# Predicting Election Outcomes from Positive and Negative Trait Assessments of Candidate Images

**Kyle Mattes**
*The University of Iowa*

**Michael Spezio**
*Scripps College*
*California Institute of Technology*

**Hackjin Kim**
*Korea University*

**Alexander Todorov**
*Princeton University*

**Ralph Adolphs**
*California Institute of Technology*

**R. Michael Alvarez**
*California Institute of Technology*

*Conventional wisdom, and a growing body of behavioral research, suggests that the nonverbal image of a candidate influences voter decision making. We presented subjects with images of political candidates and asked them to make four trait judgments based solely on viewing the photographs. Subjects were asked which of the two faces exhibited more competence, attractiveness, deceitfulness, and threat, which are arguably four of the most salient attributes that can be conveyed by faces. When we compared our subjects' choices to the actual election outcomes, we found that the candidates chosen as more likely to physically threaten the subjects actually lost 65% of the real elections. As expected, our findings support the conclusions of Todorov, Mandisodza, Goren, and Hall (2005) by showing a positive correlation between the competence judgments and the real election*

*outcomes. Surprisingly, attractiveness was correlated with losing elections, with the effect being driven by faces of candidates who looked politically incompetent yet personally attractive. Our findings have implications for future research on negative political communication, as they suggest that both threatening first impressions and fleeting impressions of attractiveness can harm a candidate's electoral chances.*

While the manipulation of candidates' facial images has always been a key strategy for national political campaigns (Schlesinger, Israel, & Frent, 1994), it has become ubiquitous now that modern technology has facilitated image manipulation and dissemination. Yet the actual effects of such manipulations remain poorly understood. A growing line of research within political science has argued for the effects of a candidate's visual image on voter evaluation. Here scholars have looked at nonverbal means of candidate presentation and have shown that a candidate's visual image can influence how potential voters evaluate a prospective candidate (Rosenberg, Bohan, McCafferty, & Harris, 1986; Rosenberg, Kahn, Tran, & Le, 1991; Rosenberg & McCafferty, 1987). The importance of a candidate's nonverbal image is of course not lost upon politicians, their advisors, and their consultants: for example, Nixon's "five-o'clock" shadow and refusal to use makeup in the 1960 debate with Kennedy is widely thought to have affected voter evaluations. More recently, the willingness of candidates to spend large sums of money on haircuts also signals the importance candidates place on their visual image.[1] The research from political science is consistent with other social and behavioral science research, which finds that rapid evaluations of faces influence social decisions (e.g., Blair, Judd, & Chapleau, 2004; Hamermesh & Biddle, 1994; Hassin & Trope, 2000; Montepare & Zebrowitz, 1998; Mueller & Mazur, 1996; Zebrowitz, Voinescu, & Collins, 1996).

A recent study (Todorov et al., 2005) tested the political impact of rapidly determined first impressions of facial images. They showed subjects paired pictures of unfamiliar Congressional political candidates and found a small but highly significant positive correlation between the subjects' judgments of political competence and the actual election outcomes of races involving these same political candidates.[2] The Todorov et al. (2005) results tie together recent work studying individual evaluation of facial images with work in political behavior and point to new directions for research designed to explain how factors like nonverbal images directly affect voter decisions in actual elections.

That subjects can make these personal assessments based on a rapid exposure to only a facial image of an individual is argued to derive from the same sorts of

---

[1] For a general discussion and history of presidential debates in the United States, see Jamieson and Birdsell (1990), and for a history of presidential advertising see Jamieson (1996). On the contemporary issue of candidates and their expensive haircuts, see coverage of Democratic presidential candidate John Edwards (http://www.msnbc.msn.com/id/18157456).

[2] A follow-up study by Ballew and Todorov (2007) finds the same for gubernatorial elections.

cues that help individuals in more general types of social encounters (Haxby, Hoffman, & Gobbini, 2000). Modern social cognition "dual-process" theories assert that some decision processes are quick and involve little cognitive effort, while other decision processes are more deliberative and cognitively demanding (Chaiken & Trope, 1999).[3] When individuals use only a facial image to make expeditious assessments about the traits of others, this conceptualizes the fast, intuitive portion of the dual-process decision-making model in which these rapid initial assessments can have lasting direct effects on decisions regarding the individual being evaluated (Todorov & Uleman, 2003; Winston, Strange, O'Douherty, & Dolan, 2002). It also seems clear that at least some social judgments based on faces can be made very rapidly, and quite likely, automatically. Willis and Todorov (2006) found in a nonpolitical study that judgments made at 100 milliseconds correlated highly with judgments with no time constraints. Bar, Maital, & Linz (2006), also in a nonpolitical study, found that people can discriminate between faces that appear threatening and faces that appear non-threatening after 38-millisecond exposure but not after 26 milliseconds masked exposure. Yet, they also found that after 38 milliseconds, subjects did not discriminate faces when judging intelligence, which is a dimension similar to competence; however, Todorov and Pakrashi (2007) found that people can make meaningful judgments of competence after only 33 milliseconds.

Here we take up two new questions: the possible effects on election outcomes of social judgments other than political competence, and the presentation and exposure time of the images on which those judgments are based. We expand upon the work of Todorov et al. (2005) by introducing two negative assessments and clarifying the subjects' instructions so that each assessment pertains to either a political or personal decision. It is well known from social psychology studies that people, when forming impressions, tend to weight negative information more strongly than positive (Fiske, 1980; Hamilton & Zanna, 1972; Kanouse & Hanson, 1972). Similarly, Kahneman and Tversky (1979, 1984) have shown the importance of loss aversion to decision making. Furthermore, political impressions formed on the basis of negative information tend to be more lasting and more resistant to change (Cobb & Kuklinski, 1997). We are interested in testing whether such negative assessments of candidates can have rapid, automatic effects on voters' judgments. For this we picked two negative traits—deceit and threat—that are

