Polling the Face: Prediction and Consensus Across Cultures

Nicholas O. Rule and Nalini Ambady
Tufts University

Reginald B. Adams, Jr.
The Pennsylvania State University

Hiroki Ozono, Satoshi Nakashima, Sakiko Yoshikawa, and Motoki Watabe
Kyoto University

Previous work has shown that individuals agree across cultures on the traits that they infer from faces. Previous work has also shown that inferences from faces can be predictive of important outcomes within cultures. The current research merges these two lines of work. In a series of cross-cultural studies, the authors asked American and Japanese participants to provide naïve inferences of traits from the faces of U.S. political candidates (Studies 1 and 3) and Japanese political candidates (Studies 2 and 4). Perceivers showed high agreement in their ratings of the faces, regardless of culture, and both sets of judgments were predictive of an important ecological outcome (the percentage of votes that each candidate received in the actual election). The traits predicting electoral success differed, however, depending on the targets' culture. Thus, when American and Japanese participants were asked to provide explicit inferences of how likely each candidate would be to win an election (Studies 3–4), judgments were predictive only for same-culture candidates. Attempts to infer the electoral success for the foreign culture showed evidence of self-projection. Therefore, perceivers can reliably infer predictive information from faces but require knowledge about the target's culture to make these predictions accurately.

Keywords: person perception, cross-cultural psychology, nonverbal behavior

Supplemental materials: http://dx.doi.org/10.1037/a0017673.supp

The old adage "Don't judge a book by its cover" compels us not to trust our first impressions of others, which can be based on superficial information, such as appearances. Yet, the psychological literature has shown robust and consistent evidence that our first impressions of others can provide meaningful information.

Much of this work has been focused on perceptions of the face. The face is believed to be a uniquely rich source of information about individuals and is the most frequent expresser of nonverbal information (e.g., Rinn, 1991; Zebrowitz, 1997).

Considerable attention in the person perception literature has been devoted to consensus in impressions from the face. For instance, one stream of work has shown that people agree in their judgments of physical traits (e.g., facial maturity and attractiveness; Zebrowitz, Montepare, & Lee, 1993) and personality attributes (e.g., extraversion; Kenny, Albright, Malloy, & Kashy, 1994) made from perceptions of faces. Moreover, this consensus extends across cultural boundaries. For example, both American and Taiwanese participants agree in their judgments of attractiveness for racially diverse targets (Cunningham, Roberts, Barbee, Druen, & Wu, 1995). Similar results have been found with regard to personality traits: Both American and Chinese individuals show high agreement in the personality traits ascribed to targets in photographs of the face (Albright et al., 1997; see also Zebrowitz et al., 1993).

A second stream of work has shown that within a culture, impressions from the face are not just consistent but that they are also predictive of outcomes. For instance, perceptions of individuals' faces affect judges' decisions in court (Zebrowitz & McDonald, 1991), employers' decisions about who to hire or

promote (Collins & Zebrowitz, 1995), and teachers' expectations about children's intellectual abilities (Clifford & Walster, 1973). In each of these cases, perceptions of faces are directly predicting an individual's behavior toward another. There are also cases, however, in which predictions of outcomes from faces are less direct. For instance, judgments of chief executive officers' faces predict their companies' profits (Rule & Ambady, 2008, 2009), and judgments of political candidates' faces predict their electoral success (Antonakis & Dalgas, 2009; Ballew & Todorov, 2007; Keating, Randall, & Kendrick, 1999; Little, Burriss, Jones, & Roberts, 2007; Martin, 1978; Poutvaara, Jordahl, & Berggren, 2009). In one study, naïve ratings of competence from the faces of U.S. congressional candidates predicted the winners of elections, independent of powerful physiognomic cues, such as physical attractiveness (Todorov, Mandisodza, Goren, & Hall, 2005). Hence, facial cues lead to consensus and also provide valuable predictive information about others.

Although previous work has examined consensus in judgments of faces across cultures, little work has examined the prediction of outcomes across cultures. Poutvaara et al. (2009) found that judgments of competence from Finnish political candidates' faces predicted the candidates' electoral success when the raters were from culturally similar nations (e.g., Sweden and the United States), and Antonakis and Dalgas (2009) have reported that the judgments of Swiss adults and children predicted the electoral success of French candidates. In the current work, we sought to further merge the two streams of work on consensus and prediction by investigating the predictive value of cross-cultural consensual perceptions from two nations that are quite culturally distinct: the United States and Japan. To do so, we explored two types of judgments from faces: naïve inferences of targets' traits and explicit inferences of targets' success in their culture. In both cases, we related the judgments to an outcome criterion: the percentage of votes that the targets received in real elections. Thus, participants made inferences about traits and electoral success for the faces of candidates from both an American and a Japanese legislative election. Based on the literature reviewed above, our first hypothesis was that naïve judgments of traits from faces would show cross-cultural agreement.

Culture and Inference

Despite the evidence for consensus in judgments across cultures, judgments of many behaviors show cross-cultural differences in interpretation. For example, although Americans and Koreans agree on the sex and age of elderly targets when observing point-light displays of their gaits, they disagree on what these behaviors mean (Montepare & Zebrowitz, 1993). Similarly, a slower rate of speech is associated with greater status in Korea, whereas in the United States the opposite is true (Peng, Zebrowitz, & Lee, 1993). In both of these cases, the perceivers agreed in their judgments (i.e., target sex, age, and speech rate), but the meaning of these cues varied according to the cultural context.

One mechanism for the different meanings attributed to these behaviors may be egocentric self-projection. Self-projection is the robust and persistent phenomenon by which individuals generalize their own perspective and beliefs to others (Krueger, 2007; Ross, Greene, & House, 1977). That is, individuals may

recognize and agree on the same behaviors cross-culturally, but when extracting the meaning of these cues, they may rely on their own interpretation of what the cues signify rather than considering the perspective of members of the other culture. Indeed, previous research has shown strong effects whereby individuals generalize their own beliefs to others, even when provided with strong indications that their own perspective may be biased or that another perspective may be more valid (Krueger & Clement, 1994, 1996). One explanation for why this occurs is that individuals have better access to their own interpretations than they do to others' interpretations (Clement & Krueger, 2000). This is naturally exacerbated when making inferences about outgroup members (such as members of other cultures) because individuals have less information about outgroups than they do about ingroups and are therefore more likely to rely on their own beliefs because there is less competing information that is needed to be displaced (Clement & Krueger, 2002; Krueger, 2007; Robbins & Krueger, 2005). Thus, we expected that individuals would show agreement in their perceptions of cultural outgroup members but would engage in self-projection when attempting to think as would a member of the cultural outgroup. In other words, perceptions of political candidates should show agreement across cultures, but the decision for whom to vote should show reliance on self-projected same-culture knowledge.

Given that perceptions of the same traits may lead to different inferences about behavior and meaning depending on culture, we also hypothesized that the relationship between trait judgments and outcome variables may differ across cultures. For instance, descriptions of leadership preferences have identified traits relating to power and warmth as important factors in electoral outcomes (e.g., Funk, 1996, 1997; see also Misumi & Peterson, 1985). But, critically, cross-cultural research has shown that power and warmth may be weighted differently in different cultures (Ayman, 1993; Burns, 1978; Den Hartog, House, Hanges, & Ruiz-Quintanilla, 1999; Dickson, Den Hartog, & Mitchelson, 2003; Hofstede, 1980; Jung & Avolio, 1998; Schwartz, 1992). Specifically, studies of leadership preferences in the United States have indicated that traits relating to power may be valued more than are traits relating to warmth (e.g., Funk, 1996, 1997; Jung & Avolio, 1998; Misumi & Peterson, 1985; Tiedens, 2001; see also Smith, Peterson, & Misumi, 1994), whereas studies of leadership preferences in Japan have indicated that traits relating to warmth may be valued more than are traits relating to power (e.g., Jung & Avolio, 1999; Jung, Bass, & Sosik, 1995; Misumi & Peterson, 1985; Smith, Misumi, Tayeb, Peterson, & Bond 1989; Yamaguchi, 1994). We therefore hypothesized that regardless of consensus in judgments, naïve inferences of traits relating to power would be predictive of electoral success in the United States, whereas naïve inferences of traits relating to warmth would be predictive of electoral success in Japan. Moreover, given these different cultural frameworks for what defines a leader, we hypothesized that perceivers' explicit judgments of electability would not predict success across cultures but might instead rely on self-projection of own-culture values.

To measure these power and warmth constructs, we asked participants to provide judgments of faces along several, previously established traits. To assess power, we chose traits that have previously been implicated as important for evaluating leadership

and that are important across cultures: competence (Martin, 1978; Todorov et al., 2005), dominance (Den Hartog et al., 1999; Mazur, 2005; Megargee, Bogart, & Anderson, 1960), and facial maturity (Zebrowitz, 1997; Zebrowitz & Montepare, 2005). To assess warmth, we also chose previously validated traits: likeability (Baum, 2005; Kinder, Peters, Abelson, & Fiske, 1980; Little et al., 2007; Todorov et al., 2005) and trustworthiness (Chemers, 1993; Jung et al., 1995; Todorov et al., 2005; Yamagishi & Yamagishi, 1994). Previous work has shown that these traits cluster together in ways that reliably compose power and warmth, respectively (Rule & Ambady, 2008).

The Current Work

To evaluate the consensual and predictive value of first impressions across cultures, we investigated both naïve inferences of faces and explicit predictions based on faces in two parts. In the first part, we asked both American and Japanese participants to provide naïve inferences of traits relating to power and warmth for the faces of candidates who ran in the 2006 election of the U.S. Senate (Study 1) and for the faces of candidates who ran in the 2000 election of the Japanese Diet's House of Representatives (Study 2). In the second part, we repeated these naïve inferences but also asked both American and Japanese participants to provide explicit judgments of likely electoral success for the same U.S. Senate candidates (Study 3) and Japanese parliament candidates (Study 4).

