait stereotypes and evaluations of leaders should be clear-cut: individuals who hold stereotypes (i.e., perceive men as more capable leaders or better suited for politics) should be less likely to support particular female politicians (see Alexander and Andersen 1993; Rosenwasser et al. 1987). In general, we expect that terrorist threat will directly and indirectly (via its effect on stereotypes) negatively affect evaluations of specific female leaders. However, the effect of crises and gender-trait stereotypes on support for female political leaders may be conditioned by the leader’s background. Existing research suggests that certain characteristics can counteract a tendency to adhere to gender stereotypes. For example, scholars have shown that while the general trend in belief-based stereotypes is to see women as more liberal than men (Koch 2002; McDermott 1997; 1998), Republican women may be able to overcome this stereotype, as the introduction of partisan information counteracts the liberal stereotype (Hayes 2005; McDermott 1997). Thus, it is conceivable that this type of counteractive effect can occur with respect to both trait-based stereotypes and political evaluations.

Given the Republican Party’s ownership of the issue of terrorism (Merolla and Zechmeister 2009; Petrocik 1996), Republican females may be able to counteract the negative effect of threat on trait stereotypes. A female leader with experience in the issues related to the given crisis may be
even better able to counteract this stereotype. One with both characteristics should be in prime shape to counter stereotyping tendencies in times of crisis. Considering the realm of terrorist threat, a female political leader with both theoretically counteracting characteristics—significant experience in national security and Republican identification—is Condoleezza Rice. Rice contrasts with Hillary Clinton, a well-known Democrat with less experience in national security (at the time of our study). We therefore analyze evaluations of these two specific female political leaders, and we expect direct, negative effects of terrorist threat and of gender stereotypes on evaluations of Clinton but not of Rice. Combined, our expectations concerning the effect of terrorist threat on gender stereotypes and the effect of the latter on evaluations of Clinton indicate a mediating model, whereby terrorist threats directly and indirectly affect evaluations of Clinton. In addition to testing these specific hypotheses, we also analyze evaluations of Rice (who we expect will be evaluated differently than Clinton) and evaluations of two male leaders: John Kerry and George W. Bush.

**EXPERIMENTAL DESIGN**

Our principal tests of the relationships among threat, gender-trait stereotypes, and evaluations are carried out using data generated from an experiment implemented in the fall of 2005 with student subjects. Participants in the studies were randomly assigned to a Status Quo group, a “Good Times” group, or a Threat group. Subjects completed a brief survey and then, following exposure to the treatments (or lack thereof, in the case of the Status Quo group), responded to a series of questions concerning gender-trait stereotypes and evaluations of political leaders.

**Participants and Design**

The computer-based study took place in a campus research lab and was in the field from late October to early November 2005. Research subjects were members of a general pool of undergraduate students, enrolled in social science classes at a large public university in Northern California, who participate in laboratory studies in exchange for class credit. Self-reported demographic data for the sample indicates that 53.5% were female; 58.0% self-classified as white; 22.0% identified as Republican, 54% as Democrat, and the remainder as Independent or Other. The fact that our sample is comprised of young college students who lean toward the left (with the majority identifying as Democrats) means our subject pool differs from the general population. To some extent—and we will elaborate on this more later—this may make for a tougher test of our expectations. To the degree that our left-leaning student sample, compared to the average individual in
the larger population, has lower levels of gender stereotypes or is more resistant to expressing such attitudes, we have some reason to expect that the relationships we find are similar, but potentially greater, within the general populace.

Participants were randomly assigned to one of four conditions: Good Times, Status Quo (Control), Terror Threat, and Economic Threat. Our arguments and analyses here concern only the first three groups. The number of subjects in each condition was 51, 48, 49, and 52, respectively.7

Procedures

Participants reported to the experimental lab in order to take part in a study about current events and opinions. Once seated, individually, in front of a computer terminal, a research assistant initiated the program and the participant was randomly assigned to either the control group or a treatment group. The study began with questions on basic demographics and political predispositions. If a subject was assigned to a treatment group, he or she was instructed by on-screen instructions to put on headphones and watch a short presentation. Those in the control group were simply forwarded to the next set of questions. After the presentations, the treatment groups answered the same set of questions, which included measures of our key dependent variables: gender-trait stereotypes and evaluations of political leaders.