---

[3] There are many different dual process models; although they fundamentally agree that there are multiple systems for information processing, they disagree on the nature of and/or interaction between these processes. Our dual-process definition follows from Koriat and Levy-Sadot (1999). As with Cognitive-Experiential Self-Theory (Epstein & Pacini, 1999), they separate processing into the experiential system (automatic, immediate, effortless processing) and the rational system (intentional, slow, effortful processing). Koriat and Levy-Sadot argue that both the analytic and nonanalytic systems have a direct effect on decision making, and that the weight an individual attaches to each process depends upon the details of the particular decision. Other research in social and political cognition has referred to similar bifurcated strategies, distinguishing between "theory-driven" and "data-driven" information processing (e.g., Fiske & Taylor, 1991; Rahn, 1993).

prominent in the study of politics. We tested these alongside competence and attractiveness.[4]

Studies indicate that competence and honesty are the character attributes most frequently ascribed to political candidates by voters (McGraw, 2003), which makes sense in a representative democracy because the elected officials must be trusted to act as efficacious agents for their constituents (Bianco, 1994). Hence trust—and the lack thereof—do correlate with gaining and losing votes, respectively (Hetherington, 1999). In evaluations of trust, negative influences are as important as—or may outstrip—the positive. Slovic (1993) finds that negative, trust-destroying events are more noticeable and more heavily weighted in impression formation, which subsequently propagates or solidifies further distrust.

Furthermore, attempts to provoke fear are particularly relevant in negative campaign communications. Well-known examples of fear-inducing presidential television advertisements include Lyndon Johnson's "Daisy" advertisement in 1964 and the 2004 George Bush advertisement "Wolves."[5] At the same time, there has been surprisingly scant academic research on how these negative associations—often evoked by negative advertising—directly influence voters' decision making. While many studies in recent decades have focused on negative campaign communications (e.g., Ansolabehere, Iyengar, Simon, & Valentino, 1994; Jamieson, 1996; Mark, 2006; Niven, 2006), much of the debate in the research literature has centered specifically on whether or not negative images suppress or boost voter participation (e.g., Ansolabehere & Iyengar, 1995; Goldstein & Freedman, 2002).[6] In contrast, we study the direct impact of reflexive negative assessments on voter decision making, which, if they adversely affect electability, are salient for future study of how negative campaigning impacts election results. But the literature on negative political communication provides us with little basis to expect the extent to which election losers can be predicted by subjects' judgments of deceitfulness or threat made only from their first impressions of a candidates' image.

Also in this study, we allow for the precise experimental manipulation of the manner and timing of subject exposure to the candidate images by using the TED protocol (Kim, O'Doherty, & Shimojo, 2007), which shows the candidate pictures one at a time rather than contemporaneously, thus forcing an encoding of the face into working memory for the comparison. This protocol also allows subjects to see each picture multiple times before making an assessment.

---

[4] Competence replicates the Todorov et al. (2005) study. Perceived attractiveness has long been shown to affect trait attribution due to the "beautiful is good" stereotype, resulting in various social benefits (see Feingold 1992 for a meta-analysis).

[5] To view all of these candidate television advertisements, see "The Living Room Candidate: Presidential Campaign Commercials, 1952–2004," http://livingroomcandidate.movingimage.us/index.php.

[6] This is not an exhaustive review of the literature on negative campaigning. A recent review of studies of negative campaigning is in Niven (2006). Additional studies of the link between negative campaigning and voter turnout include Freedman and Goldstein (1999), Kahn and Kenney (1999), Lau and Pomper (2001), and Wattenberg and Brians (1999).

In our work, we presented subjects with a subset of the candidate image pairs from Todorov et al. (2005), using image pairs which were chosen to insure that the competing candidates were of the same gender and ethnicity, that the pictures were of similar resolution, and that the candidates were facing the camera. In four separate blocks, viewers were asked to choose the one face from each pair that they judged to more strongly exhibit each of four traits: attractiveness, competence, deceitfulness, and threat. The length of time each picture was shown was either 33 milliseconds or one second. As in Todorov et al. (2005), each of our image pairs consisted of candidates who ran against each other in actual Congressional elections. When we compared our subjects' chosen photographs in the laboratory to the actual election outcomes, we found that our subjects' assessments of competence, attractiveness, and threat were significant factors in predicting the real-world election winners.

## Research Design and Method

For our study, the subjects were 43 paid graduate and undergraduate students at the California Institute of Technology, all tested in the Social Science Experimental Laboratory (http://www.ssel.caltech.edu/info) during January and February 2007. Twenty-one subjects were given one-second exposure times, while 22 subjects were given 33-millisecond exposure times. Prior to the testing, subjects were asked their party affiliation (Democrat, Republican, or Independent) and asked to place their ideology on a 5-point (liberal to conservative) scale.

As stimuli, we used 30 image pairs of political candidates who ran against each other in actual House and Senate races in 2000, 2002, and 2004. These were a subset of the 695 candidate pairs used in Todorov et al. (2005). Because of their interest in facial appearances uncontaminated by prior knowledge, the Todorov et al. (2005) dataset had already excluded races with well-known candidates. For consistency, we only used pairs of images with similar resolutions in which both candidates were facing the camera maximally head-on and were of the same race and gender.[7] All of our images were cropped to $104 \times 147$ pixels and were displayed for either 33-ms or one-second durations on a 20-inch 16:9-aspect LCD monitor centered against a grey background inside a fixation rectangle with a black border of 15-pixel width. All 30 elections pitted a Democrat against a Republican, and the winners were evenly split between the parties, with each party winning 15 of the elections.

