Assessing the Double Bind:
Public Reactions to Displays of Toughness by Male and Female Candidates
Deborah Jordan Brooks
Dartmouth College
Assistant Professor of Government
deborah.brooks@dartmouth.edu

This paper was presented at the 2010 American Political Science Association Meetings held in Washington, D.C. It will be changed substantially prior to publication. Please contact the author at the above email address if you wish to quote or cite this work.)
ASSessing the Double Bind: Public Reactions to Displays of Toughness by Male and Female Candidates

Deborah Jordan Brooks
Dartmouth College
Assistant Professor of Government
deborah.brooks@dartmouth.edu

ABSTRACT: The conventional wisdom is that female politicians are greatly constrained by a toughness-related double bind: if a woman fails to demonstrate toughness, she confirms a stereotype that women are not strong enough to lead; yet if she demonstrates toughness, she will be disliked for violating gendered expectations. Although this double-bind is commonly seen as a principal challenge facing women on the campaign trail, political scientists have neglected to evaluate this issue empirically. With a large-N, representative sample of U.S. adults, this study utilizes a survey-based experiment to thoroughly test the conventional wisdom. Contrary to the conventional wisdom, female politicians do not suffer a disproportionate penalty for acting in a tough manner; in fact, this analysis shows that tough women benefit on several key measures relative to tough men.
The conventional wisdom is that female politicians are greatly constrained by a toughness-related double bind: if a female leader fails to be tough, she confirms the common stereotype that women are not strong enough to be leaders, but if she demonstrates toughness, she will be deemed to be mean, unlikable, and unsuited to the leadership position on that basis. In the eyes of many analysts, this is the single biggest challenge facing female candidates on the campaign trail. In this regard, pollster Celinda Lake notes: “The greatest challenge for women candidates is to demonstrate ‘toughness’ in a way voters are comfortable with. ‘If you're too strong, you're labeled a bad word. If you’re not strong enough, you get run over’” (USA Today 2001).

As Kathleen Hall Jamieson stresses, "it’s gender bias, plain and simple. They [female candidates] are supposed to be warm and accessible, because that's what's perceived to be gender-appropriate. But they also need to be tough and competent. The minute they appear that way, their warmth and accessibility are called into question" (as quoted in Noveck 2008). Female candidates therefore face an impossible decision: “A woman has to choose between running as the candidate with the proper competence – and thus, being manly – or as the candidate who is properly feminine – and thus, being unqualified” (Rede 2008). According to linguistics professor Deborah Tannen, this dynamic is caused by gendered expectations: “Our image of a politician, a leader, a manager, anyone in authority, is still at odds with our expectations of a woman. To the extent that a woman is feminine, she's seen as weak. To the extent that she puts it aside and is forceful, aggressive and decisive, she’s not seen as a good woman” (as quoted in Goodman 2007).

According to the conventional wisdom, female candidates are thus stuck between a proverbial rock and a hard place: they need to be tough to convince the public they have what it takes to get the job done, but acting tough may inherently alienate people because it violates people’s expectations of appropriate feminine behavior. References to this double bind were especially prominent in discussions of Hillary Clinton’s 2008 Democratic presidential primary run. During the height of the 2008 presidential primary, Bill Clinton complained openly that Hillary Clinton was plagued by the "psychological double-bind women sometimes get caught
in….Half the time when she shows how tough she is, people say she's too tough” (Sicha 2008). Analysts also frequently highlighted this dynamic as greatly undermining Clinton’s political fortunes. For example, in an article entitled “Clinton’s Task: Being Likable AND Tough,” Mary Trigg of the Institute of Women’s Leadership at Rutgers University noted: “Research demonstrates that there’s a very narrow bandwidth of acceptable behavior for women in positions of power, because they have to be feminine, but also exhibit the kind of attributes we look for in a leader…A female leader can’t be too assertive and strident, because you get into the ‘B’ word. It’s a hard and lonely trail to walk” (Noveck 2008).

While Clinton’s presence in the race generated greatly enhanced discussion of this toughness-related double bind, it has long been widely seen as vexing nearly all female politicians. According to Mandy Greenwald, a Democratic consultant, “[Madeline] Albright is one of the few top female politicians who has succeeded in projecting ‘an image of strength without being called a bitch. Think about other prominent female politicians, women like Dianne Feinstein, Christie Whitman, Geraldine Ferraro. Sooner or later that word tends to follow them around’” (Dobbs 1997). Some argue that this is an ancient phenomenon: “A woman who displays toughness and ambition often sacrifices being liked. This is not a new discovery. It goes back 3,000 years to the first female pharaoh, Hatshepsut. Women in business, academia, science – you name it – are caught in a double bind. If they are not tough enough, they are not up to the job…If a woman is too tough, she does not fit our stereotype of female behavior…A frequent response to a strong woman leader in any sphere is, ‘I just don't like her’” (Kunin 2008).

With respect to politics, many tie the existence of the double bind specifically to national security, an area that is thought to be particularly relevant in the post-9/11 era. For example, Dee Dee Meyers, former Press Secretary for President Bill Clinton, notes that “when women in positions of authority conform to traditional stereotypes, they are perceived as ‘too soft’ to be effective (especially on issues of national security). And when they defy those norms, they are considered ‘too tough,’ too masculine, downright bitchy” (Myers 2008). Others argue that the
double bind is especially difficult for women to manage as presidential candidates: “Female candidates for all elective offices face the conundrum of having to appear strong and confident without looking unfeminine, according to specialists on women and politics. But the dilemma is more pronounced at the presidential election level, they say, because voters are more likely to be guided by their gut reactions to candidates rather than comparisons of the contenders’ records.” (Milligan 2008).

If this frequently referenced double bind does exist in the political realm, and if it is the powerful constraint that many pundits, academics, and campaign advisors assume it to be, then it is hard to understand why women have had so many electoral successes in recent years. Scholars have largely concluded that female candidates are not penalized by voters for their gender, all else being equal (see Lawless 2004 for a discussion of the findings to date). While there is some debate about whether female candidates face a higher bar than men for executive as compared to legislative offices, there is little evidence at the legislative level that the public is inherently harder on female candidates. Moreover, the world has changed a great deal with respect to gender balance in a relatively short amount of time. The percentage of female U.S. Representatives rose from 7% in 1990 to 14% in 2000 to 18% in 2010, while the percentage of female U.S. Senators rose from 2% in 1990 to 9% in 2000 to 18% in 2010. As a result, it is reasonable to expect that while female politicians may have been punished for “unfeminine” behavior in earlier decades, tough women in politics may be sufficiently prevalent now that they can be seen as leaders more than as women in the current era.

Although it is vital to understand whether the double bind exists in the political realm, the literature has neglected to examine this question. The psychology literature has established that powerful descriptive and normative gender stereotypes exist for ordinary men and women, while the business literature has established that female leaders face a powerful double bind in the business sphere. Theories have been developed outside of political science that can be used to make predictions about how people will react to stereotypical and counter-stereotypical behavior
by men and women, but no one has delineated how these theories apply to the political realm in order to test how people react to the tough behavior of male and female politicians.

