

CHAPTER 4

Evaluating Faces on Social Dimensions

Alexander Todorov

"We look at a person and immediately a certain impression of his character forms itself in us. A glance, a few spoken words are sufficient to tell us a story about a highly complex matter. We know that such impressions form with remarkable rapidity and with great ease. Subsequent observations may enrich or upset our view, but we can no more prevent its rapid growth than we can avoid perceiving a given visual object or hearing a melody" (Asch, 1948, p. 258).

Solomon Asch wrote these words 60 years ago. Since then, social psychologists have amassed evidence supporting his insights. People, indeed, are remarkably good at forming impressions of other people. First, as Asch noted, these impressions are formed from minimal information. They can originate in facial appearance (e.g., Bar, Neta, & Linz, 2006; Olson & Marshuetz, 2005; Willis & Todorov, 2006; Zebrowitz, 1999), "thin slices" of nonverbal behaviors (e.g., Albright, Kenny, & Malloy, 1988; Ambady, Hallahan, & Rosenthal, 1995; Ambady & Rosenthal, 1992), or behavioral information (e.g., Carlston & Skowronski, 1994; Todorov & Uleman, 2002, 2003, 2004; Uleman, Newman, & Moskowitz, 1996). Second, these impressions are formed rapidly and efficiently (Bar et al., 2006; Olson & Marshuetz, 2005; Willis & Todorov, 2006; Todorov, Pakrashi, & Oosterhof, 2009; Todorov & Uleman, 2003). For example, 33-millisecond exposure to a face is sufficient for people to make a trustworthiness judgment (Todorov et al., 2009). Third, these impressions

are formed spontaneously and when cognitive resources are severely limited (Uleman, Blader, & Todorov, 2005). For example, even when people are engaged in a meaningless task of counting nouns while reading behavioral information, they form person impressions (Todorov & Uleman, 2003).

We have not moved beyond the insights of Asch in any fundamental way, but we have moved closer to understanding the cognitive processes underlying impression formation and their neural basis. As the present book testifies, person perception questions are addressed by a variety of novel methods such as fMRI (Chapters 1 and 3), event-related potentials (ERPs; Chapter 6), and the study of patients with brain lesions (Chapter 11).

In this chapter, I focus on how people form person impressions from facial appearance. Faces are a particularly rich source of social information. People use dynamic changes in the face, such as expression of emotions, to understand the immediate meaning of the situation and invariant facial features to identify other people (Haxby, Hoffman, & Gobbini, 2000). Faces are also a rich source of person inferences, although the accuracy of these inferences is dubious (Hassin & Trope, 2000; Olivola & Todorov, 2010; Todorov, 2008). Nevertheless, such person inferences predict important social outcomes (Hamermesh & Biddle, 1994; Hassin & Trope, 2000; Langlois et al., 2000; Montepare & Zebrowitz, 1998;

Zebrowitz, 1999), ranging from electoral success (Hall, Goren, Chaiken, & Todorov, 2009; Little, Burriss, Jones, & Roberts, 2007) to sentencing decisions (Blair, Judd, & Chapleau, 2004; Eberhardt, Davies, Purdie-Vaughns, & Johnson, 2006; Zebrowitz & McDonald, 1991). For example, inferences of competence from faces predict electoral success (Ballew & Todorov, 2007; Todorov, Mandisodza, Goren, & Hall, 2005) and inferences of dominance predict military rank attainment (Mazur, Mazur, & Keating, 1984; Mueller & Mazur, 1996).

Research on face evaluation or how people make personality inferences from facial appearance is situated within the existing cognitive neuroscience models of face perception in the section Cognitive Neuroscience Research on Face Perception. In the sections A Dimensional Model of Evaluation of Emotionally Neutral Faces and Computer Modeling of Face Trustworthiness and Face Dominance, I outline a model of face evaluation on social dimensions. According to this model (Oosterhof & Todorov, 2008), faces are evaluated on two primary, independent dimensions: valence and dominance. Evaluation on specific trait dimensions can be derived from the combination of these two dimensions. Consistent with theories that posit that evaluation of emotionally neutral faces is constructed from facial cues that have evolutionary significance (Zebrowitz, 2004; Zebrowitz & Montepare, 2006, 2008; Zebrowitz et al., 2003), I argue that face evaluation is an overgeneralization of adaptive mechanisms for guiding appropriate social behavior. Specifically, valence evaluation of faces is based on facial cues resembling emotional expressions signaling whether the person should be avoided or approached. Dominance evaluation is based on facial cues signaling the physical strength of the person. Functionally, these types of evaluation correspond to inferences about harmful intentions and the ability to cause harm (cf., Fiske et al., 2007). In section Emotion Overgeneralization Mechanisms, I present additional evidence for the hypothesis that evaluation of faces is an overgeneralization of perception of emotional expressions, and in section The Role of the Amygdala in Evaluation

of Emotionally Neutral Faces, I review evidence for the involvement of the amygdala in valence evaluation of faces. Although people make rapid judgments from faces, they also change their minds in light of new person information. In section Beyond Facial Appearance: Impressions from Behaviors, I review evidence for the role of person knowledge in face perception, and in section Conclusions and Outstanding Questions, I conclude with a sample of outstanding questions for research on face evaluation.

COGNITIVE NEUROSCIENCE RESEARCH ON FACE PERCEPTION

Although there is a large body of cognitive neuroscience research on face perception, almost all of the studies in this tradition focus either on recognition of faces (e.g., Haxby et al., 1999, 2001; Kanwisher et al. 1997; McCarthy et al., 1997) or recognition of expressions of emotions (e.g., Adolphs, 2002, 2003; Calder et al., 2000, 2001; Phan et al., 2002; Morris et al., 1996; Philips et al., 1997). This research has led to great advances in the understanding of face perception. For example, functional imaging studies have shown that whereas areas in the fusiform gyrus are more responsive to facial identity information (Haxby et al., 1999, 2001; Kanwisher et al. 1997; McCarthy et al., 1997), areas in the superior temporal sulcus (STS) are more responsive to expression information (Allison et al., 2000; Hoffman & Haxby, 2000; Puce et al., 1998). Building on existing cognitive models of face processing (e.g., Bruce & Young, 1986), single-cell recording, and functional imaging data, a model of the neural system underlying face perception has been developed to capture the differences between processing of facial identity and emotional expressions (Haxby et al., 2000, 2002; *see also* Chapter 3).

