A Face Fit for Office?
Appearance-Based Competence Inferences in High-Information Environments

In a democratic government, citizens are expected to participate in the political process in a meaningful way—specifically, to determine where candidates stand on a number of dimensions and then to vote for the candidate who most closely aligns with their own values and preferences. It is well known that voters utilize various heuristics and shortcuts to help them with this process (Lau and Redlawsk 2006; Popkin 1991) and generally accepted that they can make reasonably good decisions by doing so. It is unclear, though, whether all shortcuts are created equal (Kuklinski and Quirk 2000). For example, candidates are often evaluated on how they look, and various aspects of their appearance lead voters to make assumptions about politically salient attributes such as personality traits, ideology, policy issue specializations, and/or political viability. All of this despite the fact that there is no guarantee that appearance cues provide any accurate information (Hassin and Trope 2000; Zebrowitz and Montepare 2008; Oliviola and Todorov 2010b).

Social psychologists have repeatedly found that individuals make assumptions about others' personality traits based on nothing more than a brief exposure to still images of their faces (e.g. Hall et al 2009; Hassin and Trope 2000). These trait judgments have proven important in the political realm and facial competence, in particular, seems to play a role in predicting election outcomes. Specifically, candidates whose faces were judged to be more competent than their opponents' by in-lab subjects are more likely to have won their races in the real world. This pattern of results has been found in elections at various levels of government and across countries (see Oliviola and Todorov 2010a for an overview of this literature).

Importantly, most of these studies have been essentially correlational, since lab subjects were shown a series of candidate photo pairs and asked to rate them on relative competence (as well as other traits and, in some instances, asked for whom they would vote) without any sort of controlled manipulation. Little is known about the process by which competence inferences lead to electoral victory. Specifically, we know little about how appearance-based competence judgments influence vote choice in

the presence of other kinds of information. This includes substantive information, such as that provided by campaigns and the news media, as well as other appearance-based, heuristic information such as a candidate's gender, age, race or ethnicity. Because competence judgments predict the outcomes of actual elections for various state and national-level races, they are clearly having an effect even in potentially very information-rich environments. How does something as seemingly inconsequential as the composition of a candidate's facial features predict election outcomes even when so much other information is available?

This study seeks to improve our understanding of the relationship between automatic trait inferences and other types of political information in voters' decision-making processes by addressing several questions. First, do competence inferences predict voting behavior even when voters access other politically-relevant information, or are the effects of competence judgments on electoral outcomes driven by individuals who possess little other information about the candidates? Will these cues influence voting behavior when subjects are placed in a high-information environment and can learn (almost) anything they want to know about a particular candidate? Further, will they function differently for those who seek out a lot of information vs. those who only access a little?

Second, do competence judgments matter in the same way for male and female candidates? While most trait inference studies control for candidate gender, several have examined whether candidate gender affects appearance-based trait inferences and their effects on elections (Herrick et al 2012; Ditonto 2017; Ditonto and Mattes forthcoming). Results from these studies have been mixed, though it seems that the inclusion of female candidates complicates the relationship between competence inferences and electoral success. In a sense, this is not surprising since candidate gender likely provides voters with other, potentially relevant information thanks to the existence of gender-bases stereotypes (e.g. Huddy and Terkildsen 1993; Kahn 1996, Sigelman, Sigelman, Walkosz and Nitz 1995). Many of these stereotypes relate to a candidate's personality traits, including competence. In particular, evidence suggests that women are considered less competent than their male counterparts (Kahn 1996, Sigelman, et al 1995) and that competence evaluations, in general, seem to be particularly important for women

candidates (Ditonto, Hamilton and Redlawsk 2014; Ditonto 2017, e.g.). For these reasons, it is important to consider candidate gender in any examination of appearance-based competence judgments and voting behavior.

In order to address these questions, this study utilizes a computer-based experiment using the Dynamic Process Tracing Environment (DPTE).[1] DPTE is a web-based computer program that allows researchers to mimic the constant flow of information in an actual campaign environment, thereby tracking subjects' information search as the campaign progresses. In particular, this study asks subjects to participate in a simulated presidential primary campaign and election in which they can learn as much or little information about the candidates in the race as they like. I manipulate the gender of the candidates in the race, as well as how competent the candidates look, and examine how these appearance cues influence subjects' evaluations of the candidates' competence, affect toward the candidates, and their ultimate vote choice, after experiencing the campaign and gathering relevant information about them. I find that that the extent to which a candidate looks competent affects post-primary competence evaluations and feeling thermometer scores for female candidates, only, but that accessing additional information may help subjects to overcome appearance based-judgments. Finally, appearance cues do not influence vote choice directly, though they may do so indirectly via explicit competence evaluations. Importantly, my results suggest that appearance-based trait inferences are persistent and consequential for women candidates (at least in the context of a primary election), even when presented along with other, substantive political information.

