# Gender and Electoral Communication: How Differences in Non-verbal and Verbal Behavior Affect Evaluations of Male and Female Politicians

**Delia Dumitrescu, Elisabeth Gidengil and Dietlind Stolle**

**Introduction**

The literature on voters' responses to female candidates reaches very different conclusions, depending on whether a study is conducted in the lab using hypothetical candidates or based on survey data about actual candidates in a real-world election. Experimental studies and survey-based studies of hypothetical or generic candidates typically find a good deal of gender stereotyping. Studies of real-world elections, on the other hand, typically uncover little evidence of gender bias.

The great advantage of experiments is that they make it possible to study gender stereotyping under very tightly controlled conditions, free of all the other confounding factors that complicate studies of real-life settings. Participants in these studies are typically asked to evaluate a hypothetical candidate based on a brief profile or a speech purportedly given by the candidate. Some participants are told that the candidate is a woman, while others are told that the candidate is a man. If the woman is ascribed different traits than the man, this must reflect gender stereotyping: except for the candidates' gender, participants are evaluating the candidates based on exactly the same information. The problem is that these results may not hold outside the lab setting. It is one thing to read a tough speech supposedly delivered by a female candidate but the response may be quite different when she is heard—and seen—delivering the same speech. The way the speech is delivered may have as much influence—and maybe more—on voters' impressions of the candidate as the content of her speech.

Survey-based studies resolve the artificiality issue but they are subject to possible social desirability bias. People want to convey a positive impression of themselves and may therefore be reluctant to express any bias toward female candidates. More importantly, survey respondents may well hold egalitarian beliefs about gender roles and the place of women in politics and yet remain under the sway of deeply socialized, unconsciously held gender stereotypes.

This paper presents the results of an experiment that was designed to maximize what McDermott (2002) terms "experimental realism." Participants are presented with a believable and engaging stimulus in the form of a candidate webcast. The gender of the candidate is varied. Participants are not merely reading the candidate's speech; they are seeing and hearing it. Physiological response technology is used to capture their unconscious reactions to the speech. The objective is to see how the verbal and non-verbal behavior of male and female candidates influences how they are evaluated. Participants view one of four different versions of a political speech delivered by male and female actors taking on the role of political candidates.
.
This experimental design enables us to answer a number of important questions. Is the way that a message is delivered more important than the message itself? How important is it for candidates to have a confident demeanour? Or is their success a function of what they say and not just how they say it? More importantly, we want to know whether the answers to these questions differ depending on the gender of the candidate. Are women and men rewarded equally for equivalent performances? Do women pay a higher price for a similarly poor performance? And do comparable performances elicit similar physiological responses?

**The Influence of Verbal and Nonverbal Communication**
After decades of research, compelling evidence exists that a candidate's nonverbal behavior affects voters' evaluations. Reactions to the Kennedy-Nixon televised debate in the 1960 US presidential debate drove home the message that exhibiting confidence matters, and maybe matters more than the substance of the candidate's speech. Voters who had listened to the debate on the radio tended to favour Nixon while those who had seen the candidates on television tended to favour Kennedy. A general consensus emerged that, in terms of substance, it was Kennedy who was the weaker of the two candidates, but in terms of demeanour, "Kennedy showed confidence and determination [while] … Nixon was perspiring noticeably and seemed ill at ease" (Patterson et al. 1992). More recently, Gentry and Duke (2009) note the negative impact of Gore's "uncomfortable" appearance in his debates with George W. Bush.

Theoretical insights and empirical findings from different fields of research (see Nagel, Maurer and Reinemann 2012) underpin the expectation that male and female candidates alike will get higher ratings when they are confident [H1]. For example, neuroscientists have shown that visual information is processed much more rapidly than verbal information, which would explain why the visual elements of a candidate's performance are likely to have more influence on voters' responses (Van Damme, Crombez and Spence 2009). Dual Coding Theory (Paivio 2007), meanwhile, implies that visual information is more easily remembered and retrieved because it is stored in both the visual and the verbal systems, unlike verbal information, which is only stored in the verbal system. Finally, recent research on conscious and unconscious information processing shows that nonverbal cues are processed unconsciously and prior to any other information (Lodge, Taber and Verhulst 2011).

However, research on verbal dimensions of communication indicates that they also can be powerful factors for overall evaluations of candidates. Accordingly, our second expectation is that both candidates will receive higher ratings for giving a strong speech [H2]. Indeed, it is possible that the verbal aspects of communication dominate the nonverbal ones. An early study of US presidential debates found that viewers of the 1976 debate between Dole and Mondale were more likely to make inferences about the candidates based on what they said rather than on how they said it (Krauss et al. 1981). Similarly, a recent analysis of responses to the 2005 German televised leaders' debate found that viewers were significantly more likely to rely on the content of the speech of the two candidates than on their visual demeanour when recording their impressions of the speakers (Nagel, Maurer and Reinemann 2012).

Finally, we can expect there to be an interactive effect between the verbal and nonverbal elements of a politician's speech: ratings will be highest when the candidates deliver a high quality speech with confidence [H3]. For example, Awamleh and Gardner (1999) found that in the organizational context a strong delivery significantly improves ratings of a leader's effectiveness and charisma, and that the best evaluations are obtained when a strong delivery is accompanied by strong content.

**Gender and Leader Evaluations**
What interests us in this paper is whether responses to the verbal and non-verbal elements of candidate speeches are gendered. In other words, do the responses differ, depending on whether the speech is being delivered by a man or by a woman? The first possibility to consider is that a

female candidate will be punished for "competent assertiveness" (Butler and Geis 1990). Women who behave assertively present a "basic schema incompatibility" (p. 48). They are violating deeply held and unconscious expectations about how women ought to behave. The effect of this gender-role incongruence is to accentuate the gender-role incongruent behaviour (Eagly, Makhijani, and Klonsky 1992). As a result, words and gestures that would be perceived as merely assertive on the part of a male candidate may be perceived as too hard-hitting on the part of a female candidate. In other words, a female candidate will be evaluated more negatively than her male counterpart when both candidates deliver a similarly strong speech with confidence [H4].

This hypothesis is grounded in role congruity theory (Eagly and Karau 2002).[1] This theory focuses on the disjuncture between gender stereotypes and the requirements of leadership roles. Women are typically perceived as possessing more *communal* characteristics. They are believed to be warm, sociable, supportive and nurturing. Men, on the other hand, tend to be perceived as having more *agentic* characteristics. They are typically thought to be assertive, competitive, self-confident and forceful. The feminine communal traits ascribed to women are seemingly incompatible with the masculine agentic traits that are often viewed as necessary for effective leadership. As a result, women are less likely than men to be perceived as having leadership ability. However, these stereotypes do not simply serve to describe how people *think* women and men behave; they also serve as *injunctive norms* (Cialdini and Trost 1998) that prescribe how women and men *ought* to behave. Thus, a woman who displays agentic behaviour is likely to be penalized for acting contrary to the injunctive norms of female behaviour. Women aspiring to leadership position thus face a classic double bind. As Eagly and Karau (2002) explain, "Conforming to their gender role would produce a failure to meet the requirements of their leader role, and conforming to their leader role would produce a failure to meet the requirements of their gender role" (p. 576).

A number of studies have documented the effects of role incongruity. Butler and Geis (1992), for example, tracked audience's nonverbal affect responses to speakers from behind one-way mirrors. They found, as hypothesized, that competent, assertive women speakers elicited more negative than positive facial expressions than did equally competent, assertive men. A role congruity-related effect was also observed in an experiment by Rudman and Glick (2001), which found that in a hypothetical hiring situation, research participants penalized agentic women for a job that was described in a feminized manner because their social skills were devalued and they did not seem "nice" enough, while agentic males' perceived hireability and social skills were relatively unaffected. The authors demonstrate that these patterns reflect implicit gender stereotypes, as measured by a gender stereotype version of the Implicit Association Test (IAT)[2]

---

[1] Note that similar expectations can be derived from expectations theory (see, for example, Ridgeway 2002). This theory emphasizes the role of status beliefs about gender that lead women to be penalized for violating the expected status order by acting assertively.

[2] The Implicit Association Test (IAT) is a (typically) computer-based measure that requires users to match an attribute with one of two target concepts, and measures the speed of these attributions to determine the strength of the associations in memory (with faster ones suggesting stronger associations and thus implicit biases). It was developed as a way to overcome the problems of social desirability biases or lack of awareness in self-report survey methods and the

developed by Greenwald, McGhee and Schwartz (1998). They conclude that agentic women can avoid negative reactions provided that they do not display agentic traits related to social dominance. Their advice to aspiring women leaders is to exhibit agentic traits associated with competence while also displaying "feminine niceness". As they acknowledge, "Treading the fine line of appearing competent, ambitious and competitive, but not at the expense of others, is a tall order, especially for women" (p.758). Indeed, it has been suggested that women's need to delicately manage these impressions in order to succeed can lead to anxieties and other issues that impair task performance (Riordian, Gross and Maloney, 1994).

Much of this research on role congruence and leadership has been conducted in the management field. However, experimental studies in political science have uncovered evidence of the attribution of stereotypically gendered traits and issue competencies to hypothetical candidates for elected office. These studies have found that voters will attribute different personality traits to hypothetical candidates who are identical in every respect save their gender (Huddy and Terkildsen 1993a, 1993b; Huddy 1994; Leeper 1991; Rosenwasser and Dean 1989; Rosenwasser and Seale 1988). Male candidates tend to be perceived as tough, aggressive, self-confident and assertive, while their female counterparts are characterized as warm, compassionate, people-oriented, kind and sensitive. Female candidates are also stereotyped as being more moral, hardworking and honest than their male counterparts (Huddy and Terkildsen 1993a, 1993b; Leeper 1991; Rosenwasser and Dean 1989). These gender-trait stereotypes influence perceptions of issue expertise (Huddy and Terkildesen 1993a): female candidates are seen as being better suited to dealing with compassion issues, such as poverty, education and health, whereas men are deemed more competent when it comes to issues like the economy, the military and defense (see, for example, Huddy and Terkildsen 1993a; Sapiro 1981/82; see also Sanbonmatsu and Dolan 2009; Dolan 2010).[3] Rosenwasser and Dean (1989) find that the stereotypically feminine issue competencies are deemed less important than the stereotypically masculine ones. Similarly, Huddy and Terkildsen (1993b) conclude that stereotypically masculine traits are considered more important in a political leader than stereotypically feminine ones. In short, these studies suggest that deeply held gender stereotypes may conflict with normative expectations about political leaders.

Role congruity theory predicts that female leaders will be penalized for acting contrary to stereotype. However, Rudman and Glick (2001) did not find any difference in the perceived hireability of agentic women and agentic men for a job that was masculine as opposed to feminized. Indeed, some recent studies have found that there seems to be more acceptance of

---

like. Mast (2004) additionally devised an IAT that involved gender and perceptions of hierarchy, which found – in line with the communal/agentic discussion above – an implicit stereotype of women as suitable to egalitarian structures and men to hierarchical ones. Note that the IAT has been the subject of some controversy and multiple critiques have been published (for an overview, see Fiedler, Messner and Bluemke, 2006).