According to this model, the major distinction in face processing is between invariant facial features and dynamic changes such as eye gaze and expressions. Whereas invariant facial features are critical for person recognition, dynamic facial changes communicate the mental states of others. This model can account for observed dissociations between processing of

facial identity and emotional expressions (but see Calder & Young, 2005, for an alternative view). For example, there are prosopagnosics who, despite their inability to recognize faces, show normal perception of emotional expressions (Bentin et al., 2007; Damasio, Tranel, & Damasio, 1990; Duchaine, Parker, & Nakayama, 2003; Humphreys, Avidan, & Behrmann, 2007; Tranel, Damasio, & Damasio, 1988).

However, it is not clear how person inferences such as trustworthiness and competence fit in this distinction. Although such inferences are based on invariant facial features, expressions of emotions affect trait judgments (Knutson, 1996; Krumhuber et al., 2007), and it is possible to observe dissociations between processing of facial identity and impression formation. For example, with Brad Duchaine, we studied four developmental prosopagnosics with severe impairments in both memory for and perception of facial identity (Todorov & Duchaine, 2008). Despite this impairment, their judgments of face trustworthiness across three different face sets were within the normal range of control judgments, and the performance of two of the prosopagnosics was typical. This dissociation suggests that different mechanisms may underlie processing of facial identity and impression formation.

Because of the focus on recognition of faces and emotional expressions, there has been little cognitive neuroscience research on how faces are evaluated on social dimensions. Social cognition research has also largely ignored this topic despite the evidence for the importance of such evaluations. As Macrae and his colleagues have noted: "Although the human face conveys a wealth of potential information, social-cognitive research has focused almost exclusively on identifying the conditions under which categorical knowledge (i.e., stereotypes) is activated in response to available stimulus cues." (Macrae et al., 2005, p. 686).

The one social dimension of face evaluation that has been studied is face trustworthiness. Adolphs, Tranel, and Damasio (1998) showed that patients with bilateral amygdala damage show impaired discrimination between trustworthy- and untrustworthy-looking faces.

Subsequent fMRI studies with normal individuals confirmed the involvement of amygdala in evaluation of faces on trustworthiness (Engell, Haxby, & Todorov, 2007; Said, Baron, & Todorov, 2009; Todorov, Baron, & Oosterhof, 2008; Winston, Strange, O'Doherty, & Dolan, 2002). But what do judgments of trustworthiness measure? As argued by Engell et al. (2007) and as I show below, the reason that the amygdala responds to face trustworthiness is that trustworthiness judgments are a good approximation of the general valence evaluation of faces (Todorov & Engell, 2008).

A DIMENSIONAL MODEL OF EVALUATION OF EMOTIONALLY NEUTRAL FACES

Although people engage in a variety of trait judgments from faces (e.g., Willis & Todorov, 2006), these judgments are highly correlated with each other. For example, for a set of standardized faces (Lundqvist et al., 1998) used in our research, judgments of trustworthiness correlated 0.83 with judgments of emotional stability, 0.75 with judgments of attractiveness, -0.76 with judgments of aggressiveness, and 0.63 with judgments of intelligence. Given the high correlations among judgments of different traits, it is almost impossible to identify (1) neural correlates specific to a trait dimension and (2) facial configurations that vary only along this dimension. For example, if the goal is to model the neural responses to faces as a function of multiple trait judgments, the high correlations among judgments introduce serious collinearity problems.

Instead of working with specific trait dimensions, we have undertaken an approach of reducing judgments on multiple trait dimensions to a few orthogonal dimensions that can account for these judgments (Oosterhof & Todorov, 2008; Todorov, 2008). In a data-driven approach, we first identified trait dimensions on which faces were spontaneously evaluated. Second, we collected judgments on these trait dimensions. Finally, we submitted these judgments to a principal components analysis (PCA) to identify the underlying dimensions of face evaluation.

The first principal component (PC) accounted for 63.3% of the variance of the mean trait judgments. All positive judgments (e.g., attractive, responsible) had positive loadings and all negative judgments (e.g., aggressive) had negative loadings on this component, suggesting that it can be interpreted as valence evaluation (Kim & Rosenberg, 1980; Rosenberg et al., 1968; cf. Osgood et al., 1957). The second PC accounted for 18.3% of the variance. Judgments of dominance, aggressiveness, and confidence had the highest loading on this component, suggesting that it can be interpreted as dominance evaluation. This two-dimensional structure of face evaluation is consistent with well-established dimensional models of social perception (Fiske

et al., 2007; Wiggins, 1979; Wiggins et al., 1989). For example, starting with a large set of traits describing interpersonal relationships, Wiggins and colleagues have shown that these traits can be represented within a two-dimensional space defined by affiliation and dominance, dimensions that are similar to the dimensions identified in our research.

The PCA also showed that the valence and dominance dimensions can be approximated by single trait judgments (Fig. 4-1). Specifically, judgments of trustworthiness had the highest loading (0.94) on the first PC and were practically uncorrelated with the second PC (-0.06). Judgments of dominance had the highest loading (0.93) on the second PC and the lowest

