Facial appearances and electoral success: Are trustworthy-looking politicians more successful in corrupt regions?
Jaeger, Bastian; Evans, Anthony; van Beest, Ilja

Document version:
Early version, also known as pre-print

Publication date:
2019

Link to publication

Citation for published version (APA):
Facial Appearance and Electoral Success:
Are Trustworthy-Looking Politicians More Successful In Corrupt Regions?
Bastian Jaeger, Anthony M. Evans, and Ilja van Beest
Tilburg University

Draft version: 6 November 2019
Word count: 6,441
This paper is currently undergoing peer review. Comments are welcome.

Author Note
Bastian Jaeger, Anthony M. Evans, and Ilja van Beest, Department of Social Psychology, Tilburg University, The Netherlands.
We thank Angelo Romano and Nils Köbis for their valuable comments.
Correspondence concerning this article should be addressed to Bastian Jaeger, Department of Social Psychology, Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands. E-mail: bxjaeger@gmail.com.
Abstract

People rely on the facial appearance of political candidates when voting. Here, we examine whether perceptions of competence, trustworthiness, or attractiveness are associated with electoral success in the 2016 Italian local elections. In line with situational leadership theory, we also test whether trait preferences for politicians vary as a function of election context. Specifically, we examine if trustworthy-looking politicians are more successful in regions where political corruption is a salient issue. We analyze electoral data of 150 mayoral candidates from 75 constituencies to test if the association between perceived trustworthiness and electoral success is stronger in Southern Italy, where corruption is more prevalent. Across three preregistered studies ($N = 470$), perceived competence and perceived trustworthiness were not associated with electoral success. Moreover, the influence of trustworthiness perceptions on electoral success did not differ between Southern Italy and the rest of the country. Instead, we found that attractive-looking politicians were more successful. A one standard deviation increase in perceived attractiveness corresponded to a 2.98 percentage point increase in vote share and a 1.91 times increase in the odds of victory. In sum, while our results support the general notion that facial appearance correlates with electoral success, we do not find evidence that corruption moderates the success of trustworthy-looking politicians.

Keywords: trait impressions, voting, corruption, trustworthiness, attractiveness, competence
Facial Appearance and Electoral Success:

Are Trustworthy-Looking Politicians More Successful In Corrupt Regions?

The functioning of democratic political systems requires citizens to elect capable leaders. However, voting decisions are complex and voters often rely on heuristics, simplified decision strategies that require fewer cognitive resources (Quattrone & Tversky, 1988). While some heuristics can lead to accurate inferences under conditions of limited knowledge (e.g., inferring a candidate’s stance on policy issues by their party affiliation), other strategies are less justifiable (Kuklinski & Quirk, 2000). For instance, even though trait impressions based on facial appearance are rarely accurate (Olivola & Todorov, 2010b; Todorov & Porter, 2014; but see Lin, Adolphs, & Alvarez, 2018), they predict the electoral success of political candidates (Olivola & Todorov, 2010). In other words, people rely on appearance-based trait impressions when voting.

A host of studies has shown that competent-looking politicians enjoy more political success (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Todorov, Mandisodza, Goren, & Hall, 2005). However, some results suggest that voters are also influenced by other apparent traits, such as the perceived attractiveness, dominance, or sociability of candidates (Berggren, Jordahl, & Poutvaara, 2010; Castelli, Carraro, Ghitti, & Pastore, 2009; Chen, Jing, & Lee, 2014; Jäckle, Metz, Wenzelburger, & König, 2019). This raises the question whether effects of specific traits systematically vary across different contexts. Evidence from controlled lab experiments provide initial support for this notion. For example, Little and colleagues (2012) showed that framing a hypothetical election as taking place during a time of war or peace influenced participants’ preference for trustworthy-looking or attractive-looking candidates. Specifically, attractive-looking leaders were favored more strongly during a time of war, while trustworthy-looking leaders were favored more strongly during a time of peace, suggesting that which (apparent) traits are favored in politicians may be influenced by the political context in which an election is taking place. However, it is thus far unclear whether the moderating role of election context generalizes to other contextual frames, and whether this effect influences real-world elections.

Here, we analyze results of the 2016 Italian local elections to examine the influence of facial appearance on voting behavior. First, we test whether the perceived attractiveness, competence, or trustworthiness of candidates is related to their electoral success. Second, we investigate whether the salience of a political issue—the regional prevalence of institutional
corruption—moderates the association between trait perceptions and electoral success. Specifically, we hypothesize that voters are more motivated to elect a trustworthy leader when corruption is a salient issue. As a consequence, trustworthy-looking candidates should be more successful in regions where corruption is more prevalent (e.g., in Southern Italy vs. the rest of the country; Linhartova & Pultarova, 2015).

**Election Context Moderates Trait Preferences**

When asked directly which personality traits a politician should possess, voters primarily mention competence (Miller, Wattenberg, & Malanchuk, 1986; Sussman, Petkova, & Todorov, 2013). In line with this explicit preference, Todorov and colleagues (2005) found that appearance-based impressions of competence, but not impressions of trustworthiness, likeability, or attractiveness, are associated with success in elections for the US Senate and House of Representatives. The notion that voters rely on the facial appearance of candidates to make voting decisions was supported by many subsequent studies, which investigated the relationship between facial appearance and political success in a wide range of countries and electoral systems (for a review, see Olivola & Todorov, 2010a). While most studies found that competent-looking politicians are more successful (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Castelli et al., 2009; Sussman et al., 2013), impressions of other traits also predict electoral success under certain conditions. For example, Berggren and colleagues (2010) found a positive effect of attractiveness on voting behavior in Finland (for similar results in the United States, see Jäckle et al., 2019). In other studies, electoral success was related to perceptions of dominance (Chen et al., 2014; Sussman et al., 2013), sociability (Castelli et al., 2009), or gender-typicality (Hehman, Carpinella, Johnson, Leitner, & Freeman, 2014).