To ensure that our sample of 30 pictures was representative of the entire set of candidate pairs, we referenced the data recorded by Todorov et al. (2005) to compare their results from their analysis for our set of 30 pictures with those for their

---

[7] Both human and primate research suggest that neural representations in face-selective areas are viewpoint dependent (Desimone, Albright, Gross, & Bruce, 1984; Lee, Matsumiya, & Wilson 2006), which supports our use of a uniform viewpoint for our stimuli. For a literature review regarding the difficultly, behaviorally, of comparing a three-fourths-view face with a frontal face, see also Burke et al. (2007).

entire original dataset. Both our subset of 30 pairs and the entire dataset had the same competence-to-winning correlation (59%). However, the elections involving our pairs were slightly more competitive than in the complete dataset; the winners' mean victory margin was by 27% of the vote, as compared to the overall average victory margin of 33%. Twenty-three of the 30 winners (77%) were incumbents, and only two of the losers (7%) were incumbents. But 25 of the 30 races we used were for 2002 House seats. In the Todorov et al. (2005) dataset, 84% of 2002 House races were won by an incumbent, while only 2% were lost by an incumbent.

The subjects in our study were first given instruction screens that explained the task and stressed the importance of taking sufficient time to make accurate decisions. Every subject participated in four blocks, each asking the participant to make a specific judgment about the candidate pairs. Two of the judgments (deceitfulness and competence) were framed in a political context, while the other two (attractiveness and threat) were framed in a personal context. Specifically, the four questions were, "Your task is to decide which of the Candidates in each A-B pair appears [trait]," and the traits were:

1. "More COMPETENT TO HOLD CONGRESSIONAL OFFICE";
2. "More ATTRACTIVE TO YOU";
3. "More likely to LIE TO THE VOTERS";
4. "More likely to act in a PHYSICALLY THREATENING MANNER TO YOU."

These four assessment blocks were arranged in one of two orders—with attractiveness and competence first or second (counterbalanced), followed by deceitfulness third and threat fourth (fixed). On one trait at a time, subjects were asked to judge all 30 image pairs; the order of the 30 pairs and the order of the two pictures comprising each pair were randomized. As shown in Figure 1, each photo was accompanied by a circle to the left (if shown first) or right (if shown second) of the picture. The circle indicated the keyboard button ("1" for left and "0" for right) to be pressed in order to choose that picture.

Subjects were shown each pair of pictures one at a time, repeating, with an interimage interval of $2 \pm 1$ seconds; within each pair, the order that the two candidates were shown was counterbalanced among subjects. Image pairs continued repeating for up to 60 seconds until the subject chose a picture, and then moved on to the next image pair after $4 \pm 1$ seconds. This method effectively equates the stimulus judgments with respect to the confidence of the judgment for each subject. In both the one-second and 33-ms trials, 99.8% of the subjects answered within the allotted time limit, and the modal decision-making time was one cycle (i.e., viewing each candidate face once).[8]

---

[8] With the 33-ms exposure, subjects viewed each face an average of 1.6 times before making a decision (standard error = 0.02), as compared to 1.2 times with one-second exposure (s. e. = 0.01).
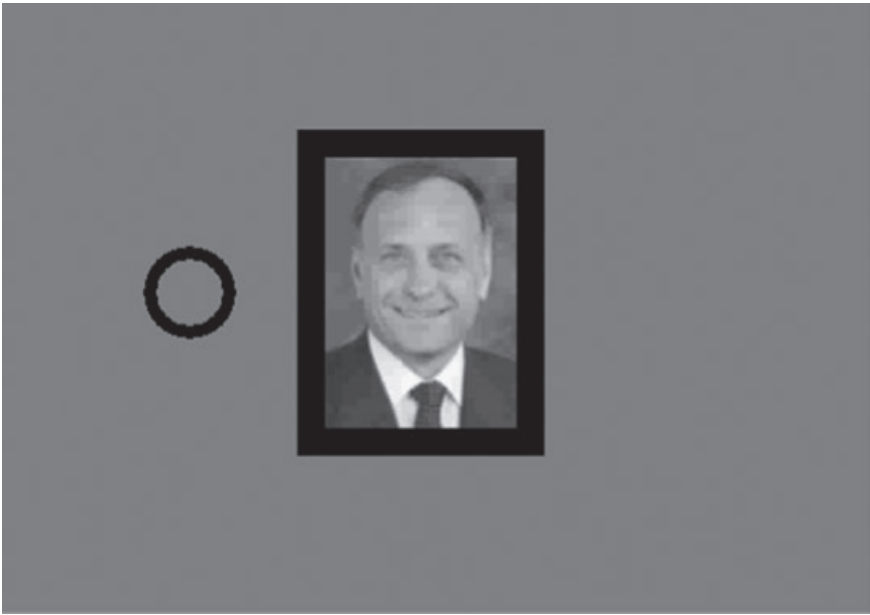
**Figure 1.** Sample Screen Shot.

As expected, some of the judgments were correlated (Table 1), although all correlations were less than 0.3, indicating that the four trait judgments we chose were relatively independent. As expected, competence and attractiveness, as well as deceit and threat, were positively correlated pairs of judgments. In general, correlations were somewhat stronger in the one-second treatment than the correlations in the 33-millisecond treatment, although both duration conditions produced very similar judgments. Across duration conditions, Cronbach's alpha (scale reliability) was 0.86 for attractiveness, 0.85 for threat and deceit, and 0.75 for competence.

## Results

We investigated whether these judgments from faces made in the laboratory have any direct relationship to the corresponding Congressional elections in which the politicians in fact participated.[9] For each image pair, we conducted logit regressions with the election winner (one of the two images of the pair) coded as

---

[9] We combined our results as four dependent draws and tested to insure that the results overall were significantly different from chance. The Kendall's W-test coefficient of concordance was 0.17, which corroborates the hypothesis that the underlying means differ; the range of this test is from 0 to 1, with "0" indicating complete disagreement across questions and "1" indicating complete agreement. Cochran's Q-test for the pooled data returned a p-value of 0.0001, which indicates that we can reject the null hypothesis that these four dependent draws have the same mean.