Spontaneous Trait Inferences

Research emerging from social psychology has found that, in all forms of human interaction, individuals constantly make rapid trait inferences based on others' appearance, and their faces, in particular (e.g. Hall et al 2009; Hassin and Trope 2000; Oliviola and Todorov 2010). It is still unclear exactly what features a face must possess in order to appear competent, trustworthy, dominant, etc., but some evidence suggests that facial features that signal maturity and attractiveness—such as distance

[1] Developed by Richard Lau and David Redlawsk with assistance from the National Science Foundation.

between the eyes, roundness of the face and angularity of the jaw—influence these types of judgments (Oliviola and Todorov 2010a; see Zebrowitz 2011 and Todorov, et al 2015). Regardless of the underlying reasons, these sorts of trait inferences are occurring quickly and frequently, and they affect individuals' decision-making in a number of domains, including politics. Researchers have found that subjects' spontaneous inferences about a candidate's traits (and, especially, competence) correctly predicted the outcomes of actual elections (Martin 1978; Todorov, Mandisotza, Goren and Hall 2005; Hall et al 2009; Ballew and Todorov 2007; Oliviola and Todorov 2010a; Antonakis and Dalgas 2013; Chen, et al 2014; Laustsen 2013; Lenz and Lawson 2011; Sussman, et al 2013; Mattes, Spezio, Kim, Todorov, Adolphs, and Alvarez 2010). In the first of these modern studies, Todorov, et al (2005) found that competence judgments predicted Congressional and Gubernatorial elections between 60% and 73% of the time, and subsequent studies have found similar results in other countries (e.g. Sussman, et al) and at different levels of government (e.g. Mattes, et al 2010). These studies showed subjects nothing more than a pair of still images of candidates' faces and asked them to determine which of the candidates was higher on a series of trait ratings. The candidate deemed most competent looking was more likely to have won the election.

Todorov et al (2005) posit that access to additional political information may weaken the relationship between appearance-based inferences and ultimate vote choice, but that the nature of the relationship should not change. Initial appearance-based impressions are an example of automatic, System 1 processing which can be difficult to overcome with more deliberate, System 2 activity since later attention-dependent behaviors like seeking out and incorporating political information are anchored on the initial impressions that are formed using intuitive, System 1 judgments (Gilbert, 1989; Chaiken and Trope 1999). In other words, spontaneous trait inferences likely function as a lens through which all subsequent political information that is learned about the candidate is considered. Because of this, it may be very difficult to change voters' perceptions based on these sorts of inferences.

Gender Stereotypes

In a separate but related literature, scholars have found substantial evidence that trait-based stereotypes are often ascribed to women candidates. Women candidates are seen as more trustworthy, warm, emotional and compassionate than men, but less assertive, rational and "tough" than men (Sapiro 1982; Alexander and Andersen 1993; Burrell 2008; Huddy and Terkildsen 1993; Kahn 1996; King and Matland 2003; Lawless 2004; Leeper 1991; Paul and Smith 2008). In particular, women also seem to be perceived as less competent and experienced than male candidates (Huddy and Terkildsen 1993a; Kahn 1996; Schneider and Bos 2014).

It is unclear whether and how much these stereotypes influence vote choice and overall candidate evaluation (Dolan 2014; Brooks 2013; Hayes 2011), though a growing body of evidence suggests that competence stereotypes may be particularly problematic for women candidates, even when other stereotypes my be less salient. Eagly and Karau (2002), for example, posit that women in leadership roles face gender-based bias and assumptions that they lack agentic qualities like competence. Similarly, Schneider and Bos (2014) find that women candidates are stereotyped as women candidates, per se, rather than as women more generally, and that part of this stereotype is a deficiency in masculine traits like competence. Ditonto, Hamilton and Redlawsk (2014) also find that subjects in a DPTE study search for more information related to competence when presented with a female candidate than when shown a man.

Further, several trait inference and voting studies have specifically considered candidate gender in their analyses. Herrick, et al (2012) find that men's and women's faces are rated differently in terms of competence, and that women's faces are seen as less mature and less competent in both real and simulated elections. Interestingly, they also find that competence is not a significant predictor of winning an election for female candidates when other variables are added to the model, such as candidate age, incumbency, and whether the candidate picture displayed a flag. Similarly, Ditonto and Mattes (forthcoming) find that, while competence inferences predict victory in a sample of state legislative races for male-male races, female-female races are decided instead by attractiveness judgments, and that competence only matters in male-female races when it is the female candidate who is seen as more competent.

