June 2008

Predicting Elections from Politicians' Faces

J. Scott Armstrong
University of Pennsylvania, armstrong@wharton.upenn.edu

Kesten C. Green
Monash University

Randall J. Jones Jr.
University of Central Oklahoma

Malcolm Wright
University of South Australia

Predicting Elections from Politicians' Faces

Abstract

Prior research found that people's assessments of relative competence predicted the outcome of Senate and Congressional races. We hypothesized that snap judgments of "facial competence" would provide useful forecasts of the popular vote in presidential primaries before the candidates become well known to the voters. We obtained facial competence ratings of 11 potential candidates for the Democratic Party nomination and of 13 for the Republican Party nomination for the 2008 U.S. Presidential election. To ensure that raters did not recognize the candidates, we relied heavily on young subjects from Australia and New Zealand. We obtained between 139 and 348 usable ratings per candidate between May and August 2007. The top-rated candidates were Clinton and Obama for the Democrats and McCain, Hunter, and Hagel for the Republicans; Giuliani was 9th and Thompson was 10th. At the time, the leading candidates in the Democratic polls were Clinton at 38% and Obama at 20%, while Giuliani was first among the Republicans at 28% followed by Thompson at 22%. McCain trailed at 15%. Voters had already linked Hillary Clinton's competent appearance with her name, so her high standing in the polls met our expectations. As voters learned the appearance of the other candidates, poll rankings moved towards facial competence rankings. At the time that Obama clinched the nomination, Clinton was ahead in the popular vote in the primaries and McCain had secured the Republican nomination with a popular vote that was twice that of Romney, the next highest vote-getter.

Predicting Elections from Politicians' Faces

J. Scott Armstrong,* The Wharton School, U. of Pennsylvania,
Philadelphia, PA 19104

Kesten C. Green, Business and Economic Forecasting Unit,
Monash University, Melbourne, Australia

Randall J. Jones, Jr, Department of Political Science,
University of Central Oklahoma, Edmond, OK 73034

Malcolm Wright, Ehrenberg-Bass Institute, University of South Australia

June 11, 2008r

Abstract

Prior research found that people's assessments of relative competence predicted the outcome of Senate and Congressional races. We hypothesized that snap judgments of "facial competence" would provide useful forecasts of the popular vote in presidential primaries before the candidates become well known to the voters. We obtained facial competence ratings of 11 potential candidates for the Democratic Party nomination and of 13 for the Republican Party nomination for the 2008 U.S. Presidential election. To ensure that raters did not recognize the candidates, we relied heavily on young subjects from Australia and New Zealand. We obtained between 139 and 348 usable ratings per candidate between May and August 2007. The top-rated candidates were Clinton and Obama for the Democrats and McCain, Hunter, and Hagel for the Republicans; Giuliani was 9th and Thompson was 10th. At the time, the leading candidates in the Democratic polls were Clinton at 38% and Obama at 20%, while Giuliani was first among the Republicans at 28% followed by Thompson at 22%. McCain trailed at 15%. Voters had already linked Hillary Clinton's competent appearance with her name, so her high standing in the polls met our expectations. As voters learned the appearance of the other candidates, poll rankings moved towards facial competence rankings. At the time that Obama clinched the nomination, Clinton was ahead in the popular vote in the primaries and McCain had secured the Republican nomination with a popular vote that was twice that of Romney, the next highest vote-getter.

Keywords: accuracy, appearance, forecasting methods, snap judgments

* Email: armstrong@wharton.upenn.edu; phone 610-622-6480

People often infer intelligence, character, and personality from appearance. Appearance can involve facial features, body build, clothing, and grooming. The idea that it is possible to know a person's character based on their looks—known as physiognomy—has its roots in the Middle Ages. Physiognomy was often frowned upon, and the Catholic Church banned its study in the 16th century (Berry and Brownlow, 1989).

Hiring people based on criteria that are irrelevant to job performance such as attractiveness, sex, religion, race, age, height, weight, or accent is widely regarded as unacceptable, and measures are sometimes taken to counter such biases. It has long been known that the most effective way to reduce bias in a typical hiring decision is to make the decision prior to seeing the candidate (Meehl 1965). Organizations can structure the hiring process to reduce or eliminate irrelevant information. For example, some orchestras audition players behind screens to reduce bias.