[3] However, a more recent study by Bernstein (2000) found that there was no significant gender difference for perceptions of policy competence for hypothetical candidates for a U.S. congressional race delivering the same messages – which varied by explicitness and feminized or masculinised issue—although women were perceived as better at education than crime, whereas men were deemed equally competent in both.

agentic behaviour on the part of women in leadership roles (see Bongiorno, Bain and David 2013). Similarly, experimental studies have found that hypothetical female candidates can play up their combative side without compromising their supposed 'feminine strengths' (Huddy and Terkildsen 1993; Leeper 1991). Indeed, the advice to female candidates that has emerged from experiments on gender stereotyping of political candidates has been to 'act tough'. If these findings about agentic behaviour hold outside the laboratory setting, H4 may fail to hold: women and men alike will be equally rewarded for "competent assertiveness" in the form of a strong speech delivered with confidence.

But what happens when a female candidate lacks "competent assertiveness"? Here the literature would lead us to expect that a female candidate will pay a higher price than her male counterpart for failing to exhibit agentic behaviour [H5]. The price should be especially high if the speech itself is tentative *and* the delivery lacks confidence [H5A]. Bongiorno, Bain and David (2013) have recently focused on responses to women in leadership roles who fail to display sufficient agency. They build on Carli's (1990) classic study in which participants responded to speeches delivered by male or female students in an agentic (assertive) or non-agentic (tentative) style. Carli's experiment found that the female speakers were more likeable and influential when their delivery was tentative whereas tentative or assertive speech made no difference to the likeability and influence of the male speakers. However, Bongiorno and her colleagues argue that agentic behaviour on women's part has become much more acceptable over the past two decades—to the point where a female leader who exhibits non-agentic behaviour can expect to be penalized. Importantly, they do not expect men to pay a similar price because "the association between being male and being a leader…. can provide legitimacy to men in leader roles, affording them the benefit of the doubt on occasions where their leader behaviour diverges from the ideal" (p. 4). Accordingly, they predict that female leaders who fail to display the expected agentic behaviour will elicit more disapproval than either non-agentic men or agentic women. They conducted two experiments, both involving a politician. Whether the experiment involved reading a transcript of the speech or listening to an audio recording, assertive women proved to be as likeable and influential as male leaders (and in some instances, even more so). Tentative women, on the other hand, were less likeable and less influential than both tentative men and assertive women.

## The Experiment

In this paper, we use multiple methods to test these hypotheses about the impact of verbal and non-verbal communication on evaluations of male and female candidates and to elucidate the results. Combining an experimental design with measures of participants' physiological responses[4] offers a powerful way of at getting at unconscious biases. There is often a disjuncture between people's conscious beliefs and their automatic responses. Even when the conscious beliefs are egalitarian, unconscious reactions to female candidates may be gendered, reflecting deeply rooted implicit biases (see, for example, Butler and Geis 1990; Levinson and Young 2010).

---

[4] There is not much work on measuring physiological responses to political communication, and the few studies that exist do not test for different possible reactions to male and female politicians. For an example of this methodology, see McHugo et al. (1991) who found evidence that prior attitudes shaped physiological responses to known politicians.

Our experimental design also overcomes some important limitations of previous studies of nonverbal communication. First, many of these studies have been based on public figures (typically, well-known presidential candidates). This approach brings an idiosyncratic element to the conclusions. It becomes difficult to determine whether the impact of non-verbal behaviour is due to the candidate's demeanour or due to the fact that it is displayed by a particular well-known politician. For example, did Reagan's positive body language in the 1984 debates induce a rating boost because positive body language helps ratings, because it helped Reagan in particular or because Reagan's opponent in the debate was Mondale. Our experiment used actors posing as candidates and thereby avoided the confounding effects of prior knowledge of the speaker. The second methodological limitation concerns the practice of isolating the visual aspect of the communication from the verbal one (for example, subjects may be asked to provide their impressions of a candidate based on an audio-only or a video-only version of the same speech). As Nagel and her colleagues (2012) note, this separation of the verbal and visual elements of communication is artificial since people's perceptions of the verbal content may differ depending on whether they see the speech or only hear it.

**The General Setup**
209 respondents were recruited from a large university via an email sent to all students enrolled in the summer session. To reduce the number of covariates influencing evaluations, only respondents whose party identification matched that of the candidate (Liberal) were selected to participate in the study. Participants were informed that they would be asked to take part in a study of campaign communication lasting about thirty minutes, during which they would be watching some materials on the computer while their vital signs were monitored by means of specialized equipment.

After coming to the lab, a trained assistant attached several devices to their body to monitor their skin conductance level (by means of electrodes placed on the tips of three of their fingers from their non-dominant hand) and their facial movements (by means of electrodes placed on their forehead and cheek). Participants were then seated at a computer wearing headphones as they watched the pre-recorded material on the screen.

The recording began with a 2 minute display of a gray screen followed by a short neutral video in order to enable the participants to become accustomed to the vital signs measurement devices. During the display of the gray screen, participants were instructed to try to relax as much as possible. After viewing the neutral video, participants again saw a gray screen for 45 seconds. This second rest period was used to derive a baseline response for their physiological reactions during the communication manipulation (we return to this point below).

The manipulation took the form of a webpage of a candidate in the recent federal election. Because the participants were attached to the physiological equipment, they watched the webpage in the form of a tutorial that navigated the various parts of the webpage for them. The manipulation contained two parts: the first part was common for everyone. It consisted of a snapshot of the candidate's experience and interest in the environment. The message listed the candidate's policy credentials and environmental initiatives. The second part contained the candidate's campaign speech on the topic of the environment (also linked from the webpage).

The speech varied by condition: some individuals saw a male candidate and others saw a female candidate[5]. There were four versions of the video for both candidates: the candidate was shown making either a strong or a weak speech in either a confident or non-confident manner.

The campaign message had been previously filmed during a single day for both candidates, using professional actors to play the part, and a professional crew to film and edit the video. The final campaign message was tailored to look very similar to campaign productions posted by Canadian candidates on Youtube.

**The Gender Manipulation**
Two professional actors were selected from a pool of available actors in the area on the basis of ratings of their pictures by 35 students who were blind to the purpose of the study. The actors were chosen on the basis of their estimated age and the similarity of their ratings of perceived attractiveness, competence, likeability and looking like a politician. Matching the actors on the basis of perceived attractiveness was particularly important. First, a number of studies have shown that when asked to make inferences based on a candidate's appearance, individuals' accuracy in identifying winners is better than chance (see, for example, Hall et al. 2012). Second, attractiveness appears to be related to both perceived electability and the stated likelihood of voting for a candidate (see, for example, Lewis and Brierly 1990; Sigelman et al. 1986; Sigelman, Sigelman and Fowler 1987). Attractive candidates receive more positive trait attributions. This applies to both communal and agentic traits. Schubert and Corran (2001) conclude that female candidates for political office are doubly disadvantaged if they are unattractive. Accordingly, we made special efforts to select actors who were of average attractiveness.

**The Verbal Manipulation**
Two speeches on environmental policies were created that varied in their strength and power. The full text for both is included in the Appendix. The *strong speech* used a number of rhetorical devices designed to convey strength and determination: it made use of the rule of three ("three priorities: to save jobs, to save on energy consumption, and to save on consumption of fossil fuels"); it stated clear positions ("I say we can and we will!"); and used repetitions to underscore the main arguments (the word "save" was repeated several times during the speech). The language was assertive: the word "I" in conjunction with direct voice in the present or future was used seven times (e.g. "I am your candidate"; "I know"; "I ask you"; "I pledge"; "I'll fight"). Moreover, phrases were quite short. The speech was further modeled according to Monroe's motivated sequence (Cook 1989): it identified a clear problem and proposed three clear solutions. In the *weak speech* the rule of three was not employed; instead, it was replaced by a long phrase ("there are several ways to improve our performance on these matters, like working harder to make our homes more energy efficient, reducing our consumption of fossil fuel, and

---

[5] Due to recruitment constraints, participants were recruited using identical procedures at different times. Participants recruited in summer 2011 viewed one of the four versions of the male candidate's speech, while participants recruited in summer 2012 viewed the parallel versions of the female candidate's speech. We have no reason to believe that the elapse of time was a confounding factor.

protecting jobs"). There were no repetitions and the main arguments were not clearly identified. The language was significantly more tentative: instead of direct appeals, the candidate stated their position several times using either the passive voice or conditional (e.g. "for me… is really important"; "I would like"). Assertive statements were nuanced by the use of "I believe" and "should". The weak speech also used more qualifiers (such as "really" and "absolutely"). The two speeches differed in their use of the passive voice (12% for the strong speech, and 21% for the weak speech).

Note that the content of both versions of the speech was restricted to policy positions and policy directions, information that is pertinent to perceptions of the candidates' competence and leadership ability. In addition, the speech made no reference to any opposing candidates in the constituency.

**The Nonverbal Manipulation**
 Each candidate used two different styles to deliver each of the two speeches. In the confident delivery condition, the candidate gave the speech while having a straight back and leaning slightly backwards, adopting an open body posture (with the hands relatively away from the body) and employing expressive and energetic gestures. This body posture was accompanied by expressive facial mimicry and more smiling. On the other hand, the non-confident delivery saw the candidate delivering the speech with their back slumped, their hands kept tight together in from of them, and without the use of gestures. There was little facial expressiveness in the non-confident delivery and less smiling. Moreover, the candidates used more inflected speech in the confident delivery than in the non-confident one.  These two styles mimic previous manipulations of strong and weak non-confident deliveries (see, for example, Awamleh and Gardner 1999; Marsh et al. 1997). These features were explained to the actors who took several takes in order to replicate the required behaviour. We retained the best take for each actor.

**Dependent variables**
After watching the speech, participants evaluated the candidate on several dimensions. They were asked to assess "how well does each of the following terms describe the candidate whose speech you have just watched?" The terms were all agentic: competence, intelligence, and leadership. Participants provided ratings on a 5-point scale ranging from "not well at all" to "extremely well".

**Physiological measurements**
Physiological responses were captured using a ProComp Infiniti encoder from Thought Technology Ltd., and purpose-built software designed at the Centre for Interdisciplinary Research in Music Media and Technology (CIRMMT) at McGill University.[6] Skin conductance (SC), reflecting the level of moisture exuded by the ecrine sweat glands, was captured by passing an infinitesimally small electrical current through a pair of electrodes on the surface of the skin — in this case, electrodes attached to the tips of the distal phalanx (outer segment) of the index and ring fingers, captured using Thought Technology's SC-Flex/Pro sensor. The current was held constant, and the electrodes monitored variations in current flow. More moisture (sweat)

---

[6] This discussion of GSL draws on Soroka and McAdams (2010).

leads to less resistance, or, conversely, more conductance. The resulting conductance data can be used to look at skin conductance levels (SCL), measured in microSiemens.

Skin conductance has been identified in numerous studies as a measure of respondents' level of arousal to various stimuli (for an overall review, see, example, Larsen et al. 2008; Simons et al. 1999; Lang et al. 1999; Bolls et al. 2001; Ravaja 2004). Skin conductance is one of the components of the orienting system and it varies with exposure to novel or unexpected stimuli (Dawson et al. 2007). Note that arousal is not the same thing as valence: arousal refers only to the degree of activation, not to the direction (positive or negative, pleasant or unpleasant) of the reaction (Larsen and Diener 1992; Russell 1980).

In this paper we examine the evolution of individuals' skin conductance levels during the candidate's performance. The signal is averaged over 1 second to minimize the noise. In addition, we subtract from each signal we observe during the manipulation the average level of arousal observed over 30 seconds during the baseline period preceding the visualisation of the webpage. During this time participants saw a gray screen and were instructed to try to relax as much as possible. The average arousal level observed during this period provides a baseline for each individual's natural arousal level when resting.