Experimental Treatments

The experimental manipulation in our study involved exposing our treated research subjects to an audiovisual news presentation that was approximately one and one-half minutes in length. The Terror Threat treatment sought to make the possibility of another terrorist attack salient in the subject’s mind. The audiovisual presentation contained frightening politically relevant images and information regarding security issues. For example, our narrator reported, “US citizens are bracing for an imminent attack at home and feeling less secure.” The Good Times treatment exposed subjects to happy images and positive information about the overall state of Americans and their country. This presentation included such information as, “Overall, more Americans report in surveys that they are healthier and happier than ever before.” The purpose was to encourage participants to refrain from thinking about national threats and envision, instead, good times.

We modeled the audiovisual presentations after media clips commonly seen on news Web sites, combining a slideshow of images with a voice-over. The texts for the narrations were primarily drawn from news and political reports and edited together by the authors. We obtained most images from news archives. In some cases, we used text boxes with quotes (which
were read within the narration). A male professional voice actor did the voice-overs, and the images and narration were edited together by a professional editor (for treatments, see supplemental materials available from the authors).

**MANIPULATION CHECK**

Our intent and expectation was to elevate national security concerns among respondents who viewed the Terror Threat treatment, and, possibly, to decrease national security concerns among those who viewed the Good Times treatment. We expected the control group to reflect the Status Quo information environment.

To test the performance of our treatment stimuli, we asked respondents to indicate how worried they felt that there would be terrorist attacks in the United States in the near future; subjects responded on a seven-point scale, where seven indicated “very worried.” Using difference of means tests, we find that those individuals exposed to the Terror Threat treatment were more worried about terror attacks than those in the control group and those in the Good Times treatment. In short, the treatments had the intended effects on respondents.

**THE EFFECT OF THREAT ON TRAIT STEREOTYPES**

Our first point of inquiry is whether conditions of terrorist threat increase the expression of gender-trait stereotypes. We measured gender-trait stereotypes with two questions that are commonly found on the General Social Survey and the World Values Survey. The General Social Survey has asked, since 1977, whether respondents agree or disagree with the following statement: “Most men are better suited emotionally for politics than are most women” (Davis and Smith 2004). The World Values Survey asks respondents to indicate whether they agree or disagree with this statement: “Men make better political leaders than women do” (World Values Survey 1981–2008). Our survey asked respondents to indicate their agreement with these statements on seven-point scales. Given evidence that the two items tap a similar construct (Cronbach’s alpha = 0.84), we created a gender stereotype composite measure by taking the average of the two measures.

We assess our first expectation by comparing mean values on the composite measure across the experimental conditions. We conduct difference of means tests between the Terror Threat condition and the other two conditions to test our argument, which is appropriate since random assignment did ensure that subjects were evenly distributed across conditions on basic demographic and political dispositional measures. The mean on the
composite gender stereotype measure by experimental condition is depicted in Figure 1.

The results of the difference in means test show that those in the Terror Threat condition expressed significantly higher levels of gender stereotypes (2.91 on the 7 point scale) than those in the control group (2.32 on the scale) \((Pr (T > t) = 0.027)\) and those in the Good Times group (2.52 on the scale) \((Pr (T > t) = 0.095)\), though the latter difference is only marginally significant.\(^{10}\) The substantive differences are moderate in size for a seven-point scale, a 0.4-unit difference between the Terror Threat and Good Times group and a 0.5-unit difference between the Terror Threat and control group. We regard this as fairly substantial, considering that subjects were exposed to only a 90-second audiovisual presentation. Such an effect might likely be larger during an actual, prolonged crisis. In short, the data support the contention that individuals display relatively more elevated levels of gender-trait stereotypes in times of national security crisis.