Lewis (2003) describes how the Oakland Athletics baseball team applies this method. In baseball, relevant performance statistics are easily available, so the General manager refused to meet with potential hires. This approach has been very successful.

In political campaigns, information that is irrelevant to a potential office-holder's job performance is difficult to avoid. An understanding of irrelevant influences on voters' decisions can increase awareness of potential bias and allow counter-measures. For example, video coverage of campaigns contains more irrelevant information than does print media coverage. Conscientious voters seeking information about candidates from debates should read transcripts of debates rather than view video footage. As a bonus, reading only takes half as long. Knowledge about irrelevant influences on voters also allows candidates to make adjustments to reduce biases against them and perhaps to assess their electoral prospects before deciding to run.

In this study, we looked at one bias that has been shown to affect voters: judgments of competence based on facial appearance.

Prior Research on Judgments Based on Facial Appearance

Two weeks before the 2004 U.S. general election, Todorov et al. (2005) presented facial photographs of candidates for the 32 Senate races to 127 subjects. Based on exposures of one second or less, the subjects rated how competent the Republican and Democratic Senate candidates looked, one pair at a time. If subjects recognized a candidate, those judgments were excluded. Competence ratings correctly predicted election results for 69% of the 32 races. Similar results were obtained for the 2004 U.S. House of Representatives elections, as well as for the 2000 and 2002 U.S. House and Senate elections. The findings of Todorov et al. would be reassuring if people who look more competent were more competent. However, we are unaware of any evidence that this is the case.

The relationship between facial appearance and voting was strong despite the many other influences on voting. Some of these are, like facial appearance, also irrelevant to job performance; for example, height, weight, dress, and accent. This would seem to leave a minor role for relevant factors such as political philosophy, positions on issues, and experience.

One explanation offered for the apparent controlling effect of facial appearance is that a candidate's babyfacedness (neoteny) is negatively related to perceptions of competence (Zebrowitz and Montepare 2005). This effect was confirmed by Nixon and Pollom (2006), although they found that it was reversed in the case of male voter perceptions of female candidates; that is, male voters preferred more babyfaced female candidates. Nixon and Pollom also found that attractiveness was not a factor influencing voter preference. In contrast, a study by Berggren, Jordahl and Poutvaara (2007) determined that beauty outperformed other traits, including competence, in explaining the share of votes in Finnish elections. Little et al. (2007) studied differences in facial features and concluded that they correlated with general election results; however, the study did not rate these features on characteristics such as competence. A second study reported in Little et al. (2007) examined the effect of a hypothetical wartime setting and found that this resulted in an increased share of votes for candidates whose faces appeared more masculine and dominant. Candidates whose faces appeared more attractive, forgiving, likeable, and intelligent performed better in hypothetical peacetime elections.

Study Design

We extended the findings on facial appearance to long-term forecasts for candidates in the U.S. Presidential primaries based on snap judgments of competence. To do this we asked subjects to rate each candidate's competence by looking briefly at a photograph of his or her face. We refer to the measure we obtained as "facial competence." We then assessed whether the ratings could be used to predict which candidates would receive the most votes for their party's nomination.

Our approach was similar to the one used by Todorov et al. (2005), as we used photographs of candidates and sought judgments of facial competence from people who did not know the identity of the candidates. However, we differed in that, whereas Todorov et al. made pairwise comparisons for a large number of individual elections, we instead sought ratings of facial competence for many candidates in a single race. We hypothesized that as voters become familiar with the candidates, relative support for the candidates would tend to move toward the relative facial competence ratings we had obtained at a much earlier date. Thus, facial competence ratings provide an early forecast of voting patterns; one that is especially valuable when voters are not already familiar with the candidates. Once the candidates become well known to voters, facial competence is incorporated into the polling information and likely offers no additional value as a predictor.

3

Candidates

We included all announced and potential candidates for the Democratic and Republican nominations of whom we were aware in May 2007. Some announced candidates later withdrew and some potential candidates never entered the race. However, by casting a wide net we made sure to include all feasible nominees. We use the term "candidates" to refer to all of the actual and potential candidates we included in our study. (The candidates are listed in Appendix B.)