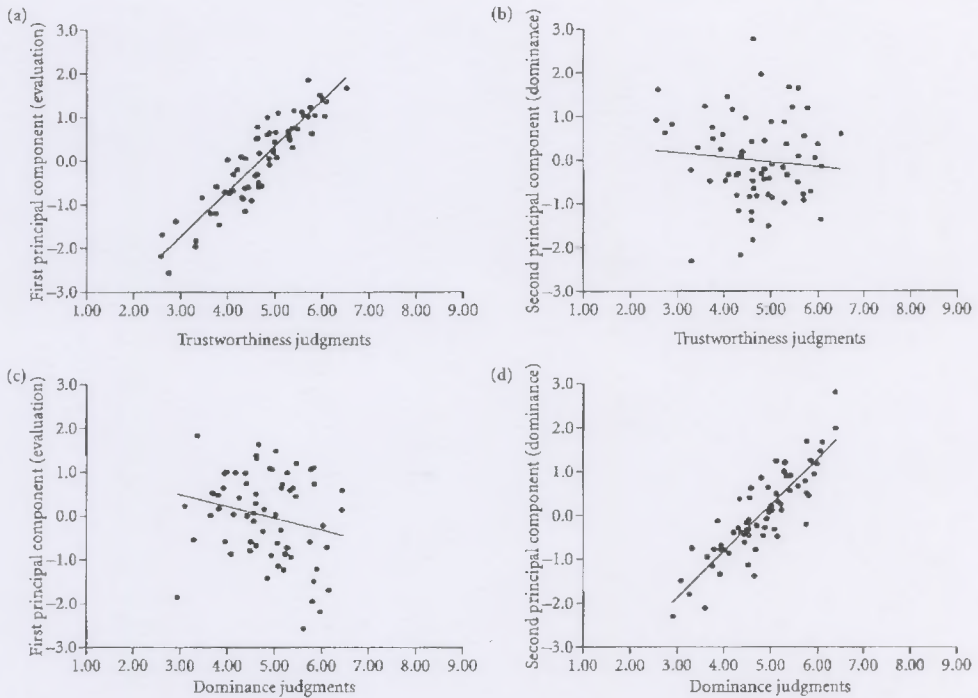


Fig. 4-1 Scatter plots of trustworthiness and dominance judgments from emotionally neutral faces and the first two principal components derived from a principal components analysis of judgments on 11 traits (other than trustworthiness and dominance) used to spontaneously characterize faces. Trustworthiness judgments and (a) the first valence component, and (b) the second dominance component. Dominance judgments and (c) the first valence component, and (d) the second dominance component. Each point is a face. The judgments were measured on a 9-point scale, ranging from 1 (not at all [trustworthy or dominant]) to 9 (extremely [trustworthy or dominant]). The lines represent the best linear fit.

loading on the first PC (-0.24). This was the case even when the principal components were obtained from an analysis excluding these two judgments to avoid biasing the PCA solution. As shown in Figures 4-1a and 4-1b, trustworthiness judgments were highly correlated with the first PC but not with the second PC. In contrast, as shown in Figures 4-1c and 4-1d, dominance judgments were highly correlated with the second PC but not with the first PC. Additional analyses showed that the two-dimensional solution is robust with respect to the set of traits used to estimate the PCs and the face stimuli (Oosterhof & Todorov, 2008).

COMPUTER MODELING OF FACE TRUSTWORTHINESS AND FACE DOMINANCE

Given the findings that judgments of trustworthiness and dominance can be used as approximations of the underlying dimensions—valence and dominance—of face evaluation, we built computer models for representing how faces vary on trustworthiness and dominance. To build trustworthiness and dominance dimensions, we used a data-driven statistical model of face representation, in which faces were represented as points in a multidimensional space (Banz & Vetter, 1999, 2003; Singular Inversions, 2006). The input to this model was a database of faces that were laser-scanned in 3D. The shape of a 3D face was represented by the vertex positions (points in 3D Euclidian space) of a polygonal model of fixed mesh topology. Finally, using PCA, the representation of each face was reduced to a limited number of independent components. We worked with 50 dimensions (50 independent principal components) representing 3D face shape.

Using the face model, we randomly generated emotionally neutral faces. We used only White faces to avoid the influence of stereotypes on trait judgments. We asked participants to judge these faces on trustworthiness and dominance and used the mean trustworthiness and dominance judgments to find vectors (representing a weighted combination of the 50 principal components) in the 50-dimensional face space whose direction was optimal in changing trustworthiness and dominance,

respectively. Specifically, these vectors were based on the best linear fit of the mean judgments as a function of the 50 shape components. Finally, to obtain an orthogonal solution (Fig. 4-2), we rotated the dominance vector to make it orthogonal to the trustworthiness vector (Oosterhof & Todorov, 2008).

To validate the computer models, first, we randomly generated faces. Second, for each face we created several versions that varied along the respective dimensions and, then, asked participants to rate the faces on these dimensions. These studies showed that the models of trustworthiness and dominance successfully manipulated trustworthiness and dominance of faces. Trustworthiness and dominance judgments of faces generated by the models tracked the

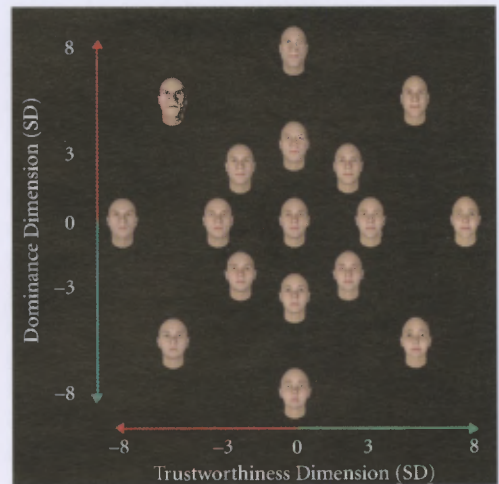


Fig. 4-2 A two-dimensional model of face evaluation. Examples of a face with exaggerated features on the two orthogonal dimensions—trustworthiness and dominance—of face evaluation. The changes in features were implemented in a computer model based on trustworthiness and dominance judgments of 300 emotionally neutral faces. The threat dimension shown on the diagonal from the 4th to the 2nd quadrant was obtained by rotating the trustworthiness dimension 45° clockwise and the dominance dimension 45° counterclockwise in the plane defined by the two dimensions. This threat dimension was practically identical to a dimension based on threat judgments of faces. The extent of face exaggeration is presented in SD units.

trustworthiness and dominance predicted by the model, respectively. Interestingly, whereas dominance judgments of faces generated by the dominance dimension were related in a linear fashion to the face dominance, trustworthiness judgments of faces generated by the trustworthiness dimension were related in a quadratic fashion to face trustworthiness. Specifically, people were more sensitive to changes at the negative end than at the positive end of the trustworthiness dimension.