**STEREOTYPES, THREATS, AND EVALUATIONS OF LEADERS**

Do crises and these gender-trait stereotypes further affect evaluations of specific female political leaders? To answer this question, we asked our respondents to evaluate Hillary Clinton and Condoleezza Rice, two well-known female leaders with the characteristics necessary to assess our expectations. To refresh, we argued that gender-trait stereotypes and terror threat would have negative effects on evaluations of Hillary Clinton. As we noted, at the time of our study Clinton had few of the characteristics (related to party affiliation and national security experience) that might counter such an effect. We expect that such factors may play an important role. To assess
this additional expectation, we examine evaluations of Condoleezza Rice. Rice’s party (Republican) has a stronger reputation on terrorism, and she also has significant experience with national security issues. We expect that these characteristics will enable her to overcome negative bias related to the terror threat condition and gender stereotypes. To assess the extent to which the effects we detect are related to gender rather than simply partisanship differences, we also look at evaluations of two male leaders, George W. Bush and John Kerry.

We measure evaluations with 100-degree feeling thermometers, which are common in the American National Election Study. Table 1 shows the results of regressions of our experimental conditions and the gender stereotype variable on evaluations of four leaders, Clinton, Rice, Bush, and Kerry. The Good Times treatment serves as the baseline. Since we expect a negative effect for the Terror Threat measure and the gender-trait stereotype measure for Clinton, we use one-tailed tests. Since we do not have strong directional expectations for the other candidates, we use two-tailed tests.

Looking first at the analysis for Clinton, as expected, we find the Terror Threat treatment has a substantial negative effect that is marginally significant, while the gender-trait stereotype measure has a significant negative effect on evaluations of Clinton. We believe these findings are quite reliable, given that our theory strongly supports a unidirectional hypothesis and, further, given their impressive substantive effects. Exposure to Terror Threat reduces feelings toward Clinton by 7.88 units. Every unit change in the stereotype measure results in a 3.97-unit decline in feelings toward Clinton.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hillary Clinton Coefficient (standard error)</th>
<th>Condoleezza Rice Coefficient (standard error)</th>
<th>George W. Bush Coefficient (standard error)</th>
<th>John Kerry Coefficient (standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>73.19*** (5.831)</td>
<td>35.67*** (5.480)</td>
<td>19.10*** (5.810)</td>
<td>64.09*** (4.829)</td>
</tr>
<tr>
<td>Status Quo/Control</td>
<td>−0.791 (5.817)</td>
<td>3.958 (5.457)</td>
<td>−6.831 (5.816)</td>
<td>−1.362 (4.823)</td>
</tr>
<tr>
<td>Terror Threat</td>
<td>−7.877* (5.814)</td>
<td>11.15*** (5.583)</td>
<td>10.50** (5.784)</td>
<td>−2.626 (4.816)</td>
</tr>
<tr>
<td>Additive Gender</td>
<td>−3.974*** (1.667)</td>
<td>1.463 (1.577)</td>
<td>4.474*** (1.664)</td>
<td>−3.967*** (1.385)</td>
</tr>
<tr>
<td>Stereotype Measure</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>148</td>
<td>142</td>
<td>147</td>
<td>146</td>
</tr>
<tr>
<td>$R^2$</td>
<td>0.060</td>
<td>0.038</td>
<td>0.117</td>
<td>0.060</td>
</tr>
</tbody>
</table>