The paper is organized as follows. This first section analyzes the psychology, business, and political science literatures to delineate the existing research on gender stereotypes that bears on why the conventional wisdom likely exists or why it might be incorrect. After identifying the general stereotype theories at stake, the next section then derives a series of hypotheses from them regarding political behavior. The remainder of the paper uses an experiment to test these hypotheses. Utilizing a large-N, national sample that approximates a random sample, this study systematically varies politician gender and behavior in order to evaluate whether acting in a tough manner disproportionately hurts or helps female candidates as compared to male candidates, and why that is the case. In contrast to the conventional wisdom on the topic, I find no evidence of a harmful double bind for female politicians. In fact, tough behavior appears to help women rather than hurt them relative to men. Further examination reveals that older people – a group that tends to hold negative gendered views of candidates – are more likely than younger people to reward female candidates for acting in a tough manner. While this finding seems counter-intuitive at first glance, it is actually in keeping with a theory called expectancy violations theory (which postulates that individuals who outperform low expectations get disproportionately rewarded for exceeding those expectations). I conclude by discussing the implications of the results.

GENDER STEREOTYPES

Stereotypes are “cognitive structures that contain the perceiver’s knowledge, beliefs, and expectancies about some human group” (Hamilton & Trolier, 1986, p. 133). Gender stereotypes serve as simplifying categories for evaluating men and women in social interactions. Within the field of psychology, scholars distinguish between “descriptive stereotypes” (in other words, how men and women do act, in the case of gender) and those that focus on normative expectations, called “prescriptive” or “proscriptive” stereotypes (i.e. how men are women are supposed to act
or not act; note that the terms “prescriptions” and “prescriptive stereotypes” typically encompass references to proscriptions as well).

As compared to other kinds of stereotypes, gender stereotypes tend to be especially prescriptive, with a strong and powerful set of normative expectations about how men and women should and should not act, combined with strong social sanctions for violating those expectations (Prentice and Carranza 2002; Prentice and Carranza 2003; Rudman and Glick 1999; Fiske, Bersoff, Borgida, Deux & Hetlman 1991). When such stereotypes are violated, the violator tends to be “devalued, and thereby marginalized” by people (male and female alike) who observe the violation (Prentice and Carranza 2003, 260).

With reference to toughness-related attributes, people hold strong descriptive stereotypes about men and women. Prentice and Carranza 2002 have the most comprehensive data on this topic, in which they measured both descriptive and prescriptive stereotypes for ordinary men and women across 100 widely ranging traits.¹ As compared to women, men were assumed to be far more aggressive, more forceful, less yielding, more assertive, more competitive, and more decisive (out of the 100 measured traits, those characteristics were in the top 15 largest gender differences). Men were also thought to have more business sense, be less weak, be less naïve, more controlling, more ambitious, have more leadership ability than women.

With reference to prescriptions about toughness-related behaviors, the gender differences are even more notable than the descriptive stereotypes. In terms of the ranking of the magnitude of gender differences on the desirability of each trait for men and women, all twelve of the toughness-related traits rank within the top 16 out of 100; moreover, “aggressive,” “weak,” and

¹ Thanks are due to Prentice and Carranza for generously sharing their data with me. That allowed me to examine the gender differences on all 100 traits rather than just the subset that was included in their published article.
“assertive,” capture the #1, #2, and #3 spots, respectively. In other words, there tend to be much stronger gender prescriptions for toughness-related traits than for most other kinds of traits.

Thus, at least for ordinary men and women, the psychology literature shows that fairly strong descriptive stereotypes are in place for many toughness-related traits and extremely strongly gendered normative expectations exist for toughness-related behavior. As compared to female leaders, ordinary women would not necessarily be expected to face a double bind because they may not need to demonstrate toughness in order to be seen as a good person. To the extent that toughness by women is penalized, it is likely that ordinary women will simply try to avoid that behavior. But for businesswomen and political women, such conformance to expectations may not be possible, because toughness may be required for success in leadership roles. Therefore, we need to look at research on leaders in order to assess the state of our knowledge about the double bind.

Political science provides surprisingly little help on that front, however. While political scientists have sometimes made assumptions that a double bind exists for female politicians on the matter of toughness (Bower 2003; Kathleen Hall Jamieson 1995; Alexander and Andersen 1993), this question has not been studied directly empirically.2

2 Leeper 1991 and Sapiro 1981 measured how people respond to male and female candidates who deliver tough speeches but neither study could directly examine the double bind because they did not use designs that would separate out preexisting descriptive stereotypes from reactions to toughness (i.e. there was not a non-tough control group). Huddy and Terkildsen 1993 come closest to studying this issue in an experiment which describes some candidates as tough and some as compassionate; however, they were not trying to analyze the double bind (their focus was on the degree to which descriptions of tough and caring predicted perceptions of domain-specific policy competence) so they did not include any overall candidate evaluations as dependent variables and thus cannot speak directly to the issue at hand.
While political scientists have not studied the effects of tough candidate behavior on overall candidate evaluations and therefore cannot address the question of a double bind directly, they have studied the starting assumptions that people make about candidates. Many studies have found that people are more likely to project “feminine” traits (compassion, caring, honesty) onto female candidates while people are more likely to project “masculine” traits (strength, resolve, toughness, etc.) onto male candidates (Brown et al. 1993; Leeper 1991; Rosenwasser and Seale 1988; Rosenwasser and Dean 1989; Sapiro 1983.) These trait stereotypes have been linked to perceived issue competencies, such that men tend to be perceived to be superior on foreign policy and defense issues and women tend to be perceived as being better on domestic issues (Alexander and Andersen 1993; Falk and Kenski 2006; Koch 1999; Lawless 2004; Huddy and Terkildsen 1993; Sanbonmatsu 2002; Sapiro 1981). Some other studies, however, have failed to find that male candidates or politicians benefit from strength-related or other stereotypes. In particular, Kahn found that fictional male and female candidates were not perceived to be significantly different on traits like the ability to handle leadership or military issues (1992, 506; 1994, 183; 1996). A recent study by Fridkin and Kenney (2009) that examines perceptions of actual Senate candidates in the 2006 elections also confirms that same general pattern: no beneficial stereotypes for male candidates on any attributes, including leadership-related traits. In short, people may have some stereotypes about male and female politicians, but the evidence is mixed for many

---

3 These issue-based gender stereotypes may have more power at some times than others: for example, Lawless 2004 argues that these stereotypes disproportionately advantage male candidates during times of heightened security concerns. Moreover, the advantages that male candidates have on strength and security issues might have relatively minimal effects on candidates for legislative office but such differences may hurt women in their quest for the presidency, given the institution’s focus on agentic leadership and defense-related issues (Dolan 1997; Lawless 2004; Paul and Smith 2008).
traits. This evidence has also been drawn from many different decades, and yet gender stereotypes may have shifted significantly over time. Most importantly, there has been no attempt to quantify the degree to which acting tough can change any of these priors.