The validation studies exemplify some of the advantages of using formal models of how faces vary on social dimensions (Todorov, 2008). First, these models can generate an unlimited number of faces that vary on the dimension of interest. Second, the variation of faces can be manipulated precisely (e.g., a face that is 3 SD above the center of the dimension vs. a face that is 3 SD below this center) and the range of differences maximized to detect subtle effects. For example, previous studies have failed to find that trait judgments are made after subliminal exposures to faces (Bar et al., 2006; Todorov et al., Exp. 2, 2009). However, the stimuli may not have been sufficiently different on the trait dimension of interest. We used faces generated by the trustworthiness dimension to test for subliminal effects. Untrustworthy (-3 SD) and trustworthy versions (3 SD) of faces were presented for 20 milliseconds and immediately masked by the neutral version of the faces (0 SD). Trustworthiness judgments of the neutral faces were more negative when these faces were preceded by untrustworthy than by trustworthy faces (Todorov et al., Exp. 3, 2009), although the recognition of the primes was at chance in a forced choice recognition task. These findings suggest that people can extract information for social judgments even when the faces are presented below their level of subjective awareness.

The third and probably most important advantage of computer models is that these models can be used as a discovery tool to identify the variations in facial cues that produce specific judgments. Although these models are holistic in the sense that they are not constrained by any set of facial features, they can be used to

discover the important features *a posteriori*. By exaggerating the features specific to an evaluative dimension, we can identify the type of facial information used for this evaluation. For example, as shown in Figure 4-2, whereas faces at the negative extreme of the trustworthiness dimension (-8 SD) were no longer neutral and looked angry, faces at the positive extreme (8 SD) looked happy. Whereas faces at the negative extreme of the dominance dimension (-8 SD) looked extremely feminine, faces at the positive extreme (8 SD) looked extremely masculine.

Subsequent experiments confirmed that the two dimensions are sensitive to different types of facial information. As in the model validation studies, we randomly generated faces and created extreme versions of the faces on the trustworthiness and dominance dimensions. First, in a study in which participants were asked to categorize these faces as neutral or as expressing one of the six basic emotions, participants classified extremely exaggerated faces in the negative direction on the trustworthiness dimension (-8 SD, see Fig. 4-2) as angry and extremely exaggerated faces in the positive direction (4 and 8 SD) as happy. Although there were fewer emotion categorizations of faces that varied on the dominance dimension, partly because of the fact that we rotated this dimension to make it orthogonal to the trustworthiness dimension, as the faces became more exaggerated in the dominance direction, they were more likely to be classified as angry; and as the faces become exaggerated in the submissiveness direction, they were more likely to be classified as fearful (see Supporting Information Table 7 in Oosterhof & Todorov, 2008). The original dominance dimension based on dominance judgments was negatively correlated with the trustworthiness dimension and would have been even more sensitive to features resembling emotional expressions, although to a lesser extent than the trustworthiness dimension.

Second, in additional nine studies (five of them reported in Oosterhof & Todorov, 2008), participants were asked to rate the faces on continuous scales on angry/happy, baby-faced/mature-faced, and feminine/masculine. We also manipulated the face information available

for the judgments. In three of the studies, participants rated the intact faces, in three studies, they rated the faces with their external features masked, and in three they rated the faces with their internal features masked. As shown in Figure 4-3, these studies showed that whereas the trustworthiness dimension was more sensitive to features resembling happy and angry expressions, the dominance dimension was more sensitive to features signaling physical strength. In particular, most of the diagnostic information for the trustworthiness dimension

was present in the internal features of the face, whereas most of the diagnostic information for the dominance dimension was present in the external (shape) features of the face.

In principle, the two-dimensional model can represent any social judgment from faces, as we have illustrated with judgments of threat (Oosterhof & Todorov, 2008). Threat judgments are particularly important from a survival point of view (Bar et al., 2006), and these judgments are highly correlated with both trustworthiness and dominance judgments. Threatening faces

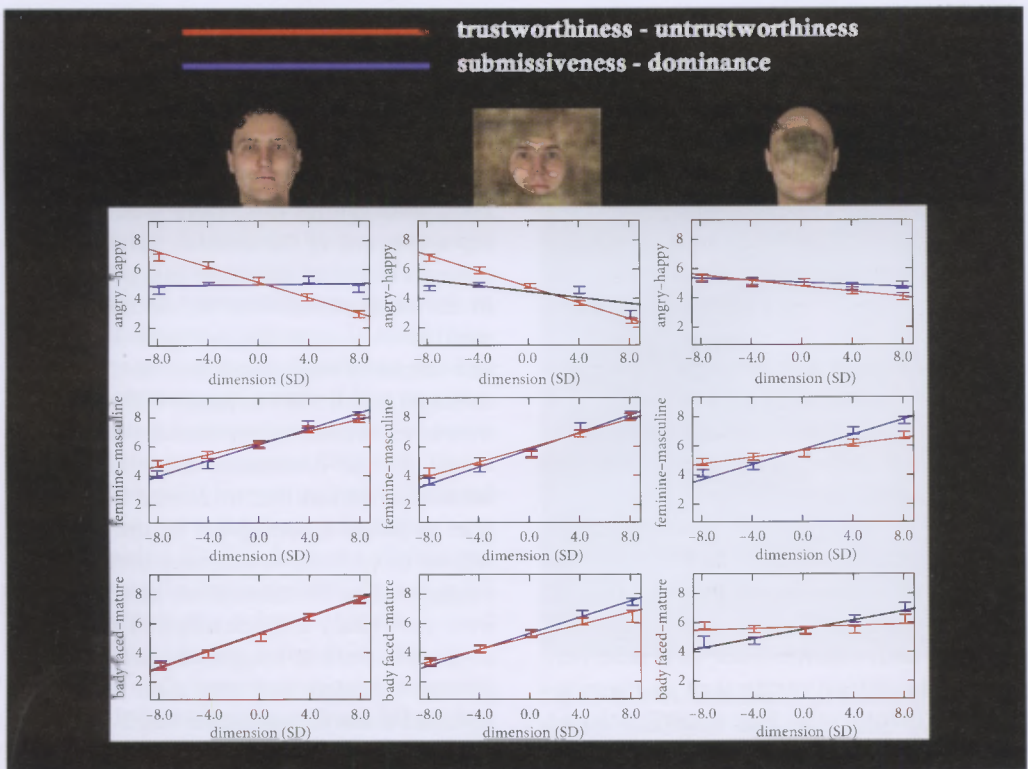


Fig. 4-3 Plots of changes in judgments of expressions of anger/happiness, femininity/masculinity, and facial maturity as a function of trustworthiness and dominance of faces. Expression judgments (1st row) were made on a 9-point scale, ranging from 1 (angry) to 5 (neutral) to 9 (happy). Femininity/masculinity judgments (2nd row) were made on a 9-point scale, ranging from 1 (feminine) to 5 (neutral) to 9 (masculine). Facial maturity judgments (3rd row) were made on a 9-point scale, ranging from 1 (baby-faced) to 5 (neutral) to 9 (mature-faced). Participants made judgments from intact faces (1st column), faces with masked external features (2nd column), and faces with masked internal features (3rd column). Error bars show standard error of the mean. The lines represent the best linear fit. The x-axis in the figures represents the extent of face exaggeration in SD units. The direction of the trustworthiness dimension was reversed for these figures to show that the slopes for the change from trustworthy to untrustworthy faces and the change from submissive to dominant faces were similar for facial maturity and femininity/masculinity judgments of intact faces and faces with masked external features.

