Predicting Elections: Child's Play!

John Antonakis* and Olaf Dalgas

*To whom correspondence should be addressed. E-mail: john.antonakis@unil.ch, Faculty of Business and Economics, University of Lausanne, 1015 Lausanne, Switzerland.

In *The Republic*, Plato states, "Imagine ... a ship in which there is a captain who is taller and stronger than any of the crew, but he is a little deaf and has a similar infirmity in sight, and his knowledge of navigation is not much better" (1). Plato argues that the crew (i.e., voters) cannot select a competent captain (i.e., ruler) because the crew is beguiled, in part, by appearances. Plato uses this allegory to suggest that voters lack the rational faculties and knowledge to elect competent rulers.

Ideally, democracies should elect politicians on their competence. Intellectual (or learning) ability predicts effective performance in complex domains (2) and in the U.S. presidency (3). Presidents, though, are not elected on ability (4). Among other factors, voters are biased by facial appearances; naïve raters can predict elections after simply rating the competence (i.e., intelligence, leadership, and competence per se) of political candidates from their photographs (5). On a general level, individuals automatically infer characteristics of social targets based on facial appearances (6). Voters acting rationally, however, should change these initial classifications as they receive information about the target individual's values, performance, political affiliation and the like. Unfortunately, voters anchored in an initial impression do not appropriately correct the initial inferences; additional information on the candidates does not change choices by much (5). Perhaps voters are acting knowledgeably if, after experience, they have learned that facial appearance correlates with competence or performance. However, intelligence of adults cannot be predicted from facial appearance (7), and there is great variation in the competence of politicians (3).

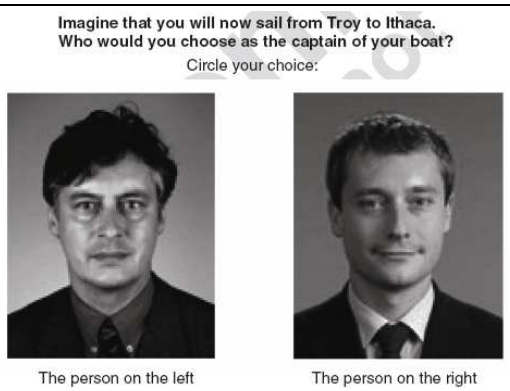

Fig. 1. An example of a pair of faces we used from the Meurthe-et-Moselle electoral district (number 1). Jean-Jacques Denis (left) lost to Laurent Hénart (right). Of the participants who rated this pair, 77% of children and 67% of adults in experiment 1 chose Denis. Over the 57 pairs of faces and across all raters, the adults in experiment 1 chose correctly 60% of the time, likelihood ratio $\chi^2(1) = 28.86$, $\phi = 0.20$, $P < 0.001$; for experiment 2, both children and adults chose correctly 64% of the time, likelihood ratio $\chi^2(1) = 68.10$, $\phi = 0.28$, $P < 0.001$. These effects become stronger when controlling for covariates and fixed effects for pairs of faces (9).

Why do naïve ratings and actual votes correlate? Are voters using the same rudimentary decision heuristics that children use? Facial stereotypes and other classification schemes are well developed in infancy (8), probably stemming from an innate template and rapid early learning (9). We hypothesized that voters might still be using the same cues that children do to categorize individuals on competency, which explains why voters may largely ignore additional information on candidates. We tested our claim by examining whether naïve voters predict actual voter preferences in the same way that children do.

We recruited adults and children in Switzerland to rate pairs of faces (the winner and runner-up) from the run-off stages of the 2002 French parliamentary election (10). In experiment 1 ($N = 684$ adults), results of a logistic regression showed that the probability of predicting an election result correctly on the basis of ratings of competence was 0.72. Ratings of competence also predicted margin of victory (standardized beta = 0.32, $P < 0.001$).