While political science has not empirically examined the double bind, business scholars have addressed the question by examining how people react to the tough behavior of male and female business leaders. The business scholarship focuses on “agentic” behavior by leaders, which is typically conceptualized by scholars as competence (traits such as ambitious, self-reliant, independent) and/or dominance (traits such as competitive, aggressive, forceful, and decisive). Research has shown that idealized leadership is generally conceptualized in agentic terms. Studies reveal that male business leaders are typically assumed to be agentic, while female business leaders are assumed to be more “communal” which reflects what is known in the business literature as the “think-leader-think-male” mindset (Sczensny 2003; Catalyst Report 2007). Female business leaders need to display agentic traits in order to be judged as being capable of the task of leadership (Glick et al 1988). Eagly and Karau 2002 argue that while male leaders also face gender stereotypes, these pose less of a challenge for most men because male leadership styles and assumptions based on descriptive stereotypes tend to overlap with the prescriptions associated with good leadership. Women, however, have to counter assumptions made about them based on descriptive stereotypes and establish competency on stereotypically male attributes in order to fulfill the prescriptions of leadership. Research has found that female leaders can successfully counteract feminine stereotypes and establish themselves as agentic if they provide strong, unambiguous signals to that effect (Rudman and Phelan 2008; Dodge et al., 1995; Glick et al. 1988; Heilman et al. 1995).

The difficulty faced by female business leaders is that negative consequences known as “backlash effects” can result from women’s attempts to counter gender stereotypes in the workplace by acting in an agentic manner. The literature shows that women who establish themselves as agentic tend to face sanctions in terms of favorability/likeability measures (see, for
example, Rudman and Glick 2001) which, in turn, can have tangible consequences such as a lower likelihood of getting hired, promoted, fairly compensated, or favorably evaluated on the job (see, for example, Rudman 1998 on hiring, Babcock and Laschever 2003 on salary negotiations, Heilman et al. 2004 on promotion, Eagly et al. 1992, and Eagly, Karau, & Makhijani 1995 on leadership evaluations). Women who act in an agentic manner while also displaying communal characteristics like warmth and niceness can minimize backlash (see, for example, Carli, 2001; Johnson, Murphy, Zewdie, and Reichard 2008). However, Rudman and Glick 2001 argue that those findings only apply to competence-related agentic behavior while backlash from dominance-related agentic behavior may be impossible for women to avoid because dominance by women inherently lowers perceptions that they are nice (758). In their review of the literature to date on the topic, Rudman and Phelan ultimately conclude that “ambitious women may have to choose between being liked but not respected (by displaying communal qualities) or being respected but not liked (by displaying agentic qualities), a dilemma not faced by men” (Rudman and Phelan 2008, 65).

These empirical results regarding female business leaders are quite suggestive that the conventional wisdom about female politicians and toughness might be correct; however, we have to be careful about assuming that findings regarding female business leaders also apply to the political realm. A critical difference is that the samples of reference are entirely different: the typical sample for a business study is comprised of business school students (that is, a male-skewed population of individuals who have self-selected into business careers) rather than the general population. Moreover, business leaders and political leaders may be held to different standards. It is entirely possible that business school students hold different opinions about the behavior of women in the business realm than ordinary people – the population of reference for this study – hold for politicians.

For these reasons and others, we cannot assume that the considerable evidence forwarded for a double bind for women in the business world will necessarily carry over to the political
world. To study the double bind, we need to identify and apply the relevant stereotype theories to political behavior in order to develop predictions about the role of gender prescriptions in campaigns. That is the purpose of the next section.

THEORIES AND HYPOTHESES

Various theories have emerged to explain how people process stereotypical and counter-stereotypical information. More specifically, several competing theories have arisen to explain how people process behavior that either conforms to, or competes with, preexisting descriptive stereotypes and/or prescriptive stereotypes. These respective theories can be used to derive predictions regarding how candidate behavior will influence voter attitudes.

The conventional wisdom, with its focus on both descriptive and prescriptive stereotypes, posits that women will be punished for violating the prescriptive stereotype that women should be non-agentic. This is precisely the prediction that emerges from the role congruity theory of prejudice developed by Eagly and Karau (2002). Role congruity theory melds the descriptive stereotype that women are less likely to have leadership characteristics with the proscriptive stereotype (and associated backlash for violating that proscription) that women are not supposed to be agentic. The theory stipulates that even when a woman can differentiate herself from the descriptive stereotype, she will still be penalized for violating the proscriptive stereotype, and thus is in a no-win situation when agency is required.

H1. Acting in a tough manner will hurt the electoral prospects of female candidates while it will help male candidates.

According to the conventional wisdom, this effect in the political sphere would work at least partially through the mechanism of depressing the caring/compassion scores of female candidates.

Somewhat differently, schema theory – with its primary focus on descriptive stereotypes rather than prescriptive stereotypes – would predict that men will benefit significantly from
toughness in the political sphere while women would be less affected by it. Schema theory predicts that information that is inconsistent with the expected schema (a schema is a mental cognitive structure that serves to order, process, and retain information) will be less powerful than schema consistent information, because inconsistent information will be harder for people to encode and recall and therefore will be discarded prior to processing (Lodge & Hamill 1986; Renn & Valvert 1993). In the case of candidate gender, a voter’s gender schema involves “their hypotheses about the beliefs, traits, and issue competencies of political candidates based on candidate gender” (Sanbonmatsu 2002, 21). If the belief that female politicians are descriptively weaker than male politicians is part of an individual’s gender schema, new information that a particular female politician acted in a tough manner will be hard for that individual to retain and will tend to be discarded; in contrast, information that a male politician acted in the expected tough manner will be readily encoded, retained, and utilized in evaluations of strength. Positive perceptions of strength, in turn, should be expected to increase overall evaluations of male candidates.

**H2. Acting in a tough manner will help the electoral prospects of male candidates while it will not have much influence on the electoral prospects of female candidates.**

Note that H1 and H2 predict the same overall result: male candidates will benefit disproportionately from displays of toughness relative to women. The hypotheses differ in that H1 posits that women will be punished by acting in a tough manner, and therefore women should avoid acting tough, whereas H2 posits that female candidates will not be affected much by acting tough while male candidates will benefit from it.

There are other possibilities, however, that would not predict a double bind for female candidates. The public may hold toughness-related descriptive stereotypes (i.e. that women are weaker leaders, less able to handle an international crisis, less able to be an effective president than men) while not holding prescriptive gender stereotypes that a female leader should act in a
traditionally feminine manner. Gendered prescriptions may be less important in the political sphere than for women in other roles, perhaps because female candidates are expected to fulfill the prescriptions of leadership (which involves toughness) more than they are expected to fulfill the prescriptions of traditional femininity. As such, positive counter-stereotypical behavior by candidates may serve to overturn the descriptive stereotype, at least as applied to a particular individual, and work solely to a candidate’s benefit as a result.