are both untrustworthy- and dominant-looking. We built a threat dimension in the space defined by the trustworthiness and dominance dimensions by giving equal weights to these two dimensions (1 and -1 for dominance and trustworthiness, respectively; the diagonal in Fig. 4-2 from the 4th to the 2nd quadrant). This dimension was practically identical to a threat dimension based on threat judgments.

EMOTION OVERGENERALIZATION MECHANISMS

The computer modeling findings suggest that trait judgments are constructed from cues that have evolutionary significance (Zebrowitz, 2004; Zebrowitz & Montepare, 2006, 2008). The primary valence dimension of face evaluation derives from cues resembling expressions of anger and happiness. As Fridlund (1994) has argued, one of the functions of emotional expressions is to signal behavioral intentions. For example, whereas expressions of happiness signal to the perceiver that the person can be approached, expressions of anger signal that the person should be avoided, and there is evidence that angry faces trigger automatic avoidance responses (Adams et al., 2006; Marsh et al., 2005). Thus, consistent with social cognition research suggesting that the valence evaluation of stimuli is directly linked to approach/avoidance behaviors (Chen & Bargh, 1999), the valence evaluation of faces may amount to an approach/avoidance decision. From an evolutionary point of view, the costs of approaching an angry individual are greater than avoiding a happy individual, and this can explain the nonlinearity of trustworthiness judgments. As described above, these judgments were more sensitive to changes at the negative than at the positive end of the trustworthiness dimension. Similarly, threat judgments of faces generated by the threat dimension were more sensitive to changes at the threatening than the nonthreatening end of the dimension.

Consistent with the emotion overgeneralization hypothesis—namely, that similarity of facial features to emotional expressions is attributed to personality traits—previous studies

have shown that emotional expressions affect trait judgments from faces (Hess et al., 2000; Knutson, 1996; Montepare & Dobish, 2003). For example, smiling faces are perceived as more trustworthy than neutral faces (Krumhuber et al., 2007) and higher on affiliation, an attribute similar to trustworthiness (Knutson, 1996; Montepare & Dobish, 2003). Moreover, judgments of anger and happiness from emotionally neutral faces are correlated with judgments of trustworthiness (Todorov & Duchaine, 2008) and judgments of affiliation (Montepare & Dobish, 2003).

To provide an extended replication of these findings, we collected emotion judgments of the emotionally neutral faces for which we had already collected judgments on trait dimensions (see section A Dimensional Model of Evaluation of Emotionally Neutral Faces). As shown in Figure 4-4a, 55 of the 84 (14 trait \times 6 emotion judgments: anger, disgust, fear, sadness, surprise, and happiness) correlations were significant (Said, Sebe, & Todorov, 2009). For example, judgments of happiness were positively correlated with all positive trait judgments and negatively correlated with all negative judgments. The pattern was reversed for judgments of anger. The valence component derived from the trait judgments was strongly correlated with judgments of happiness and anger, moderately correlated with judgments of disgust, and weakly correlated with judgments of sadness (Fig. 4-4b). Faces that were evaluated positively were perceived as happier and more surprised but less angry, less disgusted, less fearful, and less sad. The dominance component was correlated with judgments of anger, surprise, sadness, and fear (Fig. 4-4b). Dominant faces were perceived as angrier, less sad, less fearful, and less surprised than submissive faces.

Although these findings are consistent with the emotion overgeneralization hypothesis, they are also consistent with the hypothesis that these correlations can be accounted for by common semantic properties of emotion and trait judgments rather than by perceptual similarity. For example, expectations about the relation between emotional states (e.g., smiling as an expression of happiness) and personality

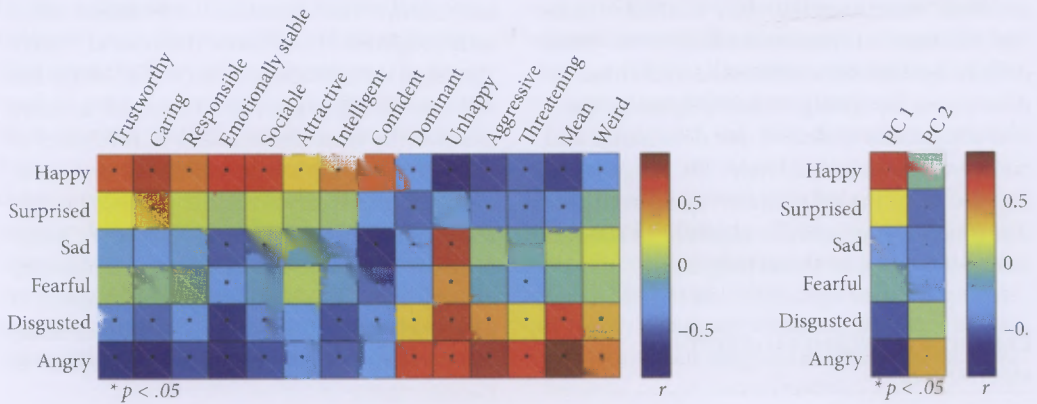


Fig. 4-4 Color maps of correlations between trait and emotion judgments of emotionally neutral faces (a). Traits are ordered by their loadings on the first principal component (PC)—valence—derived from a principal components analysis (PCA) of the trait judgments (see the section A Dimensional Model of Evaluation of Emotionally Neutral Faces). Correlations between emotion judgments and the first two PCs—valence and dominance—derived from a PCA of the trait judgments (b).

traits (e.g., sociable) may lead to strong associations between emotion and trait judgments. This hypothesis is consistent with research on implicit personality theory that shows that people hold assumptions about the relationships between various traits (Bruner & Tagiuri, 1954; Cronbach, 1955; Schneider, 1973). In fact, the dimensional structure of the emotion judgments was very similar to the dimensional structure of the trait judgments. The first PC derived from a PCA of the emotion judgments was highly correlated with the valence component of the trait judgments and uncorrelated with the dominance component. In contrast, the second PC of the emotion judgments was more strongly correlated with the dominance component than with the valence component.