**Table 1.** Pearson Correlations (p-values are Bonferroni corrected)

**1 second Exposure**

|             | Competent | Attractive | Deceitful | Threatening |
|-------------|-----------|------------|-----------|-------------|
| **Attractive**  | 0.27      | 1.00       |           |             |
|             | (0.000)   |            |           |             |
| **Deceitful**   | −0.17     | −0.23      | 1.00      |             |
|             | (0.000)   | (0.000)    |           |             |
| **Threatening** | −0.22     | −0.26      | 0.29      | 1.00        |
|             | (0.000)   | (0.000)    | (0.000)   |             |

**33 ms Exposure**

|             | Competent | Attractive | Deceitful | Threatening |
|-------------|-----------|------------|-----------|-------------|
| **Attractive**  | 0.14      | 1.00       |           |             |
|             | (0.029)   |            |           |             |
| **Deceitful**   | −0.19     | −0.09      | 1.00      |             |
|             | (0.000)   | (0.1279)   |           |             |
| **Threatening** | −0.16     | −0.07      | 0.28      | 1.00        |
|             | (0.000)   | (0.4816)   | (0.000)   |             |

**Across Exposure Times**

|                    | Competent (33 ms) | Attractive (33 ms) | Deceitful (33 ms) | Threatening (33 ms) |
|--------------------|-------------------|--------------------|-------------------|---------------------|
| **Competent** (1 sec)   | 0.59    | 0.14    | 0.13    | −0.05   |
|                    | (0.019) | (1.000) | (1.000) | (0.000) |
| **Attractive** (1 sec)  | 0.11    | 0.76    | −0.05   | 0.00    |
|                    | (1.000) | (0.000) | (1.000) | (1.000) |
| **Deceitful** (1 sec)   | −0.42   | −0.26   | 0.76    | 0.65    |
|                    | (0.564) | (1.000) | (0.000) | (0.003) |
| **Threatening** (1 sec) | −0.71   | −0.38   | 0.39    | 0.77    |
|                    | (0.000) | (1.000) | (0.992) | (0.000) |

*Note:* None of the subjects participated in both treatments, so the across exposure time correlations were instead determined by comparing, for each image pair, the percentage of subjects choosing that picture.

the dependent variable, and the face chosen on our task (for each subject, each image pair, and each of the four attributes that was judged) as each of four independent variables. We replicated the analysis also using ordinary least-squares regression on the percentage of votes received, which produced qualitatively similar results. As Table 2 shows, in the pooled dataset, competence, attractiveness, and threat were highly significant predictors of election outcomes.[10] We also ran logit regressions that included a dummy variable for the 33-ms treatment (chi-square (5 DF) = 60.30) as well as its interaction with the other variables

---

[10] To test whether the effects of the other significant variables were distinct from that of competence, we also ran the same logit regression after orthogonalizing the four independent variables. This produced essentially identical results.

**Table 2.** Logit Regression Predicting Election Outcomes

|  | 33 ms Treatment | First Diff. | 1 second Treatment | First Diff. | Pooled Data | First Diff. |
|---|---|---|---|---|---|---|
| **Competent** | 0.602 | 0.144 | 0.544 | 0.125 | 0.583 | 0.141 |
|  | (0.001) | (0.044) | (0.004) | (0.042) | (0.000) | (0.031) |
| **Attractive** | −0.954 | −0.217 | −0.717 | −0.161 | −0.844 | −0.196 |
|  | (0.000) | (0.043) | (0.000) | (0.047) | (0.000) | (0.031) |
| **Deceitful** | 0.084 | 0.019 | 0.374 | 0.089 | 0.219 | 0.053 |
|  | (0.649) | (0.044) | (0.052) | (0.048) | (0.100) | (0.032) |
| **Threatening** | −0.312 | −0.075 | −0.521 | −0.120 | −0.423 | −0.101 |
|  | (0.091) | (0.044) | (0.008) | (0.043) | (0.002) | (0.032) |
| **Constant** | −0.599 |  | −0.753 |  | −0.671 |  |
|  | (0.002) |  | (0.000) |  | (0.000) |  |
| N | 655 |  | 625 |  | 1280 |  |
| Pseudo $R^2$ | 0.0485 |  | 0.0349 |  | 0.0404 |  |
| $\chi^2$ (4) | 38.94 |  | 26.59 |  | 63.21 |  |

*Note:* Numbers in parentheses are p-values for logit results, and standard errors for first difference results. Clustering standard errors for individual subject data yields nearly identical results.
*Note:* The modal subject choices were set so that picture 0 won the election approximately half of the time. So for example, in the pooled treatment, switching the competence choice from picture 0 to picture 1 changes picture 1's probability of winning: to 64.1% from 50%. Similarly, switching the attractiveness judgment from 0 to 1 changes 1's probability of winning: to 30.4% from 50%.

(chi-square (8 DF) = 65.32). In both cases a chi-square test against the pooled model (chi-square (4 DF) = 63.21) rejected the treatment-specific model in favor of the pooled model, confirming that the quality of subject decisions was not affected by the relatively short exposure time in the 33-ms treatment. We therefore focus on the pooled data.

## Effects of Individual Social Judgments on Election Outcomes

We estimated the change in the probability that a given image would be the election winner when a hypothetical subject switches her decision on a given judgment of that image (holding the other judgments constant).[11] We found that attractiveness had the largest effect here, followed by competence and threat (Table 2). For example, switching in favor of choosing a picture in the attractive-

[11] This is commonly known as a "first difference" estimate, and we use it here because the logit estimates reported in Table 4 are nonlinear and not easily interpreted. To produce a first difference estimate, we hold all of the right-hand-side variables to fixed values (which can be thought of as applying to a hypothetical subject) and compute the estimated probability that the particular image would be the election winner; we then change the value of one of the right-hand-side variables by a set amount and recomputed the same probability. The difference between these probabilities is the first difference estimate, and we also are able to compute an estimate of statistical uncertainty for that estimate. For further discussion of the first difference methodology, see King, Tomz, and Wittenberg (2000) and Tomz, Wittenberg, and King (2003).

ness judgment makes that candidate 20% *less* likely to win, as compared to 15% *more* likely after a competence switch and 10% less likely after a threat switch.