Consistent with the literature on cross-cultural consensus, we expected that American and Japanese participants would agree in their trait judgments of targets from both cultures. However, consistent with the literature on self-projection, we expected that individuals would incorrectly infer the electoral success of candidates in the foreign culture. Specifically, given that the traits relevant to electoral success may differ for U.S. and Japanese electoral candidates (e.g., Misumi & Peterson, 1985), we expected that participants would apply their own culture's preferences when attempting to infer the electability of candidates from the foreign culture. Perceivers should therefore show prediction of their own culture's election outcomes because they have access to the information about cultural values that constitutes those judgments but should mispredict the other culture's election outcomes because, in the absence of information about cultural values, perceivers will apply their own cultural values to assess the candidates' viability.

Trait Composites

In the following studies, participants made judgments of political candidates on five personality traits: competence, dominance, facial maturity, likeability, and trustworthiness. To see whether these judgments agreed with the constructs established in the literature such as power and warmth (Fiske, Cuddy, Glick, & Xu, 2002; Judd, James-Hawkins, Yzerbyt, & Kashima, 2005; Misumi & Peterson, 1985), we conducted a principal components factor analysis with varimax rotation for the ratings of each target separately for American and Japanese participants. To achieve sufficient statistical power (e.g., Comrey & Lee, 1992), we aggregated participants' ratings for both same-culture and other-culture targets (i.e., 2006 U.S. Senate candidates and 2000 Japanese Diet candidates). This resulted in a total of 133 targets for the American

raters and 122 targets for the Japanese raters (see Studies 1 and 2 below).

Two composites resulted from the factor analysis and were very similar for participants from both cultures. Dominance and facial maturity loaded highly onto the first factor, whereas likeability and trustworthiness loaded highly onto the second factor (see Table 1). Competence loaded highly onto both factors, however. Noting the importance of competence to inferences of faces of American political candidates in past work (Ballew & Todorov, 2007; Todorov et al., 2005), we sought to model competence separately from the factors constituted by the other traits. However, due to high correlations between competence with each of these factors, this resulted in multicollinearity in the regression analyses reported in the studies below. The most common recommendation for correcting multicollinearity is to exclude one of the multicollinear variables (Allison, 1999; Hays, 1994; Kline, 1998; Meyers, Gamst, & Guarino, 2006; Tabachnick & Fidell, 2001). We therefore decided to remove competence from the main analyses.[1]

We then conducted the principal components factor analysis again, inputting only dominance, facial maturity, likeability, and trustworthiness. Two factors again emerged from the analysis, the first of which we called Power (60% of variance explained in the American sample and 45% explained in the Japanese sample), and the second of which we called Warmth (31% of variance explained in the American sample and 49% explained in the Japanese sample). As the factor solutions in Table 1 show, Power consisted of high loadings of dominance and facial maturity and low loadings of likeability and trustworthiness. Conversely, Warmth consisted of high loadings of likeability and trustworthiness and low loadings of dominance and facial maturity. We therefore created a composite score for Power by averaging each target's dominance and facial maturity scores and created a composite score for Warmth by averaging each target's likeability and trustworthiness scores.

To test the consistency of these factors cross-culturally, we then repeated the factor analyses for each of the 2 (participant culture) × 2 (target culture) combinations and calculated coefficients of congruence for the resulting factors. As in the superordinate analysis above, two factors emerged, consisting of dominance and facial maturity and of likeability and trustworthiness, respectively (see Supplemental Table 2 for factor loadings). The coefficients of congruence (CC) for American participants' ratings of U.S. targets

---

[1] It is critical to note that the results of the following analyses do not substantially differ if competence is entered into the analyses as a substitute for the composite for which it is most correlated. Specifically, competence loaded higher on Power than Warmth among the American participants. Conducting the regressions with competence in place of Power produced the same results in all of the analyses among the American samples (Studies 1A, 2A, 3A, and 4A). Similarly, competence loaded higher on Warmth than on Power among the Japanese participants, and conducting the regressions with competence in place of Warmth produced the same results in all of the analyses among the Japanese samples (Studies 1B, 2B, 3B, and 4B). That is, competence predicted votes for American candidates when Warmth, affect, and attractiveness were held constant; and competence predicted votes for Japanese candidates when Power, affect, and attractiveness were held constant. We provide correlations between the individual trait judgments and the outcome measure (the percentage of votes that candidates received in the election) in Supplemental Table 1.

Table 1

Principal Components Factor Solutions With Varimax Rotation for American and Japanese Participants

	Initial model		Final model	
Traits	Factor 1	Factor 2	Factor 1: Power	Factor 2: Warmth
American participants				
Competence	.78	.54		
Dominance	.87	−.37	**.87**	−.33
Facial maturity	.89	−.07	**.94**	.03
Likeability	−.33	.90	−.37	**.88**
Trustworthiness	.09	.95	.04	**.96**
Japanese participants				
Competence	.53	.80		
Dominance	.96	−.06	**.96**	−.11
Facial maturity	.94	.09	**.96**	.07
Likeability	−.19	.95	−.14	**.96**
Trustworthiness	.05	.97	.10	**.97**

Note. Traits loading together for use in composite formation are indicated in bold.

and Japanese targets were .97 for Power and .98 for Warmth. Similarly, the CC for Japanese participants' ratings of U.S. targets and Japanese targets were .98 for Power and .98 for Warmth. American participants and Japanese participants also showed similar levels of congruence for their ratings of U.S. targets ($CC_{Power} = .98$; $CC_{Warmth} = .98$) and Japanese targets ($CC_{Power} = .92$; $CC_{Warmth} = .94$). Thus, the factors seemed to be compatible across cultures, both in terms of target and participant nationality.

Naïve Inferences

Study 1

In Study 1, we examined naïve judgments of American political candidates by American (Study 1A) and Japanese (Study 1B) raters.

Study 1A

Faces of candidates in the 2006 U.S. Senate election were rated on 7-point scales for the Power traits (dominance and facial maturity) and the Warmth traits (likeability and trustworthiness), and the average scores for each candidate were related to the percentage of votes that the candidate received in the actual election.

Method.

Participants. Thirty-five male and female undergraduates at a U.S. university participated in the study for either partial course credit in an introductory psychology class or for monetary compensation. Six participants' data were removed from analysis because the participants recognized a politician. Thus, the final number of participants in the current study was 29.

Stimuli. Images of the candidates from the 2006 U.S. Senate election were obtained online from the Cable News Network (http://www.cnn.com/ELECTION/2006/pages/results/senate/full.list), as was information about the percentage of votes that each politician received in his state. Only the winner and runner-up

from each race were used, and female candidates were excluded because of concerns regarding the effects of gender stereotypes on judgments (e.g., Eagly & Karau, 1991). Of the remaining candidates, 1 (Joe Lieberman) was excluded because he was frequently recognized during pretesting. For 8 candidates, no picture was available and an image was obtained from the candidate's campaign website. Although the majority of candidates belonged to either the Democratic party or the Republican party, one candidate belonged to the Libertarian party, and one ran as an Independent. Following these exclusions, the total number of targets was 53, with 24 winners and 29 losers (22 matched winner-loser pairs). All images were cropped to the smallest frame that included the sides and top of the hair and the bottom of the chin, standardized in size, and gray-scaled. Independent samples of American raters coded the faces for attractiveness ($N = 9$; Cronbach's $\alpha = .76$) and affective expression ($N = 11$; Cronbach's $\alpha = .93$) along 7-point scales, and these scores were used as covariates in the analysis.

Procedure. The experiment consisted of five blocks. In each block, participants saw all 53 faces presented in a random order. The order of the blocks was random, and each block consisted of a different trait judgment: competence, dominance, facial maturity, likeability, and trustworthiness. Each trait was rated along a 7-point scale, anchored at *not at all X* (1) and *very X* (7). Participants showed high reliabilities for their ratings on each of the traits: competence (Cronbach's $\alpha = .76$), dominance (Cronbach's $\alpha = .88$), facial maturity (Cronbach's $\alpha = .93$), likeability (Cronbach's $\alpha = .91$), and trustworthiness (Cronbach's $\alpha = .82$). After completing the experiment, participants were debriefed and asked whether they had recognized any of the faces. Participants reporting recognitions were excluded from analysis.

Results and discussion. Scores for each candidate were averaged across participants for each trait. Following the results of the principal components analysis, we averaged together the Power traits (dominance and facial maturity) and Warmth traits (likeability and trustworthiness). We then tested the predictive value of these composites via a linear regression.

The percentage of votes acquired by each politician within his state served as the dependent measure, and coders' ratings of affect and attractiveness were entered as covariates. Values for the Power composite, Warmth composite, affect, and attractiveness were regressed onto politicians' vote-share simultaneously in a single-step.[2] Of the measures examined, Warmth ($\beta = -.13$), $t(48) = -0.62$, $p = .54$, affect ($\beta = -.17$), $t(48) = -0.88$, $p = .38$, and attractiveness ($\beta = .09$), $t(48) = 0.57$, $p = .57$ did not significantly predict the percentage of votes that the candidates received. The Power composite, however, did significantly predict the percentage of votes that the candidates received ($\beta = .42$), $t(48) = 2.93$, $p = .005$ (see Table 2).

Because candidates in American elections typically run in pairs, the percentage of votes acquired by the winning candidates was very often the complement of the percentage of votes acquired by the losing candidates. Thus, to account for these dependent relationships among candidates, we reanalyzed the data for the 22 pairs of matched candidates (i.e., candidates from the same state)

---

[2] For all analyses reported in this article, tests of multicollinearity indicated independence among the variables in each of the models (e.g., Kline, 1998); all variance-inflation factors <10, all tolerances >.10.