To rule out the possibility that the relations between trait and emotion judgments can be accounted entirely for by semantic similarities, we used an emotion classifier to categorize the emotionally neutral faces (Said et al., 2009). Specifically, we used a Bayesian network classifier to detect the subtle presence of features resembling emotions in the faces. The classifier accepts as input a feature vector containing the displacements between automatically chosen landmarks and the same landmarks of a prototypical neutral face and outputs a set of

probabilities corresponding to each basic emotion. Because we applied the classifier to neutral faces, the output probabilities were very low. Nevertheless, these probabilities predicted trait judgments from the faces.

The pattern of correlations was similar to the pattern of correlations for emotion and trait judgments, although the correlations were weaker (27 of the 84 probabilities—trait judgments correlations were significant). The probability of classifying faces as happy was positively correlated with all positive trait judgments and negatively correlated with all negative judgments. The probability of classifying faces as angry was positively correlated with judgments of aggressiveness, meanness, unhappiness, and dominance. The valence component was positively correlated with the classifier probabilities of happiness and negatively correlated with the probabilities of anger, disgust, and fear, although only the correlation for happiness reached significance. The dominance component was positively correlated with the classifier probabilities of anger and negatively correlated with the probabilities of surprise and fear.

Although the methods used in Oosterhof and Todorov (2008) and in Said et al. (2009) differed in a number of ways, they converged on similar solutions. Faces with positive valence

(trustworthiness) were more likely to be classified as happy and less likely to be classified as angry and disgusted than faces with negative valence. Highly dominant faces were more likely to be classified as angry and less likely to be classified as fearful than highly submissive faces.

To the extent that structural facial features signaling positive valence or trustworthiness are similar to expressions of anger and happiness, it should also be possible to demonstrate that facial features affect the perception of emotional expressions. To test this hypothesis, based on prior trustworthiness judgments, we selected trustworthy and untrustworthy faces and created dynamic stimuli in which the faces expressed either happiness or anger (Oosterhof & Todorov, 2009). Although we added the same amount of emotional intensity to faces, trustworthy faces expressing happiness were perceived as happier than untrustworthy faces. In contrast, untrustworthy faces expressing anger were perceived as angrier than trustworthy faces expressing the same emotion.

We also manipulated changes in trustworthiness during the course of the animation. For example, in incongruent animations, an untrustworthy (or a trustworthy) face gradually morphed into a trustworthy (or an

untrustworthy) face. To the extent that trait judgments are an overgeneralization of cues resembling expressions, changes that are in the direction of the expressed emotion (e.g., untrustworthy-to-trustworthy and happiness) should amplify the intensity of the perceived emotion. In contrast, changes in the opposite direction (e.g., untrustworthy-to-trustworthy and anger) should dampen this intensity. As shown in Figure 4–5, this is exactly what we found. For example, when a trustworthy face changed into an untrustworthy face, the same angry expression was perceived as angrier than when an untrustworthy face changed into another untrustworthy face or when there was no change in the identity of the face. Similarly, when an untrustworthy face changed into a trustworthy face, the same angry expression was perceived as less angry than when a trustworthy face changed into another trustworthy face or when there was no change in the identity of the face.

To test for similarities in the neural codes of perceived trustworthiness and expressions of anger and happiness, we used a behavioral adaptation paradigm (Engell, Todorov, & Haxby, in press). The adaptation paradigm has been used to investigate other dimensions of the neural

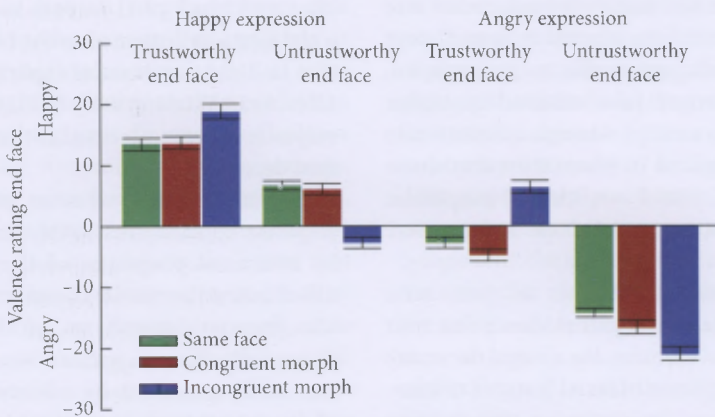


Fig. 4–5 Valence ratings of emotional expressions as a function of the type of emotion, the trustworthiness of the face, and the morphing condition: same face with no change in identity; congruent morph with no change in face trustworthiness but change in identity; and incongruent morph with changes in both face trustworthiness and identity. The ratings were made on a continuous slider ranging from angry to neutral to happy. The error bars show standard errors of the means.

representation of faces, including viewpoint invariance, gender, attractiveness, and expression (e.g., Fox & Barton, 2007; Jeffrey et al., 2006; Rhodes et al., 2006; Webster et al., 1999, 2004). The central tenet of this paradigm is that extended exposure to a stimulus dimension results in fatigue of the neural population that represents the stimulus. Thus, subsequent exposure to a stimulus along the same dimension should result in a perceptual shift away from the adapting stimulus. For example, Webster and colleagues (2004) showed that androgynous faces (i.e., faces that had an equal probability of being categorized by participants as "male" or "female") were seen as distinctly "male" after extended exposure to female faces and as distinctly "female" after extended exposure to male faces.