**Manipulation checks**
After rating the candidate's traits and election prospects, individuals rated the candidate's delivery on two different scales: confident versus not-confident and enthusiastic versus unenthusiastic. As expected, an ANOVA analysis showed only significant nonverbal effects. Confidence was rated higher in the conditions where this behaviour had been positively manipulated (F=121.23, p= 0.00 for the two candidates considered together; F=70.92, p = 0.00, for the female candidate; and F=49.34, p= 0.00 for the male candidate). Perceived enthusiasm was also higher in these conditions (F=82.72 for the two candidates; F= 31.93 for the female candidate; and F= 53.17 for the male candidate). A Candidate Gender x Verbal x Nonverbal full factorial ANOVA model revealed no effects in perceived confidence due to the candidate's gender (the highest F statistic for any direct or indirect effect of the candidate's gender on the perceptions of confidence was 1.80, p=.18; all other F-statistics were below 0.5). These results confirm that the nonverbal manipulation was successful, and that both our speakers emanated similar levels of confidence or non-confidence.

To further validate our manipulation we recruited an outside sample of 139 students from the same university to look at silent clips taken from the videos and to rate the two speeches. The students were blind to the manipulation. 103 evaluated the male candidate clips and 36 evaluated the female clips. Within each candidate, the clips in the nonverbal confident conditions were rated as expressing significantly more confidence than the clips taken from the non-confident conditions (lowest t=2.28, p=0.07 two-tailed). The students also read the two speeches. Their evaluations again confirmed our expectations: the strong speech was rated as significantly more persuasive, better structured, and using better arguments than the weak speech, irrespective of whether the politician was identified as a man (Bill Johnson) or as a woman (Mary Johnson) (lowest t=1.83, p=0.07).

Finally, we digitally analyzed one of the candidate's nonverbal behavior in each of the four

conditions. We selected a large number of frames (N=275) taken randomly every 2-seconds from each condition and used static frame analysis software (Dumitrescu, 2010) to obtain pixel-level estimates of the positioning of the candidate's head and hands in the frames. As expected, the politician held his head up higher (F=649.36, p=0.00), opened up his body more (F=85.46, p=0.00) and gesticulated significantly more in the confident conditions than in the non-confident ones. Figure 1 tracks the candidate's hand movements when confident and when lacking confidence. Larger numbers indicate greater distances between the hands, and negative numbers indicate hand crossing, illustrating visually that the candidates' hands were both further apart and more animated in the confident conditions than in the non-confident ones. In future versions of this paper we will repeat this analysis for the speeches of the female candidate.

[Figure 1 here]

**Sample checks**
The study assistant was instructed to determine whether participants had suspected that the candidate was not a real one. After viewing the video participants also reported whether they thought the individual they had seen belonged or not in politics. Eleven individuals made spoken or written comments suggesting that they knew the candidate was not real. We also asked participants if they had been able to read the candidate's message at the beginning, if they had any trouble hearing the candidate, or if they felt uncomfortable during the experiment because of the physiological equipment. Furthermore, we also checked individuals' physiological responses during the rest (baseline) period, and identified 14 outliers. Following these checks, we removed 20 respondents from the analysis. All the analyses were therefore performed on N=189. The valid sample was balanced in terms of gender (53% females). No significant differences in gender attitudes, environmental attitudes or strength of Liberal party identification were observed across conditions (highest F= 2.05, p=0.15).

**Results**

Figure 2 presents the effects of the verbal and nonverbal manipulation for evaluations of strong leadership, intelligence and competence for both candidates combined. For the sake of interpretability the marginal effects are graphed for all four conditions. The detailed ANOVA results may be found in the appendix.

Overall, no verbal effects were observed for any of the three dependent variables; nonverbal effects clearly dominate. That is, both candidates elicited much better evaluations when their delivery was confident than when it was not, but giving a strong speech did not result in better evaluations than a weak speech. Thus, H1 is confirmed but H2 is disconfirmed. The non-verbal effect is strongest for leadership and slightly weaker for intelligence and competence. As predicted by H3, significantly higher ratings were given to the candidates in the confident/strong verbal conditions (see Figures 2a, b, c). The combination of confidence and a strong speech increases their ratings by an average of nearly one full point on the 1 to 5 scale compared to the worst condition (a weak speech with a non-confident delivery). These results are consistent with much of the literature on non-verbal behavior and suggest that a confident demeanour is extremely important for evaluations of political candidates, no matter their gender. While we cannot be certain that our verbal and non-verbal manipulations are equally strong, it is telling

that there is simply no verbal effect for the two candidates.

[Figures 2a, b, c here]

The effects of candidate gender are presented in Figure 3. There are surprisingly few gender differences (see Appendix Table 2A). Indeed, overall evaluations of the female candidate are very similar to those of the male candidate, with the exception of leadership skills, where the female candidate actually scores significantly higher than her male counterpart, contrary to H4. However, the marginal effects depicted in Figure 3 tell us a fuller story. While the female candidate receives slightly higher evaluations in the two confident conditions, she receives more punishment than the male candidate in the non-confident conditions. This is particularly true for the combination of low confidence and a strong speech. Overall, Figure 3 shows that the evaluations of the male candidate do not vary much by condition, whereas the female candidate receives lower evaluations for a non-confident demeanour. This effect is particularly evident when comparing the female confident and non-confident conditions (results not shown, they vary from 0.6 to 0.9 on the 1-5 scale), but they also play a role when directly compared to the male evaluations, particularly for intelligence ($F=4$, $p=.05$) and competence ($F=2.75$, $p=.10$). These results lend support to H5, but not to H5a.

[Figure 3 here]

In Figure 4 the results from the physiological analysis are depicted. The question is not only whether we can detect different levels of arousal in reaction to the speeches given by the male and female candidate, but also whether the arousal development differs during the actual speech. Tracking the dynamics of arousal can indicate whether there are gendered reactions to the different speech scenarios. These physiological responses might work as mediating mechanisms between the gender/treatment and the evaluations of the candidates.

The first graph in Figure 4 shows the overall arousal levels by gender across the entire time of the speech. The results confirm our earlier findings and show higher levels of arousal for the female candidate during the confident conditions and slightly lower levels during the non-confident conditions. These effects are magnified during the very first part of the speech, where we find the highest level of arousal differences between the female and male candidates in some of the conditions. In particular, arousal levels drop significantly in the non-confident/strong verbal condition for the female candidate, whereas they are relatively high during her confident conditions and low confidence/weak speech conditions. The male candidate elicits different levels of arousal, which are strongest in both non-confident conditions, but very low in the high confidence/weak speech condition. So, both the female and male candidate get "punished" with low arousal indicating low attention in two different conditions: the female in the low confidence/strong verbal and the male in the high confidence/weak verbal condition. This partly confirms H5. We should remind the reader that arousal does not measure the direction of the effect, that is, whether it is a positive or negative reaction. Instead, low arousal shows that there is little engagement with the speech. The findings are similar but less significant for the remaining parts of the speeches (see Figure 4).

[Figure 4]

While we do not know the exact reasons for the differences in arousal and attention, the results for arousal are somewhat be in line with the evaluations. While the female received more arousal and better evaluations for her confident performances, there is a drop in arousal (and evaluation) for the low confidence condition. There is an exception in the low confidence/weak speech condition which elicits high arousal, and although we do not know exactly why, it seems likely for negative reasons. The male candidate does not elicit much arousal and less positive evaluations in the high confidence conditions; rather the arousal is high in the low-confidence conditions, where his evaluations are not as good.

## Conclusion

We have conducted a multifaceted experiment in order to understand better the different reactions to various political speeches, delivered by a male and female candidate. The study contributes to the literature in that it extends the study of non-verbal effects beyond facial expressions to fuller expressions. It also contributes to the literature on leadership in the political realm in that it uses an experimental but authentic scenario. Most importantly, our study speaks to the gender literature on political leadership.

Our experimental design enabled us to explore a number of important questions. We found first of all that for both of the candidates, the way in which a message is delivered seems more important than the message itself. A confident demeanor usually elicits more positive evaluations as well as more attention; whereas the verbal quality of the speech did not matter as much. However, the combination of a well-structured speech delivered with confidence led to the highest evaluations, though the level of arousal and thus attention was highest in the low confidence/weak speech condition.

Most importantly, our research attempted to understand whether these patterns are different for male versus female candidates. Both male and female politicians seem to be evaluated similarly for equivalent performances in the high confidence/strong speech condition (with a slight advantage for the female), and in the low confidence/weak speech condition. However, female leaders seem to pay a higher price for low confidence performances, especially when combined with strong speech. This pattern was also reflected in the physiological reactions. On the other hand, male politicians seem to pay a high price for the combination of high confidence/weak speech, at least in terms of leadership ratings and arousal levels—yet this effect was not as pronounced and universal as the non-verbal gendered effect for the female.

So overall, our research indicates that a female candidate who lacks "competent assertiveness" may indeed pay a slightly higher price than her male counterpart for failing to exhibit agentic behavior. These findings confirm some insights of recent studies by Bongiorno, Bain and David (2013) who argue that female leaders who exhibit non-agentic behaviour can expect to be penalized, whereas male leaders receive the benefit of the doubt in similar conditions. However our results did not confirm that compounding non-confident and poorly structured verbal speech results in more punishment for a female candidate, so there are limits to the punishment as well.

There are not just theoretical implications of our findings. Male and female political leaders cannot just focus on their ideas, but they have to think about how they deliver them. The quality of the delivery can override the quality of the speech, and that is particularly true for women.