Note: Ordinary least squares (OLS) regressions used in all analyses. Dependent variable is feeling thermometer evaluation, on a 0 to 100 scale. Significance thresholds are indicated as *p < 0.10, one-tailed test; **p < 0.10, two-tailed test; ***p < 0.05, two-tailed test.
We also find support for a mediating model, in which the condition of terror threat works directly and indirectly via gender-trait stereotypes to affect Clinton’s evaluations. MacKinnon, Fairchild, and Fritz (2007) identify four conditions that underlie a mediating model. In our case, we note first that Terror Threat affects gender stereotypes (see Figure 1 and Appendix A); second, Terror Threat has a direct effect on evaluations of Clinton without the gender stereotype measure in the model (see Appendix B); third, Terror Threat and the gender stereotype measure influence evaluations of Clinton (Table 1); and, fourth, the size of the effect of Terror Threat without the gender stereotype measure ($-9.4$), is greater than the effect when it is in the model ($-7.88$). With a direct effect of $-7.88$ and an indirect effect, via the gender stereotype measure of $-1.54$, the total effect of Terror Threat on feelings toward Clinton is $-9.42$. Overall, then, this analysis of evaluations of Hillary Clinton, a visible Democratic political leader, supports our argument that trait stereotypes and terror threat negatively affect evaluations of this female leader.

Turning to Condoleezza Rice, we find the reverse. Recall our argument that, if Republican partisanship and national security experience offset the negative effects of Terror Threat and gender-trait stereotypes, we could find muted negative effects, null effects, or even positive effects for Rice. As Table 1 shows, the Terror Threat treatment is significant and positive. Being in this condition led to an 11.15-unit increase in evaluations of Rice relative to the Good Times group. We do not find significant results for the gender stereotype measure, suggesting that Rice was able to overcome the negative effects of gender stereotyping that affected Clinton. Compared to Clinton, we do not find support for a mediating relationship. These results suggest that certain women are able to avoid some of the effects of gender stereotyping. Specifically, the ability to evade the negative effect of stereotypes and particular crises may be overcome by association with the Republican Party and a background in national security. Our research design prioritized the use of real political figures over hypothetical figures and it is therefore beyond the scope of our data to tease out which of these two factors (or both) is driving the results for Rice. Nonetheless, the results provide strong support for the notion that select characteristics of a given female political leader can offset the effects that we find in the case of Clinton.

One might question whether the effects we find are due simply to the leaders’ party affiliation (see Hayes 2005) rather than gender or other individual characteristics. To explore this possibility, we examined the effects of the treatments and gender stereotypes on evaluations of George W. Bush and John Kerry. These results are presented in the last two columns of Table 1. For Bush, we expect to find that the trait stereotype and Terror Threat variables have positive effects on his evaluations. Indeed, we find that both the Terror Threat and trait stereotype variable have positive effects on George Bush. We next turn to the feeling thermometer questions for John
Kerry. While being a male should lead to positive effects for Terror Threat and gender-trait stereotypes, being a Democrat might offset those effects. The results show that the Terror Threat condition is not significant. Thus, in our study, conditions of national security crisis hurt only the Democratic female leader. We also find that the gender stereotype measure is significant and negative. Thus, those with higher levels of gender-trait stereotypes had lower evaluations of Kerry, and the effect is similar in size to what it is for Hillary Clinton. While partisanship may be part of the reason for why we find a negative effect of this measure for Kerry, the lack of a direct effect of the Terror Threat indicates that the results are not due to simple partisan differences. We should also note that while Clinton has the highest baseline feeling thermometer among those in the Good Times condition, the predicted feeling thermometer for her ends up registering a lower value than Rice and Bush in times of terror threat with elevated levels of gender stereotypes.