Ditonto (2017) also examines gender and competence in a DPTE study and finds that women are inordinately affected by *substantive* political information that gives cues related to competence, but not by whether their *appearance* portrays competence (though, she does find some tentative evidence that Republicans are less likely to vote for incompetent-looking women). While the analysis in that study has some similarities to the one presented in this paper, there are significant differences, and this paper provides some important extensions of that research. First, Ditonto (2017) manipulates whether the information presented about the candidates in the race portrays them as competent or incompetent, which presents a very strong cue to voters and is more likely to override appearance inferences than the more neutral information presented here (discussed below). Second, the 2017 article uses data from a simulated general election campaign, which presents subjects with an influential party ID cue, whereas the data in this paper come from a simulated primary campaign, which holds party constant for both candidates. In other words, the 2017 study presents a very tough test for the effects of rapid competence inferences by including both substantive information directly related to candidate competence and an influential partisanship cue. Even in that context, there is tentative evidence that appearance inferences may matter for some voters, so it is an important extension of that research to determine whether we find a clearer pattern of results in a less explicitly competence-focused information environment, as well as one that does not pit candidate appearance against partisanship.

Method and Hypotheses

While much has been learned from studies of rapid trait inferences, it is still unclear precisely how these sorts of judgments influence electoral outcomes. Because most of the research to date has been correlational, and considered in the aggregate, it is impossible to discern whether and how individual voters are using these sorts of appearance cues in their decision-making processes. It is particularly unclear whether these appearance-based cues matter when voters take the time to learn the kinds of politically relevant information about candidates that is typically available during political campaigns. In order to explore this, I rely on dynamic process tracing (Lau and Redlawsk 2006). DPTE is a particularly useful method for examining the relationship between System 1 inferences and System 2 information

processing, since it simulates the constant flow of information in a political campaign and allows researchers to track the amount and content of information that subjects choose to access about each candidate. Further, a number of studies have used DPTE to examine the effects of candidate gender on information search and voter decision-making (e.g. Ditonto, et al 2014; Ditonto 2017; Ditonto and Andersen *forthcoming* and Ditonto *forthcoming).*

Sample

The experiment was conducted in the spring of 2012[2] and was completed by 449 total subjects. 106 of those subjects were recruited from [redacted] area and took the experiment in the lab at [redacted]. The remaining subjects were recruited through Amazon Mechanical Turk and took the study online using their own computers.[3] The sample is 58% female, 76% white, 9% African American, 9% Asian and 5% Latino. The median age is 32 and 55% had at least a college degree. 16% of the sample identified as Republican, 48% as Democrat, and 31% as independent.

Procedure and Manipulations

The experiment lasted approximately one hour and consisted of a pre-election questionnaire, a "practice" campaign, a primary campaign and election followed by a general election[4], and a post-election questionnaire. The primary "campaign" lasted approximately 11 minutes, and subjects had the opportunity to access 25 unique pieces of information about each of the two candidates in their primary, as well as 25 pieces of information about the two candidates in the other party's primary (each piece of information was available twice). Subjects could access information about the candidates' stances on a number of policy issues, their ideology, family, education, and prior experience in work and politics. After the primary campaign, subjects were asked to answer a series of candidate evaluation questions and

[2] With support from the National Science Foundation

[3] Very few differences were found between the two sub-samples or in their performance in the study. Information about these differences can be provided upon request.

[4] The manipulations in the primary and general election were slightly different, and this paper focuses on those presented in the primary campaign and election. Therefore, I do not include extensive discussion of the general election portion of the experiment. For more information on this part of the study, see Ditonto 2017.

to make a vote choice. A more thorough discussion of the experimental design, and DPTE, can be found in the online Appendix.

Subjects experienced a 2x2 manipulation, in which the picture of one of the candidates in their party's primary (Candidate A) varied by gender and whether the candidate had a competent or incompetent appearance, based on pre-tests of pictures conducted by 148 undergraduates.[5] Candidate A, then, could be male or female and either competent-looking or incompetent-looking. His or her opponent (Candidate B) was always a white man and either looked competent or incompetent, but was always the opposite of Candidate A. In other words, If Candidate A was competent-looking (regardless of the gender manipulation), Candidate B was always an incompetent-looking white man. If Candidate A was incompetent-looking (again, regardless of gender), Candidate B was always a competent-looking white man. While Candidate A allows me to analyze the interaction of gender and competence inferences, Candidate B manipulates competence, only, in a white, male candidate (still the most common type of candidate at all levels of government in the US). [6]

Measures

My dependent variables of interest include subjects' evaluations of the candidates' competence after the election, as well as their feeling thermometer scores and ultimate vote choice. Subjects were asked to indicate how competent they found each candidate to be, each on a 1-4 scale, and I take the difference between those two scores (Candidate A – Candidate B) as my dependent variable. I also use the difference score (Candidate A – Candidate B) for the feeling thermometer measure. I use difference scores rather than raw measures for each candidate for two reasons: first, to control for inter-subject variability, and second to more closely mimic the design used by Todorov and colleagues, which asked subjects to provide a relative competence rating (i.e. to specify which of two candidates looks more competent). Vote choice is measured using a dichotomous variable, for which a score of 1 signifies a vote

[5] More information, and the pictures used, can be found in the online Appendix.