Using the same materials in experiment 2, 841 individuals—of whom 681 were children aged 5 to 13 years (mean age = 10.31, $SD = 1.81$)—participated in a game involving a computer-simulated trip from Troy to Ithaca. Thereafter, participants chose from two faces the captain of their boat (Fig. 1A). For the children, results from a logistic regression showed that probability of predicting an election result correctly on the basis of choice of captain was 0.71. The results did not differ when including the other participants ($N = 160$, mean age = 30.49, $SD = 16.32$); prediction accuracy did not depend on age (fig. S1).

Next, we compared the adults in experiment 1 to the children. We used the mean (i.e., at the pair-level) predicted

probabilities for each pair of faces for both children and adults in a random-effects regression model. The variable indicating adults was unrelated to the predicted probabilities; again, child-adult response patterns were indistinguishable. Furthermore, children ratings strongly predicted the adult ratings (standardized regression beta = 0.61, $P < 0.001$). Face effects appear to be age-invariant, suggesting that adults and children use similar cues in judging competence from facial appearance.

Evidently, young children, who are less experienced than are adults in observing performance in complex domains, playing an innocuous game can predict election results retrospectively (11). These findings suggest that voters are not appropriately weighting performance-based information on political candidates when undertaking one of democracy's most important civic duties.

References and Notes

1. Plato, The Republic, B. Jowett, Transl. (Collier, New York, 1901).
2. F. L. Schmidt, J. E. Hunter, Psychol. Bull. 124, 262 (1998).
3. D. K. Simonton, Adv. Psychol. Res. 14, 163 (2002). Note, in Simonton's analysis, the standardized beta for relation between intelligence and presidential greatness is .29; however, removing an endogenous predictor (number of years in office) increases the beta of intelligence to .41.
4. If politicians were selected on ability, the correlation between ability and performance would be zero or very weak (given the range restriction in ability).
5. A. Todorov, A. N. Mandisodza, A. Goren, C. C. Hall, Science 308, 1623 (2005).
6. R. Hassin, Y. Trope, J. Pers. Soc. Psychol. 78, 837 (2000).
7. L. A. Zebrowitz, J. A. Hall, N. A. Murphy, G. Rhodes, Pers. Soc. Psychol. Bull. 28, 238 (2002).
8. J. L. Ramsey, J. H. Langlois, R. A. Hoss, A. J. Rubenstein, A. M. Griffin, Dev. Sci. 7, 201 (2004).
9. A. Slater, P. C. Quinn, Infant Child Dev. 10, 21 (2001).
10. Materials, methods, and extended results are available as supporting material on *Science* Online.
11. In experiment 2, we also presented children with two pairs of faces (one from the recent Democratic party primary and the other from the U.S. election). The children correctly predicted the Obama-Clinton [likelihood-ratio $\chi^2(1) = 3.94$, $\phi = 0.54$, $P < 0.05$] and the Obama-McCain [likelihood-ratio $\chi^2(1) = 8.45$, $\phi = 0.81$, $P < 0.01$] election results.

12. We thank our university rectorate for support and J. P. Bonardi, M. Brulhart, F. Butera, S. Faulk, U. Hoffrage, P. Jacquart, R. Lalive, G. Palazzo, B. Sattler, C. Zehnder, and the OB group for helpful comments.

Supporting Online Material
www.sciencemag.org/cgi/content/full/323/5918/PAGE/DC1
Materials and Methods
SOM Text
Fig. S1

27 October 2008; accepted 16 December 2008

10.1126/science.1167748

Supporting on-line material

Predicting election results: Child's play!

John Antonakis*, Olaf Dalgas

University of Lausanne

Faculty of Business and Economics

Internef 618, Lausanne 1015

Switzerland

*To whom correspondence should be addressed.