Figure 1: Frame analysis. Top graph: Example of a frame. Bottom graph: Candidate's hand movements during the webcast, by condition.
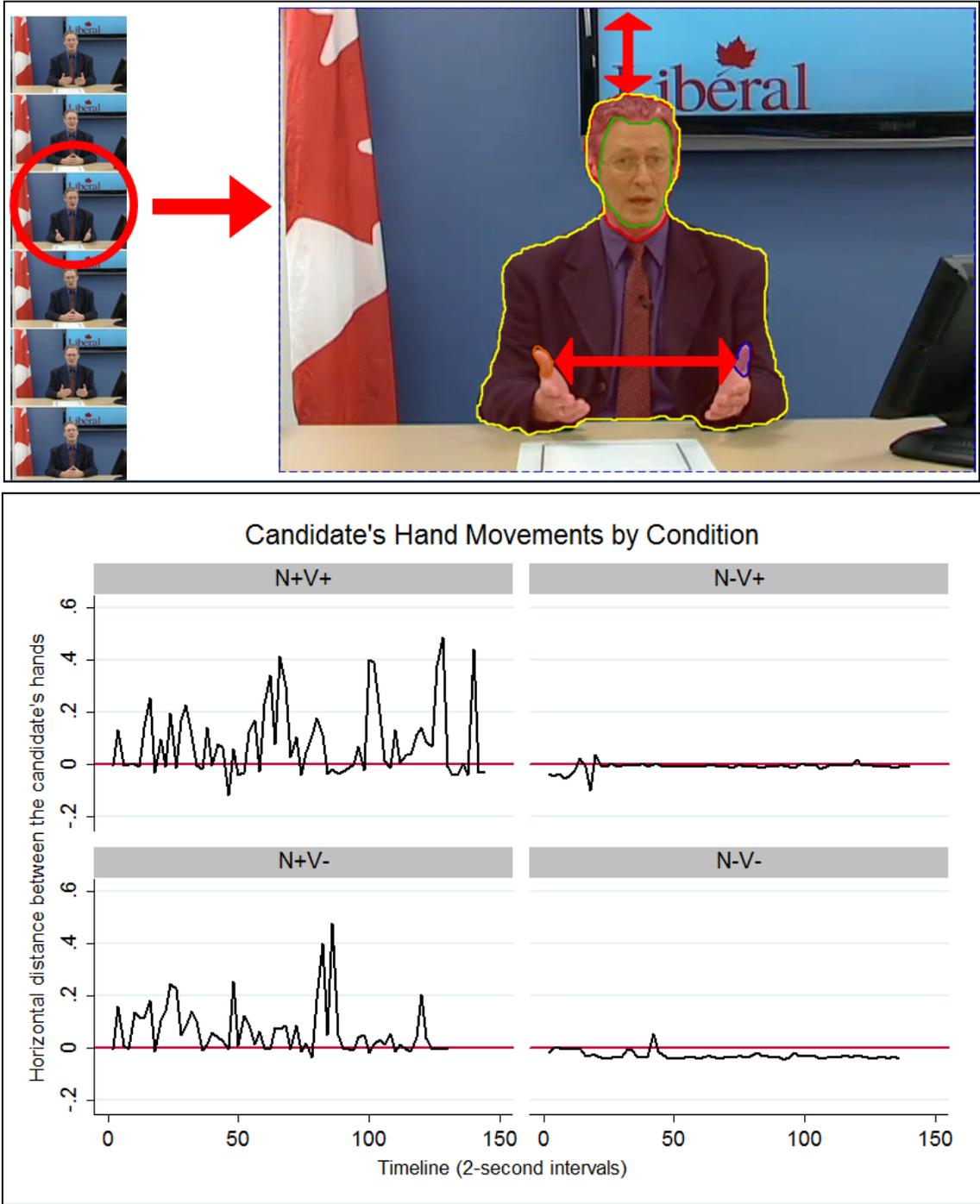
Figure 2a: Marginal effects of communication treatment on perceptions of both candidates' leadership
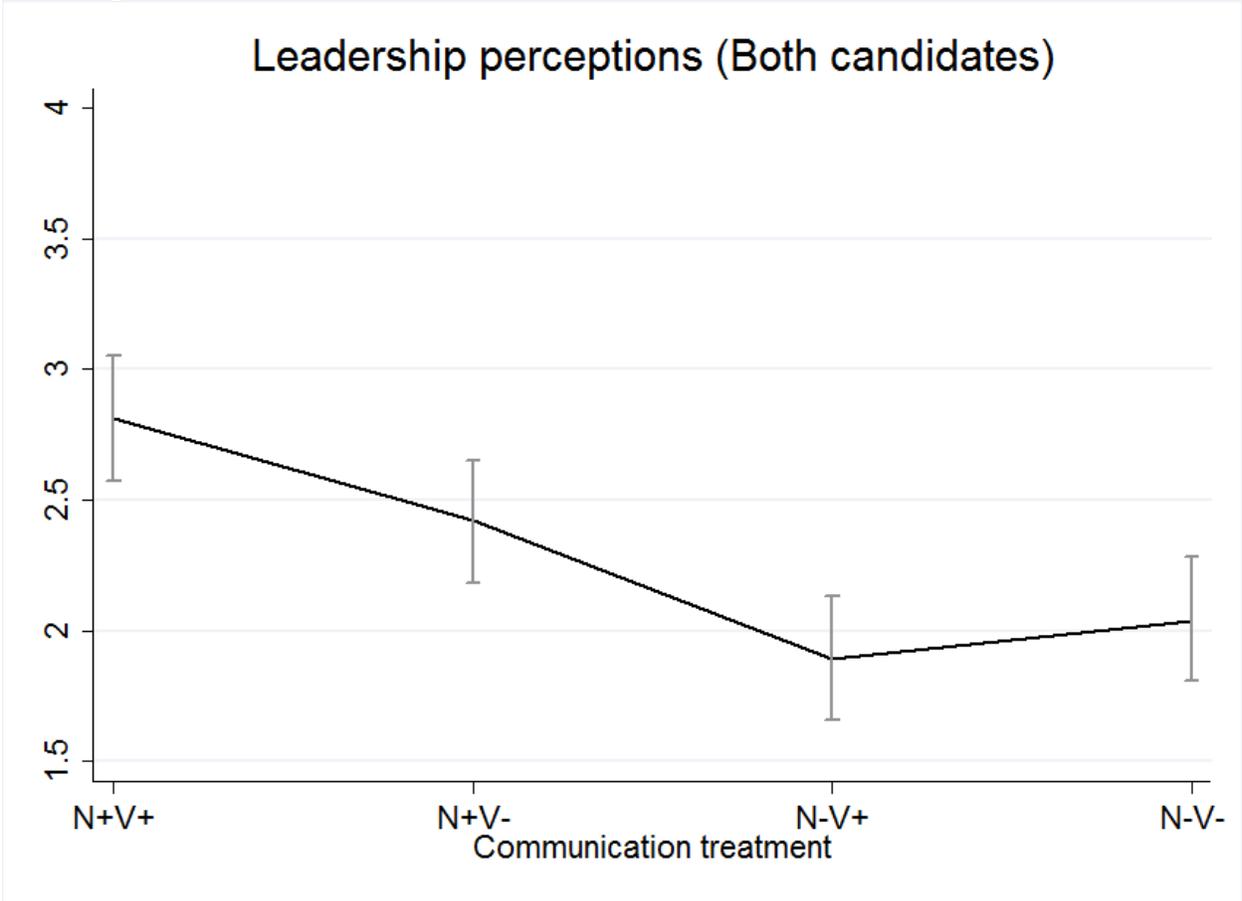


N+V+:  strong speech, confident delivery
N+V-:  weak speech, confident delivery
N-V+:  strong speech, non-confident delivery
N-V-:  weak speech, non-confident delivery

Figure 2b: Marginal effects of communication treatment on perceptions of both candidates'
Intelligence
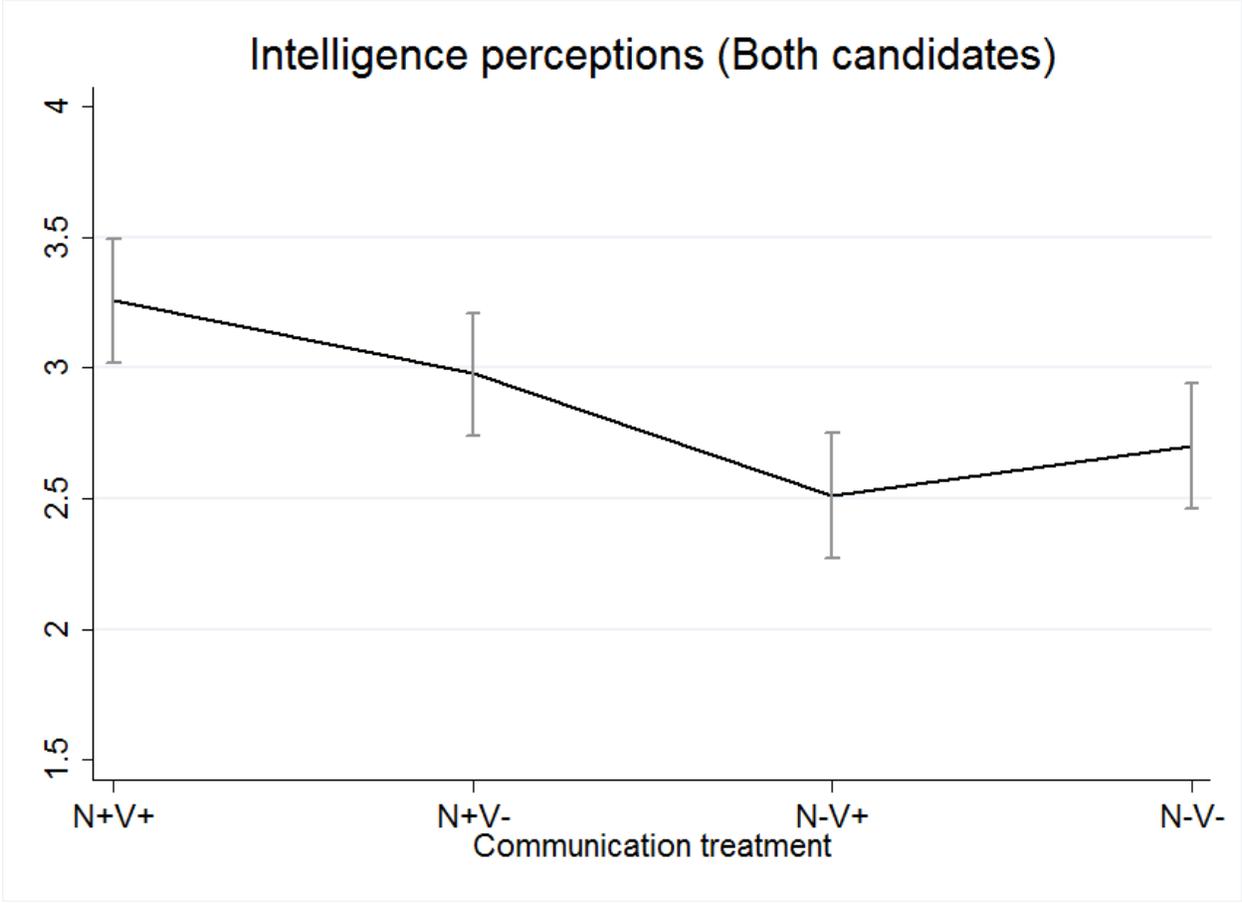


N+V+: strong speech, confident delivery
N+V-: weak speech, confident delivery
N-V+: strong speech, non-confident delivery
N-V-: weak speech, non-confident delivery