Table 3 makes election predictions in the 30 races by majority group judgment—that is, aggregating all of the subjects' choices for each social judgment, and comparing the majority winner from the experiment to the actual election winner. Here, the two strongest predictors are personal threat, which predicts 65% of election losers (p = 0.049) and personal attractiveness, which predicts 77% of election losers (p = 0.003). Interestingly, in the five open-seat elections, the lack of an incumbent did not hurt the prediction rate. Threat correctly predicts four of the five losers, attractiveness predicts all five losers, and competence predicts three of the winners.

However, because we used relatively few candidate pairs, for the remainder of the analysis we instead use individual associations—that is, the frequency in which individual choices matched the election outcomes—which yields $N = 1240$ observations for each judgment. This is provided in Table 4. Both the logit analysis (Table 2) and the percentage of the subjects' correct predictions (Table 4) corroborate that the judgments of competence, threat, and attractiveness each have significant effects on election outcomes. We next explore each of these three significant effects individually in more detail.

Our finding that competence judgments predict election outcomes (see again Table 2 and Table 4) replicates Todorov et al. (2005). We found that the individual competence judgments agreed with the election winner 55% of the time (p < 0.001, two-tailed binomial test against a mean of 50%), compared to 59% for the identical subset of stimuli taken from Todorov et al. (2005).

For threat judgments, by contrast, the face chosen as most threatening was the election loser in 55% of the individual trials (p-value of .000). Here, the personal nature of the fear question allows us to better interpret the results, because had we simply asked which of the candidates was more threatening, subjects could also have assessed a candidate as threatening for other reasons—because of the policies

**Table 3.** Predicting Election Winners from Majority Group Judgment

|  | 33 ms Treatment | 1 second Treatment | Pooled Data | Races Predicted Correctly (from pooled data) |
|---|---|---|---|---|
| **Competent** | 50% | 60% | 57% | 57% |
|  | (0.572) | (0.181) | (0.292) |  |
| **Attractive** | 35% | 28% | 23% | 77% |
|  | (0.049) | (0.008) | (0.003) |  |
| **Deceitful** | 53% | 45% | 48% | 52% |
|  | (0.428) | (0.292) | (0.707) |  |
| **Threatening** | 37% | 32% | 35% | 65% |
|  | (0.100) | (0.21) | (0.049) |  |

*Note:* P-values from a one-sided binomial test (against a mean of 50% and N = 30) are shown in parentheses.

**Table 4.** Percentage of Subjects Choosing Election Winners

| | 33 ms Treatment | 1 second Treatment | Pooled Data |
|---|---|---|---|
| **Competent** | 55% | 54% | 55% |
| | (0.031) | (0.010) | (0.001) |
| **Attractive** | 42% | 47% | 44% |
| | (0.000) | (0.201) | (0.000) |
| **Deceitful** | 49% | 52% | 51% |
| | (0.755) | (0.379) | (0.716) |
| **Threatening** | 46% | 44% | 45% |
| | (0.035) | (0.002) | (0.000) |

*Note:* P-values from a two-sided binomial test (against a mean of 50%) are shown in parentheses.

he would introduce if elected, for example. Instead, we asked subjects to gauge the risk of direct physical threat in order to elicit a rapid, automatic response of fear toward the images. And in fact, subjects' accuracy also depended on the duration of their decision-making time. When subjects were able to judge threat before the second cycle (i.e., choose after seeing each face exactly one time), they picked the election loser as more threatening 57% of the time. In comparison, slower threat decisions were no different from chance. After more than one cycle, only 49% of subjects chose the election loser. It is notable here that the images of politicians we used did not in fact exhibit any overt threat—they were all images of neutral or smiling people—yet even the forced judgment of threat in such otherwise friendly faces showed an association with election outcomes.

Deceit judgments provided no significant correlation with election outcomes, even when threat—to which deceit was positively correlated—was removed from the regression in Table 2. While facial appearance has previously been shown to affect perceived veracity (Masip, Garrido, & Herrero, 2004), and it is debated whether or not judgments of deception have a "kernel" of underlying truth about true measures of honesty (Bond, Berry, & Omar, 1994), there appears to be no effect between perceived deceptiveness and electability.

Attractiveness judgments yielded a counterintuitive result, in that they were negatively associated with election outcomes. In general, perceived physical attractiveness produces a halo effect that carries over to both positive trait attribution and positive social outcomes (Dion, Berscheid, & Walster, 1972; Kanazawa & Kovar, 2004). Furthermore, some previous work has shown that candidate "appearance" is strongly associated with election outcomes, with subjects being more likely to support candidates with "positive appearance" in mock elections (Rosenberg et al., 1986).[12] But in our study, the faces chosen as most attractive were actually far more likely to *lose* the election—56% of the chosen faces lost (p-value 0.000).

[12] Rosenberg et al. (1986) studied judgments of "appearance" and suggest that "appearance" is an aggregation of a number of other assessments, such as competence, integrity, and fitness for office.

**Table 5.** Interaction Between Significant Judgments: Percentage of Subjects Choosing
Election Winners

|                                       | Actual Success Rate | Predicted Success Rate (if additive) |
| ------------------------------------- | :-----------------: | :----------------------------------: |
| **Incompetent + Attractive**          | 37%                 | 39%                                  |
| **Incompetent + Threatening**         | 42%                 | 40%                                  |
| **Unattractive + Threatening**        | 37%                 | 39%                                  |
| **Incompetent + Attractive + Threatening** | 32%            | 34%                                  |

*Note:* Deceit is not included, as it had no significant interaction effect with any of the other
variables regarding the prediction of election outcomes.

Finally, we found no significant partisan or ideological effects. First, for all
four judgments, neither Democrats nor Republicans were more likely to choose
candidates from their own party, and Independents' choices matched those of their
partisan counterparts. Second, for all four judgments, liberals were not signifi-
cantly more likely than either moderates or conservatives to prefer Democratic
over Republican candidates. This is not surprising given that party labels were not
provided with the pictures, but confirms that subjects were not familiar with the
politicians. Third, for all four judgments, no group (i.e., Democrats, Republicans,
Independents, liberals, moderates, or conservatives) was significantly more likely
to choose election winners.