E-mail: john.antonakis@unil.ch

15 December 2008

This file includes:

Materials and methods

Estimation methods and extended results

Figure S1

1. Materials and Method

We used 57 pairs (3 women, 54 men) of photos of candidate's faces from the 2002 French parliamentary run-off elections. Photos, which we cropped and put into black-white, were official versions from parliament website. In France, run-off ballots are held in the second round, with two candidates usually competing for one parliamentary seat. We used only pairs where the loser of the run-off was the incumbent (i.e., winner) from the 1997 election (whose standardized photo was included on the parliament website; this design is conservative because losers should not appear "too" incompetent given that they had previously won). We excluded pairs that had more than two candidates or other confounds (e.g., different race, sex), as well as currently well-known individuals (e.g., Ségolène Royal, who ran for President of France). For Experiment 2, we also included two pairs of color photos: John McCain vs. Barack Obama; Barack Obama vs. Hilary Clinton.

Samples and Measures

*Experiment 1***:** Participants were 684 Swiss public university students (43.71% females). The questionnaire was on one sheet depicting one pair of faces. Participants rated which of the two individuals was more competent, more intelligent, and the better leader (Cronbach alpha .70). Participants rated the statements using a 6-point rating scale: 1 (or 6)—definitely the person on the left (or right); 2 (or 5)—the person on the left (or right); 3 (or 4)—most probably person on the left (or right). We counterbalanced the positions of faces on the experimental materials and randomized order. We showed each pair of faces to 12 participants and each participant rated one pair.

*Experiment 2***:** Participants were 681 children 13 years of age and below (mean

age = 10.31, *SD* = 1.81, age range 5-13; 44.20% girls), and 160 older participants (mean age 30.49, *SD* = 16.32, age range 14-72; 48.73% females) from Switzerland. The questionnaire was on one sheet depicting one pair of faces (see Fig. 1 of main text). After having performed in an experimental game reenacting Odysseus' trip from Troy to Ithaca with the goal of returning home as fast as possible, we asked participants to imagine that they would repeat the trip today. They then indicated who they would choose as the captain of their boat. Because the experimental game involved a choice task that required a certain amount of decision skill about the voyage, we assumed that participants would choose their captain based on how competent he/she looked. We completely randomized order of pairs and face positions within pairs. We showed each pair of faces to 11.9 participants and each participant rated one pair (for the McCain-Obama and the Clinton-Obama elections, 10 and 13 children rated these pairs respectively).

Procedure

*Experiment 1***:** We recruited students at a Swiss public university in January 2007. We simply asked students to rate the pair of faces on the criteria provided. After receiving the ratings, we asked participants whether they recognized any of the individuals. In all cases participants stated that they did not; this result is not surprising given that parliamentary candidates were from another country and that the election took place 5 years before.

*Experiment 2***:** We recruited participants during a university open house (held end of May 2008). As part of the university's public relations campaign to improve its visibility and to increase children's interest in the university, local schools were invited to attend this open house event on the first day and the general public was invited on the

second and third days. This event usually attracts a large number of visitors.

Participants entered the experimental tent in groups of about 5 individuals. Games took 20-30 minutes to complete. Each experimental table had one experimenter, who gave instructions orally to the group. At the end of the game, participants completed the questionnaire individually in front of an experimenter. When there was not much demand from children to play the game, we allowed accompanying adults and older teenagers to participate (each participant played individually, under express instructions not interfere with choices of others). Because candidates from the French elections were unknown to the adult participants in Experiment 1 and given the ages of the children who participated, it would be reasonable to assume that the children were also not familiar with the candidates that they rated for an election that took place 6 years ago in another country. As regards the U.S. data, it is possible that children may have been familiar with the faces of these candidates. Because we did not assess for this familiarity, the conclusions that we can draw regarding the U.S. data are limited.

2. Estimation methods and extended results

Experiment 1

We first estimated the following logistic regression model:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 Comp + \beta_2 Sex + \sum_{k=2}^{57} \beta_k Set_k + \varepsilon \qquad \text{Eq. 1}$$

The dependent variable is the probability of choosing the elected candidate (i.e., the winner was coded 0 when placed on the left and 1 when placed on the right), “Comp” is inference of competence, “Sex” is the sex of the rater, “Set” are dummy variables indicating rated pair (to control for unmeasured or unobserved fixed effects of pairs that might be correlated with the variables in the models). The model fit the data well,

Hosmer-Lemeshow $\chi^2(8) = 9.37$, $P > .10$ (for data divided into 10 groups). Beta 1 was significant = .64 (standardized logit estimate = .51), $SE = .10$, $Z = 6.67$, $P < .001$; Beta 2 was insignificant. Next, we calculated the predicted marginal effect for high (i.e., 5) and low levels (i.e., 2) of competence (holding the rest of the covariates at their means). On average, the probability of predicting an election result correctly was .72.