Expectancy violations theory takes this position and assumes that descriptive stereotypes matter more than prescriptive stereotypes. It predicts that people who act in a more favorable manner than expected based on a descriptive stereotype will be rewarded with higher evaluations by others as compared to people who were expected to act in a favorable manner from the start; conversely, people who act in a more negative manner than expected based on a descriptive stereotype will be sanctioned by lower evaluations as compared to people who were expected to act in a negative manner from the start (see Jussim, Coleman, & Lerch, 1987 and Jackson, Sullivan, and Hodge 1993). In other words, people who violate descriptive expectations by over-performing relative to low expectations will be disproportionately rewarded, while people who violate descriptive expectations by under-performing relative to high expectations will be disproportionately punished. By that logic, if female candidates start out lower than male candidates on agentic measures and then disprove that perception by acting tough, expectancy violations theory would predict that female candidates will benefit from that changed perception more than men.

4 See also Hayes 2005 and Kimball and Patterson 1997 for related discussions about how “expectations gaps” and “expectations-perceptions differentials” can produce disproportionate penalties for politicians who fail to meet high expectations and reward politicians who exceed low expectations.
H3. Acting in a tough manner will disproportionately help the electoral prospects of female candidates as compared to male candidates. This effect will mainly be observed among the kinds of individuals who tend to hold traditional descriptive stereotypes about female candidates (because they are the people for whom an expectancy is being violated).

An extension of H3 is that absent baseline descriptive gender stereotypes (i.e. women are weaker leaders) and absent prescriptions for political leaders (i.e. female leaders should act in an unassertive, traditionally-feminine manner), acting tough will have an equal influence on public opinion about male and female leaders.

A decade or more ago when there were fewer women leaders in national politics, H1 may very well have been the most compelling hypothesis. But in an era where women have greatly increased their proportions of leadership positions in politics and where women have taken on more visible leadership roles across many different sectors of society, the assumption that people expect female politicians to be traditionally feminine more than they expect them to display leadership traits seems less likely. H2 also depends on the assumption that people are less able to process information about tough female candidates the behavior is so unusual, which very well may not be the case in this era where there are many examples of tough political women in office. As such, there are strong reasons to expect that H3 is most compelling in the modern era, and to think that it may mainly manifest itself among a subset of people who are more likely to have traditional views of women. In my analysis, I find strong support for that hypothesis. The next section describes the experiment that produced those findings.

STUDY DESCRIPTION

To measure how tough behavior influences the perception of candidates, an experimental analysis has many advantages. There is substantial evidence that women adjust their behavior in order to conform to societal expectations (for a review of this dynamic, see Prentice and Carranza 2003). As a result, an examination of actual tough behavior by female politicians would
encounter difficulties due to selection effects. Moreover, it would be impossible to distinguish the effects of politician gender from an array of idiosyncratic behaviors and other campaign dynamics. An experiment using fictional politicians necessarily forfeits some degree of external validity, but the control it allows is necessary to directly examine the effects of different behaviors. Specifically, politician gender and behavior should be varied systematically, with all other information held constant. The purpose of the study should not be referenced and, in particular, respondents should ideally only be asked about one particular politician so they cannot discern that the study focuses on gender. Since young people are likely to have different gender stereotypes and expectations of politicians than older people, respondents should ideally represent a full range of ages rather than just undergraduates (which has been the population of study for much of the prior research on gender stereotypes in both psychology and political science). And since gender stereotypes have been found to vary geographically (Fox and Smith 1998), the sample should be geographically representative.

These goals were accomplished with an Internet-based experiment fielded by Polimetrix. The study was in the field from April 17-April 29, 2009. Polimetrix uses a technique called “sample matching” to generate a nationally representative sample (see Rivers 2006 for a description of the Polimetrix Sample Matching process). Respondents volunteer and register at a website and then become a part of an ongoing panel of respondents who take both political and non-political surveys. Modest incentives are provided to respondents for their participation on the panel. Polimetrix has a panel of over 2 million Americans, which is the population from which random samples are drawn for any given survey. The core of this sampling process is the Polimetrix “target matrix,” which is a combination of local, state, and national voter lists as well

---

5 This timing removed it from possible concerns about panel fatigue during elections. At the same time, it still retains the external validity of being a realistic time to read about a candidate’s early announcement of a 2010 election bid.
as other consumer databases which represent the American population overall. The firm draws a representative sample by matching a large number of characteristics of individuals from the panel to the characteristics determined by the target matrix and then the data are weighted on the backend to most accurately reflect the population at large. This sampling technique is becoming increasingly common for use in academic research applications, and studies have shown that its samples meet or exceed the quality of those based on more traditional telephone survey techniques (Berrens et al 2003; Sanders et al. 2007).

The experiment consisted of 737 completed interviews: 365 respondents were randomly assigned to the control group, while 372 respondents were assigned to the toughness condition. A choice existed as to whether to use videotaped stimuli (a video clip of male and female politicians acting in a tough manner) or a newspaper article for the treatment. The latter was selected to enhance internal validity by varying only gender, and not any of the countless other potential variables such as attractiveness, strength of facial features, facial expressions, tone of voice, clothing, and the like that could have otherwise reduced internal validity if male and female actors had been featured in photos and/or on video in the treatments.6

Half of the respondents were randomly assigned to read a newspaper article about a fictional candidate named “Congresswoman Karen Bailey” and half were assigned to read an

---

6 For example, several studies have shown that perceived attractiveness has very different effects for male and female candidates (Lewis and Bierly 1990; Sigelman, Sigelman, and Ribich 1986; Sigelman, Sigelman, and Fowler 1987). While it is possible that people could potentially have different reactions to viewing politicians act in a tough manner via video rather than reading about it in print, it seems less likely that people would systematically react differently to politicians of different genders via those different delivery mechanisms. As such, while this study uses print media within the experiment, it is reasonable to think that the gender differences it uncovers – the focus of this study – would not be dramatically different in the context of video.
article about “Congressman Kevin Bailey.” The article provides some baseline contextual information to respondents, which Sanbonmatsu argues is the best approach to measuring stereotypes so that respondents have the opportunity to attribute their preferences to something other than gender (2002, 28). The candidate was described as a moderate and oriented towards bipartisan solutions (due to N size limitations and other considerations, partisanship was not mentioned or systematically varied with the treatment). The article describes Bailey’s announcement that he/she will be seeking a U.S. Senate seat that will open up in 2010. It also briefly describes some of the legislative successes and challenges faced by the candidate during her/his two years and several months in office. The words “Congresswoman” and “Congressman” are utilized frequently, as are gender-indicating pronouns, in order to prominently emphasize the candidate’s gender. The treatment articles varied from the control group article with the addition of two paragraphs that described two incidents of tough behavior, a different headline, different sub-headline, and a different pull quote (the full text of each of the newspaper treatments is included in Appendix A).