[6] This study also included a randomly-assigned race manipulation in which Candidate A could have either been black or white. Including candidate race in my analyses did not change the nature of my results in any significant way, but did require that I add substantially to the front end of the paper and the explanation of my results. In the interest of parsimony and brevity, I have left the race manipulation out of the analyses that appear here.

for Candidate A and a score of 0 signifies a vote for Candidate B (subjects were required to vote for one of the two candidates).

My independent variables include my primary manipulations—whether candidate A was male or female, and whether Candidate A was competent- or incompetent-looking. Additionally, I include a measure of the number of items accessed during the campaign for Candidate A in some of my analyses in order to determine whether having more information about a candidate changes the relationship between appearance inferences and ultimate evaluations. In regression models, this a continuous measure (0- 42) and in ANOVAs, it is a trichotomous measure indicating low, medium and high search, which divides the continuous measure into thirds (with cut points at 15 and 22 based on sample search statistics).[7]

Hypotheses

Because prior research has found such a strong effect of competence inferences on voting behavior, I expect that, all else equal, competent-looking candidates will be rated as more competent than incompetent-looking candidates even after a DPTE "campaign." I also expect that they will receive higher feeling thermometer scores, relative to their opponent, and a greater likelihood of receiving a subject's vote.

Additionally, because female candidates are subject to stereotypes that cause them to be seen as less competent, I expect there to be significant interactions between my manipulations such that women candidates who are also considered incompetent-looking will be rated especially negatively in terms of competence and feeling thermometer scores. They should also be especially less likely to receive subjects' "votes."

Again, I expect these patterns to hold true even after subjects have learned a great deal of information about the candidates in the race. However, I also expect that the relationship between appearance-based judgments and post-campaign evaluations and vote choice should be stronger for

[7] I tried multiple different versions of these particular search variables, including a measure that incorporates search for both candidates as well as an ordinal version with a single cut-point in the middle, and one in which it was divided into quarters. I also ran my analyses with a simple "low search" dummy variable, in which subjects falling into the lowest 25% of the sample in terms of information search instead. None of these variations change the nature of my results.

subjects who access less other information about the candidates, since increased System 2 processing may attenuate the effects of System 1inferences.

Results

I begin with an analysis of subjects' competence ratings for each candidate, which were completed after subjects experienced the primary campaign. This will determine whether the competence/incompetence appearance manipulation affected subjects' evaluations of candidates' competence even after spending 11 minutes gathering relevant information about them. An ANOVA that includes only the main effects of candidate gender and the competence manipulation yields the expected pattern of results, but they do not reach statistical significance. The mean difference score between Candidate A and Candidate B is larger (and negative) when Candidate A is incompetent than when he or she is competent (-.1 vs. -.039) suggesting that subjects perceive the competent-looking candidate as more competent, all else equal, but the effect is not large enough to be significant.

[Figure 1 about here]

Interestingly, when the interaction between Candidate A's gender and competence is added to the model, I find a significant effect such that competent-looking female candidates are considered to be the most competent candidates of all (a mean difference score of +.043) while incompetent female candidates are considered to be the least competent, with a mean difference score of -.167. Male candidates are less affected by their appearance either way, though mean difference scores for both competent and incompetent-looking men are negative (-.105 for competent-looking men and -.033 for incompetent looking men).

The extent to which a female candidate appears competent seems to matter to subjects, even after they experience a simulated campaign replete with a large amount of politically-relevant information. But does the amount of information accessed by subjects influence this relationship? The short answer is that is does not. Adding the information search variable to the model yields no significant differences by level of information accessed. The interaction between candidate gender and competence remains significant

with the same pattern of results, suggesting that subjects are particularly influenced by a female candidate's appearance regardless of how much information they access.[8]

Table 1 presents the results of a series of OLS regressions analyzing the effects of my manipulations on feeling thermometer difference scores (again, Candidate A – Candidate B). Model 1 includes only the two primary manipulations and shows that, while both seeing a female candidate and an incompetent looking candidate result in lower feeling thermometer scores for Candidate A, neither reaches statistical significance. Again, though, Model 2 includes the interaction between these two manipulations and the result is a statistically significant and negative effect on difference scores ($b=-10.763$, $p<.05$) suggesting that subjects who see an incompetent-looking female candidate like her approximately 10 points less than competent male candidates (relative to Candidate B). Not only are female candidates whose appearance is judged to be incompetent seen as more incompetent after an entire "campaign," but they are liked less than other candidates, as well.