Prior research shows that choice of clothing and the setting (e.g., pictured alongside an American flag) affect evaluations of people (Stuart & Fuller 1991). Therefore, we used photographs that included only the faces of the candidates. The task of finding neutral photos with uninformative backgrounds proved more difficult than we expected, and we needed to retouch the backgrounds in most of the photographs to create a standard appearance.

To guard against picking an unrepresentative picture, we selected what we judged to be the two most representative color photos of each candidate from among those we found. These photos were rotated so that participants saw one photograph of each candidate.

Ratings

Our raters were students in supervised classroom settings at Victoria University of Wellington, the University of South Australia, Wellington Girls College (a New Zealand high school), and the University of Central Oklahoma. From May through mid-August, 2007, the students were shown one photograph of each of the candidates. The order in which the photographs were presented to the raters was rotated using two orderings. The combinations of photographs and ordering gave four treatments.

Once candidates become known, familiarity is likely to influence judgments of competence. Thus, for each candidate we used ratings only from subjects who could not identify the candidate. We asked the raters if they recognized the people in the photos. If they said they did, we eliminated those responses even if they wrongly identified the candidate. As we anticipated, raters in Australia and New Zealand recognized fewer candidates than raters in the United States. However, some candidates, particularly Hillary Clinton, were quite often recognized.

Using our questionnaire (reproduced in Appendix A), the raters assessed each candidate's competence on a scale of 0 to 10, with zero denoting total incompetence and 10 the highest possible level of competence.

Findings on the Predictive Validity of Facial Competence Ratings

The findings are presented in Table 1 for Democratic candidates and in Table 2 for Republican candidates. The sample sizes ranged from 139 for Clinton to 348 for Tancredo.

The smaller sample sizes mean that more subjects were excluded because they said they recognized the candidates. In addition to Clinton, other widely recognized candidates (i.e., those with lower sample sizes) were John Kerry (n=199), Al Gore (n=208) and Barack Obama (n=225). In general, there was much higher recognition of the Democratic candidates than of the Republican candidates.

In columns 2 – 5 of both tables, we report the average ratings from each of the four samples of raters for each candidate. In the right-most column of each table are the overall average ratings for each candidate.

In general, the Democratic candidates had the highest facial competence ratings. The six highest rated candidates were all Democrats. Of the six, Clinton, with a 7.2 rating, had a clear lead. Clark was second at 6.9, closely followed by Obama at 6.8, Edwards at 6.5, and Gore at 6.4, and Dodd at 6.3. On the Republican side, McCain, Hunter, and Hagel were tied for first at 6.2. Paul, Romney, and Gilmore were close behind at 6.1.

Table 1: **Average facial competence ratings of Democratic candidates**

	USA	NZ Study One*	NZ Study Two	Australia	Total n	*All Raters*
Hillary Clinton	6.9	7.5	6.5	6.8	139	*7.2*
Wesley Clark	6.6	7.1	6.7	7.2	336	*6.9*
Barack Obama	6.1	7.2	6.5	6.5	225	*6.8*
John Edwards	6.4	6.6	6.4	6.8	295	*6.5*
Al Gore	6.0	6.7	5.9	6.4	208	*6.4*
Chris Dodd	5.9	6.5	6.2	6.9	334	*6.3*
Joe Biden	5.8	6.6	6.0	6.4	334	*6.2*
Dennis Kucinich	5.9	6.1	5.5	6.5	331	*6.0*
Mike Gravel	5.8	5.8	5.8	6.0	344	*5.8*
John Kerry	6.0	5.7	5.7	5.3	199	*5.7*
Bill Richardson	4.9	5.5	4.5	5.9	339	*5.2*