A choice existed as to what kind of tough behavior to use for the treatment. In their studies of perceptions of male and female business leaders, Rudman and Glick 2001 specifically identified dominance – not competence – as being the agentic trait that puts women in the double bind, since dominance inherently lowers perceptions of niceness for women. As such, behaviors related to dominance were selected. The headline for the toughness treatment was “Congresswoman Karen Bailey threatens to hold up bills: ‘I will be heard,’ Bailey says,” while

7 A substantial proportion of the findings about descriptive stereotypes for male and female candidates are drawn from generic “would a male or female candidate be better on the following issues/traits” kinds of questions (see, for example, Dolan 1997; Dolan 2010; Hansen and Otero 2007; Kenski and Falk 2004; Lawless 2004; Sanbonmatsu and Dolan 2008; Rossenthal 1995; Simmons 2001), in which candidate gender is the only information provided to respondents.
the pull quote was “I won’t let anyone stand in my way.” The treatment situations involved refusing to yield the floor and continuing to give a speech when leaders ask the politician to stop. Bailey’s response to the situation was “I don’t take ‘no’ for an answer, and I won’t let anyone stand in my way. I will be heard.” The article also reports that an email was anonymously leaked in which Bailey threatened to “hold up the House leaders’ favored bills in committee unless they agreed to hold hearings to consider the provisions [he/she] was demanding.” The hypotheses derived from all of the theories are generalizable and thus should apply to a wide variety of toughness-related scenarios, including the one I selected. Regardless, this one is a fair representation of the kinds of rough-and-tumble legislative situations that policymakers may have to navigate. It is neither entirely admirable behavior nor entirely appalling behavior, which allows respondents to project gendered expectations onto the candidates rather than just gravitate towards a generalized assessment that the behavior is universally good or bad. In other words, this scenario is plausible for evaluating whether the double bind exists.

The key dependent variables are referred to, collectively, as the “outcome measures” (see Appendix 1 for a list of the variables and question wordings). I opted against asking a direct “vote preference” question, which would be overly artificial given that the announcement concerned a run for a primary more than a year away; moreover, particular competitors were not mentioned in the article, so asking for a vote selection would have been meaningless. Instead, I use three measures: overall favorability, likely effectiveness in the Senate, and likely effectiveness as U.S. President about 10 years from now. Favorability is included as an indication of satisfaction with the candidate in his or her current position in the House, and contains an affective component lacking in the other two measures. With respect to the likely Senate effectiveness measure, if a respondent thinks that the candidate will do a good job as Senator, it is a reasonable early indicator of likely vote preference. Likely presidential effectiveness is included because some scholars have hypothesized that female candidates are disadvantaged by gender stereotypes in their quest for the presidency but not for legislative office, and that the
presidential level might be particularly sensitive to strength-related stereotypes (Adams 1975; Huddy and Terkildsen 1993; Smith, Paul, and Paul; Dolan 2007; see also Lawless 2004). All of the measures utilize 7-point response scales, on which a “7” represents the most positive possible rating by a respondent. In all cases where there is a commonly accepted normative ideal (“strong,” “compassionate,” etc.), the “better” / more favorable option is associated with a high score. Exact question wording can be seen in Appendix B. “Don’t Know” responses needed to be minimized since item nonresponse bias could potentially mask a greater-than-average tendency to hold gendered views due to exclusion bias (see Berinsky 2004 for findings on this topic regarding racial policy preferences). More generally, for a study about the public’s tendencies to make assumptions, encouraging people to relay their first impressions and gut feelings was core to this analysis. As such, respondents were encouraged to make their best guess on responses and item nonresponse levels were below 1%.8

Several direct measures of strength were included. One is the standard “provides strong leadership” measure that is included in many surveys of candidate traits, including the NES, and has been found to correlate highly with outcome measures. Because “provides strong leadership” is as much a measure of good leadership as it is of “strength,“ per se, I also measured “strong” (vs. “weak”) as a separate measure. Additionally, since strength-related traits are likely to have issue-based implications, I included a measure of “ability to handle an international crisis.”

Two measures of caring were included. I included the standard NES measure “really cares about people like you” measure that captures a combination of empathy and personal identification with the candidate. I also included “compassionate” as a measure since it removes

Because respondents on Internet studies always have the option of registering a “don’t know” by “skipping” a question (i.e. not providing a response) a “don’t know” response was not explicitly provided for this study. Respondents were encouraged to “make your best guess even if you feel you do not have much information on which to base your answer.”
the personal identification element to focus more precisely on perceived warmth and empathy-related communal values.

Additionally, I used two measures to capture assessments of the appropriateness of the behavior and attributions for that behavior. I asked about whether the candidate’s behavior was “appropriate” or “inappropriate.” For the treatment group only, I also asked whether the candidate’s behavior could best be explained by “the nature of the situation” (“7”) or with “the personal characteristics of the candidate” (“1”). This is a question that stems from attribution theory, in which stereotypes are thought to influence perceptions of people by changing perceptions of causal mechanisms for behaviors (Deaux, 1976; Jones & McGillis, 1976; Jackson 1993). Behavior that is generally positive in nature but inconsistent with a stereotype will be tend to be attributed to temporary, external, “unstable” factors like the nature of a particular situation, while positive behavior that is consistent with a stereotype will be attributed to permanent, internal, “stable” factors like personality (see, for example, Heilman, 1983; Nieva & Gutek, 1980). This measure should provide a means by which different assessments of the cause of the behavior can be explored.

Finally, I measured the degree to which candidates were thought to be “assertive” (vs. “unassertive”) as a manipulation check.

ANOVA was utilized for the analysis of each model. Due to the randomized nature of all of the factors in this experiment, control variables are not necessary.

ANALYSIS

The first question is whether the experimental design worked; that is, did respondents react to the toughness treatment? The answer is yes. Setting candidate gender aside (that is, assessing the overall means based on half of respondents randomly viewing a male candidate and half viewing a female candidate), acting tough significantly increased the degree to which a candidate is viewed as being assertive (a 1-point difference on a 7-point scale, see Table 1 for the
main effects for “tough” as well as the means). Tough-acting candidates were also viewed as being “stronger” (.4 difference) as compared to the control group.

Overall, tough behavior is seen as being less “appropriate” than the ordinary behavior of the control group candidates (a difference of .3, with a p-value of .07). There was a significant difference between the control group candidates and the treatment candidates for “compassionate” – the measure that most directly taps into the feminine trait of empathy. Toughness takes a significant toll on perceptions of candidate compassion, with a .3 difference between the tough and control group candidates.

Given the push/pull effects operating under the surface, with toughness increasing perceptions of candidate strength but decreasing perceptions of compassion, it is perhaps not surprising that there is no overall effect of tough behavior on the three outcome measures (favorability, likely effectiveness in the Senate, or likely presidential effectiveness 10 years in the future). These results should perhaps be taken with a grain of salt, however; to the extent that there is an interaction between the “tough” condition and candidate gender, the main effects of toughness are not particularly meaningful.

Contrary to the expectations of the conventional wisdom and role congruity theory, there are no interactions between candidate gender and toughness that indicate disproportionate damage to female candidates for acting in a tough manner. In fact, on several measures the opposite is the case (see Table 1 for the full list of results.)

With reference to the most critical measures – the outcome measures – women are not penalized for being tough. On likely effectiveness in the Senate, there is no interaction between candidate gender and toughness. On likely presidential effectiveness – an especially key measure because of the assumed linkage between perceived strength and executive leadership – women who acted in a tough manner were actually viewed more favorably than women who did not act in a tough manner, while the opposite was true for men. In short, women improved their prospects for being viewed favorably for the presidency by acting tough, while men hurt their prospects
with the same behavior. On favorability itself – a key measure given that violations of prescriptions have so often been thought to result in being “disliked” – there is a suggestive effect (p=.08) that men are viewed less favorably if they act tough, but not women, who remain unchanged on favorability regardless of their behavior.