Figure 2c: Marginal effects of communication treatment on perceptions of both candidates' competence
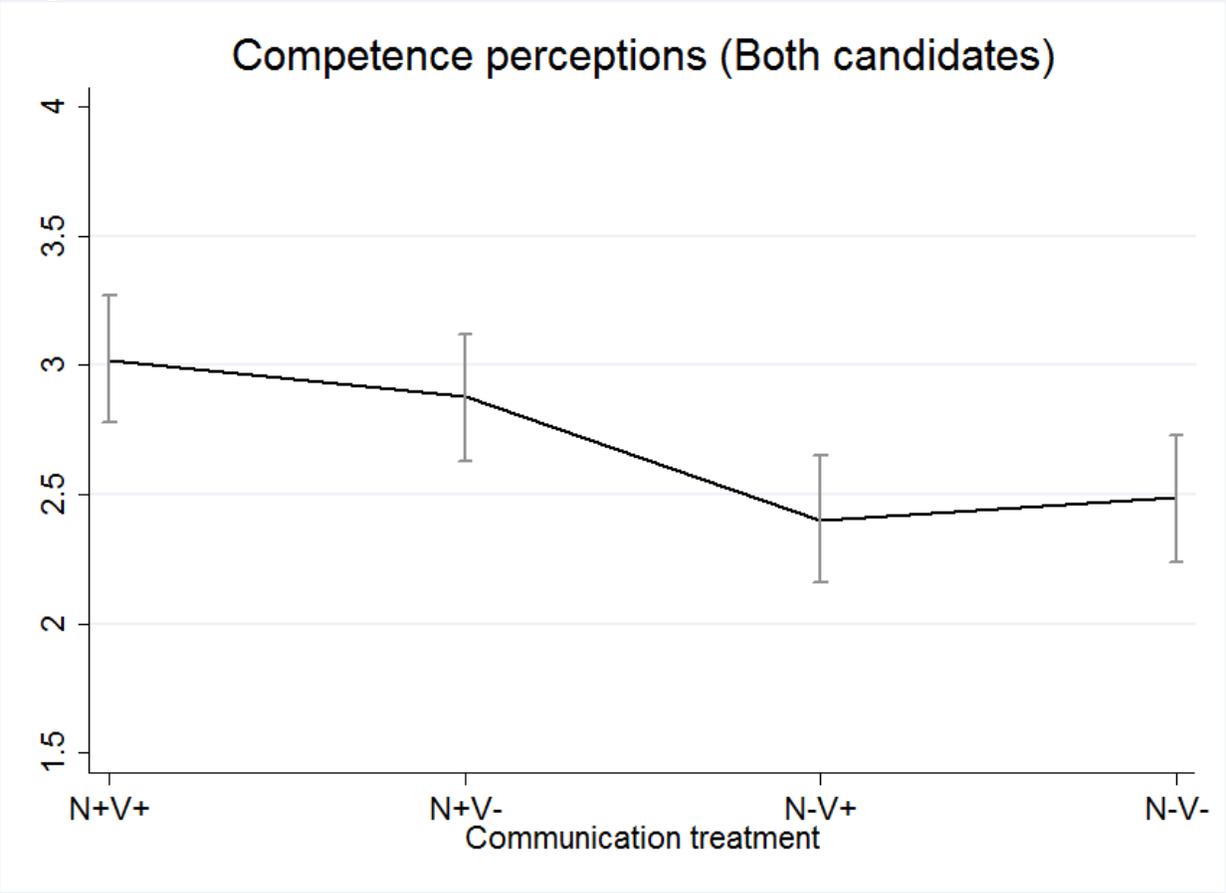


N+V+: strong speech, confident delivery
N+V-: weak speech, confident delivery
N-V+: strong speech, non-confident delivery
N-V-: weak speech, non-confident delivery

Figure 3: Marginal effects of candidate gender and communication treatment on perceptions of the candidate's leadership, intelligence and competence.
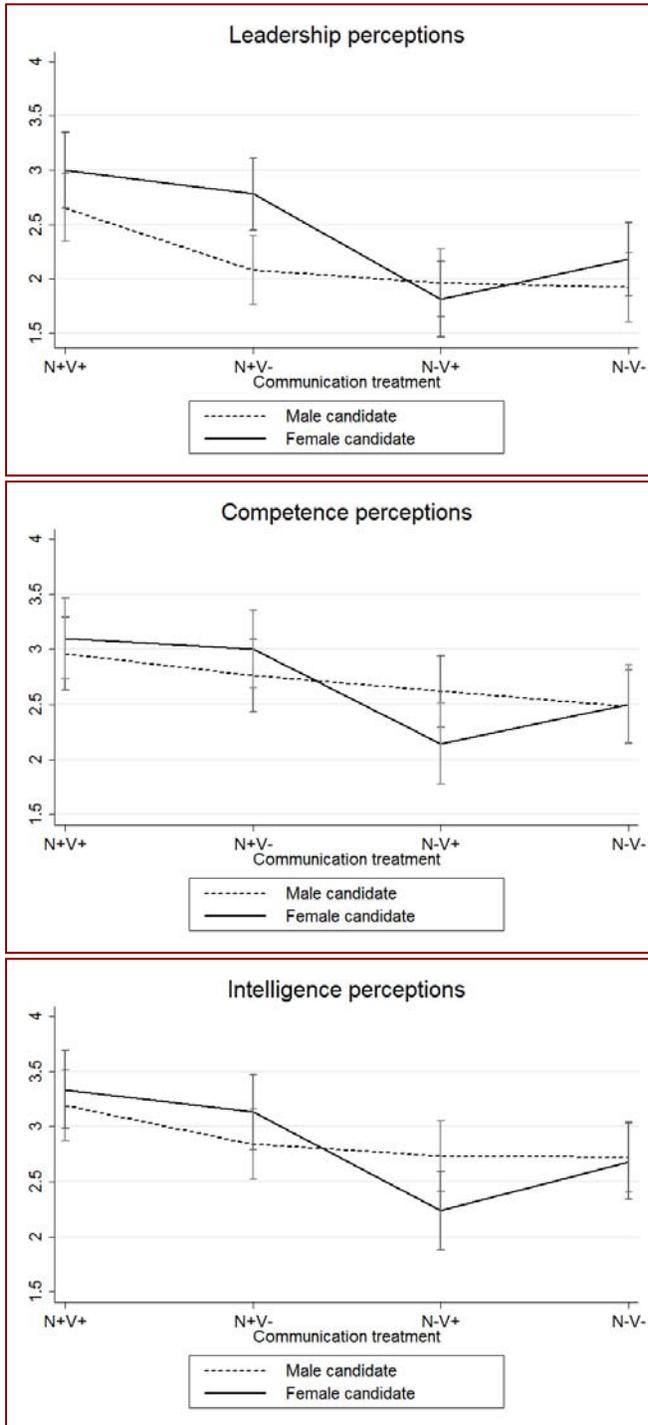


N+V+: strong speech, confident delivery
N+V-:  weak speech, confident delivery
N-V+:  strong speech, non-confident delivery

N-V-:   weak speech, non-confident delivery

Figure 4: Skin conductance levels by candidate gender and communication treatment. Top panel: full speech. Second and third row: speech moments, in order of occurrence.
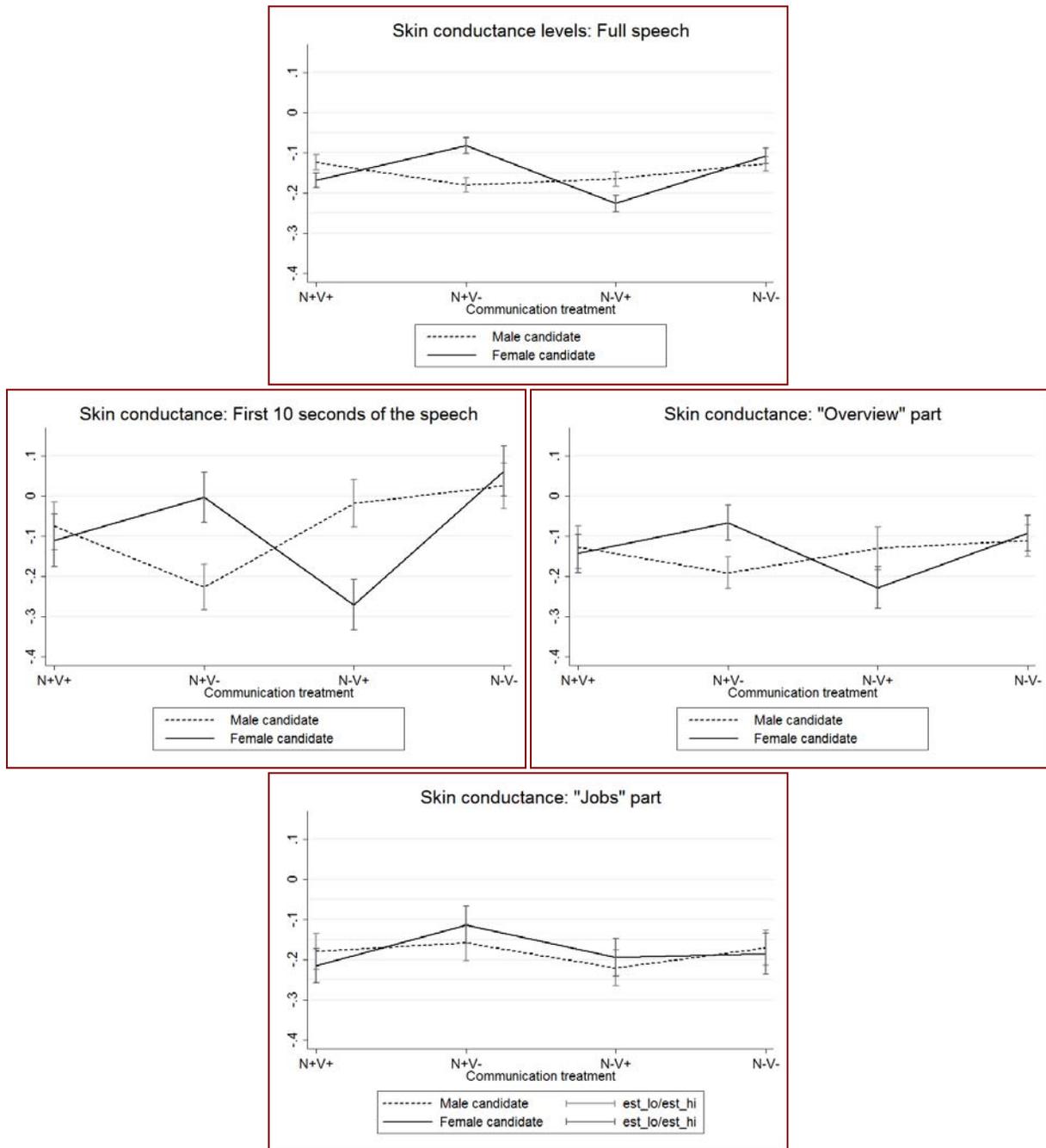


N+V+: strong speech, confident delivery
N+V-:  weak speech, confident delivery
N-V+:  strong speech, non-confident delivery
N-V-:   weak speech, non-confident delivery
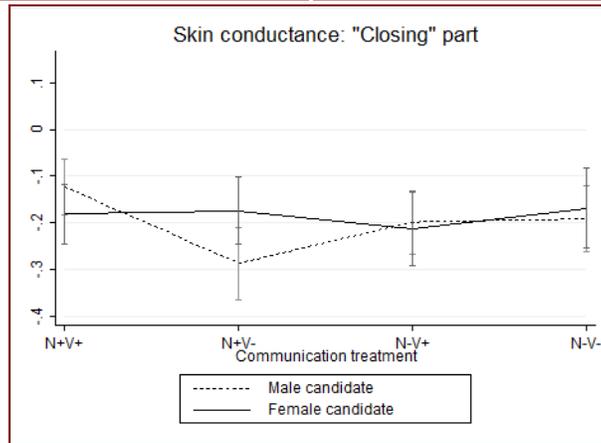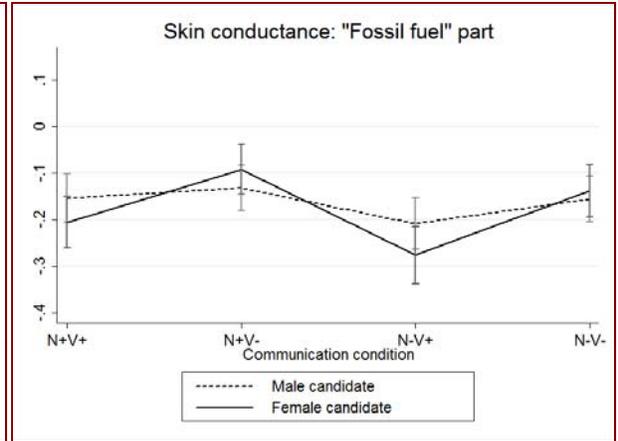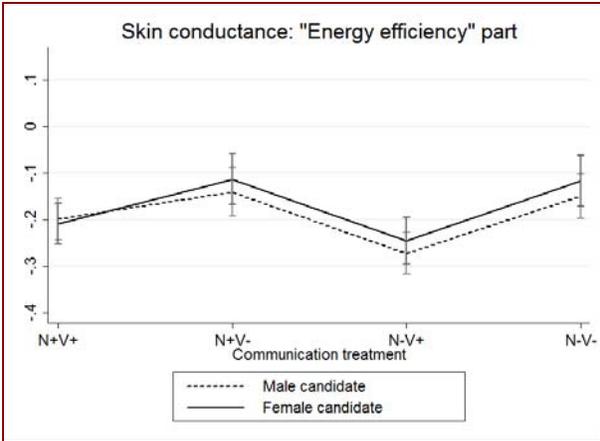
N+V+: strong speech, confident delivery
N+V-:  weak speech, confident delivery
N-V+:  strong speech, non-confident delivery
N-V-:   weal speech, non-confident delivery