* Mostly high school girls

Table 2:   **Average facial competence ratings of Republican candidates**

	USA	NZ Study One	NZ Study Two	Australia	Total n	*All Raters*
John McCain	6.0	6.4	5.8	6.3	289	*6.2*
Duncan Hunter	6.0	6.4	6.0	6.3	342	*6.2*
Chuck Hagel	5.6	6.6	6.1	6.4	346	*6.2*
Ron Paul	5.7	6.6	5.5	6.5	346	*6.1*
Mitt Romney	6.0	6.2	5.7	6.3	309	*6.1*
Jim Gilmore	5.6	6.4	5.8	6.6	347	*6.1*
Newt Gingrich	5.6	6.1	5.7	6.2	310	*5.9*
Tommy Thompson	5.6	5.9	5.6	6.5	342	*5.8*
Tom Tancredo	5.5	5.9	5.4	6.5	348	*5.8*
Sam Brownback	5.4	6.0	5.2	6.5	342	*5.7*
Rudy Giuliani	5.2	5.9	5.3	5.8	263	*5.7*
Fred Thompson	5.2	5.6	5.2	5.6	309	*5.4*
Mike Huckabee	4.9	5.5	4.7	6.0	333	*5.3*

Facial competence and performance in the polls

After obtaining facial competence ratings of the candidates, we compared them with voter preferences measured by opinion polls (reported by PollingReport.com). We assessed the link between candidates' facial competences and their standings in the polls during two time periods: first, the week following Labor Day 2007, as the fall campaign in the early primary states was gaining increased candidate attention; second, winter and spring 2008, from the early January primary and caucus contests to the time when the party nominees were decided in March for the Republicans and June for the Democrats.

Poll Results – Period 1: September 2007

The first period observed was September 4-12, 2007, using polls of likely voters conducted during that period. For this analysis we omitted candidates for whom poll data were not reported. We also excluded polls that did not measure support for Al Gore who, although undeclared, was at that time an important potential candidate. Exclusion of these polls did not change the rankings of the other Democratic candidates. The data for nine Democratic and 11 Republican candidates are reported in Table 3. As can be seen from the polling results (as well as from the betting markets), there was considerable uncertainty as to which candidates would win the party nominations.

Table 3: Facial competence ratings compared to polling percentages, September 4-12, 2007

Democratic Nomination	Average Competence	Average Polling	Republican Nomination	Average Competence	Average Polling
Hillary Clinton	7.2	37.8	John McCain	6.2	14.9
Barack Obama	6.8	19.8	Duncan Hunter	6.2	1.0
John Edwards	6.5	15.0	Chuck Hagel	6.2	0.1
Al Gore	6.4	11.5	Ron Paul	6.1	1.0
Chris Dodd	6.3	0.8	Mitt Romney	6.1	9.0
Joe Biden	6.2	2.0	Newt Gingrich	5.9	4.9
Dennis Kucinich	6.0	1.9	Tom Tancredo	5.8	1.2
Mike Gravel	5.8	0.5	Sam Brownback	5.7	1.6
Bill Richardson	5.2	2.5	Rudy Giuliani	5.7	28.1
			Fred Thompson	5.4	21.9
			Mike Huckabee	5.3	3.7

We are grateful to Christopher Wlezien for assistance with these poll data.

For early September, Table 3 shows a strong association between perceived competence and standing in the polls for the Democratic candidates, but not for Republican candidates. One possible explanation is that people who participated in these early polls did not know what many of the Republican candidates looked like (as noted above). That is, they had not linked candidates' names with their appearance. Indeed, this is why we think that facial competence is better for long-term prediction than polls are. Thus, over time, we expect that the polling results will tend to move toward the same ordering as facial competence ratings if voters have the opportunity to become familiar with the candidates' appearances. A poorly funded campaign is less likely to provide this opportunity. For example, Hunter received the same facial competence rating as McCain but had only very modest funding for his campaign. Because his campaign was unable to afford much advertising, voters had little chance to become familiar with his competent appearance. This hampered his ability to attract voter support and hence his poll ratings remained low. On the other hand, Giuliani's facial competence score was low and he eventually had to quit the race despite the fact that he spent the second-highest amount of money of any Republican candidate. Giuliani's large advertising expenditures may have been counterproductive as the public became *more* aware of his appearance.