Table 2

Parameter Estimates (βs) and t-Statistics for the Influence of Power, Warmth, Affect, and Attractiveness on the Percentage of Votes That U.S. Candidates Received in the 2006 Senate Election

	American participants				Japanese participants			
	Study 1A		Study 3A		Study 1B		Study 3B	
Factor	β	t	β	t	β	t	β	t
Multiple regression								
Power	.42	2.93**	.48	3.47***	.49	3.45***	.45	2.83**
Warmth	−.13	−0.62	−.09	−0.40	.00	0.01	−.27	−1.25
Affect	−.17	−0.88	−.13	−0.62	−.22	−1.32	−.24	−1.28
Attractiveness	.09	0.57	.03	0.16	.12	0.53	.42	1.77
Multilevel model								
Power	.50	3.18**	.56	3.82***	.51	3.40***	.46	2.63**
Warmth	−.11	−0.49	−.04	−0.19	.11	0.49	−.12	−0.51
Affect	−.10	−0.49	−.07	−0.34	−.20	−1.16	−.21	−0.98
Attractiveness	.12	0.77	.04	0.25	.12	0.51	.23	0.90

** $p < .01$.   *** $p < .001$.

using multilevel modeling, with each winning and losing candidate nested within state or electoral race. Analysis of the predictor variables paralleled those of the multiple regression. Warmth ($\beta = -.11$), $t(44) = -0.49$, $p = .63$, affect ($\beta = -.10$), $t(44) = -0.49$, $p = .63$, and attractiveness ($\beta = .12$), $t(44) = 0.77$, $p = .45$, did not significantly predict the percentage of votes that the candidates received, whereas Power did significantly predict the percentage of votes that the candidates received ($\beta = .50$), $t(44) = 3.18$, $p = .003$.

Thus, consistent with previous accounts of evaluations of leadership in the United States (e.g., Funk, 1996, 1997), Power emerged as a significant predictor of the percentage of votes that politicians' received. That is, participants' inferences of personality traits from the faces of political candidates from the 2006 U.S. Senate race predicted their electoral success.

Study 1B

Both Study 1A and previous work have shown that individuals' naïve judgments of personality traits from the faces of U.S. politicians predict electoral success. However, this work has only shown this to be the case when the raters and targets are members of the same culture. Noting that some judgments of nonverbal cues are cross-culturally consistent (e.g., Albright et al., 1997; Zebrowitz et al., 1993), though others are culturally specific (e.g., Montepare & Zebrowitz, 1993), we sought to test whether inferences of politicians' faces by members of a different culture would predict U.S. electoral outcomes, as well. We therefore repeated Study 1A with judges from a second culture: Japan.

Method. Thirty-nine male and female undergraduates at a Japanese university rated the same 53 U.S. politicians, following the same procedures as in Study 1A. Instructions and trait terms were translated into Japanese and then back-translated to assure consistency in meaning: *yuno* (competence; Cronbach's $\alpha = .91$), *shihaiteki* (dominance; Cronbach's $\alpha = .94$), *otonappoi/ kodomoppoi kaodachi* (mature/babyish face; Cronbach's $\alpha = .94$), *konomashisa* (likeability; Cronbach's $\alpha = .91$), and *sinraidekiru teido* (trustworthiness; Cronbach's $\alpha = .88$). None of the participants recognized any of the candidates. In addition, independent

samples of Japanese raters coded the faces for attractiveness ($N = 15$; Cronbach's $\alpha = .84$) and affective expression ($N = 8$; Cronbach's $\alpha = .91$) along 7-point scales, and these scores were used as covariates in the analysis.

Results and discussion. As in Study 1A, ratings of each target were averaged across judges and combined into Power and Warmth composites. Correlations between these mean scores with those from Study 1A showed cross-cultural consensus in American and Japanese perceivers' ratings of the U.S. politicians' faces (r^*s > .42; see Table 3).[3] It is interesting to note that judgments of Power were significantly more correlated than were judgments of Warmth ($Z = 4.52$, $p < .001$), suggesting that the power-related traits (dominance and facial maturity) may be more cross-culturally legible than the warmth-related traits (likeability and trustworthiness). However, further work would be needed to test this hypothesis directly.

Mean scores for Power, Warmth, affect, and attractiveness were simultaneously regressed onto the percentage of votes that each politician received in the election. Power again emerged as the only significant predictor in the model ($\beta = .49$), $t(48) = 3.45$, $p = .001$; Warmth ($\beta < .01$), $t(48) = 0.01$, $p = .99$, affect ($\beta = -.23$), $t(48) = -1.32$, $p = .19$, and attractiveness ($\beta = .12$), $t(48) = 0.53$, $p = .60$, were all nonsignificant (see Table 2). Comparison of the parameter estimate for Power with the American participants' parameter estimate for Power in Study 1A showed no significant difference[4] ($Z = 0.47$, $p = .32$).

This pattern was repeated in the multilevel model. Power significantly predicted the percentage of votes that candidates received ($\beta = .51$), $t(44) = 3.40$, $p = .001$, whereas Warmth ($\beta = .11$), $t(44) =$

---

[3] We use r^* to refer to the disattenuated correlation coefficient (e.g., Kenny & West, 2009). The standard (attenuated) correlations are provided in Supplemental Table 3, and all meta-analytic comparisons of these correlations are based on the attenuated correlations.

[4] Throughout the article, parameter estimates (βs) were converted to rs for meta-analytic analysis following the formula recommended by Peterson and Brown (2005): $r = \beta + .05\lambda$, where $\lambda = 0$ if β is negative and $\lambda = 1$ if β is positive.

Table 3

Disattenuated Correlations for Traits and Composites by Targets Across Cultures Based on Ratings in Studies 1–4

Traits	U.S. candidates			Japanese candidates		
	Study 1	Study 3	Aggregate	Study 2	Study 4	Aggregate
Power	.96	.90	.92	.93	.95	.92
Dominance	.89	.87	.86	.83	.80	.79
Facial maturity	.96	.91	.94	.97	.98	.95
Warmth	.58	.64	.64	.53	.45	.47
Likeability	.70	.74	.75	.48	.38	.41
Trustworthiness	.42	.47	.51	.60	.63	.60

Note. All of the standard (attenuated) correlations significant at $p < .01$.

0.49, $p = .63$, affect ($\beta = -.20$), $t(44) = -1.16$, $p = .25$, and attractiveness ($\beta = .12$), $t(44) = 0.51$, $p = .61$, did not (see Table 2). Comparison of the parameter estimate for Power with that from the American participants' multilevel model in Study 1A again showed no significant difference ($Z = 0.07$, $p = .47$).

Similar to American raters' inferences of American politicians, then, Japanese raters' inferences of American politicians also corresponded to the targets' actual electoral success. These data therefore showed evidence for a relationship between judgments of faces from one culture and an ecologically valid outcome criterion in a second culture. Previous studies have shown that judgments of physiognomic and personality traits agree across cultures for some traits (Albright et al., 1997; Zebrowitz et al., 1993). Here, we extend these findings by showing cross-cultural consensus for additional traits. More important, we show that these naïve judgments are predictive of important social outcomes. In Study 2, we extend these findings by examining another sample of judgments.

Study 2

Inferences of personality traits from faces predict electoral outcomes for American political candidates. Whereas previous work has demonstrated this for judgments of various types of candidates (e.g., gubernatorial: Ballew & Todorov, 2007; congressional: Todorov et al., 2005), it has focused only on Americans' judgments of American politicians. Study 1B extended these effects to show that judgments made by perceivers from a different culture also predicted the outcomes of U.S. political elections. Despite this cross-cultural consistency among perceivers, a remaining question is whether these judgments might vary according to the targets. To test this, in Study 2, we asked both same-culture (Japanese) and other-culture (American) raters to make inferences of the same traits from the faces of political candidates for the candidates from the 2000 election of the Japanese House of Representatives (*shugi-in*).

Study 2A

Previous work has shown that electoral success is related to naïve inferences of personality from the faces of American politicians. This is true when the raters are both from the same culture (i.e., other Americans; Study 1A) and from a different culture (i.e., Japanese; Study 1B). Is this relationship limited to American politicians or does it extend across cultures? In Study 2A, we investigated this question by measuring whether American partic-

ipants' inferences of personality traits would predict electoral outcomes in a Japanese election.

Method. Forty American undergraduates participated for partial course credit in an introductory psychology class. No participants were Japanese or of Japanese descent, none had previously lived in Japan, and none had visited Japan for more than 2 weeks.

Images of candidates from the 2000 election of the Japanese House of Representatives (*shugi-in*) were obtained from an online database of political flyers used in the election (http://politics.j.u-tokyo.ac.jp/data/data12.html). Nearly 700 candidates ran in the election, so a sample of the male candidates was chosen on the basis of image resolution (i.e., those images with the highest picture quality, as measured by the number of pixels per inch) for a total of 80 images (40 winners and 40 losers), equally distributed from the electoral districts and parties. Candidates' heads were cropped from the flyers, standardized in size, and gray-scaled. As with the U.S. candidates, the faces were coded for attractiveness ($N = 8$; Cronbach's $\alpha = .83$) and affective expression ($N = 13$; Cronbach's $\alpha = .94$) along 7-point scales for use as covariates in the analysis. Information about the percentage of votes that the candidates received was acquired from an Internet news site that reported the election outcomes (http://www.senkyo.janjan.jp/bin/search/election.php?election_type=1).

Procedures were identical to those of Study 1A, and the participants' judgments showed high interrater reliability for all of the five traits: competence (Cronbach's $\alpha = .88$), dominance (Cronbach's $\alpha = .94$), facial maturity (Cronbach's $\alpha = .93$), likeability (Cronbach's $\alpha = .95$), and trustworthiness (Cronbach's $\alpha = .84$).