If trustworthiness evaluation is an overgeneralization of perceiving features resembling angry and happy facial expressions, then we should be able to influence this evaluation by first adapting the neural populations that support the perception of those expressions. In the pre-adaptation stage of the experiment, participants rated the trustworthiness of faces. After the pre-adaptation stage, participants were randomly assigned to one of three adapting conditions: passive viewing of angry, fearful, or happy expressions for 66 seconds. After the adaptation, participants rated the trustworthiness of faces again. The test faces were reduced in size to 80% of the size of the adapter faces to disrupt any low-level adaptation effects. As expected, adaptation to angry faces resulted in higher trustworthiness ratings, whereas adaptation to happy faces resulted in lower trustworthiness ratings. In the control condition of adaptation to fearful faces, trustworthiness ratings were not influenced.

To conclude, several lines of behavioral research provide convergent evidence that trait inferences from emotionally neutral faces are based on resemblance of facial features to emotional expressions. In particular, the evidence suggests that the primary, valence dimension of face evaluation is derived from similarity of facial features to expressions of anger and happiness, expressions that signal potential behavioral intentions.

THE ROLE OF THE AMYGDALA IN EVALUATION OF EMOTIONALLY NEUTRAL FACES

As noted earlier, the amygdala, a subcortical brain region critical for evaluation of novel stimuli, fear conditioning, and consolidation of emotional memories (Amaral, 2002; Davis & Whalen, 2001; Phelps & LeDoux, 2005; Vuilleumier, 2005), has been implicated in the evaluation of face trustworthiness (Adolphs et al., 1998; Engell et al., 2007; Todorov et al., 2008; Winston et al., 2002). Following the Adolphs et al. (1998) findings from patients with bilateral amygdala damage, in an fMRI study with normal participants, Winston and colleagues showed that the amygdala's response to faces increased as their subjectively perceived trustworthiness decreased (Winston et al., 2002). This was the case independent of whether the evaluation task was explicit (judging the trustworthiness of faces) or implicit (judging the age of faces).

We replicated Winston et al.'s findings, using a single implicit task to rule out the possibility that the performance on implicit evaluation trials was influenced by prior performance on explicit evaluation trials (Engell et al., 2007). Participants were presented with a series of faces in an ostensibly memory task and asked after each block of 11 faces to indicate whether a test face was presented in the block. Although this task did not demand explicit evaluation of faces, as in Winston et al. (2002), the amygdala response to faces increased as their trustworthiness decreased.

We also tested whether the amygdala's response to face trustworthiness was driven by structural properties of the face that signaled untrustworthiness across observers or by idiosyncratic components of trustworthiness judgments. The amygdala's response to faces was better predicted by consensus judgments of trustworthiness—aggregated across a large number of participants separate from the fMRI participants—than by the fMRI participants' individual judgments. When the analysis controlled for the shared variance of individual and consensus judgments, there was little residual

variance accounted for by individual judgments in the amygdala.

In a subsequent study, we first built a computer model of face trustworthiness based on behavioral judgments (this work preceded Oosterhof & Todorov, 2008). Second, we generated novel faces based on this model. Third, we used these novel faces in an fMRI study, using the same implicit task as in Engell et al. (2007). As in the previous studies, we found that the right amygdala's response to faces increased as their trustworthiness decreased (Todorov et al., 2008). However, we also found a nonlinear response in the left amygdala so that extremely trustworthy faces evoked a stronger response than faces at the middle of the dimension. This finding is discussed at the end of the section.

Across three different studies, the response to faces in the amygdala was linearly related to judgments of face trustworthiness. However, as described in section A Dimensional Model of Evaluation of Emotionally Neutral Faces, trustworthiness judgments are highly correlated with other social judgments and approximate the valence evaluation of emotionally neutral faces (Fig. 4-1). Because we used the same set of faces in one of our prior fMRI studies (Engell et al., 2007) as in the behavioral studies in which we collected trait judgments (see section A Dimensional Model of Evaluation of Emotionally Neutral Faces), we were able to test the hypothesis that the amygdala is involved in general valence evaluation of emotionally neutral faces rather than in evaluation of faces on specific trait dimensions (Todorov & Engell, 2008). According to this hypothesis, face variations on any social dimension (e.g., trustworthiness, attractiveness, aggressiveness) should engage the amygdala to the extent that this dimension has a valence content. In other words, variations on dimensions with clear valence connotations (e.g., trustworthiness and meanness) should engage the amygdala more strongly than variations on dimensions with less clear valence connotations (e.g., dominance).

To test the valence hypothesis, we derived the response to each of the faces in face responsive voxels in the amygdala and then correlated this response with the mean trait evaluations of

the faces. Consistent with this hypothesis and as shown in Figure 4-6, the amygdala activation correlated negatively with all judgments on positive traits (e.g., caring, trustworthy, attractive) and positively with all judgments on negative traits (e.g., mean, weird). That is, across trait dimensions, the amygdala responded more strongly to faces that were evaluated negatively. Although all trait judgments (except for dominance) correlated significantly with the amygdala's response, there was considerable variation in the magnitude of the correlations. According to the valence hypothesis, this variation should be predicted by the valence content of the specific judgments.

We used the valence component from the PCA of the trait judgments (see section A Dimensional Model of Evaluation of Emotionally Neutral Faces) as a measure of general valence evaluation. This valence component was correlated with both the response in the right and left amygdala ($r = -0.50$ and -0.48 , respectively, $p < 0.001$, Fig. 4-6c & 4-6d). For comparison, the amygdala's response was uncorrelated with the dominance component ($r = 0.06$ and 0.07 , for right and left amygdala, respectively).

We used the variance accounted for by the valence component for each trait judgment as an estimate of the valence content of the trait dimension. For example, the valence component accounted for 90% of the variance of trustworthiness judgments and 9% of the variance of dominance judgments. This variance was strongly correlated with the variance accounted for by each judgment in the amygdala's response to faces ($r = 0.90$ and 0.79 for right and left amygdala, respectively, $p < 0.001$). The stronger the association of a trait judgment with the valence component, the stronger this judgment engaged the amygdala. Moreover, after controlling for the valence content of the trait judgments, there were no significant relationships between any of the judgments and the amygdala's response (Fig. 4-6b).