We then estimated the following ordinary least squares regression model:

$$\text{Margin} = \beta_0 + \beta_1 Comp + \beta_2 Sex + \sum_{k=2}^{57} \beta_k Set_k + \varepsilon \qquad \text{Eq. 2}$$

The dependent variable is the margin of victory and the rest of the variables are as in Eq. 1. Beta 1 was significant = .03 (standardized beta estimate = .31), $SE = .00$, $T = 7.90$, $P < .001$.

Experiment 2

For children below 13 years (n = 681), we estimated the following logistic regression model:

$$\ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 Choice + \beta_2 Sex + \beta_3 Age + \beta_4 Choice * Age + \sum_{k=2}^{57} \beta_k Set_k$$

$$+ \beta_5 ManipA + \beta_6 ManipB + \sum_{l=2}^{173} \beta_l Group_l + \sum_{m=2}^{3} \beta_m Day_m + \sum_{n=2}^{8} \beta_n Seat_n + e \qquad \text{Eq. 3}$$

The dependent variable is the probability of choosing the elected candidate (i.e., the winner was coded 0 when placed on the left and 1 when placed on the right), “Choice” is choice of captain (i.e., 0 if raters chose left, 1 if raters chose right), “Sex” is sex of the rater, “Age” is age of the rater (we mean-centered age so as to reduce collinearity with the interaction term), “Choice*Age” is the interaction of Choice and Age, “Set” is a dummy variable indicating rated pair. We also controlled for potential

confounds emanating from the experimental game played before participants made their choice of captain: “ManipA” and “ManipB” were randomly manipulated dummy variables for a deferred choice task, “Group” is a dummy variables indicating the group in which participant participated in the game, “Day” is a dummy variable indicating the day visited, “Seat” is the seating position around the experimental table. As expected, a Hausman test indicated that the estimated parameters of the variables Choice, Sex, Age, Choice*Age and Set did not differ significantly with the inclusion of the control variables, $\chi^2(60) = 62.26$, $P > .10$.

Results indicated that the model fit the data well, Hosmer-Lemeshow $\chi^2(8) = 12.01$, $P > .10$ (for data divided into 10 groups). Beta 1 was significant = 1.80 (standardized logit estimate = .37), $SE = .28$, $Z = 6.45$, $P < .001$. Sex and the Choice*Age were unrelated to the dependent variable[1]. Again, we calculated the predicted marginal effect using choice of captain. On average, the probability of predicting an election result correctly was .71.

We then added the rest of the sample (n = 160) and re-estimated the model. The model fit the data well, Hosmer-Lemeshow $\chi^2(8) = 5.44$, $P > .10$ (for data divided into 10 groups). Beta 1 was still significant = 1.74 (standardized logit estimate = .39), $SE = .22$, $Z = 7.75$, $P < .001$, and basically unchanged. Sex and the Choice*Age interaction were unrelated to the dependent variable. The predicted marginal effects for choice of captain were unchanged (i.e., .71). Even though the Choice*Age interaction was not significant, we plotted the predicted marginal effects across all levels of age (for choice of captain) to

[1] For this sample, and for the sample including the adults, we also re-estimated the model without the fixed effects controls; the choice*age interaction remained insignificant. Because coefficients and standard errors of interaction terms could be incorrectly estimated in the case of binary-dependent models, we also estimated the model using the procedures recommended by Ai and Norton—again, the interaction was far from being significant and very similar to the original estimate. See C. Ai, E. C. Norton, *Econ. Lett.* **80**, 123 (2003).

demonstrate the extent to which effects were age invariant. As indicated in Figure S1, prediction accuracy did not change much across age, though it did taper downwards somewhat as age increased. Also, as an alternative estimation procedure, we created a binary variable (combining information on the winner-choice pair) indicating whether the participant was correct (coded 1) in identifying the winner or not (coded 0). A logistic regression, using age as an independent variable and the other controls showed that age was unrelated to prediction accuracy.