Results and discussion. Ratings for each candidate's face were averaged across participants for each trait and combined into Power and Warmth composites, following the principal components analysis reported above. Scores for Power, Warmth, affect, and attractiveness were then simultaneously regressed onto the percentage of votes that the candidates received. Scores for Warmth uniquely predicted the percentage of votes that Japanese candidates received in the election ($\beta = .37$), $t(75) = 2.01$, $p < .05$, whereas Power ($\beta = .04$), $t(75) = 0.30$, $p = .76$, affect ($\beta = -.37$), $t(75) = -1.85$, $p = .07$, and attractiveness ($\beta = .05$), $t(75) = 0.42$, $p = .68$ were nonsignificant predictors (see Table 4).

Similar to Study 1B, then, political candidates' electoral success was predicted by judgments from individuals belonging to a different culture. Notably, however, the variable predicting Japanese candidates' electoral success was distinct from that predicting U.S.

Table 4
Standardized Parameter Estimates (βs) and t-Statistics for the Influence of Power, Warmth, Affect, and Attractiveness on the Percentage of Votes That Japanese Candidates Received in the 2000 Diet Election

	American participants				Japanese participants			
	Study 2A		Study 4A		Study 2B		Study 4B	
Factor	β	t	β	t	β	t	β	t
Power	.04	0.30	.03	0.20	.06	0.53	.04	0.38
Warmth	.37	2.01*	.39	2.18*	.27	2.29*	.31	2.64**
Affect	−.37	−1.85	−.39	−1.90	.04	0.33	.02	0.20
Attractiveness	.05	0.42	.10	0.84	−.20	−1.71	−.22	−1.87

* $p < .05$. ** $p < .01$.

candidates' electoral success. Whereas Power was significantly related to electoral outcomes in the United States, Warmth was significantly related to electoral outcomes in Japan. This may be consistent with previous, descriptive accounts of leadership styles and leadership preferences in Japan (e.g., Misumi & Peterson, 1985). In Study 2B, we therefore examined whether the same result would emerge when inferences were made by same-culture (i.e., Japanese) raters.

Study 2B

Study 2A showed that inferences of personality traits from the faces of Japanese political candidates significantly predicted electoral outcomes. Similar to previous work, Power predicted electoral success in the United States. But distinct from this previous work, Warmth predicted electoral success in Japan. Although Japanese raters' inferences of U.S. political candidates in Study 1B agreed with American raters' inferences of U.S. political candidates in Study 1A, perhaps it is possible that American raters' inferences of Japanese political candidates in Study 2A are anomalous. That is, perhaps Japanese raters will show different judgments of same-culture political candidates. To test this, in Study 2B we asked Japanese participants to make the same trait ratings of the Japanese candidates that the Americans did in Study 2A.

Method. Forty-five male and female undergraduates at a Japanese university participated for partial credit in an introductory psychology course. Stimuli and procedures were the same as in Study 2A, except that instructions and materials were presented in Japanese, as in Study 1B. Pretesting of the targets indicated that 11 of the winning candidates were frequently recognized. Therefore, these targets were excluded for a total of 29 winning candidates.[5] Participants showed high interrater reliability for each of the traits: competence (Cronbach's $\alpha = .89$), dominance (Cronbach's $\alpha = .91$), facial maturity (Cronbach's $\alpha = .93$), likeability (Cronbach's $\alpha = .91$), and trustworthiness (Cronbach's $\alpha = .84$). As above, independent samples of Japanese raters coded the faces for attractiveness ($N = 14$; Cronbach's $\alpha = .71$) and affective expression ($N = 8$; Cronbach's $\alpha = .82$) along 7-point scales for use as covariates in the analysis. Four participants recognized one or more candidates and were removed from analysis, and a fifth participant's data were lost due to a mechanical failure (final $N = 40$).

Results and discussion. Scores for each candidate were averaged across participants and combined into Power and Warmth composites, following the principal components analysis. As with ratings of U.S. candidates in Study 1, American and Japanese perceivers' ratings were strongly correlated (r^*s > .48; see Table 3), indicating cross-cultural agreement for the traits and composites. As with the U.S. targets, agreement for Power was significantly greater than agreement for Warmth ($Z = 2.89$, $p = .002$). In addition, the strength of consensus for the Power and Warmth scores of U.S. targets in Study 1 did not significantly differ from the strength of consensus for Power and Warmth scores of Japanese targets here ($Z_{\text{Power}} = 0.58$, $p = .28$; $Z_{\text{Warmth}} = 0.37$, $p = .35$).

Scores for Power, Warmth, affect, and attractiveness were then simultaneously regressed onto the percentage of votes that each candidate received. As in Study 2A, the only variable significantly predicting the percentage of votes was Warmth ($\beta = .27$), $t(64) = 2.29$, $p < .03$. Power ($\beta = .06$), $t(64) = 0.53$, $p = .59$, affect ($\beta = .04$,) $t(64) = 0.33$, $p = .75$, and attractiveness ($\beta = −.20$), $t(64) = −1.71$, $p = .09$, were all nonsignificant predictors (see Table 4), and the parameter estimate for Warmth did not significantly differ from American participants' parameter estimate for Warmth in Study 2A ($Z = 0.69$, $p = .24$).

As stated above, this outcome is sensible given previous accounts of leadership in Japan, which emphasized interpersonal relationships among leaders and followers (e.g., Jung et al., 1995). In addition, these data show cross-cultural consensus and predictive validity, as naïve inferences of personality from the faces of Japanese political candidates predicted the percentage of votes that the candidates received when judged by both members of their own culture (current study) and members of another culture (Study 2A). Thus, naïve inferences of politicians' faces predicted electoral outcomes when both the perceivers and the targets were from different cultures.

Explicit Inferences

Studies 1 and 2 showed that naïve, indirect judgments predicted electoral success across cultures. None of the participants in any of these studies knew that they were rating political candidates.

---

[5] Notably, the results of Study 2A do not change when these famous candidates are also excluded, ($\beta_{\text{Warmth}} = .43$), $t(64) = 2.22$, $p = .03$; ($\beta_{\text{Power}} = .05$), $t(64) = 0.36$, $p = .72$; ($\beta_{\text{affect}} = −.39$), $t(64) = −1.89$, $p = .06$; ($\beta_{\text{attractiveness}} = .05$), $t(64) = 0.35$, $p = .73$.

Individuals reporting any recognition of a candidate, familiarity with the candidates, or suspicion that the targets were political candidates were excluded from analysis. Thus, naïve inferences of personality traits predicted electoral success in both cultures. This suggests that individuals may extract information from faces similarly, regardless of their cultural orientation and regardless of the targets' cultural orientation. If cultural knowledge is not needed to form predictive inferences about candidates' electoral success from their faces, how might the explicit application of one's own cultural values affect electoral predictions based on face judgments?

Considering self-projection theory (e.g., Robbins & Krueger, 2005), one might predict that individuals would be able to infer the outcomes of the elections in the opposite culture, as long as their cultural values were congruent (i.e., as long as the same traits were predictive in both cultures). Given that Studies 1 and 2 and the previous literature on leadership all suggest cross-cultural differences in the traits related to electoral success in the United States and Japan, we therefore expected that self-projection would lead to accurate same-culture judgments of electoral success but erroneous other-culture judgments of electoral success.

To test this, in Studies 3 and 4, we repeated the trait judgments from Studies 1 and 2 but, in addition, asked both American and Japanese participants to indicate how likely they believed people would be to vote for U.S. and Japanese political candidates in the respective culture. By doing so, we sought to evaluate the participants' predictions about votes for political candidates from both their own culture and the other's culture and sought to relate these explicit judgments to the participants' previous, naïve attribution of traits to the same candidates.

Study 3

In Study 1, American and Japanese participants' judgments of personality traits from the faces of U.S. political candidates predicted the percentage of votes that the candidates received. Might their judgments predict electoral success when they are required to explicitly apply their own knowledge about electoral values? In Study 3, we tested this in two parts.

Study 3A

In Study 3A, American participants provided both ratings of traits as in Study 1A and ratings of how likely they believed Americans would be to vote for each of the candidates from the 2006 U.S. Senate election along an 8-point scale, ranging from *not at all likely* (1) to *very likely* (8). These judgments were then related to the actual percentage of votes that the candidates received and the participants' trait judgments of the candidates.

Method. To test whether Americans' explicit judgments of political candidates' faces would predict electoral outcomes, 31 male and female American undergraduates rated the photos of each candidate from the 2006 U.S. Senate election in exchange for partial course credit. Targets consisted of the same 53 images used in Study 1, and participants repeated the same procedures as in Study 1A (Cronbach's αs > .81). After rating each of the traits, however, the experiment continued with participants asked to provide explicit ratings of how likely they believed Americans would be to vote for each of the targets along an 8-point scale

anchored at *not at all likely* (1) and *very likely* (8). Participants showed high interrater reliability for their explicit judgments of the candidates (Cronbach's α = .87).

Participants were informed that they were rating actual political candidates and were asked to identify any politicians that they recognized during debriefing. Five participants indicated that they recognized one or more candidates and provided the names of the individuals that they recognized. Although our intention was for participants to provide explicit judgments in the last section of the experiment, specific knowledge about a given candidate would likely relate to his status relative to the election's outcome (i.e., whether he is currently an elected official). These participants' data were therefore removed from analysis, as in Study 1A (final N = 26).

Results.

Trait inferences. Similar to Study 1A, American participants' judgments of Power, Warmth, affect, and attractiveness were simultaneously regressed onto the percentage of votes that the candidates received. Replicating Study 1A, Power was the only significant predictor of the percentage of votes that candidates received (β = .48), $t(48)$ = 3.47, p = .001 (see Table 2). Warmth (β = −.09), $t(48)$ = −0.40, p = .69, affect (β = −.13), $t(48)$ = −0.62, p = .54, and attractiveness (β = .03), $t(48)$ = 0.16, p = .87, were all nonsignificant.