[Table 1 about here]

Models 3-5 incorporate total information search for Candidate A into the equation, along with the 2-way and 3-way interactions between information search and my manipulations. The search measure on its own (Model 3) is positive and significant, while the interaction between gender and competence remains negative and significant, suggesting that seeking out more information about candidate A leads to higher feeling thermometer scores, all else equal. Neither 2-way interaction (Model 4) reaches statistical significance, but Model 5 shows that the 3-way interaction between search, gender and competence is positive and significant at $p<.065$. The more information subjects seek out specifically for incompetent-looking female candidates, the more they seem to like them. Each new item accessed raises their score (relative to their opponent) by approximately 1 point. Interestingly, the marginal effect for the 2-way interaction between candidate gender and competence in the model is negative, significant, and very large. This suggests that the expected difference score for subjects who saw an incompetent female

[8] It is also worth noting that the amount of information subjects seek out is not affected by my manipulations.

candidate and searched for no information about her would be approximately -29 points ($p<.05$), or they would rate Candidate B a full 29 points higher than Candidate A on a 1-100 scale.

[Table 2 about here]

I have so far found fairly substantial evidence that whether a female candidate "looks competent" can have a major effect on how subjects in my experiment evaluate her. I turn now to a series of logistic regressions to determine whether those effects transfer to vote choice, as well. Table 2 shows the results of 6 separate models, each predicting a vote for Candidate A. Model 1 includes the two dummy variables indicating that Candidate A is incompetent and female, and Model 2 adds the interaction. On their own, neither of these variables is a significant predictor of vote choice, nor is the interaction between gender and competence, which has driven my findings previously. While appearance seems to influence evaluations, then, it seems that this does not transfer to subjects' choice about which candidate they ultimately prefer. Adding variables for the total number of information boxes opened for Candidate A in Model 3 adds significant explanatory power to the model. Searching for more information about Candidate A leads to a greater likelihood of voting for him or her. Even controlling for information search, though, the appearance of the candidates does not have a direct effect on vote choice. Model 4 adds the two-way interactions between my manipulations and information search to the model, while Model 5 adds the 3-way interaction. None of these have any statistically significant effect on vote choice.

Model 6, finally, adds the competence evaluation difference score between Candidate A and Candidate B to the model in order to determine whether post-election competence evaluations have a role in driving vote choice. If they do, that may mean that candidate appearance variables (and, specifically, the competence and gender interaction) have an indirect effect on vote choice via explicit competence evaluations. Indeed, the coefficient for the competence difference score is positive and significant (2.086, $p<.001$), which suggests that the more subjects see Candidate A as competent, the more likely they are to vote for him or her. Since the combination of being female and having an incompetent appearance has a deleterious effect on competence ratings, it seems that appearance cues have an indirect effect on vote choice, even though they do not affect the vote directly.

Discussion and Conclusions

Clearly, the relationship between appearance-based trait inferences, gender, and voting behavior is a complicated one. The analysis of the effects of appearance manipulations on competence ratings shows that whether a female candidate appears competent or not does seem to play a role in how she is evaluated, even after subjects are given access to a host of political and personal information about the candidates. This lends further support to previous findings that these sorts of judgments are persistent, even in the face of other information. That a competent (or incompetent) appearance only has a significant effect on competence ratings for women candidates, and not for men, was an unexpected finding, but not necessarily surprising. While I expected appearance-based competence to matter for both men and women candidates, I expected the effects to be stronger for women, since they are also subject to competence-related stereotypes and seem to be evaluated more heavily on competence-related information than men are (e.g. Schneider and Bos 2014; Ditonto, et al 2014; Ditonto 2017). Of course it is impossible to tell from these data precisely why women candidates are affected by these kinds of appearance cues while men are not, but it seems likely that competence-related stereotypes have a role to play.

A major contribution of this paper is to show that these sorts of appearance cues matter at all when more deliberate information search also takes place. Importantly, the differences in competence scores and feeling thermometer ratings hold regardless of the amount of information that subjects search for. Recall that subjects could learn up to 25 unique pieces of information about each candidate in the race, each of which conveys something about the candidate's background, character, or policy stances. That appearance cues affected what subjects thought of female candidates over and above what they learned in the campaign may have important implications for real-world politics. This is particularly true given that subjects' competence ratings predicted their ultimate vote choice. It would seem that appearance-based impressions of female candidates matter a great deal and that they can be difficult to overcome, regardless of home much information voters acquire about them. On the other hand, there is some evidence that feeling thermometer scores for incompetent-appearing women went up as subjects

accessed more information about them. While deliberate, System 2 processing does not seem to override automatic, System 1 inferences, then, this may suggest that substantive information can at least attenuate the effects of appearance-based competence inferences.