To account for these findings, researchers have started investigating how the context in which an election is taking place influences the association between specific trait impressions and electoral success. For example, cross-cultural data suggests that competence-related traits are more predictive of electoral success in Western societies (e.g., the United States) than in East Asian societies (e.g., Japan or Taiwan), whereas the opposite pattern holds for trustworthiness-related traits (Chen, Jing, Lee, & Bai, 2016; Rule et al., 2010). Next to cultural differences, trait preferences may also vary as a function of the political context in which an election is taking place. In general, voters may prefer different traits in political leaders depending on which political issues are particularly salient. This idea follows from situational leadership theory,
which stress that leader selection is context-sensitive, with different leader types being favored depending on which tasks they are expected to perform (Epitropaki & Martin, 2004; Hollander & Julian, 1969; Olivola, Eubanks, & Lovelace, 2014; Yukl, 1989).

Following this reasoning, Little and colleagues (2014; 2007; 2012) demonstrated that participants’ hypothetical voting behavior can be influenced by manipulating the political context of an election. They found that participants had a stronger preference for individuals whose facial trustworthiness had been digitally enhanced in a time of peace, whereas individuals whose facial attractiveness had been digitally enhanced were more strongly favored in a time of war. This pattern suggests that prosocial traits (i.e., trustworthiness) are favored in leaders when the political context is characterized by collaboration, but traits related to health and formidability (i.e., attractiveness) are favored when the political climate is characterized by conflict (for similar results, see Ferguson et al., 2019; Laustsen & Petersen, 2015; Spisak, Dekker, Kru¨ger, & Van Vugt, 2012).

The influence of war vs. peace frames on hypothetical votes provides initial evidence for the context-sensitive nature of face-based leader choice in the political domain. It also suggests that impressions of trustworthiness—a trait which should be highly desirable in a politician (Miller et al., 1986)—may influence voting behavior under some (but not all) conditions. It is unclear though (a) whether the effect of political context generalizes to issues other than a country’s state of war or peace and (b) whether it extends to real-world elections. Here, we examine if the salience of a different political issue—the prevalence of institutional corruption—moderates which trait perceptions predict electoral success. Corruption is a recurring issue for political systems around the world and considerable resources are devoted to monitoring and diminishing corrupt practices (Jain, 2001). It is also a salient issue for voters as corruption charges lead to a substantial loss in votes (Peters & Welch, 1980; Welch & Hibbing, 1997). Building on these observations, we hypothesize that voters are more motivated to elect a trustworthy candidate when corruption is a salient issue. As a consequence, trustworthy-looking politicians should be more successful in constituencies with high levels of institutional corruption.

The Current Studies

In the current investigation, we attempt to replicate the finding that voters rely on trait impressions from faces when deciding whom to elect. Specifically, we test which traits are
associated with electoral success and whether the effect of trait impressions on voting varies as a function of political context. To this end, we examine the effect of facial appearance on electoral success in the 2016 Italian local elections. We focus on Italy because Italy exhibits large regional differences in the prevalence of corruption, with substantially higher levels in the south (Del Monte & Papagni, 2007; Fiorino, Galli, & Petrarca, 2012; Linhartova & Pultarova, 2015). We therefore test whether trustworthy-looking politicians are more successful in Southern Italy (compared to the rest of the country).

In all three studies, we measured the perceived competence, trustworthiness, and attractiveness of 150 mayoral candidates from 75 constituencies where two-candidate elections took place. We used ratings on the three trait dimensions to predict (a) the winner of the election and (b) the margin of victory. To probe the robustness of our results, we varied whether trait impressions were obtained from American (Study 1) or Dutch participants (Studies 2 and 3). We also varied whether trait perceptions were assessed with a single trait item (e.g., ratings of trustworthiness; Studies 1 and 3) or with multiple trait items (e.g., ratings of trustworthiness, honesty, and fairness for measuring trustworthiness; Study 2). Finally, we varied whether participants indicated their trait impressions by selecting the candidate scoring higher on a given trait dimension in a two-alternative forced-choice format (Study 1) or by rating all candidates sequentially on a continuous scale (Studies 2 and 3).

Methods

The studies were preregistered and all data, preregistration documents, and analysis scripts are available at the Open Science Framework (https://osf.io/jdqn2/). We report how our sample sizes were determined, all data exclusions, and all measures in the studies.

---

1 We recruited participants from outside of Italy to ensure that most, if not all, participants are unfamiliar with the political candidates. In general, there is substantial cross-cultural agreement in trait ratings from faces (Cunningham, Roberts, Barbee, & Druen, 1995; Rule et al., 2010). Moreover, previous studies have shown that, for example, trait ratings by American participants predict election outcomes in Bulgaria (Sussman et al., 2013) and trait ratings by Germany participants prediction election outcomes in the United States (Jäckle et al., 2019).