With respect to caring and compassion, there is not a significant interaction between candidate gender and toughness. Much of the conventional wisdom was predicated on the idea that toughness would specifically cause scores for feminine traits such as compassion to fall, and that they would fall farther for women. This was also the dynamic predicted and found by Rudman and Glick with reference to business leaders (2001, 753). However, for political candidates, there is no evidence of a backlash effect occurring at all. Perceptions of compassion are compromised somewhat by toughness for all candidates, but not for women in particular.

Strength does appear to be responsible for some of the differences in the outcome measures. “Strong” (vs. “weak”) was the only individual trait to approach significance for the gender interaction (p=.07). Male candidates increase on perceived strength when they act tough (+.3), but female candidates increase to an even greater degree (+.7). There is not a significant gender interaction for leadership strength, although the means are directionally consistent with those of the “strong” vs. “weak” measures. There is not a significant interaction effect for ability to handle an international crisis.

Contrary to the conventional wisdom and role congruity theory (which would posit that tough behavior would be seen as inappropriate for female candidates), tough behavior was actually seen as being slightly more appropriate for female candidates than for male candidates (p=.07). This may be due, at least in part, to the fact that people are more likely to attribute the behavior to situational factors for women and dispositional factors for men (a difference of .4). Because toughness has both positive (higher strength scores) and negative (lower compassion scores) components, it is hard to reconcile this finding with attribution theory entirely. If toughness was entirely positive, then attribution theory would predict that men would benefit
from the fact that it is attributed to their personality; if toughness was entirely negative, then the
theory would predict that women would benefit from the fact that it is attributed to situational
factors. In this case, the fact that tough behavior by women is more likely attributed to the
situation seems to cause the behavior to be viewed as more appropriate. The relationship between
appropriateness and the outcome measures is not clear in this analysis, however; it may be
contributing to the higher scores for women, but the benefit to women on the outcome measures
may be driven in whole or in part by other factors.

Overall, we can conclude from the analysis that there is no penalty for female candidates
who act tough; in fact, they can benefit from it in some important ways. In many ways, the fact
that there is no penalty for acting tough is not very surprising; each of the stereotype theories and
the conventional wisdom are predicated on there being descriptive gender stereotypes in place
that hurt women. In fact, my data show little evidence of such stereotypes on agency-related
traits. While negative gender stereotypes may well have existed in earlier times, people overall
did not apply traditional stereotypes to female candidates in mid-2009 when my study was
conducted. People may now consider candidate gender in a new light after the historic
presidential primary run by Hillary Clinton, the vice presidential run by Sarah Palin, Nancy
Pelosi’s election as the first female Speaker of the House in 2007, and/or in response to the
prominent roles of many other women leaders at various levels of government in recent years.

Further leverage can be gained on the theories at stake by examining the population which
is most likely to still hold traditional stereotypes of female candidates: older individuals.9 The

9 There is a different possible approach to this same general dynamic: comparing respondents
who have more feminist views to those who have less feminist views. The upside of that
approach is that it more directly assesses individual rather than group-based stereotypes; the
downside to that approach is that it incorporates measurement error (especially error associated
with social desirability effects), while age is a more straightforward measure. The results for
analysis of older individuals is very illuminating. 60+ year olds were analyzed separately on the basis that their socially and politically formative years took place before the social upheaval of the 1960s era in which traditional gender roles and expectations became sharply contested. The results (see Table 2) show that younger people (18-59 year olds) hold no significant gender stereotypes about the control group male and female candidates. For 60+ year olds, two results are substantively large enough that we can conclude, albeit just shy of the p=.05 level, that they hold traditional descriptive stereotypes about female candidates concerning two key matters: presidential effectiveness (where female candidates are rated a striking .7 lower than male candidates) and ability to effectively handle an international crisis (where female candidates are .8 lower than male candidates.) The N sizes for older individuals in the control group are sufficiently small that it would be inadvisable to rule out the presence of gender stereotypes on many other attributes based on these results (the results of several traits are directionally suggestive of traditional stereotypes for this population group).

In short, older individuals start out with at least some traditional stereotypes about female candidates, while younger individuals do not. But what happens when female candidates demonstrate that they are agentic and tough? Are they penalized disproportionately by the older individuals who are more likely to have traditional views on gender? Or are they rewarded by them, in keeping with expectancy violations theory? The answer is the latter. A 3-way ANOVA between toughness, candidate gender, and respondent age reveals that older individuals disproportionately reward women for acting tough across a range of attributes (favorability, likely senate effectiveness, likely presidential effectiveness, strong leadership, and effectively handle an international crisis.) Older people are no more likely than younger people to attribute the feminism related measures (both self-ID and measures that tap into feminist views) do not achieve statistical significance but the results are directionally consistent with the results for age shown here.
behavior to situational vs. behavioral factors, which rules out attribution bias as a possible explanation for the dynamic.

Thus, it is those people who are most likely to hold traditionally gendered assumptions about candidates who are according the most disproportionate benefits to female candidates who act tough. This appears to be a clear repudiation of the conventional wisdom and role congruity theory, at least as applied to politics. Those theories would presumably expect the older generation to hold the strongest prescriptive stereotypes for women; yet older individuals are not only refraining from penalizing stereotype-deviant candidates, they are actually rewarding such candidates for the deviant behavior. These results provide confirmation of expectation violations theory which claims that when people outperform negative stereotypes and associated low expectations, they will be disproportionately rewarded. From this standpoint, when men act in a tough manner, they are simply doing what they should do. But when female candidates act in a tough manner, she is showing that she is more capable of directing international affairs and being presidential than expected. Tough behavior by female candidates appears to come as a surprise, albeit a pleasant one, to those who might be generationally inclined towards traditional gender assumptions.10

CONCLUSION

The results of this study fail to confirm the conventional wisdom about the double bind for female politicians. Acting tough does not harm female candidates; instead, it helps them in some

10 Note that while some scholars have noted that men and women respondents might hold different stereotypes of male and female candidates (see, for example, Kahn 1994 and 1996), I do not find an interaction for candidate gender*toughness*respondent gender. Overall, female respondents tend to reward tough behavior by candidates far less than do male respondents, but male and female respondents are no more or less likely to reward male or female candidates for tough behavior.
ways. Regarding the more specific hypotheses, H1 was not confirmed; these results show that the prediction of role congruity theory (that female politicians will be penalized for violating traditional prescriptions of the female role) is not valid in the political context. H2 was also not confirmed: the prediction of schema theory—that people are more likely to give credit to the male politicians for schema-consistent tough behavior—received no support from this analysis. In contrast, H3 was confirmed by this analysis. The public rewards female candidates for acting tough in a couple of specific ways. This is especially the case among older individuals who tend to hold more traditional views on female politicians, which suggests that some specific people may be rewarding tough women for exceeding low expectations. Rather than producing the double bind that is postulated by the conventional wisdom, acting tough thus represents an opportunity for female candidates to positively distinguish themselves.