Like all studies, this one has important limitations that must be considered. For example, there is a trade-off in using actual candidate images in a study like this. Using real pictures increases external validity, since the individuals in the pictures were actually elected to public office and therefore relatively realistic. The study was also designed to mimic studies such as Todorov (2005) and Ditonto and Mattes (forthcoming), which show subjects actual candidate photos. However, while the images were pretested on a number of traits and selected specifically for their competence ratings (which never correlated with other trait ratings higher than .3-.4), it is possible that the effects of the candidates' appearance on voting behavior is a result of some other aspect of how they look that was not included in the pretest. A study that creates candidate images and manipulates facial features specifically to appear more or less competent (and that holds other features such as hair and wardrobe constant) would give a better indication that it is a candidate's face, per se, that is driving results. Of course, using those sorts of manipulated images would decrease external validity since they are not actual candidates and they can look less than "real."

Nevertheless, the results of this study add to our understanding of how and when appearance-based trait judgments may influence voting behavior and it would be useful to expand on these findings in a number of ways. In particular, while this study considers the amount of substantive political information that voters access about candidates, in order to better understand why competence inferences are persistent in the face of real-world campaigns, it would be instructive to consider other campaign factors that may mediate the relationship between appearance cues and electoral success. For example, are competent-looking candidates better able to secure more campaign funding than less competent-looking candidates? Do they receive more positive media coverage? Are voters just generally more enthusiastic about them? While the nature of the information environment certainly has a role to play in this story, it is likely that there are other important factors to consider, as well.

Works Cited

Alexander, Deborah and Kristi Andersen.1993. Gender as a Factor in the Attribution of Leadership Traits. *Political Research Quarterly* 46 (3): 527-545.

Ballew, C. C., & Todorov, A. 2007. Predicting political elections from rapid and unreflective face judgments. Proceedings of the National Academy of Sciences of the USA, 104, 17948– 17953.

Brooks, Deborah Jordan. 2013. *He Runs, She Runs: Why Gender Stereotypes do not Harm Women Candidates.* Princeton, NJ: Princeton University Press.

Burrell, Barbara. 1994. *A Woman's Place is in the House: Campaigning for Congress in the Feminist Era.* Ann Arbor: University of Michigan Press.

Ditonto, Tessa. Forthcoming. The Direct and Indirect Effects of Prejudice: Sexism and Information Search in Political Campaigns. *Politics, Groups and Identities*

Ditonto, Tessa. 2017. "A High Bar or a Double Standard? Gender, Competence, and Information in Political Campaigns." *Political Behavior.* DOI: 10.1007/s11109-016-9357-5

Ditonto, Tessa and David J. Andersen. Forthcoming. Two's A Crowd: Women Candidates in Concurrent Elections. *Journal of Women, Politics and Policy*

Ditonto, Tessa, Allison J. Hamilton, and David P. Redlawsk. 2014. Gender Stereotypes, Information Search, and Voting Behavior in Political Campaigns. *Political Behavior* 36(2): 335-358.

Ditonto, Tessa and Kyle Mattes. Forthcoming. Gender Differences in Automatic Inferences of Competence, Attractiveness and Threat *Journal of Women, Politics and Policy*

Dolan, Kathleen. 2004. *Voting for Women How the Public Evaluates Women Candidates.* Boulder, CO: Westview Press.

Dolan, Kathleen. 2014. *When does Gender Matter? Women Candidates and Gender Stereotypes In American Elections.* New York: Oxford.

Eagly, Alice H. and Karau, Steven J., 2002. Role congruity theory of prejudice toward female leaders. *Psychological review*, *109*(3), p.573.

Hassin, R. and Y. Trope. 2000. Facing Faces: Studies on the Cognitive Aspects of Physiognomy. *Journal of Personality and Social Psychology,* 78: 837-852.

Hall, C. C., Goren, A., Chaiken, S., & Todorov, A. (2009). Shallow cues with deep effects:Trait judgments from faces and voting decisions. In E. Borgida, J. L. Sullivan, & C. M. Federico (Eds.), The political psychology of democratic citizenship (pp. 73–99). New York: Oxford University Press.

Hayes, Danny. 2011. When Gender and Party Collide: Stereotyping in Candidate Trait Attribution. *Politics and Gender*, 7(2): 133-165

Huddy, Leonie and Nayda Terkildsen. 1993. Gender Stereotypes and Perceptions of Male and

Female Candidates. *American Journal of Political Science,* 37: 119-147.

Gilbert, Daniel. 1989. "Thinking Lightly About Others: Automatic Components of the Social Inference Process" in James Uleman and John Bargh (eds.) *Unintended Thought.* New York Guilford: pp 189-211.

Hassin, Ran and Yaacov Trope. 2000. "Facing Faces: Studies on the Cognitive Aspects of Physiognomy." *Journal of Personality and Social Psychology,* 78(5): 837-852.