## Interactions of Attractiveness, Competence, Deceit, and Threat

We also checked for the cumulative effect of the subjects' judgments on their
predictive accuracy. Surprisingly, the three significant judgments (attractiveness,
competence, and threat) had additive effects on subjects' success rates in choosing
election winners. We had assumed a negative interaction effect between trait
judgments—a drop-off in subjects' prediction accuracy when combining these
judgments—largely because of the difficulty of the task, the unfamiliarity of the
politicians, the short exposure times, and the binary nature of the four decisions.[13]
For the latter, since every subject necessarily had to make at least two of these
three judgments in favor of one candidate, it seemed likely that this would dampen
the prediction accuracy of the combined judgments. However, that was not the
case, and Table 5 reports the cumulative prediction rates along with the expected
success rates from positing additive effects. It shows the strong relationship
between these three traits—whenever a subject judged a candidate to be attractive,

Thus, it is not clear that their results are about a candidate's attractiveness, as we have used the
concept in this study.

[13] However, we did not approach the study with the intent of predicting the existence (or extent) of
these interactions, and we believe that our data are limited for a strong statistical test of a formal
interaction. Here, we note their effects and suggest that future experimental studies should consider
these interactions in more detail.

threatening, and incompetent, only 32% of the time was that candidate an election winner, which is an extraordinarily large effect.

Table 5 also shows that when a subject judged a candidate as more attractive but less competent than the opponent, that candidate won only 37% of the time. This is another strong relationship as it is actually slightly *more* accurate than the additive expectation (39%), and thus suggests an interesting interaction between competence and attractiveness. Rather than counteracting incompetence, the incongruency of attractive yet incompetent looking candidates may make them even less likely to be elected—an interaction that we believe should be looked at more formally in future experimental studies. Note that attractiveness and competence judgments had a significantly positive correlation in our study (see Table 1), and in fact subjects chose the same face as both more competent and more attractive 60% of the time (57% under the 33-ms treatment and 64% under the one-second treatment). Yet, despite this correlation, competence and attractiveness yielded opposite predictions for election outcomes. For faces chosen as both more competent and more attractive, the results were more ambiguous, predicting election winners 49% of the time (47% at 33 ms, and 51% at one second).

What might explain the surprising association between attractiveness judgments and election outcomes? As suggested above, one possibility is that in politics, incompetence is augmented by attractiveness—at least by the rapid, fleeting attractiveness for which we found the largest effect. That the effect of attractiveness is context-specific—and sometimes detrimental—has been shown in the extant literature. Examples include studies of court verdicts (Piehl, 1977; Sigall & Ostrove, 1975; Thornton, 1977) and studies of the beauty penalty when building trust or setting performance expectations (Andreoni & Petrie, 2008; Solnick & Schweitzer, 1999; Wilson & Eckel, 2006).

The attractiveness judgments may also have been affected by the very short exposure times, because in our results, the duration of the exposure correlates with their predictive accuracy. The attractiveness judgments were significant for the 33-ms treatment, where the losers were chosen 59% of the time (p-value 0.000), but not in the one-second exposure time (53%, p = 0.20). But, within the one-second treatment, the faster attractiveness judgments (one-cycle trials) correlated with election loss (55% after one-cycle trials, compared to 49% after multiple-cycle trials).

Thus we suspect that in many of the quick attractiveness assessments (i.e., one-cycle decisions at one second and all decisions at 33 ms) subjects were judging something fundamentally different from what they were judging as attractiveness in the slower assessments. It could be that, perhaps especially with shorter exposure times, attractiveness might be confused with threat, since both are driven by high emotional arousal; candidates who elicit fleeting judgments of attractiveness might also be those who elicit more permanent feelings of threat and aversion when it comes to elections. Indeed, the emotional arousal caused by threat has been shown to have an automatic and positive effect on perceived attractiveness,

but only if the person is unaware of this connection (Foster, Witcher, Campbell, & Green, 1998). In line with this idea, it is interesting to note that Table 1 showed a negative correlation between attractiveness and threat for the one-second treatment, but no significant correlation for the 33-ms treatment. Furthermore, candidates judged to be both more attractive and more threatening were very likely to have lost elections (Table 5) and were much more likely to be seen as deceitful (58%, $\rho = 0.15$, p-value 0.001). In contrast, though attractive and incompetent-looking candidates could also conceivably be judged as dishonest and frightening, neither of these effects emerged in our study. When judging one candidate as both more attractive and less competent, subjects were not significantly more likely to choose that candidate as more deceitful ($\rho = 0.02$, p-value = 0.52) or as more threatening ($\rho = -0.03$, p-value = 0.54).

## Conclusion

To our knowledge, this is the first study in which laboratory subjects' judgments about negative traits imputed from images of unfamiliar politicians who have run against each other for office are used to predict the actual election outcomes involving pairs of these same candidates. In fact, the majority group results from the threat and attractiveness judgments (Table 3) could each be used to predict the winners of 65% and 77%, respectively, of the Congressional elections—though counterintuitively, the *less* attractive faces of each pair were the winners. It is important to stress that even though the images were shown for short durations (33 ms or one second), subjects produced judgments that correlated with which candidate wins the election; this we take as support for the idea that these assessments are quick and relatively effortless processes, as argued by Todorov et al. (2005).