2 For Study 1 and Study 2, we preregistered to conduct multilevel regression analyses to account for the fact that individual candidates are nested within different municipalities. However, this analysis is not suitable given the dyadic structure of the data in which one candidate’s vote share and election outcome is perfectly mirrored by the other candidate’s vote share and election outcome. We therefore chose to follow a different analysis strategy to account for the dependencies in our data (see the Results section for more details) and conducted a third study for which the correct analyses were specified a priori.
Participants. For each trait dimension, ratings from at least 29 independent raters were collected, as previous studies have shown that this provides relatively stable average ratings (Hehman, Xie, Ofosu, & Nespoli, 2018). We asked participants at the end of each study whether they had recognized any of the individuals that were shown in the photos and, in case they answered affirmatively, who they had recognized. While some participants claimed to have recognized at least one candidate (Study 1: 6.12%, Study 2: 8.64%, Study 3: 14.29%), none provided correct names or mentioned the fact that the depicted individuals are Italian politicians.3

Study 1. Participants were 160 workers from Amazon Mechanical Turk (Paolacci & Chandler, 2014) who completed the study in return for $1. Thirteen participants (8.13%) who failed an attention check at the end of the study were excluded, leaving a final sample of 147 participants (43.54% female; $M_{age} = 32.07$, $SD_{age} = 8.25$). On average, candidates were rated by 47 participants ($Min = 43$, $Max = 54$) on each of the three trait dimension (competence, trustworthiness, and attractiveness).

Study 2. Participants were 223 Dutch undergraduate psychology students from Tilburg University who participated in return for partial course credit. Three participants (1.35%) who provided the same response across all trials were excluded, leaving a final sample of 231 participants (76.92% female; $M_{age} = 20.18$, $SD_{age} = 2.31$). The final sample size was based on the number of students that participated in the study within two weeks. On average, candidates were rated by 31 participants ($Min = 29$, $Max = 33$) on each of the three trait dimension.

Study 3. Participants were 93 Dutch undergraduate psychology students from Tilburg University who participated in return for partial course credit. One participant (1.08%) who provided the same response across all trials was excluded, leaving a final sample of 92 participants (49.45% female; $M_{age} = 20.87$, $SD_{age} = 2.22$). The final sample size was based on the number of students that participated in the study within two weeks. On average, candidates were rated by 30 participants ($Min = 30$, $Max = 31$) on each trait dimension.

Materials. We retrieved the results of the 2016 Italian local elections. Residents of cities with a population greater than 15,000 could directly vote for different mayoral candidates in a multi-candidate two-round system. For the current analysis, we focused on 126 constituencies in which no candidate received the absolute majority in the first round. In that case, the two

---

3 Among the people that were purportedly recognized were the actor James Franco, the boxer Nassem Hamad, and a participant’s dentist.
candidates who received the most votes competed in a second round which was held two weeks later. In line with our preregistered exclusion criteria, elections with at least one female candidate (36 elections, 28.57%) were excluded to remove the confounding role of gender (Chiao, Bowman, & Gill, 2008). Next, images of the candidates were downloaded from the internet. We selected photos in which candidates faced the camera with their faces being completely visible. For most candidates, we selected the photo from their election poster, as this was the photo that most voters were exposed to prior to the election. If the election poster could not be retrieved, another photo was selected. The wide majority of candidates were smiling in their photos and we tried to ensure that differences in affective expression between the two candidates were minimal. If one candidate showed a broad smile while the other looked neutral and no other photos could be found for the latter, then the election was excluded from analysis (15 elections, 16.67%). This resulted in a final sample of 75 elections with a total of 150 candidates. The photos were converted to grayscale, cropped so that only the candidate’s face and hair were visible, and resized to a height of 300 pixels. For each election, we recorded which of the two candidates won and their margin of victory, which constituted our dependent variables. We also recorded and whether the candidate was the incumbent or running against the incumbent and whether the constituency was located in the south (n = 34). Southern Italy encompasses the administrative regions of Abruzzo, Apulia, Basilicata, Campania, Calabria, Molise, and Sicily.

**Procedure.** To measure candidates’ perceived competence, trustworthiness, and attractiveness, participants, who were unaware of the context of the study and the identity of the people shown in the photos, evaluated all candidates on one specific trait dimensions. Each participant rated the candidates on only one trait in order to avoid consistency effects in ratings (Penton-Voak, Pound, Little, & Perrett, 2006).

**Study 1.** In Study 1, binary trait ratings on three dimensions were collected. Participants were randomly allocated to one of three conditions which determined whether they would rate the candidates’ competence, trustworthiness, or attractiveness. They saw the 75 pairs of candidates in a random order and were asked to select the candidate that looks more competent, trustworthy, or attractive depending on the condition. The percentage of participants who selected a given candidate as scoring higher than his opponent served as our measure of perceived competence, trustworthiness, and attractiveness.
**Study 2.** In Study 2, trait ratings were assessed with Likert scales. Participants were randomly allocated to one of the seven trait conditions: Perceived competence was measured via ratings of competence, capability, and intelligence, perceived trustworthiness via ratings of trustworthiness, honesty, and fairness, and perceived attractiveness via ratings of attractiveness. The 150 photos were rated in a random order on a 9-point scale ranging from not at all [trait] (1) to extremely [trait] (9). We calculated the intraclass correlation coefficients (ICC) as a measure of inter-rater consistency (Shrout & Fleiss, 1979). Participants showed significant consensus in their ratings (competence: ICC(2, 1) = .089, capability: ICC(2, 1) = .110, intelligence: ICC(2, 1) = .176, trustworthiness: ICC(2, 1) = .106, honesty: ICC(2, 1) = .111, fairness: ICC(2, 1) = .151, attractiveness: ICC(2, 1) = .331; all ps < .001). A confirmatory factor analysis indicated that a three-factor structure adequately fit ratings on the seven trait dimensions: $\chi^2 = 882.54$ ($df = 21$), RMSEA = 0.87, SRMR = 0.36, CFI = 0.98. Therefore, ratings of competence, capability, and intelligence were averaged to form a competence score; ratings of trustworthiness, honesty, and fairness were averaged to form a trustworthiness score; and ratings of attractiveness constituted a candidate’s attractiveness score. We created relative trait scores per election by subtracting the runner-up’s trait score from the winner’s trait score. In other words, each candidate’s trait scores reflected their perceived trustworthiness, competence, or attractiveness relative to their opponent.