These results were again reflected when candidates were paired with their opponents, and the percentage of votes obtained was estimated in a multilevel model. Warmth (β = −.04,) $t(44)$ = −0.19, p = .85, affect (β = −.07), $t(44)$ = −0.34, p = .74, and attractiveness (β = .04), $t(44)$ = 0.25, p = .81, were all nonsignificant, but Power significantly predicted the percentage of votes that candidates received (β = .56), $t(44)$ = 3.82, p < .001. Thus, American participants' inferences of Power from the faces of American political candidates significantly predicted the percentage of votes that they received, replicating Study 1A (see Table 2).

Voting judgments. As with the trait ratings, scores on the voting judgment for each candidate were averaged across participants. These mean scores were then regressed onto the percentage of votes that each candidate received, with ratings of affect and attractiveness entered simultaneously as covariates. Participants' voting judgments significantly predicted the percentage of votes that candidates actually received in the election (β = .35), $t(49)$ = 2.48, p = .02. Attractiveness was not a significant predictor of the percentage of votes that candidates received (β = −.07), $t(49)$ = −0.46, p = .65, but ratings of affect did emerge as a significant predictor in the model (β = −.48), $t(49)$ = −3.31, p = .002 (see Table 5). The negative relationship between affect and percentage of votes received indicated that candidates with more neutral expressions in their photos received significantly more votes than did those displaying happier expressions.

We then conducted a second analysis with multilevel modeling to estimate the relationship between the voting judgments and the actual votes that the candidates received, while accounting for the relationships between winning candidates and losing candidates from the same states. As in the ordinary least-squares regression with the full set of candidates, participants' voting judgments significantly predicted the percentage of votes that the candidates received (β = .47), $t(44)$ = 2.83, p = .007. Attractiveness ratings did not significantly predict the percentage of votes that candidates received (β = −.10), $t(44)$ = −0.55, p = .59, but, as in the above analysis, ratings of

Table 5

Standardized Parameter Estimates (βs) and t-Statistics for the Influence of Participants' Voting Judgments in Studies 3 and 4 on the Percentage of Votes That U.S. and Japanese Candidates Received

| | U.S. candidates | | | | Japanese candidates | | | |
| | American participants: Study 3A | | Japanese participants: Study 3B | | American participants: Study 4A | | Japanese participants: Study 4B | |
Model	β	t	β	t	β	t	β	t
Multiple regression								
Voting	.35	2.48*	−.15	−0.76	.18	1.36	.38	2.41*
Affect	−.48	−3.31**	−.59	−3.86***	−.12	−1.04	−.01	−0.07
Attractiveness	−.07	−0.46	.41	1.78	.10	0.79	−.10	−0.64
Multilevel model								
Voting	.47	2.83**	−.05	−0.24				
Affect	−.49	−2.85**	−.57	−3.57***				
Attractiveness	−.10	−0.55	.34	1.61				

* *p* < .05. ** *p* < .01. *** *p* < .001.

affect did (β = −.49), *t*(44) = −2.85, *p* = .007. Again, this indicated that candidates with more neutral expressions in their photos received significantly more votes than did those displaying happier expressions, even when accounting for relationships between opposing candidates.

We were also interested in the relationships between participants' trait ratings of the candidates' faces and their voting judgments of the faces. We therefore calculated partial correlations (controlling for the covariates, affect and attractiveness) between participants' voting judgments and their scores on the Power and Warmth composites. As expected, there was a significant positive relationship between voting judgments and Power, *r*(49) = .65, *p* < .001 (see Table 6). It is interesting to note that participants' ratings of Warmth were also significantly correlated with their voting judgments, *r*(49) = .38, *p* = .006. However, meta-analytic comparisons of the effect sizes (e.g., Rosenthal & Rosnow, 2007) showed that the relationship between voting judgments and Power was significantly greater than the relationship between voting judgments and Warmth (*Z* = 1.92, *p* = .03).

Discussion. Study 3A contributed several results of interest. First, American participants' trait judgments showed that Power uniquely predicted the percentage of votes that American political candidates received, replicating Study 1A. Second, participants' Power and Warmth judgments were related to their voting judgments of the candidates, and congruent with the trait judgments, the effect for Power was significantly stronger than was the effect

for Warmth when the covariates were controlled. Finally, participants' voting judgments significantly predicted the percentage of votes that the candidates received in the actual election. Thus, Americans' indirect judgments of personality traits and direct, explicit judgments of voting likelihood significantly predicted the amount of votes that the U.S. candidates received. This is reasonable, as Americans likely possess explicit knowledge about their culture's values, which may contribute to whether individuals will vote for a particular candidate. It is unlikely that members of a different culture would necessarily possess this insight, however. In Study 3B, we explored this hypothesis further.

Study 3B

In Study 3B, we tested whether Japanese participants' explicit judgments would predict votes received by U.S. political candidates and examined how these voting judgments would relate to participants' naïve inferences of traits from the candidates' faces.

Method. Twenty-four male and female undergraduates from a Japanese university rated each of the faces from the 2006 U.S. Senate election in exchange for partial course credit. Instructions and procedures were the same as in Study 3A except that they were translated into Japanese and then back-translated to English to assure the consistency of their meaning. Thus, Japanese participants completed the same task as in Study 1B (Cronbach's αs > .73) and were subsequently asked to indicate how likely they

Table 6

Partial Correlations (Controlling for Affect and Attractiveness) and Meta-Analytic Comparisons for the Relationships Between the Trait Composites (Power and Warmth) With Participants' Voting Judgments in Studies 3 and 4

| | U.S. candidates | | Japanese candidates | |
| | American participants: Study 3A | Japanese participants: Study 3B | American participants: Study 4A | Japanese participants: Study 4B |
Factor				
Power	.65***	.19	.66***	.29**
Warmth	.38**	.78***	.64***	.88***
Z	1.92*	4.14***	0.18	6.10***

* *p* < .05. ** *p* < .01. *** *p* < .001.

believed it was that most Americans would vote for each target in a political election on a scale from *not at all likely* (1) to *very likely* (8). Participants showed high interrater reliability (Cronbach's α = .87). None of the participants recognized any of the U.S. candidates.

Results.

Trait inferences.   As in Study 1, American and Japanese participants showed high levels of agreement in their ratings of the U.S. candidates (r^*s > .47), and the Power composite again showed significantly higher cross-cultural agreement than did the Warmth composite ($Z = 2.89$, $p < .01$; see Table 3). In addition, we aggregated across participants' ratings of U.S. candidates in both Studies 1 and 3 to measure the overall agreement between American (Studies 1A and 3A) and Japanese (Studies 1B and 3B) raters. These relationships showed significant cross-cultural consensus (r^*s > .51), and the consensus for Power was significantly greater than the consensus for Warmth ($Z = 3.88$, $p < .001$).

Scores for each candidate were averaged across participants, and composites for Power and Warmth were formed. We then regressed the scores for Power and Warmth onto the percentage of votes that candidates received in the election, with the ratings for affect and attractiveness from Study 1B included as covariates. As in the previous judgments of American candidates, Power emerged as the only significant predictor of the percentage of votes that candidates received ($\beta = .45$), $t(48) = 2.83$, $p = .007$, whereas Warmth ($\beta = -.27$), $t(48) = -1.25$, $p = .22$, affect ($\beta = -.24$), $t(48) = -1.28$, $p = .21$, and attractiveness ($\beta = .42$), $t(48) = 1.77$, $p = .08$, were all nonsignificant. As in Study 1, American and Japanese participants' parameter estimates for Power did not significantly differ ($Z = 0.07$, $p = .47$).

Controlling for candidates' shared electoral races with multilevel modeling produced similar results. Power emerged as the only significant predictor of the percentage of votes that candidates received ($\beta = .46$), $t(44) = 2.63$, $p = .01$. Warmth ($\beta = -.12$), $t(44) = -0.51$, $p = .61$, affect ($\beta = -.21$), $t(44) = -0.98$, $p = .33$, and attractiveness ($\beta = .23$), $t(44) = 0.90$, $p = .37$, were all nonsignificant, replicating the results of Study 1B and extending the results of Studies 1A and 3A to a sample of foreign judges (see Table 2).

Voting judgments.   Scores for participants' voting judgments were averaged to yield a mean score for each candidate and were simultaneously regressed onto the percentage of votes that the candidates actually received, along with affect and attractiveness scores from Study 1B. Unlike the results for American participants in Study 3A, neither Japanese participants' voting judgments ($\beta = -.15$), $t(49) = -0.76$, $p = .45$, nor their ratings of attractiveness ($\beta = .41$), $t(49) = 1.78$, $p = .08$, significantly predicted the percentage of votes that the U.S. candidates received. However, ratings of affect ($\beta = -.59$), $t(49) = -3.86$, $p < .001$, emerged as a significant predictor and were negatively related to electoral success, as they were for American participants in Study 3A (see Table 5). Meta-analytic comparisons of the parameter estimates showed that the predictive power of American participants' voting judgments in Study 3A was significantly greater than that for the Japanese participants' voting judgments in Study 3B ($Z = 2.87$, $p = .002$).

The results were similar when accounting for the candidates' shared electoral races. As above, neither Japanese participants' voting judgments ($\beta = -.05$), $t(44) = -0.24$, $p = .81$, nor ratings

of attractiveness ($\beta = .34$), $t(44) = 1.61$, $p = .11$, predicted the actual percentage of votes that the candidates received. Ratings of affect, however, again significantly predicted the candidates' actual votes ($\beta = -.57$), $t(44) = -3.57$, $p < .001$. Thus, when considering the relationships among the subset of candidates who ran in opposing pairs, Japanese participants' judgments of affect negatively predicted the percentage of votes that candidates received (i.e., candidates perceived to have more neutral expressions acquired a greater percentage of votes). Moreover, the regression coefficient for American participants' own-culture voting judgments in Study 3A was again significantly greater than was the regression coefficient for Japanese participants' other-culture voting judgments here ($Z = 2.88$, $p = .002$).