**ROBUSTNESS CHECK: SURVEY DATA ANALYSIS**

To test the robustness of the findings for female political leaders in a more general sample of the population, we turn to the 2000–2002–2004 American National Election Study. The survey included a feeling thermometer question on Hillary Clinton for 2002 and 2004, though a similar question was not asked of Condoleezza Rice. We compare the feelings about Hillary Clinton to those about John Kerry. With respect to national security threat, the 2002 and 2004 waves asked respondents to indicate on a four-point scale how likely it was that the United States would suffer an attack as serious as the 9/11 attacks in New York and Washington in the next 12 months. We coded the threat measure such that higher values are more negative assessments. The study did not include any of our key measures of trait stereotypes of female political leaders. As a rough proxy, we include a feeling thermometer toward feminists. We also include a five-point scale that asked respondents to place themselves on a scale that ranged from “Women and men should have equal roles” to “A woman’s place is in the home.” We would expect that individuals with warmer feelings toward feminists should also have warmer feelings toward Clinton and possibly Kerry, while thinking a woman’s place is in the home should lead to cooler feelings toward Clinton and, again, possibly Kerry. Since we are now dealing with a public opinion survey rather than an experiment in which potential confounding factors are controlled by virtue of random assignment, we also control for ideology (higher values are more conservative), partisan identification (higher values are more Republican), gender (1 = female, 0 = male), education, income, age, and race (1 = white, 0 = nonwhite). The results for both years for Clinton and 2004 for Kerry are presented in Table 2.
### TABLE 2 OLS Regressions on Feelings Toward Hillary Clinton and John Kerry, ANES Panel

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hillary Clinton ANES 2002 Coefficient (standard error)</th>
<th>Hillary Clinton ANES 2004 Coefficient (standard error)</th>
<th>John Kerry ANES 2004 Coefficient (standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>78.11*** (5.910)</td>
<td>76.00*** (6.688)</td>
<td>58.90*** (6.452)</td>
</tr>
<tr>
<td>Fear of a Terror Attack</td>
<td>−2.010*** (0.909)</td>
<td>−0.776 (1.007)</td>
<td>0.159 (0.973)</td>
</tr>
<tr>
<td>Feminist</td>
<td>0.253*** (0.0388)</td>
<td>0.273*** (0.0405)</td>
<td>0.219*** (0.0392)</td>
</tr>
<tr>
<td>Equal Roles</td>
<td>−0.0328 (0.804)</td>
<td>−0.992 (0.884)</td>
<td>−1.512** (0.854)</td>
</tr>
<tr>
<td>Party</td>
<td>−6.457*** (0.444)</td>
<td>−5.671*** (0.459)</td>
<td>−5.916*** (0.442)</td>
</tr>
<tr>
<td>Ideology</td>
<td>−2.716*** (0.621)</td>
<td>−3.154*** (0.692)</td>
<td>−0.764 (0.671)</td>
</tr>
<tr>
<td>Female</td>
<td>2.001* (1.517)</td>
<td>3.580*** (1.632)</td>
<td>−0.0298 (1.570)</td>
</tr>
<tr>
<td>Education</td>
<td>0.174 (0.520)</td>
<td>−1.223*** (0.592)</td>
<td>0.575 (0.568)</td>
</tr>
<tr>
<td>Age</td>
<td>−0.0183 (0.0502)</td>
<td>0.0406 (0.0564)</td>
<td>0.0661 (0.0545)</td>
</tr>
<tr>
<td>White</td>
<td>−5.445*** (2.051)</td>
<td>−5.394*** (2.516)</td>
<td>−0.603 (2.431)</td>
</tr>
<tr>
<td>Income</td>
<td>−1.561*** (0.479)</td>
<td>−0.150 (0.513)</td>
<td>−0.223 (0.451)</td>
</tr>
<tr>
<td>N</td>
<td>796</td>
<td>633</td>
<td>630</td>
</tr>
<tr>
<td>R²</td>
<td>0.479</td>
<td>0.513</td>
<td>0.451</td>
</tr>
</tbody>
</table>

*Note: OLS = Ordinary least squares; ANES = American National Election Study. Significance thresholds are indicated as *p < 0.10, one-tailed test; **p < 0.10, two-tailed test; ***p < 0.05, two-tailed test.*