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

Kahn, Kim Fridkiin. 1996. *The Political Consquences of Being a Woman: How Stereotypes Influence the Conduct and Consequences of Political Campaigns.* New York: Columbia.

King, David C. and Richard E. Matland. 2001. "Sex and the Grand Old Party: An Experimental Investigation of the Effect of Candidate Sex on Support for a Republican Candidate." *American Politics Research*, 31(6): 595-612.

Kuklinski, James H. and Paul J. Quirk. 2001. Conceptual Foundations of Citizen Competence. *Political Behavior,* 23: 285-311.

Lau, Richard R. and David P. Redlawsk. 1997. Voting Correctly. *American Political Science Review,* 91: 585-599.

Lau, Richard R. and David P. Redlawsk. 2006. *How Voters Decide: Information Processing During Political Campaigns.* New York: Cambridge University Press.

Lau, Richard R. and David P. Redlawsk. 2001. Advantages and Disadvantages of Cognitive Heuristics in Political Decision-Making. *American Journal of Political Science,* 45: 951-971.

Lawless, Jennifer L. 2004. "Women, War, and Winning Elections: Gender Stereotyping in the Post-September 11th Era." *Political Research Quarterly* 57 (3): 479-490.

Leeper, Mark S. 1991. "The Impact of Prejudice on Female Candidates: An Experimental Look at Voter Inference." *American Politics Research,* 19(2): 248-261.

Lenz, Gabriel S. and Chappell Lawson. 2007. Looking the Part: Television Leads Less Informed Citizens to Vote Based on Candidates' Appearance. *American Journal of Political Science.*

Mattes, Kyle, Michael Spezio, Hackjin Kim, Alexander Todorov, Ralph Adolph, and R. Michael Alvarez. 2010. Predicting Election Outcomes from Positive and Negative Trait Assessments of Candidate Images. *Political Psychology,* 31(1): 41-58.

Oliviola, Christopher and Alexander Todorov. 2010a. Elected in 100 Milliseconds: Apperance-Based Trait Inferences and Voting. *The Journal of Nonverbal Behavior,* 34: 83-110.

Oliviola, Christopher and Alexander Todorov. 2010b. Fooled by First Impressions? Reexamining the Diagnostic Value of Appearance-Based Inferences. *Journal of Experimental Psychology,*40: 267-280.

Popkin, Samuel. 1991. *The Reasoning Voter: Communication and Persuasion in Presidential Campaigns.* Chicago: The University of Chicago Press.

Sapiro, Virginia. 1982. If US Senator Baker were a Woman: An Experimental Study of Candidate Images. *Political Psychology*, 2: 61-83.

Schneider, Monica C. and Bos, Angela L., 2014. Measuring stereotypes of female politicians. *Political Psychology*, *35*(2), pp.245-266.

Todorov, A. 2008. Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. In A. Kingstone & M. Miller (Eds.), The Year in Cognitive Neuroscience 2008: Annals of the New York Academy of Sciences (Vol. 1124, pp. 208–224).

Todorov, A. 2010. Evaluating faces on social dimensions. In A. Todorov, S.T. Fiske, & D. Prentice (Eds.), Social neuroscience: Toward understanding the underpinnings of the social mind. Oxford University Press.

Todorov, A. and A. Engell. 2010. The Role of the Amygdala in Implicit Evaluation of Emotionally Neutral Faces. *Social Cognitive and Affective Neuroscience*, 3: 303-312.

Todorov, A., & Uleman, J. S. 2002. Spontaneous trait inferences are bound to actor's faces: Evidence from a false recognition paradigm. Journal of Personality and Social Psychology, 83, 1051–1065.

Todorov, A., & Uleman, J. S. 2003. The efficiency of binding spontaneous trait inferences to actor's faces. Journal of Experimental Social Psychology, 39, 549–562.

Todorov, A., Chaiken, S., & Henderson, M. D. 2002. The heuristic-systematic model of social information processing. In J. P. Dillard & M. Pfau (Eds.), The persuasion handbook: Developments in theory and practice (pp. 195–211). Thousand Oaks, CA: Sage.

Todorov, A., Mandisotza, A. N., Goren, A., & Hall, C. 2005. Inferences of competence from faces predict election outcomes. Science, 308, 1623–1626.

Todorov, A., Said, C. P., Engell, A. D., & Oosterhof, N. N. 2008. Understanding evaluation of faces on social dimensions. Trends in Cognitive Sciences, 12, 455–460.

Todorov, A., Pakrashi, M., & Oosterhof, N. N. 2009. Evaluating faces on trustworthiness after minimal time exposure. Social Cognition, 27, 813–833.