Appendix: Detailed Results

| Table A1. Communication effects on perceptions of the candidate's strong leadership, intelligence and competence | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Strong leader | | | | Intelligent | | | | Competence | | |
| | Partial SS | df | F | P > F | Partial SS | df | F | P > F | Partial SS | df | F | P > F |
| **Model** | **23.67** | **3** | **7.89** | **0.00** | **14.93** | **3** | **4.98** | **0.00** | **12.52** | **3** | **5.71** | **0.00** |
| **Verbal** | 0.70 | 1 | 0.70 | 0.32 | 0.08 | 1 | 0.08 | 0.73 | 0.04 | 1 | 0.06 | 0.81 |
| **Nonverbal** | **19.63** | **1** | **19.63** | **0.00** | **12.33** | **1** | **12.33** | **0.00** | **11.87** | **1** | **16.24** | **0.00** |
| **Verbal* Nonverbal** | **3.45** | **1** | **3.45** | **0.03** | **2.58** | **1** | **2.58** | **0.05** | 0.63 | 1 | 0.86 | 0.35 |
| **Residual** | 127.33 | 185 | 0.69 | | 127.49 | 185 | 0.69 | | 135.29 | 185 | | |
| **Total** | 150.99 | 188 | 0.80 | | 142.42 | 188 | 0.76 | | 147.81 | 188 | | |

**Estimated marginal effects**

| | Margin | 95% Conf. Interval | | Margin | 95% Conf. Interval | | Margin | 95% Conf. Interval | |
|---|---|---|---|---|---|---|---|---|---|
| **Condition** | | | | | | | | | |
| **N-V-** | 2.04 | 1.81 | 2.28 | 2.70 | 2.46 | 2.94 | 2.49 | 2.24 | 2.73 |
| **N+V-** | 2.42 | 2.18 | 2.65 | 2.98 | 2.74 | 3.21 | 2.88 | 2.63 | 3.12 |
| **N-V+** | 1.89 | 1.66 | 2.13 | 2.51 | 2.27 | 2.75 | 2.40 | 2.16 | 2.65 |
| **N+V+** | 2.81 | 2.57 | 3.05 | 3.26 | 3.02 | 3.49 | 3.02 | 2.78 | 3.27 |
| | | | | | | | | | |
| R-squared | | 189 | | | 189 | | | 189 | |
| N | | 0.16 | | | 0.10 | | | 0.08 | |

Note: One-way ANOVA between-subjects results.

| Table A2. Gender and communication effects on perceptions of the candidate's strong leadership, intelligence and competence | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | Strong leader perceptions | | | | Intelligence perceptions | | | | Competence perceptions | | | |
| | Partial SS | df | F | P > F | Partial SS | df | F | P > F | Partial SS | df | F | P > F |
| Model | **32.04** | **7** | **6.97** | **0.00** | 19.01 | 7 | 3.98 | 0 | 16.01 | 7 | 3.14 | 0.00 |
| Female candidate | 3.94 | 1 | 5.99 | **0.02** | 0.03 | 1 | 0.04 | 0.84 | 0.02 | 1 | 0.03 | 0.87 |
| Verbal | 0.62 | 1 | 0.95 | 0.33 | 0.04 | 1 | 0.06 | 0.80 | 0.02 | 1 | 0.02 | 0.88 |
| Female candidate * Verbal | 1.74 | 1 | 2.65 | 0.11 | 1.07 | 1 | 1.57 | 0.21 | 1.05 | 1 | 1.44 | 0.23 |
| Nonverbal | **20.49** | **1** | **31.18** | **0.00** | 13.24 | 1 | 19.42 | 0.00 | 12.67 | 1 | 17.39 | 0.00 |
| Female candidate * Nonverbal | **2.58** | **1** | **3.93** | **0.05** | 2.71 | 1 | 3.98 | 0.05 | 2.00 | 1 | 2.75 | 0.10 |
| Verbal * Nonverbal | **3.69** | **1** | **5.62** | **0.02** | 2.86 | 1 | 4.20 | 0.04 | 0.79 | 1 | 1.08 | 0.30 |
| Female candidate * Verbal * Nonverbal | 0.01 | 1 | 0.01 | 0.90 | 0.27 | 1 | 0.40 | 0.53 | 0.44 | 1 | 0.60 | 0.44 |
| Residual | 118.95 | 181 | | | 123.41 | 181 | | | 131.8 | 181 | | |
| Total | 150.99 | 188 | | | 142.42 | 188 | | | 147.81 | 188 | | |

**Estimated marginal effects**

| Condition | Margin | 95% Conf. Interval | | Margin | 95% Conf. Interval | | Margin | 95% Conf. Interval | |
|---|---|---|---|---|---|---|---|---|---|
| Male candidate | | | | | | | | | |
| N-V- | 1.92 | 1.60 | 2.24 | 2.72 | 2.40 | 3.04 | 2.48 | 2.15 | 2.81 |
| N+V- | 2.08 | 1.76 | 2.40 | 2.84 | 2.52 | 3.16 | 2.76 | 2.43 | 3.09 |
| N-V+ | 1.96 | 1.65 | 2.27 | 2.73 | 2.41 | 3.05 | 2.62 | 2.29 | 2.94 |
| N+V+ | 2.65 | 2.34 | 2.97 | 3.19 | 2.87 | 3.51 | 2.96 | 2.63 | 3.29 |
| Female candidate | | | | | | | | | |
| N-V- | 2.18 | 1.84 | 2.52 | 2.68 | 2.34 | 3.03 | 2.50 | 2.14 | 2.86 |
| N+V- | 2.78 | 2.45 | 3.11 | 3.13 | 2.79 | 3.47 | 3.00 | 2.65 | 3.35 |
| N-V+ | 1.81 | 1.46 | 2.16 | 2.24 | 1.88 | 2.59 | 2.14 | 1.78 | 2.51 |

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **N+V+** | | 3.00 | 2.65 | 3.35 | | 3.33 | 2.98 | 3.69 | 3.10 | 2.73 | 3.46 |
| | R-squared | 0.21 | | | | 0.13 | | | 0.11 | | |
| | N | 189 | | | | 189 | | | 189 | | |

Note: One-way ANOVA between-subjects results.

| Table A3. Communication and gender treatment effects on skin conductance levels during the full speech, the first 10 seconds only, and the overview part only | | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | **Full speech** | | | | **First 10 seconds** | | | | **Overview** | | | |
| Source | Partial SS | df | F | P>F | Partial SS | df | F | P>F | Partial SS | df | F | Prob > F |
| Model | **739.27** | **8** | **305.11** | **0.00** | **190.66** | **8** | **116.90** | **0.00** | **326.63** | **8** | **133.48** | **0.00** |
| Prior skin conductance level | **626.11** | **1** | **2067.25** | **0.00** | **144.21** | **1** | **707.33** | **0.00** | **288.77** | **1** | **944.03** | **0.00** |
| Female candidate | 0.05 | 1 | 0.16 | 0.69 | 0.02 | 1 | 0.09 | 0.76 | 0.05 | 1 | 0.18 | 0.67 |
| Verbal | **14.10** | **1** | **46.55** | **0.00** | **2.98** | **1** | **14.63** | **0.00** | **1.87** | **1** | **6.12** | **0.01** |
| Female candidate * Verbal | **20.04** | **1** | **66.17** | **0.00** | **8.09** | **1** | **39.67** | **0.00** | **4.42** | **1** | **14.45** | **0.00** |
| Nonverbal | **1.95** | **1** | **6.45** | **0.01** | **1.20** | **1** | **5.86** | **0.02** | 0.07 | 1 | 0.24 | 0.62 |
| Female candidate * Nonverbal | **3.70** | **1** | **12.21** | **0.00** | **4.38** | **1** | **21.47** | **0.00** | **2.37** | **1** | **7.73** | **0.01** |
| Verbal * Nonverbal | **6.63** | **1** | **21.89** | **0.00** | **4.74** | **1** | **23.26** | **0.00** | **1.37** | **1** | **4.46** | **0.03** |
| Female candidate * Verbal * Nonverbal | **1.48** | **1** | **4.90** | **0.03** | 0.02 | 1 | 0.12 | 0.73 | 0.04 | 1 | 0.12 | 0.72 |
| Residual | 7952.15 | 26256 | | | 354.55 | 1739 | | | 1405.26 | 4594 | | |
| Total | 8691.42 | 26264 | | | 545.21 | 1747 | | | 1731.89 | 4602 | | |
| Number of | 26265 | | | | | 1748 | | | 4603 | | | |
| R-squared | 0.0851 | | | | | 0.35 | | | 0.1886 | | | |

Note: One-way ANOVA between subjects results.    The prior skin conductance level is the average skin conductance observed during the biography part of the video stimulus.

Table A4. Communication and gender treatment effects on skin conductance levels during the three arguments and the c

| | Jobs argument | | | | Home energy argument | | | | Fuel argument | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Partial SS | df | F | Prob > F | Partial SS | df | F | Prob > F | Partial SS | df | F | Prob > F |
| Model | **80.45** | **8** | **34.82** | **0.00** | **129.80** | **8** | **53.14** | **0.00** | **14.97** | **8** | **5.27** | **0.00** |
| Prior skin conductance level | **72.36** | **1** | **250.60** | **0.00** | **111.42** | **1** | **364.94** | **0.00** | **5.07** | **1** | **14.27** | **0.00** |
| Female candidate | 0.03 | 1 | 0.11 | 0.74 | 0.34 | 1 | 1.12 | 0.29 | 0.21 | 1 | 0.58 | 0.45 |
| Verbal | **2.27** | **1** | **7.85** | **0.01** | **10.16** | **1** | **33.27** | **0.00** | **6.28** | **1** | **17.70** | **0.00** |
| Femal candidate * Verbal | 0.10 | 1 | 0.35 | 0.56 | 0.11 | 1 | 0.38 | 0.54 | **1.83** | **1** | **5.17** | **0.02** |
| Nonverbal | 0.72 | 1 | 2.50 | 0.11 | **0.91** | **1** | **2.99** | **0.08** | **2.17** | **1** | **6.12** | **0.01** |
| Female candidate * Nonverbal | 0.00 | 1 | 0.00 | 0.96 | 0.11 | 1 | 0.35 | 0.56 | 0.09 | 1 | 0.25 | 0.62 |
| Verbal * Nonverbal | 0.28 | 1 | 0.98 | 0.32 | 0.60 | 1 | 1.95 | 0.16 | 0.18 | 1 | 0.50 | **0.48** |
| Female candidate * Verbal * Nonverbal | **1.01** | **1** | **3.49** | **0.06** | 0.06 | 1 | 0.21 | 0.65 | 0.00 | 1 | 0.00 | 0.95 |
| Residual | 1271.99 | 4405 | | | 1240.18 | 4062 | | | 1361.55 | 3835 | | |
| Total | 1352.43 | 4413 | | | 1369.97 | 4070 | | | 1376.52 | 3843 | | |
| Number of | 4414 | | | | 4071 | | | | 3844 | | | |
| R-squared | 0.0595 | | | | 0.0947 | | | | 0.0109 | | | |

Note: One-way ANOVA between subjects results. ¨The prior skin conductance level is the average skin conductane observed during the biogra

**References**

Aguinis, Herman, and Christine A. Henle (2001) Effects of Nonverbal Behavior on Perceptions of a Female Employee's Power Bases. *The Journal of Social Psychology* 141(4): 537-549

Aguinis H., M. M. Simonsen and C. A. Pierce (1998) Effects of nonverbal behavior on perceptions of power bases. *Journal of Social Psychology* 138(4): 455-69.