Turning first to the national security crisis proxy, we see that it is negative as expected in both models predicting feelings toward Hillary Clinton. Considering the year in which 9/11 ought to have been relatively more salient, 2002, the variable is significant. It is in the correct direction but loses both size and significance in 2004 for Clinton. The substantive impact of the national security measure in 2002 is quite meaningful. In 2002, moving from not thinking future attacks are likely to thinking they are very likely decreases feelings toward Hillary Clinton by 2.01 units on the thermometer rating. In both 2002 and 2004, feelings toward feminists exert a positive and significant effect on feelings toward Hillary Clinton, while the equal roles measure is insignificant. With respect to controls, those who are more Republican, more conservative, and are white have less positive feelings toward Clinton in both years. In 2002, those with higher levels of income have lower feelings toward Hillary, while in 2004 those with higher levels of education have lower feelings, while women have warmer feelings. Looking at the feelings toward Kerry in 2004, concern about a terror attack is an insignificant predictor. However, feelings toward feminists is significant...
and positive, while the equal roles measure is significant and negative. The party measure is also significant in the Kerry model.

Overall, analysis of the ANES data supports the generalizability of our findings across subject types. Analysis of the survey data does not provide as strict a test of our hypotheses, given that we sacrifice claims of causality in favor of correlation and use proxies for our key variables. The fact that we find results in the survey data that are consistent with our experimental results bolsters our claim that environmental factors have the potential to affect evaluations of candidates. Specifically, we find again that fears of terrorism in the general population can be associated with lower evaluations of Democratic female leadership (as embodied by the person of Hillary Clinton).

**DISCUSSION AND CONCLUSION**

Are trait-based gender stereotypes fixed or contextually influenced? Our argument and analysis supports the notion that the expression of trait stereotypes can shift across times of terrorist threat compared to other conditions. As we expected, a context of terror threat caused our subjects to perceive men as better leaders and as better able to deal emotionally with politics.

But then we asked: If terrorist threat conditions generally have negative implications for attitudes supporting female leadership, can some female leaders avoid the effects such conditions have on the public’s evaluations? We argued that conditions of threat and trait stereotypes affect evaluations of existing leaders and this will be conditional on the type of threat and the characteristics of the leader. Namely, when considering terrorist threat, Republicans and those with experience in national security may be able to counteract negative stereotypes. As expected, we found that gender stereotypes and the Terror Threat condition had negative effects on evaluations of Hillary Clinton. Meanwhile, Condoleezza Rice escaped the negative effects of trait stereotypes and terror threat. By exploring evaluations of Kerry and Bush, we demonstrated that the differences we observed cannot solely be explained by partisanship; rather, gender stereotypes have a role to play but can be moderated by partisanship.

Our experiment arguably understates the effects of conditions of terrorist threat on gender stereotyping and evaluations of female political leaders. At least two factors are relevant. First, as we noted earlier, our student subject pool differs from the general US population. When using convenience samples recruited from university classrooms, it is important to consider how findings with this group would theoretically compare to those we would expect to find among the larger population. The average student in our sample is younger and more left-leaning than the average US citizen. Most of the students in our study are between the ages of 18 and 22, while
the average respondent in the ANES sample was 50 years old in the 2002 wave and 54 years old in the 2004 wave. In addition, the experimental sample self-identifies as, on average, Democrat, while those in the ANES sample place themselves, on average, in the center (leaning right) between Democrats and Republicans. If we make the plausible assumption that it is more difficult to increase stereotypes and related attitudes among our student sample, then we can conclude that this tough test of our expectations likely underestimates the effects we would find in the general population.

Second, our experiment induced contextual effects by way of a very short audiovisual presentation. Conditions of crisis and/or campaigns that reference threats typically bombard citizens with greater quantities of information. As such, it is reasonable to suspect that the effects we find here linking crisis conditions to gender-trait stereotypes to evaluations of (certain) female political leaders would be more pronounced during an actual crisis and/or during a campaign that highlighted such conditions. While our study cannot speak to the duration of the effects, we suspect that they influence public opinion when the crisis is made initially salient but may diminish over time as the salience of the threat recedes. Our ANES analyses, which show that concerns about a future terrorist attack affect evaluations of Clinton in 2002 but not in 2004, provide some support for this speculation. During the 2004 election, issues other than terrorism were on voters’ minds. Of course, an examination of duration is beyond the scope of our study.