In some respects, it is not especially surprising that the conventional wisdom is wrong on this topic. The foundation of the conventional wisdom was built at a time when gender discrimination very well may have existed for female politicians. There are more women in national politics than ever before now, and tough female politicians have been especially visible in recent years. Nancy Pelosi, Condoleeza Rice, and Hillary Clinton are just some of the more prominent examples of tough female political leaders who are at the forefront of national decision making today. Tough behavior by female politicians may have now become sufficiently commonplace that it has lost any prescriptive component that might have been there in earlier days.

These findings diverge from those in the business literature, where research indicates that women face substantial challenges in the corporate world. Evidence exists that female leaders in non-business realms also face a double bind. In our own business—that is, the business of higher education—female instructors have been found to be subject to far greater penalties for being tough (i.e. tough graders) than male instructors on teaching evaluations and overall perceptions of
competence (Sinclair and Kunda 2000). In short, there seem to be many examples of disproportionate penalties for tough behavior by women in the world at large, yet this dynamic does not operate in the sphere of politics. Why the difference?

There are important distinctions between the other literatures and the present one. The sample for business and education studies is typically comprised of either business school students or undergraduates, all of whom may very well have different expectations of women than the U.S. population as a whole. Moreover, many of the most relevant business studies were conducted in the 1970s, 1980s, and 1990s, and that data may no longer reflect the state of gender stereotypes at the present time. But there may also be a more fundamental difference at the root of the apparent discrepancy between the findings in this study and those in non-political science literatures: female politicians may simply be subject to different expectations than ordinary women, business women, and female educators. Setting aside the matter of gender, each of those roles involves different overall prescriptions: business executives “should be” agentic nearly all the time in order to lead their companies, while professors “should be” tough only when absolutely necessary (at least in the eyes of our students). As for ordinary people, they rarely need to be agentic in everyday life. In contrast, politicians – male and female – have always faced the need to act tough when necessary while being caring and compassionate nearly all the time. They need to feel your pain and be able to stand up to those who would try to cause you pain. Male politicians are expected to display communal traits and female politicians are

11 For example, using both evaluations of actual professors and an experimental design using actors acting as management instructors, Sinclair and Kundar found that undergraduates who receive negative feedback from instructors derogate the competence of female instructors but not male instructors. In turn, the final evaluation scores of female instructors were found to be far more dependent upon the amount of positive feedback they give to students than those of male instructors.
expected to display agentic traits, and that may be arguably less true with respect to business leadership.

Moreover, the ascendance of women as political leaders in recent years has been far more rapid than within the business world. If one considers U.S. legislators to be roughly equivalent to CEOs in terms of status, the gender imbalance is over six times greater in the business world than in U.S. national politics. As a result, the average person today is likely to have far more exposure to executive-level political leadership by women than the average worker has to executive-level business leadership by women. Studies have shown that gender stereotypes are exacerbated in conditions where there are relatively fewer women leaders, in part because perceptions of illegitimacy are reduced and because women attract less gender-based attention and are subject to fewer stereotypes when they move from “token” status to occupy a greater proportion of leadership roles within an organization (see reviews by Ely 1995, Eagly and Karau 2002, Rudman and Phelan 2008). As such, gender may be relatively less salient in the political realm than in the business realm in the current day due to relatively more equitable gender representation in the former as compared to the latter.

Barack Obama’s election indicates that racial stereotypes regarding candidates can and have changed dramatically in recent years, and my analysis may suggest that a similar shift may have occurred with respect to female candidates. Indeed, Obama’s election does not simply illustrate the general point that stereotypes regarding candidates can shift dramatically over time,

12 18% of members of the U.S. House and Senate were female in 2009, while only 2.8% of CEOs of Fortune 1000 companies in 2009 were female.

13 Interestingly, Ely 1994 uses 15% or more female (in that case, for the proportion of female partners in law firms) as the definition of a sex-integrated firm and finds significant differences with that distinction, which suggests that even differences between extreme underrepresentation of female leaders and better-but-still-poor representation of female leaders can be important.
but may have also contributed to the more specific result I find here: having a non-white male become President for the first time may well have caused many voters to update their views regarding all other kinds of politicians – racial minorities and women, alike.

Reasonable questions remain about whether the media treats tough female and tough male politicians equally. Anecdotal evidence suggests that some people in the media will say things about female candidates that they would never say of male candidates; certainly a male politician would not be called “castrating, overbearing, and scary” as Tucker Carlson referred to Hillary Clinton on MSNBC during the 2008 primary season (Carlson 2007). With respect to the bulk of more traditional, less “editorial” coverage, Kahn’s research shows that male and female politicians in the 1980s received different types of media coverage, and that receiving “male” versus “female” types of coverage can affect public perceptions of candidates, especially for Senate races where such coverage differences were found to be more prominent than in gubernatorial races (Kahn 1992, 1994, 1996; Kahn and Goldenberg 1991; also see Bystrom et al. 2001; Kropf and Boiney 2001). Some other scholars, however, have found evidence that the coverage differences based on gender might be more modest than initially suspected (see, for example, Banwart et al. 2003; Fowler and Lawless 2009; Smith 1997). We should be conscious that a media effect could exist with reference to tough behavior by female politicians, and future research along these lines would be very beneficial. But the conventional wisdom is focused on direct reactions by people to female toughness – i.e. what the public expects of male and female politicians – and this study shows that those concerns are unfounded in the modern era.
<table>
<thead>
<tr>
<th></th>
<th>Favorability</th>
<th>Senate Effectiveness</th>
<th>Pres Effectiveness</th>
<th>Strong Leader</th>
<th>Strong</th>
<th>International Crisis</th>
<th>Cares about people like you</th>
<th>Compassionate</th>
<th>Assertive</th>
<th>Appropriate</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Tough (vs. the control group)</strong></td>
<td>0.85</td>
<td>0.04</td>
<td>0.11</td>
<td>5.53</td>
<td>26.66</td>
<td>0.03</td>
<td>0.94</td>
<td>4.94</td>
<td>105.69</td>
<td>3.34</td>
</tr>
<tr>
<td><strong>Candidate Gender</strong></td>
<td>3.05</td>
<td>0.93</td>
<td>4.70</td>
<td>0.73</td>
<td>3.39</td>
<td>1.35</td>
<td>0.17</td>
<td>0.34</td>
<td>0.01</td>
<td>0.53</td>
</tr>
<tr>
<td><strong>p-val</strong></td>
<td>0.08</td>
<td>0.34</td>
<td>0.03</td>
<td>0.39</td>
<td>0.07</td>
<td>0.25</td>
<td>0.68</td>
<td>0.56</td>
<td>0.91</td>
<td>0.47</td>
</tr>
<tr>
<td><strong>Fp-val</strong></td>
<td>4.5</td>
<td>4.5</td>
<td>3.8</td>
<td>4.5</td>
<td>4.4</td>
<td>4.0</td>
<td>4.3</td>
<td>4.3</td>
<td>4.5</td>
<td>4.3</td>
</tr>
<tr>
<td><strong>Overall</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>CONTROL (male + female candidates)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOUGH (male + female candidates)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TOUGH vs. CONTROL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>GENDER &amp;. TOUGH vs. CONTROL</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