**Study 3.** In Study 3, trait ratings were assessed with Likert scales. Participants were randomly allocated to one of three conditions which determined whether they would rate candidates’ competence, trustworthiness, or attractiveness. The 150 images were rated in a random order on a 9-point scale ranging from not at all [trait] (1) to extremely [trait] (9). We calculated the intraclass correlation coefficient as a measure of inter-rater consistency. Participants showed significant consensus in their ratings (competence: ICC(2, 1) = .113, trustworthiness: ICC(2, 1) = .123, attractiveness: ICC(2, 1) = .222; all ps < .001). For each candidate, ratings on the three dimensions were averaged across all participants and this served as our measure of perceived competence, trustworthiness, and attractiveness. We again created relative trait scores per election by subtracting the runner-up’s trait score from the winner’s trait score.

**Analysis plan & sensitivity analysis.** All trait scores were z-standardized to allow for comparisons between the studies. In all three studies, we tested for the effects of facial
appearance on election outcomes by predicting in separate models (a) the winner of the election and (b) the margin of victory with candidates’ perceived facial competence, trustworthiness, and attractiveness. We also tested whether the effect of trait perceptions varied as a function of the geographical location of a constituency (south vs. rest of the country). We control for incumbency status of the two candidates in all regression analyses as incumbents often have the advantage over political challengers (Cox & Katz, 1996). All analyses were conducted in R (R Core Team, 2018).

For each effect of interest, a sensitivity analysis was conducted to determine the minimum effect size we were able to detect with 80% power (and \( \alpha = 0.05 \)). As software commonly used for sensitivity analyses, such as G*Power (Faul, Erdfelder, Lang, & Buchner, 2007), does not support dyadic data, we relied on the simr package in R (Green & Macleod, 2016). The package does not provide a function specifically designed for conducting sensitivity analyses. However, it can provide estimates of observed power for coefficients in regression models. For each of our models, we systematically varied the effect size for all effects of interest and computed observed power. Performing power calculations across a range of effect sizes allowed us to determine the minimum effect size at which our model had at least 80% power to detect a significant effect.

Regarding the effects of perceived competence, trustworthiness, and attractiveness on the percentage of received votes (i.e., the winner’s margin of victory), analyses showed that we had 80% power to detect an increase of 2.10 percentage points, 2.19 percentage points, and 2.08 percentage points, respectively. We had 80% power to detect a difference of 2.14 percentage points for the interaction effect between perceived trustworthiness and geographical location of the constituency. Regarding the effects of perceived competence, trustworthiness, and attractiveness on the likelihood of a candidate’s success, analyses showed that we had 80% power to detect odds ratios of 1.60, 1.69, and 1.62, respectively. We had 80% power to detect an odds ratio of 1.59 for the interaction effect between perceived trustworthiness and geographical location of the constituency.

Results

The average margin of victory was 7.84 percentage points with a median of 5.72 percentage points (\( SD = 6.54, Min = 0.14, Max = 25.05 \)). Twenty-five elections (33%) featured an incumbent. For each study, we computed correlations between candidates’ perceived
attractiveness, competence, and trustworthiness and incumbency status. A random-effects meta-analysis across the three studies showed that incumbency status was positively correlated with competence ratings, $r = .10, p = .029$. There were no significant correlation between incumbency status and attractiveness ratings, $r = -.07, p = .12$, or trustworthiness ratings, $r = .06, p = .21$. Trustworthiness ratings were moderately correlated with competence ratings, $r = .42, p < .001$, and attractiveness ratings, $r = .47, p < .001$. There was no significant correlation between competence ratings and attractiveness ratings, $r = .07, p = .33$.

We also conducted an exploratory analysis of Google Trends data to test if political corruption is a more salient issue in Southern Italy. Google Trends provides access to the number of search queries for specific terms across different time frames and geographical locations (Choi & Varian, 2012). We recorded the number of searches that contained the word “corruption” (in Italian) across different Italian regions for four time windows: one month, three months, six months, and twelve months prior to the election. For each time window, the number of searches were rescaled to range from 0 to 100. Data for two southern regions was unavailable for the one-month time window. Corruption-related searches were more prevalent in southern regions (vs. the rest of the country) one month prior to the election (south: $M = 42.36, SD = 23.65$, rest: $M = 24.54, SD = 9.47$), $t(24.67) = 3.39, p = .002, d = 0.90$, three months prior to the election (south: $M = 58.18, SD = 14.48$, rest: $M = 41.59, SD = 14.24$), $t(70.01) = 4.98, p < .001, d = 1.15$, six months prior to the election (south: $M = 64.50, SD = 13.04$, rest: $M = 53.85, SD = 15.37$), $t(72.95) = 3.25, p = .002, d = 0.75$, and twelve months prior to the election (south: $M = 64.82, SD = 14.86$, rest: $M = 56.27, SD = 14.10$), $t(68.94) = 2.54, p = .013, d = 0.59$. These results lend support to our assumption that corruption is a more salient issue in the south of Italy.