As in Study 3A, we examined the relationships between the participants' trait judgments of the candidates and their voting judgments (see Table 6). When controlling for affect and attractiveness, scores on the Warmth composite were significantly correlated with participants' voting judgments, $r(49) = .78$, $p < .001$, but scores on the Power composite showed no significant relationship with participants' voting judgments, $r(49) = .19$, $p = .19$. Moreover, Warmth was significantly more related to participants' voting judgments than was Power ($Z = 4.14$, $p < .001$).

Discussion.   Japanese participants' Power scores significantly predicted the percentage of votes that the American candidates received, replicating Study 1B. In distinction, however, Japanese participants' explicit voting judgments did not predict the American candidates' electoral success. Rather, consistent with self-projection theory, Japanese participants' Warmth scores were significantly correlated with their voting judgments, suggesting that they may have attempted to apply their own cultural values toward predicting Americans' voting behaviors.

In relation, affective expression was negatively related to the percentage of votes that candidates received in the current model, as it was for American participants in Study 3A. This indicates that candidates with more neutral expressions obtained a higher percentage of votes than did candidates with happier expressions. However, judgments of affect were not simply a proxy for participants' voting judgments, otherwise participants' ratings of voting likelihood would have also predicted electoral success, as in Study 3A.

Cultural knowledge therefore seems to affect participants' ability to make explicit predictions of U.S. candidates' electoral success. This suggests that experience with the targets' culture is necessary to make direct judgments about electoral success, perhaps because cultural knowledge is required to do so. Applying one's own cultural schema to judgments of electoral success in a different culture (self-projection) therefore seems to lead to misguided predictions.

Study 4

Study 3A showed that American perceivers' impressions of U.S. candidates' likely electoral success were significantly related to their actual success, but Study 3B showed that Japanese perceivers' impressions of U.S. candidates' likely electoral success did not relate to their actual success. This suggests that some level of cultural knowledge may be necessary to make these explicit judgments. To test this hypothesis, in Study 4 we examined the complement of these effects: American participants provided both trait inferences and explicit judgments of the electability of the candi-

dates from the 2000 election of the Japanese House of Representatives (*shugi-in*) in Study 4A, and Japanese participants provided trait inferences and explicit judgments of the electability of the Japanese candidates in Study 4B.

Study 4A

American participants' explicit judgments of the electoral viability of politicians predicted their actual electoral success in Study 3A when the candidates were from their own culture. Might Americans' explicit judgments also predict electoral success when the candidates are from a different culture? We tested this question in Study 4A.

Method. Twenty-two male and female American undergraduates provided trait inferences (Cronbach's αs > .75) and explicit judgments of electability for the same Japanese political candidates as in Study 2A. Participants first followed the same procedures as in Study 2A. After the trait inferences were complete, the experiment continued with the participants asked to indicate how likely they believed it was that most Japanese people would vote for each candidate along an 8-point scale anchored at *not at all likely* (1) and *very likely* (8). Participants' judgments showed high interrater reliability (Cronbach's α = .77). No participants recognized any of the candidates.

Results.

Trait inferences. Ratings of each trait were averaged across participants for each target, and composites for Power and Warmth were formed, as above. Power, Warmth, affect, and attractiveness were simultaneously regressed onto the percentage of votes that candidates received. Replicating Study 2A, Warmth significantly predicted the percentage of votes that candidates received (β = .39), $t(75)$ = 2.18, p = .03. Attractiveness (β = .10), $t(75)$ = 0.84, p = .40, affect (β = −.39), $t(75)$ = −1.90, p = .06, and Power (β = .03), $t(75)$ = 0.20, p = .84, were all nonsignificant (see Table 4).

Voting judgments. To assess whether American participants' voting judgments might predict the percentage of votes that the Japanese candidates received, we regressed the targets' mean scores for the voting judgments simultaneously with affect and attractiveness; however, none of the predictors was significant (βs < |.18|, ts < 1.36, ps > .18; see Table 5).

Partial correlations (controlling for affect and attractiveness) between participants' voting judgments and their scores on the Power and Warmth composites showed that both Power, $r(76)$ = .66, p < .001, and Warmth, $r(76)$ = .64, p < .001, were significantly related to participants' voting judgments, but the correlations were not significantly different (Z = 0.18, p = .43; see Table 6).

Discussion. Although American participants' naïve and indirect trait inferences predicted the outcomes of the Japanese elections, their explicit judgments did not. The absence of a relationship between the American participants' voting judgments and the percentage of votes acquired by the Japanese candidates may have to do, in part, with Americans' overweighting of the importance of Power in determining election outcomes, consistent with self-projection theory.

Study 4B

In Study 4A, American participants' explicit judgments of the electability of Japanese politicians did not correspond to the candidates' actual success. Consistent with the findings of Study 3B, in which Japanese participants' explicit ratings did not predict the electoral success of U.S. candidates, this suggests that participants' cultural orientation is relevant in making explicit judgments about a candidate's potential for electoral success in a given culture. We further tested this possibility by asking Japanese participants in Study 4B to provide explicit ratings of Japanese political candidates from their own culture, that is the 2000 election of the House of Representatives (*shugi-in*).

Method. Twenty-four male and female undergraduates at a Japanese university repeated the same procedures as in Study 2B (Cronbach's αs > .77) followed by a final block of ratings in which they indicated how likely most people in Japan would be to vote for each target along an 8-point scale anchored at *not at all likely* (1) and *very likely* (8); Cronbach's α = .83. No participants reported recognizing any of the candidates.

Results.

Trait inferences. American and Japanese participants showed high levels of agreement in their ratings of the Japanese candidates, as in Study 2 (r^*s > .38), and the Power composite again showed significantly higher cross-cultural agreement than did the Warmth composite (Z = 4.92, p < .001; see Table 3). We then aggregated across participants' ratings of Japanese candidates in both Studies 2 and 4 to measure the overall agreement between American (Studies 2A and 4A) and Japanese (Studies 2B and 4B) raters. These relationships showed significant cross-cultural consensus (r^*s > .41; see Table 3), and the consensus for Power was significantly greater than the consensus for Warmth (Z = 5.51, p < .001). In addition, comparisons of the strength of consensus for U.S. candidates did not significantly differ from the strength of consensus for Japanese candidates for Power (Z = 0.20, p = .42) or for Warmth (Z = 1.23, p = .11) in the aggregated samples of all of the American and Japanese participants.

Scores for each candidate were averaged across participants and composites for Power and Warmth were created. We then regressed the scores for the Power and Warmth composites onto the percentage of votes that the candidates received in the election with the affect and attractiveness ratings from Study 2B entered simultaneously as covariates. Replicating Study 2B, participants' scores on the Warmth composite were the only significant predictor of the percentage of votes that the candidates received (β = .31), $t(64)$ = 2.64, p = .01; Power (β = .04), $t(64)$ = 0.38, p = .70, affect (β = .02), $t(64)$ = 0.20, p = .84, and attractiveness (β = −.22), $t(64)$ = −1.87, p = .07, were not significant (see Table 5). The parameter estimates for Warmth did not significantly differ between American (Study 4A) and Japanese participants (Z = 0.64, p = .26).

Voting judgments. Scores for participants' voting judgments for each candidate were averaged across participants. These mean voting scores were then simultaneously regressed onto the percentage of votes that the candidates received in the election along with the affect and attractiveness ratings. Participants' voting judgments significantly predicted the percentage of votes that the candidates received (β = .38), $t(65)$ = 2.41, p = .02, whereas affect (β = −.01), $t(65)$ = −0.70, p = .95, and attractiveness (β = −.10), $t(65)$ = −0.64, p = .52, were both nonsignificant. Japanese participants' parameter estimate for the voting judgments of Japanese candidates was greater than Americans participants' parameter estimate for voting judgments of Japanese candidates (Z = 1.69, p < .05).

Partial correlations (controlling for affect and attractiveness) showed that participants' voting judgments were significantly correlated with their scores on the Warmth, $r(65) = .88$, $p < .001$, and Power composites, $r(65) = .29$, $p = .02$ (see Table 6). However, consistent with the relationship between Warmth judgments and actual votes that the candidates received, meta-analytic comparisons showed that the Warmth composite was significantly more strongly related to the participants' subjective voting judgments than was the Power composite ($Z = 6.10$, $p < .001$).

Discussion.   As with American participants judging U.S. candidates in Study 3A, Japanese participants' indirect and explicit judgments of same-culture targets predicted the percentage of votes that the candidates received. This provides further evidence that cultural knowledge may be necessary to make explicit inferences about the electoral success of candidates within a given culture and that when this information is absent, individuals will use self-projection of their own cultural knowledge in an attempt to make the inference.

General Discussion

Perceivers extract information about others from their facial appearance, which is predictive both within and across cultures. In Studies 1–4, perceivers' naïve judgments of traits from political candidates' faces significantly predicted the percentage of votes that the candidates actually received, regardless of the perceiver's culture. The traits that predicted electoral success differed across the two cultures, however. In Studies 3 and 4, perceivers' explicit judgments of electability from political candidates' faces significantly predicted the percentage of votes that the candidates received, but only for those candidates from their own culture. When participants attempted to infer the electoral success of candidates from the other culture, they showed evidence of self-projection.