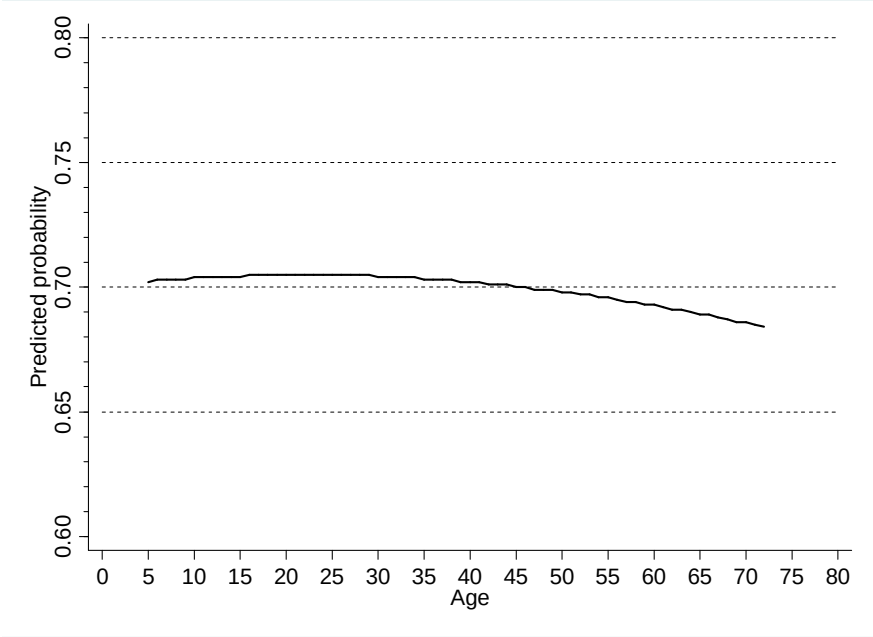

Fig. S1: Estimated marginal (predicted) probability as a function of age for Experiment 2 including adults and children.

Combined data

We then compared the adults in Experiment 1 to the children and estimated a random-effects regression model:

$$\text{Predicted prob}_{ij} = \beta_0 + \beta_{1j} Adult_j + \sum_{k=2}^{57} \beta_k Set_k + \varepsilon_j + \delta_{ij} \qquad \text{Eq. 4}$$

The dependent variable is the predicted probabilities of adults and children aggregated at the pair-level, "Adult" is a dummy variable indicating child or adult group. Note, the panel variable was set-winner pair (the adult and child rated winner or loser pair). We estimated the model for the i^{th} set-winner pair in the j^{th} adult-child group, with group and panel specific residual variances. Although Beta 1 was negative (replicating the negative trend reported in Fig. 1), it was not significant whether using conventional, cluster robust, or jackknifed standard errors.

Next, we regressed the predicted probabilities of the adults on the predicted probabilities of the children:

$$\text{Predicted prob(Adults)} = \beta_0 + \beta_1 Children + \sum_{k=2}^{57} \beta_k Set_k + \varepsilon \qquad \text{Eq. 5}$$

The dependent variable is the adults' predicted probability and "Children" is the children's predicted probability. Beta 1 was significant = .66 (standardized beta estimate = .61), SE = .07, T = 8.88, $P < .001$.

Finally, we combined the data of the children and the adults to determine whether the ROC (receiver operating characteristics) curves differed. For the adults, we dichotomized the data at the midpoint (i.e., 4 and above or 3 and below) of the competence ratings essentially to test if the child and adult logistic models differed using choice of individual as the independent variable. The child and adult models did not differ significantly, $\chi^2(1) = 2.02$, $P > 0.10$.