Awamleh, Raed (2003) Towards A Model of Charismatic Non-Verbal Impression Management. *International Business and Economics Research Journal* 2(7).

Awamleh, Raed, and William L. Garner (1999) Perceptions of Leader Charisma and Effectiveness: The Effects of Vision Content, Delivery and Organizational Performance. *The Leadership Quarterly* 10(3): 345-73.

Bernstein, Arla G. (2000) The Effects of Message Theme, Policy Explicitness and Candidate Gender. *Communication Quarterly* 48(2): 159-73.

Bongiorno, Renata, Paul Bain and Barbara David (2013) If you're going to be a leader, at least act like it! Prejudice towards women who are tentative in leader roles. *British Journal of Social Psychology*. online only.

Braun, Susanne, Claudia Peus and Dieter Frey. 2012. "Is Beauty Beastly?" *Zeitschrift für Psychologie* 220: 98-108.

Brody, L.R. and Hall, J.A. 1993. *Gender and Emotion*. In M. Lewis and J. Haviland, eds., *Handbook of emotion* (pp. 447-461). New York, NY: Guilford.

Brunswik, E. (1956). Perception and the representative design of psychological experiments. Berkeley, CA: University of California Press.

Bucy, Erik P. (2000) Emotional and Evaluative Consequences of Inappropriate Leader Displays *Communication Research* 27: 194-226, doi:10.1177/0093650000027002004

Bucy, E. P., and S. D. Bradley (2004) Presidential expressions and viewer emotion: Counterempathic responses to televised leader displays. *Social Science Information* 43(1): 59-94.

Butler, Doré and Florence Geis. 1990. "Nonverbal affect responses to male and female leaders: Implications for leadership evaluations." *Journal of Personality and Social Psychology* 58(1): 48-59.

Cacioppo, J.T., and R. E. Petty (1981) Electromyograms as measures of extent and affectivity of information processing. *American Psychologist* 36: 441-456.

Carli, L. L. (1990) Gender, language, and influence. *Journal of Personality and Social*

*Psychology* 59: 941–951.  doi:10.1037/0022-3514.59.5.941

Carli, L. L., S. J. LaFleur and C. C. Loeber (1995) Nonverbal behavior, gender, and influence. *Journal of Personality and Social Psychology* 68(6): 1030-1041. doi:10.1037/0022-3514.68.6.1030

Carney, D. R., J. A. Hall and L. S. LeBeau (2005) Beliefs about the nonverbal expression of social power. *Journal of Nonverbal Behavior* 29: 105–123.

Chiao J. Y., N. E. Bowman and H. Gill (2008) The Political Gender Gap: Gender Bias in Facial Inferences that Predict Voting Behavior. PLoS ONE 3(10): e3666. doi:10.1371/journal.pone.0003666

Cialdini, Robert B., and Melanie R. Trost (1998) Social influence: Social norms, conformity and compliance. In Daniel T. Gilbert, Susan T. Fiske, and Gardner Lindzey, eds., *The Handbook of Social Psychology* Volumes 1 and 2 (4th ed.), pp. 151-192. New York: McGraw-Hill.

Devine, P. G. (1989). Stereotypes and prejudice: Their automatic and controlled components. *Journal of Personality and Social Psychology* 56: 5–18.

Diekman, Amanda B., and Alice H. Eagly (2000) Stereotypes as Dynamic Constructs: Women and Men of the Past, Present, and Future. *Pers Soc Psychol Bull* 26(10): 1171-1188.

Dimberg, U. (1988a) Facial electromyography and the experience of emotion. Journal of *Psychophysiology* 3: 277-282.

Dimberg, U., and L. O. Lundquist (1990) Gender differences in facial reactions to facial expressions. *Biological Psychology* 30(2): 151-159.

Dimberg, Ulf, and Arne Öhman (1996) Behold the wrath: Psychophysiological responses to facial stimuli. *Motivation and Emotion* 20(2): 149-182.

Dolan, Kathleen (2010) The Impact of Gender Stereotyped Evaluations on Spport for Women. *Political Behaviour*

Donovan, W. L., and A. Leavitt (1980) Physiologic correlates of direct and averted gaze. *Biol Psychol*. 10(3):189-99.

Duehr, E. E., and Bono, J. E. (2006) Men, Women, and Managers: Are Stereotypes Finally Changing? *Personnel Psychology* 59: 815–846. doi: 10.1111/j.1744-6570.2006.00055.x

Eagly, Alice, Steven Karau and Mona Makhijani (1995) "Gender and Effectiveness of Leaders: A Meta-Analysis." *Psychological Bulletin* 117(1): 125-145.

Eagly, Alice, Mona Makhijani and Bruce Klonsky (1992) "Gender and the evaluation of leaders: A meta-analysis." *Psychological Bulletin* 111(1): 3-22

Eagly, Alice and Karau, Steven (2002) "Role congruity theory of prejudice toward female leaders." *Psychology Review* 109(3): 573-598.

Eagly, Alice H.; Sczesny, Sabine. Stereotypes about women, men, and leaders: Have times changed? Barreto, Manuela (Ed); Ryan, Michelle K. (Ed); Schmitt, Michael T. (Ed), (2009). *The glass ceiling in the 21st century: Understanding barriers to gender equality.* Psychology of women book series., (pp. 21-47). Washington, DC, US: American Psychological Association, xvii, 334 pp. doi: 10.1037/11863-002.

Fabes, R.A. and C.L Martin. Gender and age stereotypes of emotionality. Personality and Social Psychology Bulletin, 17 (1991), pp. 532–540

Fiedler, Klaus, Messner, Claude and Bluemke, Matthias (2006) Unresolved problems with the "I", the "A", and the "T": A logical and psychometric critique of the Implicit Association Test (IAT). *European Review of Social Psychology* 17(1): 74-147.

Fiske, S. T., & Neuberg, S. L. (1990). A continuum of impression formation, from category-based to individuating processes: Influences of information and motivation on attention and interpretation. In M. Zanna (Ed.), *Advances in experimental social psychology* (pp. 1–74). San Diego, CA: Academic Press.

Foschi, Martha. (1996) Double Standards in the Evaluation of Men and Women. *Social Psychology Quarterly* 59(3) Special Issue: Gender and Social Interaction: 237-254
http://www.jstor.org/stable/2787021

French, J. R. P., Jr., and B. Raven (1959). The bases of social power. In D. Cartwright, ed., *Studies in Social Power* (pp. 150–167). Oxford: University of Michigan Press.

Fridlund, A. J., & Izard, C. E. (1983). Electromyographic studies of facial expressions of emotions and patterns of emotions. In J. T. Cacioppo & R. E. Petty (Eds.), Social psychophysiology: A sourcebook (pp. 243-286). New York: Guilford Press.

Stefano Gagliarducci and M. Daniele Paserman. Gender Interactions within Hierarchies: Evidence from the Political Arena *Review of Economic Studies (2012) 79(3): 1021-1052 first published online December 7, 2011*

Gati, Itamar, Gershon Ben-Shakar and S. Avni-Liberty (1996) Stimulus novelty and significance in electrodermal orienting responses: The effects of adding versus deleting stimulus components. Psychophysiology, 33: 637–643. doi: 10.1111/j.1469-8986.1996.tb02358.x

Gati, Itamar, and Gershon Ben-Shakar (1990) Novelty and significance in orientation and habituation: A feature-matching approach. *Journal of Experimental Psychology* 119(3): 251-263. doi: 10.1037/0096-3445.119.3.251

Gentry, W. A., and M. P. Duke (2009) A historical perspective on nonverbal communication in debates: Implications for elections and leadership. *Journal of Leadership Studies* 2: 36–47.

Gidengil, Elisabeth and Everitt, Joanna (2003) Conventional Coverage/Unconventional Politicians: Gender and Media Coverage of Canadian Leaders' Debates, 1993, 1997, 2000. *Canadian Journal of Political Science* 36(3): 559-577

Greenwald, A. G., McGhee, D., & Schwartz, J. L. K. (1998). Measuring individual differences in implicit cognition: The Implicit Association Test. Journal of Personality and Social Psychology, 74, 1464–1480.

Hall, J. A., Carter, J. D., & Horgan, T. G. (2000). Gender differences in nonverbal communication of emotion. In A. H. Fischer (Ed.), Gender and emotion: Social psychological perspectives (pp. 97–117). Cambridge, MA: Cambridge University Press.

Hall, J. A. (2006). How big are nonverbal sex differences? The case of smiling and nonverbal sensitivity. In K. Dindia & D. J. Canary (Eds.), Sex differences and similarities in communication (pp. 59–81). Mahwah, NJ: Erlbaum.

Hall, J.A., 1985. Male and female nonverbal behavior. In: Siegman, A.W., Feldstein, S. (Eds.), Multichannel Integrations of Nonverbal Behavior. Lawrence Erlbaum, Hillsdale, NJ, pp. 195–225.

Henley, N. M. (1977). Body politics: Power, sex, and nonverbal communication. Englewood Cliffs, NJ: Prentice Hall.

Henley, Nancy and Harmon, Sean. Chapter 8: The Nonverbal Semantics of Power and Gender: A Perceptual Study. S. L. Ellyson et al. (eds.), *Power, Dominance, and Nonverbal Behavior*. Springer-Verlag New York Inc. 1985

Heilman, M. E. (2001), Description and Prescription: How Gender Stereotypes Prevent Women's Ascent Up the Organizational Ladder. *Journal of Social Issues* 57: 657–674. doi: 10.1111/0022-4537.00234

Heilman, M. E., & Haynes, M. C. (2005). No Credit Where Credit Is Due: Attributional Rationalization of Women's Success in Male-Female Teams. *Journal of Applied Psychology, 90*(5), 905-916. doi:10.1037/0021-9010.90.5.905

Heo, N., & Sundar, S. S. (1998, July). Source perception and electrodermal activity. Paper presented to the Communication Theory & Methodology division at the annual conference of the Association for Education in Journalism and Mass Communication (AEJMC), Baltimore, MD.

Herrick, Rebekah, Jeanette Mendez , Sue Thomas and Amanda Wilkerson (2012): Gender and Perceptions of Candidate Competency. *Journal of Women, Politics & Policy* 33(2): 126-150.

Hinkin, T. R., and C. A. Schriesheim (1989) Development and application of new scales to measure the French and Raven (1959) bases of social power. *Journal of Applied Psychology* 74(4): 561-567.