Previous studies have shown both cross-cultural consensus among perceivers and predictive validity for judgments based on facial cues. In terms of consensus, previous work has shown cross-cultural agreement for judgments of faces for a restricted set of traits (Albright et al., 1997; Zebrowitz et al., 1993). In terms of prediction, previous work has shown that judgments of faces predict ecologically valid outcome criteria, but only cases in which both the perceivers and targets were from very similar cultures (e.g., Poutvaara et al., 2009; Rule & Ambady, 2008, 2009; Todorov et al., 2005). We extend the findings on consensus by showing that there is cross-cultural agreement for judgments of politicians' personality traits and show that some traits (such as those relating to power) show significantly higher agreement across cultures than do others (such as those relating to warmth). Additionally, we extend the findings on prediction by showing that inferences of personality traits by raters in one culture can predict the outcomes of faces from another, unrelated culture.

Although there was high agreement between raters in the United States and raters in Japan for both American and Japanese faces, these ratings related to the outcome criterion differently, depending on the targets' culture. For American political candidates, both American and Japanese perceivers' judgments of Power predicted the percentage of votes that the candidates received. But for Japanese political candidates, American and Japanese perceivers' judgments of Warmth predicted the percentage of votes that the candidates received. These differences are reflective of previously described, qualitative and theoretical accounts of leadership preferences in the United States and Japan (e.g., Den Hartog et al., 1999; Funk, 1996, 1997; Jung & Avolio, 1998; Jung et al., 1995; Misumi & Peterson, 1985). Thus, naïve judgments of personality traits from the faces of American and Japanese political candidates predicted leadership outcomes in a manner congruent with decades of descriptive work. First impressions of traits from candidates' faces are therefore highly predictive, regardless of the culture of the perceiver or the culture of the target.

That perceivers' naïve judgments predicted electoral outcomes that concur with cultural values does not suggest that the perceivers possess intuitive knowledge about the foreign culture. Rather, we found that perceivers' attempts to explicitly judge the electoral success of political candidates' faces were predictive only for their own culture. American perceivers' judgments predicted the percentage of votes that American candidates received but not the percentage of votes that Japanese candidates received. Japanese perceivers' judgments predicted the percentage of votes that Japanese candidates received but not the percentage of votes that American candidates received. These effects are consistent with the naïve trait inferences, as well: Although perceivers agreed in their judgments of the faces, the targets' culture influenced how the traits related to the outcome criterion (electoral success). Thus, Studies 3 and 4 conflated the face-based judgments with cultural knowledge by asking the participants to make inferences that required an understanding of the values of the other culture. Together, the two sets of findings therefore illustrate that inferences across cultures may not apply to characteristics that require specific cultural knowledge. Instead, they present a distinction between cross-cultural agreement in perceptions of traits from targets' faces and cross-cultural differences in how these traits are weighted when making a decision about for whom to vote.

Indeed, the absence of perceivers' knowledge of the foreign culture led them to use self-projection strategies. That is, participants attempted to apply the traits that predicted electoral success in their own culture to the other culture, resulting in a misprediction of the outcome that revealed ethnocentric reasoning. Likely due to an absence of knowledge about the other culture (Robbins & Krueger, 2005), participants essentially predicted that candidates characterized by the traits valued in their own culture would be successful in the other culture. Specifically, American participants showed evidence of assuming that candidates rated high on the Power composite would be more likely to get elected in Japan, whereas Japanese participants showed evidence of assuming that candidates rated high on the Warmth composite would be more likely to get elected in the United States. This occurred even under the specific instructions to rate the targets according to how most people in the candidate's home culture (not the perceiver's culture or the perceiver's personal opinion) would evaluate him.

Notably, this observation is not fully aligned with previous accounts of self-projection. In the self-projection literature, it has typically been suggested that individuals' projection to ingroup members is stronger than their projection to outgroup members (Robbins & Krueger, 2005). Additionally, evidence for self-projection tends to show projection from one individual to another individual, whereas here we essentially observed projection from one group to another group. Although the underlying mechanisms behind these effects might be similar, the latter seems to represent an extension of self-projection theory to the group level. These

findings may also then relate to alternative accounts of inferences about others, such as Ames's (2004) report of "mindreading" behaviors, whereby individuals self-project to similar others but use stereotypes to make inferences about dissimilar others.

In the present work, we have therefore identified a difference between the traits that targets express and what perceivers do with that trait information. The current data have various practical and theoretical applications to these two levels of trait expression and trait perception.

These data may provide insights to cultural values. Perceptions of Power (dominance and facial maturity) are significantly related to electoral success in the United States, and perceptions of Warmth (likeability and trustworthiness) are significantly related to electoral success in Japan. Even though both American and Japanese perceivers agree in their perceptions of faces, regardless of the targets' culture, they appear to possess different ideas about how these traits correspond to what is desirable in an elected leader. If Americans elect leaders who exhibit cross-culturally consistent expressions of power and Japanese elect leaders who exhibit cross-culturally consistent expressions of warmth, it is logical to infer that power and warmth fit the respective cultures' definitions for what constitutes good leadership. Indeed, previous survey and self-report data from the two cultures suggest that these traits roughly correspond to American and Japanese cultural values for leadership.

Theoretically, the data contribute to a greater understanding of the processes that constitute our perceptions of others' traits and how this information is used differently depending on the culture. Funder's (1995) realistic accuracy model (RAM), for instance, suggests that accurate social perceptions are transmuted from a target's actual disposition (Stage 1: relevance) into a set of cues (Stage 2: availability) that then allow for the detection of the information (Stage 3: detection) with the subsequent application of the perceived information (Stage 4: utilization), which is verified by a criterion or outcome. Thus, although we cannot know from these data whether the targets actually possess the traits that they are perceived to, we have evidence that information about traits is reliably detected from facial cues. Previous research has shown that the use and application of trait inferences varies cross-culturally (e.g., Kashima, Kashima, Kim, & Gelfand, 2006; Maass, Karasawa, Politi, & Suga, 2006), suggesting a cross-cultural difference at RAM's Stage 3, the detection or attribution of traits. The current work may provide some suggestion that there are also cultural differences at Stage 4, as American and Japanese participants perceived targets similarly but related different traits to the outcome criterion, electoral success. In that sense, these data may connect this final stage of RAM to a classic Gibsonian (Gibson, 1979) account of person perception. Specifically, if "perception is for doing" it is at this last stage of RAM (utilization) at which "perception is done." Thus, the earlier stages of the RAM may be uninfluenced by social or cultural variables (here perception is not for the perceiver's doing but for the target's expressing), and it is only at the latter stages of the RAM that the Gibsonian perspective applies.

Future Directions

The current studies are limited in a number of ways that may provide interesting opportunities for future research. First, due to the relative paucity of female candidates in the 2000 election of the Japanese House of Representatives, in the present work we examined only male targets. Recent research has suggested that male and female political leaders may be evaluated differently, however (Chiao, Bowman, & Gill, 2008), and it would therefore be interesting to test how these effects might differ for female candidates. Similarly, we used only a small percentage of the total candidates that ran in the Japanese election. Although it would have been difficult for participants to provide trait ratings for all of the more than 700 candidates that ran in the 2000 election (particularly in a repeated-measures design), this presents the possibility for variance across the random sample of winning and losing targets that we selected. Repetition of these studies with other random samples from the same election or subsequent elections would therefore be desirable to bolster the reliability of our effects. Finally, it must be acknowledged that although face-based judgments may predict electoral success, they are unlikely to be the determinants of election outcomes. Indeed, many other factors influence candidates' success and voting decisions. The current work and previous related work merely show that candidates' facial appearances, as illustrated by their public photographs, may be one contributing factor. Although in the present and previous studies relating facial appearance to electoral outcomes, researchers have attempted to control for confounding variables (such as affective expression and facial attractiveness), in future work researchers should seek to explore whether there are other factors that covary with facial appearance that may be contributing to these effects. Finally, although we have assumed that cultural values may be critical in mediating these processes, it should also be acknowledged that we have not tested the role of values, in particular. Rather, other forms of cultural knowledge may be responsible for the effects that we have observed, and one may therefore wish to test the role of values in these effects more specifically in future studies.

Conclusion

The present findings highlight an important breaking point at which culture and person perception are dissociated. We agree across cultures in our perceptions of individuals' qualities from their faces. However, we differ across cultures in how we weight and value these qualities. Thus, our naïve inferences predict different outcomes depending on the target culture, and our explicit inferences are predictive only for the culture to which we possess cultural knowledge. These data therefore provide important insight to the varied nature of first impressions and how they are influenced by and immune to cultural differences.

References

Albright, L., Malloy, T. E., Dong, Q., Kenny, D. A., Fang, X., Winquist, L., et al. (1997). Cross-cultural consensus in personality judgments. *Journal of Personality and Social Psychology, 72,* 558–569.

Allison, P. D. (1999). *Multiple regression: A primer.* Thousand Oaks, CA: Pine Forge Press.

Ames, D. R. (2004). Inside the mid reader's tool kit: Projection and stereotyping in mental state inference. *Journal of Personality and Social Psychology, 87,* 340–353.

Antonakis, J., & Dalgas, O. (2009, February 27). *Predicting elections: Child's play! Science, 323,* 1183.

Ayman, R. (1993). Leadership perception: The role of gender and culture.

In R. M. Chemers & R. Ayman (Eds.), *Leadership theory and research* (pp. 137–165). San Diego, CA: Academic Press.

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

Baum, M. A. (2005). Talking the vote: Why presidential candidates hit the talk show circuit. *American Journal of Political Science, 49,* 213–234.

Burns, J. M. (1978). *Leadership.* New York: Harper & Row.

Chemers, M. M. (1993). An integrative theory of leadership. In R. M. Chemers & R. Ayman (Eds.), *Leadership theory and research* (pp. 293–319). San Diego, CA: Academic Press.

Chiao, J. Y., Bowman, N. E., & Gill, H. (2008). The political gender gap: Gender bias in facial inferences that predict voting behavior. *PLoS ONE, 3,* 3666.

Clement, R. W., & Krueger, J. (2000). The primacy of self-referent information in perceptions of social consensus. *British Journal of Social Psychology, 39,* 279–299.