Poll Results – Period 2: Spring 2008

The second period, most of which coincided with voting in primary elections and caucuses, was a time of change in candidates' standings in the polls as people began to recognize the candidates. As evident in data from PollingReport.com, this change began in mid-December among the Republicans and in early January for the Democrats. Among Republicans, leader Giuliani began a steep decline from about

30% in November, to the low 20s in December, to less than 15% in January. Similarly, support for Thompson's , who ranked second in early Republican polls also declined precipitously. Since these candidates had lower appearance ratings, their drop in the polls was consistent with the facial competence thesis.

Also consistent with the facial competence hypothesis was the increase in standings in the polls of two candidates with high competence scores. In January, McCain's poll numbers had nearly doubled from the mid-teens to 29%. By mid-February, McCain's poll standing exceeded 50%, as he became the presumptive Republican nominee. Poll numbers for Romney, whose competence rating was only slightly lower than McCain's, increased from about 10% to 15%.

Changes in poll standing among other Republican candidates were generally consistent with their facial competence scores. However, due to an early win in the primaries, Huckabee gained in the polls despite the fact that his 5.3 competence rating was the lowest among all Republican candidates. He rose from less than 10% at the end of November to more than 18% in December and January and then to 27% in February. However, the increased visibility meant that more people related his name with his facial appearance, and his popularity leveled off. Huckabee withdrew on March 4.

For the Democratic candidates, Clinton's poll numbers remained constant in the low-to-mid 40s. However, Obama's standing rose markedly in January, from the low-20s to the low 30s, and then to the mid-40s by February. By March, Obama had surpassed Clinton in the polls despite his somewhat lower 6.8 competence rating (compared to Clinton's 7.2). Edwards, whose 6.5 competence rating was lower than that of either Clinton or Obama, failed to break out of the low teens in the polls and, consistent with the competence hypothesis, dropped out of the race on January 30.

Prediction Market

Intrade launched markets for candidates in the Republican primaries on October 25, 2007. Traders assessed McCain's chances of winning the nomination as less than 5% until January 2008 when his contract began a steady increase in value. In contrast, the market's assessment of Giuliani's probability of becoming the Republican candidate was about 70% from November to January after which time it began a steady decline. In the end, McCain, whose facial competence rating was higher, prevailed in the market, while Giuliani, who had a low rating, faded. Traders who bet early on the candidate with the highest facial competence would have made money.

Among leading Democratic candidates, Clinton was the early leader in the Intrade nomination market, with her contract fluctuating mostly between probabilities of 0.4 and 0.5 from early 2005 until mid-2007 when the contract value began rising to a peak of more than 70 in fall 2007. However, after Obama won the Iowa caucuses in early January 2008, the Clinton contract dropped precipitously to

about 30. Though it rose again briefly, it then dropped to below 30 after Super Tuesday in early February and never recovered. The history of Obama's contract value is roughly the converse of Clinton's. After languishing around 30 and below through the end of 2007, Obama's contract spiked up to 70 after winning in Iowa, dropped to the 30-40 range, and then began a long ascent to more than 90, where it was in late May. As with McCain, traders who recognized early that Obama's contract price was too low given his high facial competence would have made money.

Clinton versus Obama result

Facial competence ratings predicted that Clinton would defeat Obama in the popular vote count of the primaries and caucuses. At the time that Obama clinched the nomination, Clinton was slightly ahead in the popular vote.[1] McCain had already secured the Republican nomination as he had 47% of the popular vote vs. 22 for Romney and 20 for Huckabee.

   We wondered whether the difference in the advertising budgets of the candidates made the result closer than was indicated by the 7.2 versus 6.8 competence ratings of the candidates. In the two weeks following Super Tuesday, Obama outspent Clinton on advertising by 3.7 to 1.0, and then continued to outspend her by 1.6 to 1.0 for the rest of the campaign.[2] This matches the time at which Obama overtook Clinton in the polls.

<div align="center">

Discussion

</div>

These findings imply that political parties should select competent-looking candidates to increase their chances of winning elections and implementing their policies. As with marketing of commercial products, packaging counts.

   Conversely, voters should ignore the appearance of candidates so they can choose a candidate based on policies and record, rather than potentially misleading snap judgments of appearances. This is difficult for major elections, of course. One way to focus on the candidates' positions is to read about debates rather than watching them. In addition, voters can use the web sites that report on candidates' positions. Impartial sources that are trying to help voters should not provide pictures of the candidates for political office.