However, while our results provide intriguing support for the hypothesis that subjects are using rapid evaluation processes as the dual-process model predicts, we in fact have much more to learn in general about the specific mechanisms through which these brief political and personal judgments are produced from simple facial cues and then carried forward to the process of voter decision making. We believe that using the new tools of social neuroscience will be especially effective in gaining a better understanding of how these evaluation processes underscore the nature of political decision making.[14]

[14] For example, studies using lesion or functional imaging, focusing on the facial images of political candidates, and the same positive and negative evaluations, would help delineate the mechanisms by which political faces are evaluated (Spezio et al., 2008). It would also help better understand the extent to which political decisions are reached through mechanisms that are similar to, or distinct from, the mechanisms by which individuals react to generic facial images (Adams, Gordon, Baird, Ambady, & Kleck, 2003; Adolphs, Tranel, Damasio, & Damasio, 1994; Adolphs et al., 2005). Furthermore, the TED protocol, which shows each candidate separately, is optimized for the analysis of fMRI results, as it enables the isolation of the subjects' response to each candidate throughout the course of each decision-making process (Kim et al., 2007).

Our work has broad implications for the study of political campaigns, especially negative campaigning, for which we offer an approach for gauging the importance of facial imagery that is different from that of recent literature. Even the relative perception of threat from faces that are not overtly threatening (but in fact smiling) is sufficient to bias election outcomes. And attempting to make a candidate look more attractive may well backfire if the candidate looks otherwise incompetent. What these findings suggest to us is that how we personally feel about a candidate may be as important as our objective judgment of political competence and that voters are neither easily fooled by fleetingly attractive faces nor by smiles that hide threat.

Furthermore, we know that the subjects are responding to the risk of direct personal threat. Yet, the images were not provided in a negative context, but instead were devoid of any threatening (or disturbing) voice-over, text, music, or supporting imagery. Future work should focus on how (and if) campaigns can choose images to evoke the fear response, and also the extent to which negative manipulations of a candidate's image, along with other negative information and sensory input, can accentuate this perception of personal threat, and thus affect not only electoral outcomes but also broader concerns like political participation and voter trust in government.

## ACKNOWLEDGMENTS

## REFERENCES

Adams, R. B., Gordon, H. L., Baird, A. A., Ambady, N., & Kleck, R. E. (2003). Effects of gaze on amygdala sensitivity to anger and fear faces. *Science, 300*, 1536.

Adolphs, R., Gosselin, F., Buchanan, T. W., Tranel, D., Schyns, P., & Damasio, A. R. (2005). A mechanism for impaired fear recognition after amygdala damage. *Nature, 433*, 68–72.

Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. (1994). Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature, 372*, 669–672.

Andreoni, J., & Petrie, R. (2008). Beauty, gender, and stereotypes: Evidence from laboratory experiments. *Journal of Economic Psychology, 29*, 73–93.

Ansolabehere, S., & Iyengar, S. (1995). *Going negative: How political advertisements shrink and polarize the electorate*. New York: Free Press.

Ansolabehere, S., Iyengar, S., Simon, A., & Valentino, N. (1994). Does attack advertising immobilize the electorate? *The American Political Science Review, 88*(4), 829–838.

Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the United States of America, 104*(46), 17948–17953.

Bar, M., Maital, N., & Linz, H. (2006). Very first impressions. *Emotion, 6*, 269–278.

Bianco, W. T. (1994). *Trust: Representatives and constituents*. Ann Arbor: University of Michigan Press.

Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science, 15*, 674–679.

Bond Jr., C. F., Berry, D. S., & Omar, A. (1994). The kernel of truth in judgments of deceptiveness. *Basic and Applied Social Psychology, 15*(4), 523–534.

Burke, D. et al. (2007). Are face representations viewpoint dependent? A stereo advantage for generalising across different views of faces. *Vision Research*, doi:10.1016/j.visres.2007.04.018.

Chaiken, S., & Trope, Y. (Eds.) (1999). *Dual process theories in social psychology*. New York: Guilford.

Cobb, M. D., & Kuklinski, J. H. (1997). Changing minds: Political arguments and political persuasion. *American Journal of Political Science, 41*(1), 88–121.

Desimone, R., Albright, D., Gross, G., & Bruce, C. (1984). Stimulus-selective properties of inferior temple neurons in the Macaque. *Journal of Neuroscience, 4*(8), 2051–2062.

Dion, K., Berscheid, E., & Walster, E. (1972). What is beautiful is good. *Journal of Personality and Social Psychology, 24*, 285–290.

Epstein, S., & Pacini, S. (1999). Some basic issues regarding dual-process theories from the perspective of cognitive-experiential self-theory. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 462–482). New York: Guilford.

Feingold, A. (1992). Good-looking people are not what we think. *Psychological Bulletin, 111*(2), 304–341.

Fiske, S. T. (1980). Attention and weight in person perception: The impact of negative and extreme behavior. *Journal of Personality and Social Psychology, 38*, 889–906.

Fiske, S. T., & Taylor, S. E. (1991). *Social cognition* (2nd ed.). New York: McGraw-Hill.

Foster, C. A., Witcher, B. S., Campbell, W. K., & Green, J. D. (1998). Arousal and attraction: Evidence for automatic and controlled processes. *Journal of Personality and Social Psychology, 74*(1), 86–101.

Freedman, P., & Goldstein, K. (1999). Measuring media exposure and the effects of negative campaign ads. *American Journal of Political Science, 43*(4), 1189–1208.

Goldstein, K., & Freedman, P. (2002). Campaign advertising and voter turnout: New evidence for a stimulation effect. *Journal of Politics, 64*, 721–740.

Hamermesh, D., & Biddle, J. (1994). Beauty and the labor market. *The American Economic Review, 84*, 1174–1194.

Hamilton, D. L., & Zanna, M. P. (1972). Differential weighting of favorable and unfavorable attributes in impressions of personality. *Journal of Experimental Results in Personality, 6*, 204–212.

Hassin, R., & Trope, Y. (2000). Facing faces: Studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology, 78*, 837–852.

Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Neuroscience, 4*, 223–233.

Hetherington, M. J. (1999). The effect of political trust on the presidential vote. *American Political Science Review, 93*(2), 311–326.

Jamieson, K. H. (1996). *Packaging the Presidency: A history and criticism of presidential campaign advertising*. New York: Oxford University Press.