Clement, R. W., & Krueger, J. (2002). Social categorization moderates social projection. *Journal of Experimental Social Psychology, 38,* 219–231.

Clifford, M. M., & Walster, E. H. (1973). The effect of physical attractiveness on teacher expectation. *Sociology of Education, 46,* 248–258.

Collins, M. A., & Zebrowitz, L. A. (1995). The contributions of appearance to occupational outcomes in civilian and military settings. *Journal of Applied Social Psychology, 25,* 129–163.

Comrey, A. L., & Lee, H. B. (1992). *A first course in factor analysis* (2nd ed.). Hillsdale, NJ: Erlbaum.

Cunningham, M. R., Roberts, A. R., Barbee, A. P., Druen, P. B., & Wu, C. (1995). "Their ideas of beauty are, on the whole, the same as ours": Consistency and variability in the cross-cultural perception of female physical attractiveness. *Journal of Personality and Social Psychology, 68,* 261–279.

Den Hartog, D. N., House, R. J., Hanges, P. J., & Ruiz-Quintanilla, S. A. (1999). Culture specific and cross-culturally generalizable implicit leadership theories: Are attributes of charismatic/transformational leadership universally endorsed? *Leadership Quarterly, 10,* 219–256.

Dickson, M. W., Den Hartog, D. N., & Mitchelson, J. K. (2003). Research on leadership in a cross-cultural context: Making progress, and raising new questions. *Leadership Quarterly, 14,* 729–768.

Eagly, A. H., & Karau, S. S. (1991). Gender and the emergence of leaders: A meta-analysis. *Journal of Personality and Social Psychology, 60,* 685–710.

Fiske, S. T., Cuddy, A. J. C., Glick, P., & Xu, J. (2002). A model of (often mixed) stereotype content: Competence and warmth respectively follow from perceived status and competition. *Journal of Personality and Social Psychology, 82,* 878–902.

Funder, D. C. (1995). On the accuracy of personality judgment: A realistic approach. *Psychological Review, 102,* 652–670.

Funk, C. L. (1996). The impact of scandal on candidate evaluations: An experimental test of the role of candidate traits. *Political Behavior, 18,* 1–24.

Funk, C. L. (1997). Implications of political expertise in candidate trait evaluations. *Political Research Quarterly, 50,* 675–697.

Gibson, J. J. (1979). *The ecological approach to visual perception.* New York: Harper & Row.

Hays, W. L. (1994). *Statistics* (5th ed.). Belmont, CA: Wadsworth.

Hofstede, G. (1980). *Culture's consequences: International differences in work-related values.* Newbury Park, CA: Sage.

Judd, C. M., James-Hawkins, L., Yzerbyt, V., & Kashima, Y. (2005). Fundamental dimensions of social judgment: Understanding the relations between judgments of competence and warmth. *Journal of Personality and Social Psychology, 89,* 899–913.

Jung, D. I., & Avolio, B. J. (1998). Examination of transformational leadership and group process among Caucasian- and Asian-Americans.

Research in International Business and International Relations, 7, 29–66.

Jung, D. I., & Avolio, B. J. (1999). Effects of leadership style and followers' cultural orientation on performance in group and individual task conditions. *Academy of Management Journal, 42,* 208–218.

Jung, D. I., Bass, B. M., & Sosik, J. J. (1995). Bridging leadership and culture: A theoretical consideration of transformational leadership and collectivistic cultures. *The Journal of Leadership Studies, 2,* 4–18.

Kashima, Y., Kashima, E. S., Kim, U., & Gelfand, M. (2006). Describing the social world: How is a person, a group, and a relationship described in the East and the West? *Journal of Experimental Social Psychology, 42,* 388–396.

Keating, C. F., Randall, D., & Kendrick, T. (1999). Presidential physiognomies: Altered images, altered perceptions. *Political Psychology, 20,* 593–610.

Kenny, D. A., Albright, L., Malloy, T. E., & Kashy, D. A. (1994). Consensus in interpersonal perception: Acquaintance and the Big Five. *Psychological Bulletin, 116,* 245–258.

Kenny, D. A., & West, T. V. (2009). Zero acquaintance: Definitions, statistical model, findings, and process. In N. Ambady & J. J. Skowronski (Eds.), *First impressions* (pp. 129–146). New York: Guilford Press.

Kinder, D. R., Peters, M. D., Abelson, R. P., & Fiske, S. T. (1980). Presidential prototypes. *Political Behavior, 2,* 315–337.

Kline, R. B. (1998). *Principles and practice of structural equation modeling.* New York: Guilford Press.

Krueger, J. I. (2007). From social projection to social behavior. *European Review of Social Psychology, 18,* 1–35.

Krueger, J., & Clement, R. W. (1994). The truly false consensus effect: An ineradicable and egocentric bias in social perception. *Journal of Personality and Social Psychology, 67,* 596–610.

Krueger, J., & Clement, R. W. (1996). Inferring category characteristics from sample characteristics: Inductive reasoning and social projection. *Journal of Experimental Psychology: General, 125,* 52–68.

Little, A. C., Burriss, R. P., Jones, B. C., & Roberts, S. C. (2007). Facial appearance affects voting decisions. *Evolution and Human Behavior, 28,* 18–27.

Maass, A., Karasawa, M., Politi, F., & Suga, S. (2006). Do verbs and adjectives play different roles in different cultures? A cross-linguistic analysis of person representation. *Journal of Personality and Social Psychology, 90,* 734–750.

Martin, D. S. (1978). Person perception and real-life electoral behaviour. *Australian Journal of Pyschology, 30,* 255–262.

Mazur, A. (2005). *The biosociology of dominance and deference.* Lanham, MD: Rowman & Littlefield.

Megargee, E. I., Bogart, P., & Anderson, B. J. (1960). The prediction of leadership in a simulated industrial task. *Journal of Applied Psychology, 50,* 292–295.

Meyers, L. S., Gamst, G., & Guarino, A. J. (2006). *Applied multivariate research: Design and interpretation.* Thousand Oaks, CA: Sage.

Misumi, J., & Peterson, M. F. (1985). The performance-maintenance (PM) theory of leadership: Review of a Japanese research program. *Administrative Science Quarterly, 30,* 198–223.

Montepare, J. M., & Zebrowitz, L. A. (1993). A cross-cultural comparison of impressions created by age-related variations in gait. *Journal of Nonverbal Behavior, 17,* 55–68.

Peng, Y., Zebrowitz, L. A., & Lee, H. K. (1993). The impact of cultural background and cross-cultural experience on impressions of American and Korean male speakers. *Journal of Cross-Cultural Psychology, 24,* 203–220.

Peterson, R. A., & Brown, S. P. (2005). On the use of beta coefficients in meta-analysis. *Journal of Applied Psychology, 90,* 175–181.

Poutvaara, P., Jordahl, H., & Berggren, N. (2009). Faces of politicians:

Babyfacedness predicts inferred competence but not electoral success. *Journal of Experimental Social Psychology, 45,* 1132–1135.

Rinn, W. E. (1991). Neuropsychology of facial expression. In R. S. Feldman & B. Rime (Eds.), *Fundamentals of nonverbal behavior* (pp. 3–30). New York: Oxford University Press.

Robbins, J. M., & Krueger, J. I. (2005). Social projection to ingroups and outgroups: A review and meta-analysis. *Personality and Social Psychology Review, 9,* 32–47.

Rosenthal, R., & Rosnow, R. L. (2007). *Essentials of behavioral research* (3rd ed.). Boston: McGraw Hill.

Ross, L., Greene, D., & House, P. (1977). The "false consensus effect": An egocentric bias in social perception and attribution processes. *Journal of Experimental Social Psychology, 13,* 279–301.

Rule, N. O., & Ambady, N. (2008). The face of success: Inferences from chief executive officers' appearance predict company profits. *Psychological Science, 19,* 109–111.

Rule, N. O., & Ambady, N. (2009). She's got the look: Inferences from female chief executive officers' faces predict their success. *Sex Roles.* DOI: 10.1007/s11199-009-9658-9.

Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 25, pp. 1–65). Orlando, FL: Academic Press.

Smith, P. B., Misumi, J., Tayeb, M., Peterson, M. F., & Bond, M. (1989). On the generality of leadership style measures across cultures. *Journal of Occupational Psychology, 62,* 97–109.

Smith, P. B., Peterson, M. F., & Misumi, J. (1994). Event management and work team effectiveness in Japan, Britain and USA. *Journal of Occupational and Organizational Psychology, 67,* 33–43.

Tabachnick, B. G., & Fidell, L. S. (2001). *Using multivariate statistics* (4th ed.). Boston: Allyn & Bacon.

Tiedens, L. Z. (2001). Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral. *Journal of Personality and Social Psychology, 80,* 86–94.

Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005, June 10). Inferences of competence from faces predict election outcomes. *Science, 308,* 1623–1626.

Yamagishi, T., & Yamagishi, M. (1994). Trust and commitment in the United States and Japan. *Motivation and Emotion, 18,* 129–166.

Yamaguchi, S. (1994). Collectivism among the Japanese: A perspective from the self. In U. Kim, H. C. Triandis, C. Kagitcibasi, S. C. Choi, & G. Yoon (Eds.), *Individualism and collectivism: Theory, method, and applications* (pp. 175–188). Thousand Oaks, CA.: Sage.

Zebrowitz, L. A. (1997). *Reading faces: Window to the soul?* Boulder, CO: Westview.

Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' baby-facedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior, 15,* 603–623.

Zebrowitz, L. A., & Montepare, J. M. (2005, June 10). Psychology. Appearance does matter. *Science, 308,* 1565–1566.

Zebrowitz, L. A., Montepare, J. M., & Lee, H. K. (1993). They don't all look alike: Individuated impressions of other racial groups. *Journal of Personality and Social Psychology, 65,* 85–101.