**Margin of victory.** First, we examined whether candidates’ facial appearance predicted the margin of victory (i.e., candidates’ relative vote share). We estimated OLS regression models in which vote share was simultaneously regressed on candidates’ perceived attractiveness, competence, and trustworthiness. Due to the dyadic structure of our data, for any given election, one candidate’s data (e.g., their vote share, their relative attractiveness) always perfectly mirrored their opponent’s data. To account for this dependency, we randomly selected one candidate from each election and conducted our analyses on this sample of 75 candidates. However, results of these analyses vary depending on the specific combination of winners and runner-ups that are sampled. Therefore, we selected and analyzed 100,000 random samples of 75
candidates that included one candidate from each election. We calculated mean estimates for our predictors across all randomly drawn samples. This bootstrapping procedure was performed for each study. Finally, the results of the three studies were aggregated in a random-effects meta-analysis (see Figure 1).

Results showed that, across the three studies, perceived attractiveness positively predicted vote share, $\beta = 2.977$, $SE = 0.765$, $z = 3.89$, $p < .001$. Surprisingly, perceived competence negatively predicted vote share, $\beta = -1.676$, $SE = 0.745$, $z = 2.25$, $p = .025$. Candidates who scored one standard deviation higher on attractiveness received 2.98 percentage points more votes whereas candidates who scored one standard deviation higher on competence received 1.68 percentage points fewer votes. We did not find evidence that perceived trustworthiness was related to vote share, $\beta = 0.102$, $SE = 0.818$, $z = 0.12$, $p = .90$.

We also examined the influence of regional variation in corruption by estimating a second model in which we added an interaction effect between the geographical region in which an election took place (dummy-coded 0.5 for the south and -0.5 for all other regions) and candidates’ perceived trustworthiness. This interaction effect was significant but in the opposite direction of our hypothesis (i.e., trustworthiness had less of an impact in the south), $\beta = -2.907$, $SE = 1.379$, $z = 2.12$, $p = .034$. Perceived trustworthiness was negatively associated with vote share in the south, $\beta = -3.851$, $SE = 1.281$, $z = 3.01$, $p = .003$, and positively associated with vote share in the north, $\beta = 3.114$, $SE = 1.048$, $z = 2.97$, $p = .003$. We also explored whether associations between vote share and perceived attractiveness or competence differed between Southern Italy and the rest of the country. There was no significant interaction effect between region and perceived attractiveness, $\beta = 0.190$, $SE = 1.412$, $z = 0.13$, $p = .89$, but the interaction effect between region and perceived competence was significant, $\beta = 4.361$, $SE = 1.359$, $z = 3.21$, $p = .001$. Perceived competence was positively associated with vote share in the south, $\beta = 2.832$, $SE = 1.406$, $z = 2.01$, $p = .044$, but negatively associated with vote share in the north, $\beta = -4.356$, $SE = 0.851$, $z = 5.12$, $p < .001$.

The pattern of results did not change when analyzing median estimates, instead of mean estimates. Detailed statistics for each study can be found in the Supplemental Materials.

We also regressed vote share on each trait dimension separately. This again yielded a positive effect of perceived attractiveness, $\beta = 2.823$, $SE = 0.681$, $z = 4.15$, $p < .001$, and no effect of perceived trustworthiness, $\beta = 0.751$, $SE = 0.697$, $z = 1.08$, $p = .28$. The effect of perceived competence was negative but only marginally significant, $\beta = -1.295$, $SE = 0.703$, $z = 1.84$, $p = .066$. 

---

4 The pattern of results did not change when analyzing median estimates, instead of mean estimates. Detailed statistics for each study can be found in the Supplemental Materials.

5 We also regressed vote share on each trait dimension separately. This again yielded a positive effect of perceived attractiveness, $\beta = 2.823$, $SE = 0.681$, $z = 4.15$, $p < .001$, and no effect of perceived trustworthiness, $\beta = 0.751$, $SE = 0.697$, $z = 1.08$, $p = .28$. The effect of perceived competence was negative but only marginally significant, $\beta = -1.295$, $SE = 0.703$, $z = 1.84$, $p = .066$. 

In sum, candidates’ perceived attractiveness and competence, but not their perceived trustworthiness, predicted their vote share with attractive-looking candidates receiving more votes and competent-looking candidates receiving fewer votes. The association between perceived trustworthiness and vote share significantly differed between Southern Italy and the rest of the country. However, the observed pattern was opposite to our prediction: More trustworthy-looking politicians received fewer votes in the south, but more votes in the north.

**Figure 1.** The influence of winners’ facial appearance on the margin of victory. The graph displays the results of the three studies (starting with Study 1 on the left) and the meta-analytic estimates. In a first model, vote share was regressed on candidates’ perceived trustworthiness, competence, and attractiveness while controlling for incumbency status. In a second model, an interaction term between the region in which an election took place (dummy-coded 0.5 for the south and -0.5 for all other regions) and perceived trustworthiness was added.

**Electoral success.** Next, we examined whether candidates’ facial appearance predicted their likelihood of winning the election. We estimated logistic regression models in which election outcome (0 = candidate lost, 1 = candidate won) was regressed on the candidates’ perceived attractiveness, competence, and trustworthiness. We followed the same bootstrapping
CORRUPTION, FACIAL APPEARANCE, AND ELECTORAL SUCCESS

procedure as described before and the results of the three studies were again aggregated in a random-effects meta-analysis (see Figure 2).