---

[1] See http://www.realclearpolitics.com/epolls/2008/president/democratic_vote_count.html
[2] http://wiscadproject.wisc.edu/wiscads_pressrelease_060208_table2.pdf

Conclusions

This study provides is a successful extension of the study by Todorov, et al. (2005). Snap judgments of facial competence by unbiased raters provided useful predictions of the popular vote winners for the Republican and Democratic nominees about one year in advance. This supports prior research showing that voters are heavily influenced by the irrelevant information provided by people's facial appearances. This information was not fully incorporated into early polling results because voters lacked recognition of the facial appearances of many of the candidates. People who used this information would have made money by participating in betting markets prior to the time that the candidates became familiar to the voters.

References

Berggren, Niclas, Henrik Jordhal & Panu Poutvaara (2007), "The looks of a winner: beauty, gender and electoral success," CESinfo Area Conference, Munich, 20-22, April 2007.

Knutson, K. M, J. N. Wood, M. V. Spampinato & J. Grafman (2006), "Politics on the brain: An fMRI investigation", *Social Neuroscience*, 1 (1), 25-40.

Lewis, Michael (2003), *Moneyball*. New York: W.W. Norton.

Little, Anthony, C., Robert P. Burriss, Benedict, C. Jones & S. Craig Roberts (2007), "Facial appearance affects voting decisions", *Evolution and Human Behavior*, 28, 18-27.

Meehl, Paul E. (1965), "Seer over sign: The first good example", *Journal of Experimental Research in Personality*, 1, 27-32.

Nixon, Becky & Sarah Pollom (2006), "Effects of facial maturity on voting preference," Hanover College Working Paper, PSY 401, Winter 2006.

Stuart, Elnora W. & Barbara K. Fuller (1991), "Clothing as communication in two business-to-business sales settings," *Journal of Business Research*, 23, 264-290.

Todorov, Alexander, Anesu N. Mandisodza, Amir Gorem & Crystal C. Hall (2005), "Inferences of competence from faces predict election outcomes," *Science*, 308, June 10, 1623-1626.

Zebrowitz, Leslie A. & Joann M. Montepare (2005), "Appearance DOES matter," *Science,* 308, June 10, 1565-1566.

Acknowledgments

Appendix A

Questionnaire on political candidates (sample)

Please look at the following pictures. Assume that each person is a political candidate. Rate each of them from 0 to 10 as to how competent they look. We are interested in your first impressions, so please work quickly.

Rate the candidate's competence

1. How competent is the person in the photo? [____] {0…10}
{0 = completely incompetent… 10 = extremely competent}

2. Do you recognize the person in the photo?   ☐ Yes
                                                 ☐ No

3. If "Yes", who is it?   [_____]

Appendix B

Candidates for 2008 Presidential Nomination Included in this Study
(Announced and Unannounced)

Democratic Candidates

Joe Biden – Senator from Delaware
Wesley Clark – retired General, U.S. Army; former NATO commander
Hillary Clinton – Senator from New York
Chris Dodd – Senator from Connecticut
John Edwards – former Senator from North Carolina
Al Gore – former Vice President and former Senator from Tennessee
Mike Gravel – former Senator from Alaska
John Kerry – Senator from Massachusetts
Dennis Kucinich – Congressman from Ohio
Barack Obama – Senator from Illinois
Bill Richardson – Governor of New Mexico; former UN Ambassador

Republican candidates

Sam Brownback – Senator from Kansas
Jim Gilmore – former Governor of Virginia
Newt Gingrich – former Speaker of House of Representatives and Congressman from Georgia
Rudy Giuliani – former Mayor of New York City
Chuck Hagel – Senator from Nebraska
Mike Huckabee – former Governor of Arkansas
Duncan Hunter -- Congressman from California
John McCain – Senator from Arizona
Ron Paul – Congressman from Texas
Mitt Romney -- former Governor of Massachusetts
Tom Tancredo – Congressman from Colorado
Fred Thompson – former Senator from Tennessee and current TV actor
Tommy Thompson -- former Governor of Wisconsin