Zebrowitz, L.A. and J.M, Montepare. 2008. Social Psychological Face Perception: Why Appearance Matters. *Social and Personality Psychology Compass*, 2: 1497-151

Table 1. OLS Regression Models, Feeling Thermometer Difference Scores

	Model 1	Model 2	Model 3	Model 4	Model 5
Constant	1.962 (2.012)	-.504 (2.281)	-10.142 (3.484)	-6.833 (5.044)	-11.399* (5.603)
Candidate A Female	-1.086 (2.376)	4.953 (3.569)	4.398 (3.526)	6.151 (6.547)	17.017† (8.782)
Candidate A Incompetent	-.652 (2.382)	4.197 (3.198)	3.994 (3.164)	-3.700 (6.267)	5.736 (8.067)
Female X Incompetent	-	-10.763* (4.764)	-9.469* (4.721)	-9.485* (4.728)	-29.437* (11.772)
Total Search	-	-	.497*** (.137)	.326 (.233)	.561* (.265)
Female X Search	-	-	-	-.076 (.275)	-.618 (.402)
Incompetent X Search	-	-	-	.391 .275	-.089 (.378)
3-Way	-	-	-	-	1.017† (.550)
R^2	.001	.013	.033	.034	.039

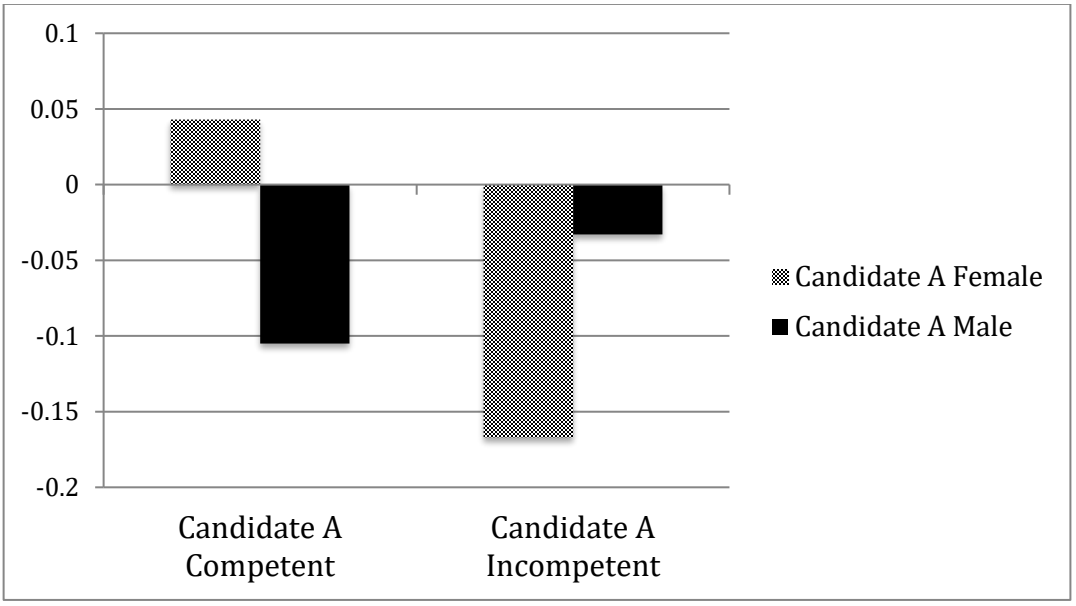
Notes: Regression coefficients are shown, with standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001, †p<.1.

Table 2. Logistic Regression Models, Vote Choice for Candidate A

	Model 1	Model 2	Model 3	Model 4	Model 5	Model 6
Constant	-.158 (.333)	-.176 (.188)	-.983 (.298)	-.525 (.428)	-.824† (.484)	-.658† (.346)
Candidate A Female	-.034 (.190)	.005 (.280)	.022 (.284)	-.318 (.534)	.300 (.691)	-.247 (.338)
Candidate A Incompetent	.208 (.190)	.243 (.262)	.261 (.267)	-.329 (.528)	.282 (.683)	.194 (.310)
Female X Incompetent	-	-.072 (.381)	-.053 (.387)	-.018 (.388)	-1.282 (.978)	.433 (.455)
Total Search	-	-	.041*** (.012)	.018 (.020)	.033 (.023)	.031* (.013)
Female X Search	-	-	-	.017 (.023)	-.015 (.033)	-
Incompetent X Search	-	-	-	.030 (.023)	-.001 (.032)	-
3-Way	-	-	-	-	.066 (.047)	-
Competence Difference Score	-	-	-	-	-	2.086*** (.242)
Pseudo-R^2	.004	.004	.043	.049	.055	.378

Notes: Regressioin coefficients are shown, with standard errors in parentheses. * p<0.05, ** p<0.01, *** p<0.001, †p<.1.

Figure 1. Competence Difference Scores (Candidate A – Candidate B)



Note: Figure from ANOVA. $F=3.866$, $p<.05$