Results showed that, across the three studies, perceived attractiveness predicted electoral success, $\beta = 0.647$, $SE = 0.179$, $z = 3.62$, $p < .001$. Candidates who scored one standard deviation higher on attractiveness were 1.91 times more likely to win their election. We did not find evidence that perceived competence, $\beta = -0.260$, $SE = 0.162$, $z = 1.61$, $p = .11$, or perceived trustworthiness were related to electoral success, $\beta = -0.100$, $SE = 0.178$, $z = 0.56$, $p = .58$.

We also examined the influence of regional variation in corruption on the predictive power of perceived trustworthiness. A second model was estimated in which we added an interaction effect between the geographical region in which an election took place (dummy-coded 0.5 for the south and -0.5 for all other regions) and candidates’ perceived trustworthiness. This interaction effect was not significant, $\beta = -0.505$, $SE = 0.310$, $z = 1.63$, $p = .10$. We also explored whether associations between electoral success and perceived attractiveness or competence differed between Southern Italy and the rest of the country. There were no significant interaction effects between region and perceived attractiveness, $\beta = -0.140$, $SE = 0.331$, $z = 0.42$, $p = .67$, or region and perceived competence, $\beta = 0.450$, $SE = 0.317$, $z = 1.41$, $p = .16$.

In sum, candidates’ perceived attractiveness, but not their perceived competence or trustworthiness, predicted their likelihood of winning the election with attractive-looking candidates being more successful. Crucially, we did not find evidence that perceived trustworthiness is more predictive of electoral success in Southern Italy compared to the rest of the country.

---

6 We also regressed election outcome on each trait dimension separately. This again yielded a positive effect of perceived attractiveness, $\beta = 0.544$, $SE = 0.154$, $z = 3.52$, $p < .001$, no effect of perceived competence, $\beta = -0.223$, $SE = 0.142$, $z = 1.57$, $p = .12$, and no effect of perceived trustworthiness, $\beta = 0.076$, $SE = 0.138$, $z = 0.55$, $p = .58$. 
Figure 2. The influence of candidates’ facial appearance on their likelihood of winning the election. The graph displays the results of the three studies (starting with Study 1 on the left) and the meta-analytic estimates. In a first model, election outcome was regressed on candidates’ perceived trustworthiness, competence, and attractiveness while controlling for incumbency status. In a second model, an interaction term between the geographical region in which an election took place (dummy-coded 0.5 for the south and -0.5 for all other regions) and perceived trustworthiness was added.

**Exploratory robustness checks.** We conducted several exploratory analyses to probe the robustness of our results. First, we re-ran our regression models with additional control variables. We included population size, voter turnout, and dummy variables indicating whether the candidate was a member of the Five Star Movement or running against one. The Five Star Movement is a recently established party whose political agenda includes a strong anti-establishment and anti-corruption stance (Mosca, 2014). Perceived attractiveness still positively predicted vote share, $\beta = 2.678, SE = 0.769, z = 3.48, p < .001$, and there was no effect of perceived trustworthiness, $\beta = -0.609, SE = 0.809, z = 0.75, p = .45$. The effect of perceived competence was no longer significant, $\beta = -0.928, SE = 0.754, z = 1.23, p = .22$. There was no significant interaction effect between region and perceived trustworthiness on vote share, $\beta = 0.001, SE = 3.248, z < 0.01, p > .99$. Moreover, perceived attractiveness still positively predicted electoral success, $\beta = 0.599, SE = 0.191, z = 3.14, p = .002$, but perceived competence, $\beta = -0.181, SE = 0.180, z = 1.01, p = .31$, and perceived trustworthiness, $\beta = -0.122, SE = 0.193, z =$
0.63, \( p = .53 \), did not. There was also no significant interaction effect between region and perceived trustworthiness on electoral success, \( \beta = 0.007, SE = 0.757, z = 0.01, p = .99 \).

Second, our analyses included two municipalities located on the island of Sardinia. While Sardinia is sometimes treated as a separate region altogether (e.g., Bigoni, Bortolotti, Casari, Gambetta, & Pancotto, 2016), it is similar to the island of Sicily and the southern part of mainland Italy—both of which were coded as being part of the south—in regards to the prevalence of corruption (Fiorino et al., 2012). We therefore recoded the two Sardinian municipalities as belonging to the south (i.e., the region where we expected to find stronger effects of perceived trustworthiness). The interaction effect between region and perceived trustworthiness was only marginally significant for predicting the vote share, \( \beta = -2.653, SE = 1.382, z = 1.92, p = .055 \), and not significant for predicting the likelihood of success, \( \beta = -0.314, SE = 0.308, z = 1.02, p = .31 \).

Finally, we analyzed Google Trends data for a more fine-grained analysis of how regional variation in corruption salience influence the success of trustworthy-looking candidates. Thus, instead of including a dummy variable for region (south vs. rest of the country), we included a variable indicating the relative frequency of Google searches that included the word “corruption” in the specific region (ranging from 0 indicating no searches to 100 indicating the number of searches in the region with the highest search frequency). We analyzed four time windows: one month, three months, six months, and twelve months prior to the election. We found no significant interaction effect between the number of Google searches and perceived trustworthiness on vote share (one-month window: \( \beta = -0.047, SE = 0.065, z = 0.72, p = .47 \), three-month window: \( \beta = -0.036, SE = 0.054, z = 0.68, p = .50 \), six-month window: \( \beta = 0.025, SE = 0.057, z = 0.44, p = .66 \), twelve-month window: \( \beta = 0.004, SE = 0.054, z = 0.07, p = .94 \)) or electoral success (one-month window: \( \beta = 0.009, SE = 0.016, z = 0.56, p = .58 \), three-month window: \( \beta = -0.005, SE = 0.012, z = 0.45, p = .66 \), six-month window: \( \beta = 0.009, SE = 0.012, z = 0.68, p = .49 \), twelve-month window: \( \beta = -0.009, SE = 0.012, z = 0.73, p = .47 \)).