N=372-735 for all 4 conditions combined

**Candidate Gender**

<table>
<thead>
<tr>
<th>Situational Attribution</th>
<th>Fp-val</th>
<th>TOUGH Male cand</th>
<th>TOUGH Female cand</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Situational = &quot;7&quot;</strong></td>
<td>4.59 **</td>
<td>NA</td>
<td>NA</td>
</tr>
<tr>
<td><strong>Dispositional = &quot;1&quot;</strong></td>
<td>0.03</td>
<td>NA</td>
<td>NA</td>
</tr>
</tbody>
</table>

N=372 for both conditions combined
<table>
<thead>
<tr>
<th>Candidate Gender * Tough * Age60+</th>
<th>ANOVA</th>
<th>AGE60+ * Candidate Gender</th>
<th>A PRIORI STEREOTYPES - Sig control group diffs for male vs. female candidates</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>F</td>
<td>p-val</td>
<td>CONTROL Male Cand</td>
</tr>
<tr>
<td>Favorability</td>
<td>5.41 **</td>
<td>0.02</td>
<td>4.5</td>
</tr>
<tr>
<td>Senate Effectiveness</td>
<td>4.10 **</td>
<td>0.04</td>
<td>4.5</td>
</tr>
<tr>
<td>Pres Effectiveness</td>
<td>5.55 **</td>
<td>0.02</td>
<td>3.9</td>
</tr>
<tr>
<td>Strong Leader</td>
<td>3.65 *</td>
<td>0.06</td>
<td>4.4</td>
</tr>
<tr>
<td>Strong</td>
<td>1.55</td>
<td>0.21</td>
<td>4.3</td>
</tr>
<tr>
<td>International Crisis</td>
<td>4.77 **</td>
<td>0.03</td>
<td>4.0</td>
</tr>
<tr>
<td>Cares about people like you</td>
<td>2.37</td>
<td>0.12</td>
<td>4.2</td>
</tr>
<tr>
<td>Compassionate</td>
<td>1.94</td>
<td>0.16</td>
<td>4.3</td>
</tr>
<tr>
<td>Assertive</td>
<td>1.05</td>
<td>0.31</td>
<td>4.4</td>
</tr>
<tr>
<td>Appropriate Behavior</td>
<td>2.93</td>
<td>0.09</td>
<td>4.4</td>
</tr>
<tr>
<td>N averages approx. 143 per condition</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>N averages approx. 41 per condition</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Candidate Gender * Age60+ (ONLY TOUGH)</th>
<th>Situational Attribution</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(Situational = &quot;7&quot;, Dispositional = &quot;1&quot;)</td>
</tr>
<tr>
<td></td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>0.93</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3.4</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3.8</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3.5</td>
</tr>
<tr>
<td></td>
<td>NA</td>
</tr>
<tr>
<td></td>
<td>3.8</td>
</tr>
</tbody>
</table>

* p≤.1, ** p≤ .05, *** p≤ .01, 2-tailed
APPENDIX A

News Article Text

Text for the male candidates was identical and simply substituted “Kevin” for “Karen,” and used male rather than female pronouns where appropriate.

CONTROL GROUP ARTICLE:

HEADLINE: Congresswoman Karen (/Kevin) Bailey announces Senate bid

PULL QUOTE: None

Congresswoman Karen Bailey recently announced she will be seeking a U.S. Senate seat that will be vacated in 2010 due to a retirement. [TREATMENT INSERT]

Congresswoman Bailey is expected to be a major contender in the Senate race due to her strong appeal to voters from both political parties. “I am a solutions candidate,” Bailey says, “with a focus on bringing the parties together to solve challenges for our state and for our country.”

With a preference for moderate policies, she is “one of the few remaining politicians in the sensible center – someone who tries to find bipartisan solutions to the country’s problems,” says Martin Druckman of the non-partisan Wilson Institute of American Politics.

With one full term under her belt and now several months into her second term, Congresswoman Bailey has experienced many legislative successes as well as some legislative challenges. On the success front the state now has one of the highest per capita levels of road improvement spending in the country and has dramatically increased its total number of small businesses with the help of federal business improvement grants. Congresswoman Bailey has been less successful as the co-author of two bills to increase spending on safety equipment for soldiers in Iraq and Afghanistan and to make federal bureaucracies more cost-efficient, both of which failed to pass in the House earlier this year.

TREATMENT ARTICLE:

HEADLINE: Congresswoman Karen Bailey threatens to hold up bills: “I will be heard,” Bailey says

PULL QUOTE: “I won’t let anyone stand in my way”

TREATMENT INSERT: The announcement has been overshadowed by other events, however. In an incident earlier this week, the congresswoman refused to yield time she had remaining on the House floor to a colleague and is seen on a video clip calmly but very firmly stating, “You’ve had your turn, and now it’s my turn. I will not yield the floor until I have said what I need to say,” prior to finishing the full length of her speech. And to House leaders who refused to let the congresswoman add provisions to a bill, she said, “I don’t take no for an answer and I won’t let anyone stand in my way. I will be heard.”

Additionally, in an e-mail that was anonymously provided to the media, Representative Bailey threatened to hold up the House leaders’ favored bills in committee unless they agreed to hold hearings to consider the provisions she was demanding. The congresswoman’s office declined to comment on the incidents.
APPENDIX B
Question Wording

Q1. How favorable or unfavorable do you feel toward the candidate? (Extremely Unfavorable = 1, Extremely Favorable = 7)

Q2. How good of a job do you think that the candidate would probably do with the following… Be an effective U.S. Senator (A poor job = 1, An excellent job = 7)

Q3. How good of a job do you think that the candidate would probably do with the following…Be an effective U.S. President about 10 years from now (A poor job = 1, An excellent job = 7)

Q4. How good of a job do you think that the candidate would probably do with the following… Effectively handle an international crisis (A poor job = 1, An excellent job = 7)

Q5. How well do you think this word describes the candidate?...Provides strong leadership (Not well at all = 1, Extremely Well = 7)

Q6. How well do you think this word describes the candidate?...Really cares about people like you (Not well at all = 1, Extremely Well = 7)

Q7. How well do you think this word describes the candidate?...Compassionate (Not well at all = 1, Extremely Well = 7)

Q8. Please rate the candidate on the following characteristics…
   A) Weak= 1          Strong=7
   B) Unassertive= 1    Assertive=7
   C) Acts inappropriately= 1    Acts appropriately=7

Q9. I think the behavior of the candidate in the article can probably best be explained by (the personal characteristics of the candidate = 1, the difficulty of the situation = 7)
BIBLIOGRAPHY


Falk and Kenski 2006. Issue saliency and gender stereotypes: Support for women as presidents in times of war and terrorism. Social Science Quarterly 87 (1); 1-18.