Our study and its findings suggest several avenues for future research. First, in counteracting against the effects of a terrorist threat condition, is partisanship or security experience more consequential? Second, are there other relevant counteracting characteristics beyond the two our study highlights? Third, are there particular strategies that female leaders could employ to avoid the negative assessments that accompany some conditions of threat?

On their own, the findings we present here have important implications for understanding the fate of women in the political arena and the strategies they may seek to employ as they campaign for office. First, while women will frequently have to confront the existence of trait stereotypes, the particular expression of these may vary depending on the type of threat present (compared to no threat). During times of national security crisis (terrorist threat), our study suggests that trait stereotypes will be elevated as masculine traits and issue competencies are more highly valued, which could translate into difficulty for existing female leaders in maintaining favorable evaluations. Consistent with some existing research, our study suggests that it might be in the interest of certain female candidates to refrain from highlighting conditions of terrorist threat and national security crisis more generally. At the same time, our research goes further to suggest that Republican leaders and female candidates with experience in defense or military-related areas may be able to avoid some of the pitfalls of these gender-trait stereotypes.
NOTES

1. At the aggregate level, studies have demonstrated that certain factors can disadvantage women who are running for elected office, such as particular institutional arrangements, a small pool of female candidates in many states, and running in more conservative and Republican states. Meanwhile, states with a history of electing women to office are generally more open to female leaders (e.g., Arceneaux 2001; Fox and Oxley 2003; 2004; Necchemias 1987; Norrander and Wilcox 1998; Rule 1981; 1990; Wilcox and Norrander 2005). Other research has found that even when women overcome challenges, they still may be disadvantaged by their gender (e.g., Duerst-Lahiti 2005 and chapter 4 in Seltzer, Newman, and Leighton 1997). At the individual level, scholars have identified a number of variables that increase the likelihood of voting for female candidates: for example, being a woman, a Democrat, a liberal, and having liberal opinions on women's issues (e.g., Cook 1994; Dolan 1998; Huddy and Terkildsen 1993a; 1993b; McDermott 1997; Paolino 1995; Schroedel and Snyder 1994).

2. Belief stereotypes refer to the ideology and policy preferences that individuals ascribe to males and females, while trait stereotypes refer to the personal qualities and characteristics that people infer about men and women (Huddy and Terkildsen 1993b).

3. Male duties are also generally perceived as more important than female duties, which results in a larger effect at higher levels of office (Rosenwasser and Dean 1989).

4. Scholars also examine factors that decrease these stereotypes such as females emphasizing masculine traits and issues, differences by partisanship, type of office, and the information environment (e.g., Fox and Oxley 2003; Hedlund et al. 1979; Huddy and Terkildsen 1993a; Kahn 1992; Leeper 1991; Little et al. 2007; Matland and King 2002; McDermott 1997; Rosenwasser and Dean 1989).

5. In related work on the importance of masculine traits in times of war, Hansen and Otero (2007) argue that the trait of strong leadership should become more consequential to presidential voting decisions. However, in their analysis of ANES data from 1984 to 2004, they do not find a stronger effect for leadership on voting decisions in 2004 compared to earlier years, a result that may be due to not separately analyzing those concerned about war or terrorism. Other scholars have found that leadership becomes more consequential to presidential voting decisions among those who think a future attack is more likely (Merolla and Zechmeister 2009).

6. We refer to the group that did not receive a treatment as both the Control and the Status Quo group, to acknowledge that subjects entering the lab carried with them understandings of current security threat levels, perceptions which we increased with the Terror Threat condition and decreased through the Good Times condition.

7. There were no significant differences in the distribution of subjects' potentially relevant variables (party identification, ideology, political sophistication, gender, education, and race) across groups.