**General Discussion**

The current set of studies investigated the association between appearance-based trait impressions and voting decisions. Analyzing data from the 2016 Italian local elections, we found that more attractive-looking candidates received more votes and were more likely to win their election. The size of this relationship was not trivial. Candidates who scored one standard
deviation higher on perceived attractiveness received 2.98 percentage points more votes and were 1.91 times more likely to win. To put the increase in vote share in perspective, 22 of the 75 elections we studied here were decided by fewer than 3 percentage points. We found no evidence that the perceived trustworthiness of candidates was related to their electoral success. Our results are therefore in line with previous studies showing that attractive politicians are more successful (Berggren et al., 2010; Jäckle et al., 2019; King & Leigh, 2009; Rosar, Klein, & Beckers, 2008), whereas perceived trustworthiness seems to be unrelated to electoral success (Berggren et al., 2010; Todorov et al., 2005).

The results for perceived competence were more ambiguous. Contrary to previous studies (Ballew & Todorov, 2007; Todorov et al., 2005), we found a negative effect of perceived competence on vote share. However, the effect was smaller than the minimum effect size we were able to detect with 80% power, making a false positive result more likely. Exploratory analyses also showed that the effect was not robust to controlling for additional variables (e.g., turnout) and we found no significant effect of competence on the likelihood of success. We therefore conclude that the present results do not provide clear support for the notion that perceived competence is positively or negatively related to electoral success.

While our findings replicate the effect of facial appearance on voting behavior, our main aim was to test predictions from situational leadership theory, which emphasizes the context-specific nature of leader selection (Epitropaki & Martin, 2004; Hollander & Julian, 1969; Olivola et al., 2014; Yukl, 1989). To this end, we examined whether the effect of trait perceptions were moderated by the political context in which the elections took place. Framing an election as taking place during a time of war vs. a time of peace has been shown to influence participants’ preferences for attractive-looking and trustworthy-looking candidates in a simulated voting environment (Little et al., 2012). Here, we tested if a different political issue, the prevalence of institutional corruption, influences voters’ trait preferences in real life elections. We reasoned that trustworthy-looking politicians would be more successful Southern Italy where corruption is more prevalent than in the rest of the country and voters may be more motivated to elect a trustworthy leader.

Across three studies, we found no evidence that trustworthy-looking candidates were more successful in Southern Italy. Some analyses even revealed the opposite pattern: Perceived trustworthiness was negatively related to vote share in Southern Italy, but positively related to
vote share in the rest of the country. However, this pattern of results did not emerge when including additional control variables or when analyzing candidates’ likelihood of winning an election. We also analyzed the frequency of Google searches that included the term corruption as a proxy for the salience of political corruption. Search queries were more frequent in the south, suggesting that corruption is indeed a more salient issue in Southern Italy, but the salience of corruption (as measured with Google searches) did not moderate the electoral success of trustworthy-looking politicians. Overall, the present results do not provide support for the hypothesis that trustworthy-looking candidates are more successful in regions where political corruption is more salient.

In fact, we did not find any evidence that apparent trustworthiness positively affected vote share or the likelihood of electoral success. This may seem surprising as morality judgments are a strong predictor of overall person evaluations (Brambilla, Rusconi, Sacchi, & Cherubini, 2011; Goodwin, Piazza, & Rozin, 2014) and people name morality-related traits such as honesty and incorruptibility as traits that a politician should possess (Miller et al., 1986; Sussman et al., 2013). Previously reported null results for the effect of perceived trustworthiness on electoral success may have been due to the fact that the majority of studies were conducted in countries with relatively low levels of institutional corruption, such as Finland (Berggren et al., 2010) and the United States (Todorov et al., 2005). In these countries, voters may be less concerned with electing a potentially corrupt leader and place more weight on other traits such as competence. However, investigations in countries with higher levels of corruption such as Bulgaria (Sussman et al., 2013) or Italy (Castelli et al., 2009) did not find a positive relationship between trustworthiness-related traits (e.g., morality, honesty/incorruptibility) and electoral success either. The current results are in line with these findings and suggest that the apparent trustworthiness of political candidates is not associated with their electoral success, even in countries where levels of institutional corruption are relatively high.

Limitations and Future Directions

Given the higher levels of corruption in Southern Italy, some candidates in our sample might have been accused of (or even directly involved in) corrupt practices. Even if voters in the south are more concerned with electing a trustworthy candidate, corruption allegations probably constitute a stronger indicator of a candidate’s trustworthiness, overriding any effect of facial appearance. In other words, given the prevalence of corruption in Italy, voters might be more
knowledgeable about the trustworthiness of candidates because of their (alleged) link to corrupt practices and rely on this knowledge, rather than appearance-base impressions, when making voting decision. Future studies could circumvent this issues by investigating the effect of corruption under more controlled conditions in the lab. Following the procedure of Little and colleagues (2012), it could be tested whether participants vote more often for individuals whose perceived trustworthiness was digitally enhanced when institutional corruption is made salient.

An alternative interpretation of the current null results is that trait preferences for politicians are in fact stable and that the same trait perceptions affect voting decisions across different contexts. However, this view cannot explain why different traits have been linked to electoral success in previous studies (Berggren et al., 2010; Castelli et al., 2009; Chen et al., 2014; Todorov et al., 2005).