Kahn, Kim Fridkin (1992) Does Being Male Help? An Investigation of the Effects of Candidate Gender and Campaign Coverage on Evaluations of U.S. Senate Candidates. *The Journal of Politics* 54(2): 497-517.

Heo, N., and S. S. Sundar (1998) Source perception and electrodermal activity. Paper presented to the Communication Theory & Methodology division at the annual conference of the Association for Education in Journalism and Mass Communication (AEJMC), Baltimore, MD.

Hess, Ursula, Sylvie Blairy, and Robert E. Kleck. (2000) Influence of Facial Emotion Displays, Gender and Ethnicity on Judgments of Dominance and Affiliation. *Journal of Nonverbal Behavior* 24(4).

Huddy, Leonie (1994) The Political Significance of Voters' Gender Stereotypes. *Research in Micropolitics* 4: 169-193.

Huddy, Leonie, and Nayda Terkildson (1993a) Gender Stereotypes and the Perception of Male and Female Candidates. *American Journal of Political Science* 37: 119-47.

Huddy, Leonie, and Nayda Terkildson (1993b) The Consequences of Gender Stereotypes for Women Candidates at Different Levels and Types of Office. *Political Research Quarterly* 46: 503-25.

Koch, Sabine C. (2004) Constructing Gender: A Lens-Model Inspired Gender Communication Approach. *Sex Roles* 51(3/4)

Krauss, Robert M., William Apple, Nancy Morency, Charlotte Wenzel and Ward Winton. 1981. Verbal, Vocal and Visible Factors in Judgments of Another's Affect. *Journal of Personality and Social Psychology* 40(2): 312-20

LaFrance, M., and M. A. Hecht (1999) Option or obligation to smile: the effects of power and gender on facial expression. In P. Philippot, R. Feldman and E. Coats, eds.,, *The Social Context of Nonverbal Behavior*. Cambridge University Press, Cambridge, pp. 45–70.

Leeper, Mark Stephen (1991) The impact of prejudice on female candidates: An experimental look at voter inference. *American Politics Quarterly* 19: 248-61.

Levinson, Justin D., and Danielle Young (2010) Implicit Gender Bias in the Legal Profession: An Empirical Study. *Duke Journal of Gender, Law & Policy* 18(1)

Lewis, Kathryn and Margaret Bierly (1990) Toward a Profile of the Female Voter: Sex Differences in Perceived Physical Attractiveness and Competence of Political Candidates. *Sex Roles* 22(1/2): 1-12.

Lodge, M., C. Taber and B. Verhulst (2011) Conscious and Unconscious Information Processing with Implications for Experimental Political Science. In J. N. Druckman, D. P. Green, J. H. Kuklinski and A. Lupia, eds., *Cambridge Handbook of Experimental Political Science* (pp. 155-170), Cambridge: Cambridge University Press.

Mast, M. S., & Hall, J. A. (2004). Who is the boss and who is not? Accuracy of judging status. *Journal of Nonverbal Behavior* 28: 145–165.

Mast, M.S. (2004) Men Are Hierarchical, Women Are Egalitarian: An Implict Gender Stereotype. Swiss Journal of Psychology Vol. 63 Number 2.

Mast, M. S., Hall, J. A., Murphy, N. A., & Colvin, C. R. (2003) Judging assertiveness in female and male targets. *Facta Universitatis* 2: 731–743.

 Mast, Marianne Schmid, and Sabine Sczesny (2010) Gender, Power, and Nonverbal Behavior. *Handbook of Gender Research in Psychology*, pp 411-425.

Masters, R. D. (1991) Individual and Cultural Differences in Response to Leaders' Nonverbal Displays. *Journal of Social Issues* 47(3): 151-165.
http://onlinelibrary.wiley.com/doi/10.1111/j.1540-4560.1991.tb01828.x/abstract

McDermott, Rose (2002) Experimental Methodology in Political Science. *Political Analysis* 10: 325-42.

McElroy, Gail and Michael Marsh (2010) Candidate Gender and Voter Choice: Analysis from a Multimember Preferential Voting System. *Political Research Quarterly* 63(4): 822-833.

McHugo, Gregory J., John T. Lanzetta, Lauren K. Bush (1991) The effect of attitudes on emotional reactions to expressive displays of political leaders. *Journal of Nonverbal Behavior* 15(1): 19-41

McHugo, G.J., Lanzetta, J.T., Sullivan, D.G., Masters, R.D., Englis, B.G. (1985) Emotional Reactions to a Political Leader's Expressive Displays. *Journal of Personality and Social Psychology* 49(6): 1513-1529
10.1037/0022-3514.49.6.1513

Meeks, L. (2012) Is She "Man Enough"? Women Candidates, Executive Political Offices, and News Coverage. *Journal of Communication* 62(1): 175-193.

Mendes, W.B., S. B. Hunter, J. Blascovich, B. Lickel and J. T. Jost (2007) Threatened by the Unexpected: Physiological Responses During Social Interactions With Expectancy-Violating Partners. *Journal of Personality and Social Psychology* 92(4): 698–716.

Nagel, Friederike, Marcus Maurer and Carsten Reinemann (2012) Is There a Visual Dominance in Political Communication? How Verbal, Visual and Vocal Communication Shape Viewers' Impressions of Political Candidates. *Journal of Communication* 65(5): 833-50.

Osterhout, Lee, Michael Bersick and Judith M Laughlin (1997) Brain potentials reflect violations of gender stereotypes. *Memory & Cognition* 25 (3): 273-285

Ottati, Victor C., and Megan Deiger (2002) Visual Cues and the Candidate Evaluation Process: The Social Psychology of Politics. *Social Psychological Applications to Social Issues*, pp 75-87.

Patterson, M. L., et al. (1992) Verbal and nonverbal modality effects on impressions of political candidates: Analysis from the 1984 presidential debates. *Communication Monographs* 59(3): 231-242.

Paivio, A. (2007) *Mind and its Evolution: A Dual Coding Theoretical Approach*. Mahwah, NJ: Psychology Press.

Porter, Natalie, and Florence Geis (1981) Women and Nonverbal Leadership Cues: When Seeing Is Not Believing. In C. Mayo, et al., eds., *Gender and Nonverbal Behavior*. New York: Springer-Verlag, pp. 39-61.

Ridgeway, C. (1987) Nonverbal behavior, dominance, and the basis of status in task groups. *American Sociological Review* 52: 683-694.

Ridgeway, C. L., J. Berger and L. Smith (1985) Nonverbal cues and status: An expectation states approach. *American Journal of Sociology* 90: 955-978.

Ridgeway, C. L. (2002) Gender, Status, and Leadership. *Journal of Social Issues* 57: 637–655. doi: 10.1111/0022-4537.00233

Ritter, B. A., and J. D. Yoder (2004) Gender differences in leader emergence persist even for dominant women: an updated confirmation of role congruity theory. *Psychology of Women Quarterly* 28: 187–193.

Riordan, C. A., T. Gross and C. C. Maloney (1994) Self-monitoring, gender, and the personal consequences of impression management. *American Behavioral Scientist* 37: 715–725.

Rosenwasser, Shirley, and Norma Dean (1989) Gender Role and Political Office: Effects of Perceived Masculinity/Femininity of Candidate and Political Office. *Psychology of Women Quarterly* 13: 77-85.

Rosenwasser, Shirley, and Jana Seale (1988) Attitudes toward a Hypothetical Male or Female Presidential Candidate--A Research Note. *Political Psychology* 9: 591-98.

Rudman, Laurie A. and Glick, Peter (2001) Prescriptive Gender Stereotypes and Backlash

Toward Agentic Women. *Journal of Social Issues* 57(4): 743–762.

Rule, Nicholas O., and Nalini Ambady (2009) She's Got the Look: Inferences from Female Chief Executive Officers' Faces Predict their Success. *Sex Roles* 61(9-10): 644-652.

Sanbonmatsu, Kira, and Kathleen Dolan (2009) Do Gender Stereotypes Transcend Party? *Political Research Quarterly* 62(3): 485-94.

Schubert, James N., and Margaret Ann Curran (2001) Stereotyping Effects in Candidate Evaluation: The Interaction of Gender and Attractiveness Bias**.** Paper prepared for presentation at the 2001 Annual Meeting of the Midwestern Political Science Association, Chicago, April 19-22.

Sczesny, Sabine, Sandra Spreemann and Dagmar Stahlberg (2006) "Masculine = Competent? Physical Apperance and Sex as Sources of Gender Stereotypic Attributions." *Swiss Journal of Psychology* 65(1): 15-23.

Sapiro, Virginia (1981/82) If US Senator Baker Were a Woman: An Experimental Study of Candidate Images. *Political Psychology* 3(1/2): 61-83.

Shearer, D., & Mikulka, P. (1996) Effect of facial familiarity and   task requirement on electrodermal activity.  *American Journal of Psychology* 109(10): 131-137.

Sigelman, Carol K., Lee Sigelman, Dan B. Thomas, and Frederick D. Ribich (1986) Gender, Physical Attractiveness, and Electability: An Experimental Investigation of Voter Biases. *Journal of Applied Social Psychology* 16: 229-48.

Sigelman, Lee, Carol K. Sigleman, Christopher Fowler (1987) A Bird of a Different Feather? An Experimental Investigation of Physical Attractiveness and the Electability of Female Candidates. *Social Psychology Quarterly* 50:1 32-43.

Smith, Eric R.A.N., and Richard L. Fox. (2001) A Research Note: The Electoral Fortunes of Women Candidates for Congress, *Political Research Quarterly* 54(1): 205-21.

Soroka, Stuart N. and McAdams, Stephen (2012) *News, Politics, and Negativity*, CIRANO - Scientific Publications 2012s-14. Available at SSRN: http://ssrn.com/abstract=2075941 or http://dx.doi.org/10.2139/ssrn.2075941

Stewart, Patrick A. Frank K. Salter, and Marc Mehu (2009) Taking leaders at face value: Ethology and the analysis of televised leader displays. Politics and the Life Sciences 28(1):48-74. doi: http://dx.doi.org/10.2990/28_1_48

Stewart, Patrick A., F. K. Salter, and M. Mehu (2011) The face as a focus of political communication: Evolutionary perspectives and the ethological methods. In Erik P. Bucy and R. Lance Holbert, eds., *The Sourcebook for Political Communication Research* New York: Routledge, pp. 165-193.

Stewart, Patrick A., Pearl K. Ford Dowe (2013) Interpreting President Barack Obama's Facial Displays of Emotion: Revisiting the Dartmouth Group. *Political Psychology* 34(3): 369–385.

Thayer, J. and Johnsen, B. H. (2000) Sex differences in judgement of facial affect: A multivariate analysis of recognition errors. *Scandinavian Journal of Psychology* 41: 243–246. doi: 10.1111/1467-9450.00193

Thomas, Sue. (1998) Introduction: Women and Elective Office: Past, Present, and Future, In Sue Thomas and Clyde Wilcox, eds., *Women and Elective Office*, New York: Oxford University Press.

Van Damme, S., G. Crombez and C. Spence (2009) Is Visual Dominance Modulated by the Threat value of Visual and Auditory Stimuli? *Experimental Brain Research* 193: 197-204.

Vrana, S.R., and D. Rollock. (2002) The role of ethnicity, gender, emotional content and contextual differences in physiological, expressive, and self-reported emotional responses to imagery. *Cognition and Emotion* 16: 165–192

Way, Baldwin M., and Roger D. Masters (2006) Emotion and Cognition in Political Information Processing. *Journal of Communication* 46(3)