8. Those in the Terror Threat treatment expressed higher levels of worry about a terror attack than those in the control group (\( Pr (T > t) = 0.0262 \)) and those in the Good Times treatment (\( Pr (T > t) = 0.0264 \)).

9. We coded the data so that a higher value indicates a higher level of agreement, on the seven-point scales, with the statement; thus, on the combined measure higher values indicate higher gender-trait stereotypes.

10. Hanushek and Jackson (1977, 68–69), among others, note that one-tailed hypothesis tests are appropriate when one has a strong theory pointing to a unidirectional hypothesis, as we do in this case. The results are the same if we run a multivariate regression with good times as the baseline (see Appendix A) or with the control group as the baseline.

11. We tested for, but did not find, interactions between gender and the conditions enhanced the model for either female political leader. If we use the control group as the baseline instead of Good Times, we also find support for a mediating relationship.

12. Refer to note 11.

13. We ran tests to see if partisanship moderated the effects of the treatments and found this to be the case only for John Kerry. In Kerry's model, the inclusion of party identification and interactions between party and the treatments results in the Terror Threat treatment achieving significance, as well as the gender-trait stereotype and the party identification variable. In this model, strong Democrats appraise Kerry more favorably in response to the terror threat condition, while Republicans in the condition have a more negative evaluation of Kerry.
14. This finding occurs once the gender stereotype measure is at 4.0 for Rice and at 5.0 for Bush.
15. The ANES sample averages a 2.9 (in 2002) and a 3.0 (in 2004) on the seven-point partisan scale, which is more conservative than the mean of 2.01 in the experimental sample.

REFERENCES


### APPENDIX A

#### Experimental Effect of Crisis on Gender Stereotypes

<table>
<thead>
<tr>
<th>Variable</th>
<th>Coefficient (standard error)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant</td>
<td>2.520*** (0.201)</td>
</tr>
<tr>
<td>Status Quo/Control</td>
<td>−0.197 (0.289)</td>
</tr>
<tr>
<td>Terror Threat</td>
<td>0.389* (0.288)</td>
</tr>
</tbody>
</table>

N = 148

$R^2$ = 0.028

Note: Ordinary least squares (OLS) regressions used in all analyses. Dependent variable is gender stereotype measure, on a scale of 1 to 7. Significance thresholds are indicated as *$p < 0.10$, one-tailed test; **$p < 0.10$, two-tailed test; ***$p < 0.05$, two-tailed test.

### APPENDIX B

#### Experimental Effect of Crisis on Evaluations of Political Leaders

<table>
<thead>
<tr>
<th>Variable</th>
<th>Hillary Clinton</th>
<th>Condoleezza Rice</th>
<th>George W. Bush</th>
<th>John Kerry</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Coefficient (standard error)</td>
<td>Coefficient (standard error)</td>
<td>Coefficient (standard error)</td>
<td>Coefficient (standard error)</td>
</tr>
<tr>
<td>Constant</td>
<td>63.18*** (4.109)</td>
<td>39.35*** (3.773)</td>
<td>30.37*** (4.109)</td>
<td>54.10*** (3.419)</td>
</tr>
<tr>
<td>Status Quo/Control</td>
<td>−0.00980 (5.901)</td>
<td>3.711 (5.448)</td>
<td>−7.585 (5.934)</td>
<td>−0.694 (4.937)</td>
</tr>
<tr>
<td>Terror Threat</td>
<td>−9.421* (5.870)</td>
<td>11.74*** (5.544)</td>
<td>12.24*** (5.871)</td>
<td>−4.036 (4.910)</td>
</tr>
</tbody>
</table>

N = 148

$R^2$ = 0.023, 0.032, 0.072, 0.005

Note: Ordinary least squares (OLS) regressions used in all analyses. Dependent variable is feeling thermometer evaluation, on a 0 to 100 scale. Significance thresholds are indicated as *$p < 0.10$, one-tailed test; **$p < 0.10$, two-tailed test; ***$p < 0.05$, two-tailed test.