Another possibility is that context only moderates explicitly stated preferences. Trait judgments from faces occur spontaneously and quickly (Klapper, Dotsch, van Rooij, & Wigboldus, 2016; Willis & Todorov, 2006) and voters may be unaware of their influence. This suggests that the effect of trait impressions on voting may not be susceptible to voters’ context-specific leader preferences. We do not think that this explanation for the current null results is likely though, as there is ample evidence showing that the effect of facial appearance on decisions varies across different contexts. For example, preferences for dominant-looking partners are stronger when intergroup conflict is made salient (Hehman, Leitner, Deegan, & Gaertner, 2015). In the political domain, cultural differences (Chen, Jing, & Lee, 2012; Rule et al., 2010) and the political knowledge of voters (Berggren, Jordahl, & Poutvaara, 2017; Lenz & Lawson, 2011) have been shown to moderate how much voters rely on the facial appearance of candidates when making voting decisions. Previously mentioned work by Little and colleagues (2014; 2007, 2012) also demonstrates that leaders with different facial appearance are favored when an election is framed as taking place during a time of war or peace. In sum, these findings suggest that contextual factors can moderate the effect of trait perceptions on decision-making. More studies are needed to explore under what conditions election context moderates trait preferences in politicians.

Going beyond the political domain, there is ample evidence showing that trait impressions impact decision-making in various domains such as criminal sentencing (Wilson & Rule, 2015), personnel selection (Gomulya, Wong, Ormiston, & Boeker, 2017), and consumer
behavior (Jaeger, Sleegers, Evans, Stel, & van Beest, in press). However, there is often conflicting evidence regarding which trait impressions people rely on. For example, Ert and colleagues (2016) analyzed Airbnb prices in Stockholm and found that people favor trustworthy-looking, but not attractive-looking hosts. The opposite pattern was found in a larger set of listing in New York City (Jaeger et al., in press). Duarte and colleagues (2012) found that trustworthy-looking, but not attractive-looking loan applicants were more likely to be funded, whereas Ravina and colleagues (2008) did find an effect of attractiveness. More research is needed to understand which trait perceptions from faces predict different real-world outcomes. Future studies should consider a wide range of traits and test for potential moderating factors.

It is also plausible that different traits are preferred at different stages of the decision-making process. For example, Re and Rule (2017) found that faces of mafia members were perceived to be more powerful but less socially skilled than faces of lawyers, suggesting that different traits are valued in these two groups. However, this pattern reversed when analyzing rank attainment within groups: Perceived social skills were correlated with the rank of mafia members, while perceived power was correlated with the rank of lawyers. Thus, distinct traits were related to selection into a group and rank attainment within the group. In the context of political elections, it may be the case that certain traits are required to become a politician, to be nominated as a candidate, or to survive a preliminary round, whereas other traits are related to electoral success. It is unlikely though, that this feature can explain diverging results between the current study and previous investigations. Similar to previous studies showing that competent-looking politicians are more successful (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Todorov et al., 2005), we analyzed results from the second round of run-off elections, but found no effect of perceived competence.

One shortcoming of the current set of studies was the limited number of constituencies. Our sample size was constrained by the number of constituencies that met our predefined inclusion criteria. For example, we decided to discard all elections for which we were not able to find a suitable photo for both candidates and all elections that featured at least one female candidate to avoid introducing gender stereotypes as a confound (Chiao et al., 2008). Our sample of 75 constituencies was larger than that of most studies which previously examined (and found) effects of facial appearance on election outcomes (e.g., Castelli et al., 2009; Chen, Jing, & Lee, 2012, 2014; Chen et al., 2016; Rule et al., 2010). Nonetheless, it might have been insufficient to
detect regional differences in the effect of perceived trustworthiness. Future studies should consider a wider set of constituencies to ensure that the failure to find support for our central hypothesis was not due to insufficient power.

**Conclusion**

Even though we did not find any evidence that corruption moderates the success of trustworthy-looking politicians, our results do support the general notion that the facial appearance of political candidates is associated with voting behavior. In the context of the 2016 Italian local elections, attractive-looking candidates had a non-trivial advantage over their less attractive-looking opponents. Specifically, candidates who scored one standard deviation higher on perceived attractiveness received 2.98 percentage points more votes and were 1.91 times more likely to win. Thus, our results suggest that people rely on trait impressions from faces even when making such consequential decisions as whom to elect as their political leader.
 References
https://doi.org/10.1126/science.1167748

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. https://doi.org/10.1073/pnas.0705435104

https://doi.org/10.1016/j.jpubeco.2009.11.002

better and voters reward it. *Journal of Public Economics, 146*, 79–86.
https://doi.org/10.1016/j.jpubeco.2016.12.008

Bigoni, M., Bortolotti, S., Casari, M., Gambetta, D., & Pancotto, F. (2016). Amoral familism,
social capital, or trust? The behavioural foundations of the Italian North–South divide.

primary role of morality (vs. sociability and competence) in information gathering.

and sociability on electoral outcomes. *Journal of Experimental Social Psychology, 45*(5),
1152–1155. https://doi.org/10.1016/j.jesp.2009.06.018

Chen, F. F., Jing, Y., & Lee, J. M. (2012). “I” value competence but “we” value social
competence: The moderating role of voters’ individualistic and collectivistic orientation in
https://doi.org/10.1016/j.jesp.2012.07.006

https://doi.org/10.1016/j.jesp.2013.10.008

all the same. *Social Psychological and Personality Science, 7*(6), 570–578.
https://doi.org/10.1177/1948550616644962


https://doi.org/10.1177/0956797615590992