Jamieson, K. H., & Birdsell, D. S. (1990). *Presidential debates: The challenge of creating an informed electorate*. New York: Oxford University Press.

Kahn, K. F., & Kenney, P. J. (1999). Do negative campaigns mobilize or suppress turnout? Clarifying the relationship between negativity and participation. *American Political Science Review, 93*(4), 877–889.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. *Econometrica, 47*, 263–291.

Kahneman, D., & Tversky, A. (1984). Choices, values, and frames. *American Psychologist, 39*, 341–350.

Kanazawa, S. & Kovar, J. L. (2004). Why beautiful people are more intelligent. *Intelligence, 32*, 227–243.

Kanouse, D. E., & Hanson, L. R. (1972). Negativity in evaluation. In E. E. Jones et al. (Eds.), *Attribution: Perceiving the causes of behavior* (pp. 47–62). Morristown, NJ: General Learning Press.

Kim, H., Adolphs, R., O'Doherty, J. P., & Shimojo, S. (2007). Temporal dissociation of neural processes underlying face preference decisions. *Proceedings of the National Academy of Sciences of the United States of America, 104*, 1823–1825.

King, G., Tomz, M., & Wittenberg, J. (2000). Making the most of statistical analyses: Improving interpretation and presentation. *American Journal of Political Science, 44*, 347–361.

Koriat, A., & Levy-Sadot, R. (1999). Processes underlying metacognitive judgments: Information-based and experience-based monitoring of one's own knowledge. In S. Chaiken & Y. Trope (Eds.), *Dual process theories in social psychology* (pp. 483–502). New York: Guilford.

Lau, R., & Pomper, G. M. (2001). Effectiveness of negative campaigning in U.S. Senate elections. *American Journal of Political Science, 46*(1), 47–66.

Lee, Y., Matsumiya, K., & Wilson, H. R. (2006). Size-invariant but viewpoint-dependent representation of faces. *Vision Research, 46*, 1901–1910.

Mark, D. (2006). *Going dirty: The art of negative campaigning*. New York: Rowman and Littlefield.

Masip, J., Garrido, E., & Herrero, C. (2004). Facial appearance and impressions of credibility: The effects of facial babyishness and age on person perception. *International Journal of Psychology, 39*(4), 276–289.

McGraw, K. M. (2003). Political impressions: Formation and management. In D. O. Sears, L. Huddy, & R. Jervis (Eds.), *Oxford handbook of political psychology* (pp. 394–432). New York: Oxford University Press.

Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: The salience and significance of age in social judgments. *Advances in Experimental Social Psychology, 30*, 93–161.

Mueller, U., & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces, 74*, 823–850.

Niven, D. (2006). A field experiment on the effects of negative campaign mail on voter turnout in a municipal election. *Political Research Quarterly, 59*(2), 203–210.

Piehl, J. (1977). Integration of information in the "courts": Influence of physical attractiveness on amount of punishment for a traffic offender. *Psychological Reports, 41*, 551–556.

Rahn, W. M. (1993). The role of partisan stereotypes in information processing about political candidates. *American Journal of Political Science, 37*, 472–496.

Rosenberg, S. W., Bohan, L., McCafferty, P., & Harris, K. (1986). The image and the vote: The effect of candidate presentation on voter preferences. *American Journal of Political Science, 30*(1), 108–127.

Rosenberg, S. W., Kahn, S., Tran, T., & Le, M. (1991). Creating a political image: Shaping appearance and manipulating the vote. *Political Behavior, 13*(4), 345–367.

Rosenberg, S. W., & McCafferty, P. (1987). The image and the vote: Manipulating voters' preferences. *Public Opinion Quarterly, 51*(1), 31–47.

Schlesinger, A. M., Jr., Israel, F. L., & Frent, D. J. (Eds.) (1994). *Running for president: The candidates and their images, 1789–1896*. New York: Macmillan.

Sigall, H., & Ostrove, N. (1975). Beautiful but dangerous: Effects of offender attractiveness and nature of crime on juridic judgment. *Journal of Personality and Social Psychology, 31*, 410–414.

Slovic, P. (1993). Perceived risk, trust, and democracy. *Risk Analysis, 13*, 675–682.

Solnick, S. J., & Schweitzer, M. E. (1999). The influence of physical attractiveness and gender on ultimatum game decisions. *Organizational Behavior and Human Decision Processes, 79*(3), 199–215.

Spezio, M. L., Rangel, A., Alvarez, R. M., O'Doherty, J. P., Mattes, K., Todorov, A., Kim, H., & Adolphs, R. A. (2008). A neural basis for the effect of candidate appearance on election outcomes. *Social Cognitive and Affective Neuroscience, 3*(4), 344–352.

Thornton, B. (1977). Effect of rape victim's attractiveness in a jury simulation. *Personality and Social Psychology Bulletin, 3*, 666–669.

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

Todorov, A., & Pakrashi, M. (2007). *Trait judgments from faces: Rapid, unreflective, and robust*. Unpublished manuscript. Princeton University.

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

Tomz, M., Wittenberg, J., & King, G. (2003). CLARIFY: Software for interpreting and presenting statistical results. Version 2.1. Stanford University, University of Wisconsin, and Harvard University. Available at http://gking.harvard.edu/.

Wattenberg, M. P., & Brians, C. L. (1999). Negative campaign advertising: Demobilizer or mobilizer? *American Political Science Review, 93*(4), 891–899.

Willis, J., & Todorov, A. (2006). First impressions: Making up your mind after 100-ms exposure to a face. *Psychological Science, 17*(7), 592–598.

Wilson, R. K., & Eckel, C. C. (2006). Judging a book by its cover: Beauty and expectations in the trust game. *Political Research Quarterly, 59*(2), 189–202.

Winston, J. S., Strange, B. A., O'Douherty, J., & Dolan, R. J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nature Neuroscience, 5*, 277–283.

Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). "Wide-eyed" and "crooked-faced": Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin, 22*, 1258–1269.