We found the same pattern of responses in face responsive regions in temporal and occipital cortices—specifically in the right superior occipital gyrus, bilateral fusiform gyri, and the right middle temporal/occipital gyrus. In all

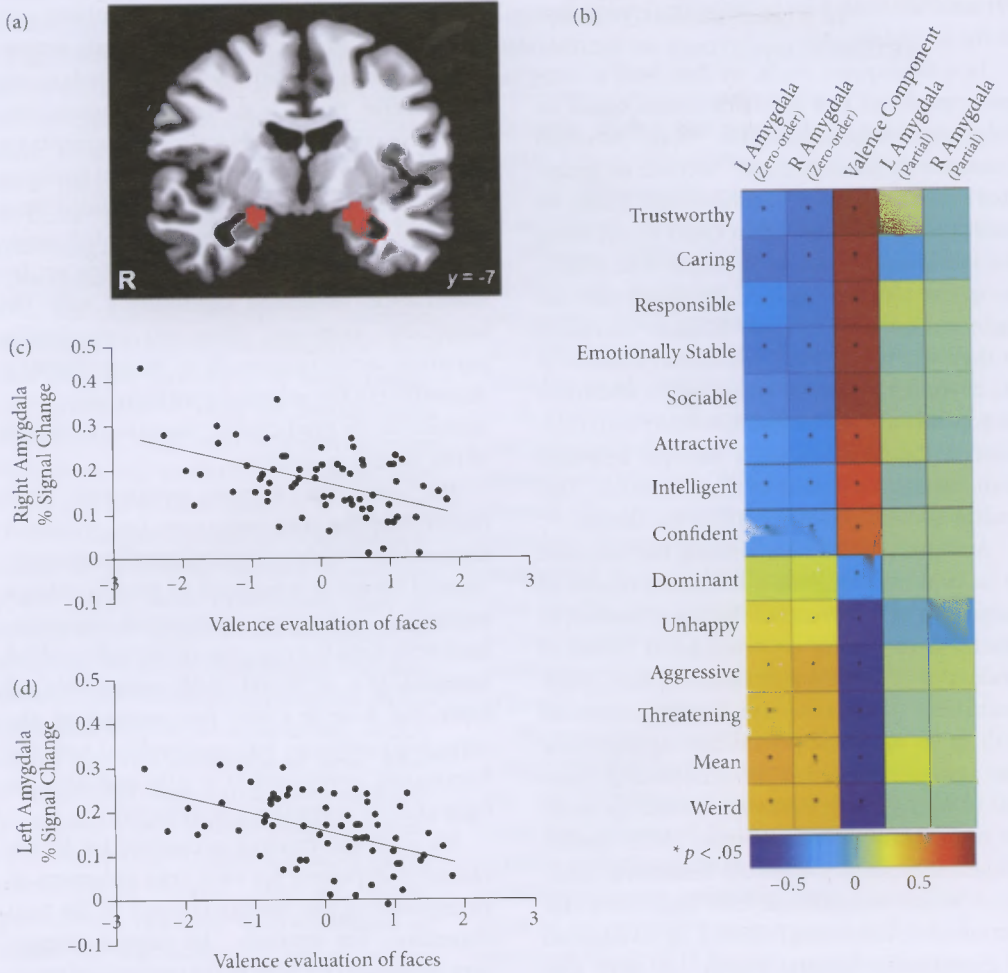


Fig. 4-6 The relation between the amygdala’s response to emotionally neutral faces and variations of these faces on trait dimensions. A coronal brain slice showing face responsive voxels in bilateral amygdala (a). An intensity color plot showing correlations between the response in left and right amygdalae to faces and trait judgments of these faces (b). The first two columns show zero-order correlations and the fourth and fifth columns show partial correlations controlling for the valence content of the judgments. The third column shows the correlations between trait judgments and a valence component derived from a principal components analysis of the judgments. The traits are ordered according to their correlations with the valence component (see the section A Dimensional Model of Evaluation of Emotionally Neutral Faces). Scatter plots of the amygdala’s response to faces (c for right and d for left) and their values on the valence component. Each point represents a face.

these regions, the response to faces was correlated with their valence, and after controlling for the valence content of specific trait judgments, there were no significant relationships between judgments and brain activation. These findings suggest that the valence evaluation of faces recruits a network of perceptual regions

in temporal and occipital cortices. Additional analyses have suggested that the response in these regions is modulated by the amygdala. Specifically, controlling for the amygdala’s response to faces, the relationship between the activation in these regions and face valence was no longer significant. In contrast, the

relationship between the activation in the amygdala and face valence remained significant after controlling for the activation in these regions.

These findings are consistent with the hypothesis that the amygdala amplifies attention to emotionally salient stimuli in perceptual regions (Vuilleumier, 2005). Although such correlational findings cannot be conclusive for a causal influence of the amygdala on perceptual regions in temporal and occipital cortex, Vuilleumier et al. (2004) showed that whereas patients with hippocampal lesions show enhanced responses in regions in occipital and inferotemporal cortex to emotionally salient but unattended stimuli, patients with amygdala lesions do not show such enhanced responses. These regions included the same regions observed in our study. In addition, anatomical evidence from tracing studies of the macaque brain shows that the projections from the amygdala to visual cortex are more extensive than those from visual cortex to the amygdala (Amaral et al., 2003). Whereas the amygdala receives visual input only from temporal visual areas, it projects to multiple areas in both temporal and occipital visual areas, including early visual areas.

The findings suggest that the amygdala automatically evaluates novel faces along a general valence dimension and that it modulates a face responsive network of regions in occipital and temporal cortices recruited for this evaluation. The extent to which the amygdala is engaged in tracking variations of faces on social dimensions is a function of the valence content of these dimensions. Given the high correlation between trustworthiness judgments and valence evaluation of faces (Fig. 4-1a), it is not surprising that previous studies have found that the amygdala is engaged in the evaluation of face trustworthiness. However, in light of the current findings, it would be misleading to describe the amygdala's response to emotionally neutral faces as driven by their trustworthiness. In terms of practical implications, it would often be unfeasible to collect multiple social judgments of faces to estimate their valence evaluation, although some of our analyses suggest that a robust estimation

of face valence can be achieved with as few as five different social judgments (Oosterhof & Todorov, 2008). If, in fact, it is unfeasible to collect multiple judgments, then it would be best to collect judgments of trustworthiness as an *approximation* of general valence evaluation.

As shown in Figure 4-6c and 4-6d, the response of the amygdala to face valence was linear. However, there have been three recent studies—two coming from our lab—reporting a nonlinear amygdala response to face trustworthiness (Said et al., 2009; Todorov et al., 2008) and face attractiveness (Winston et al., 2007). Specifically, as noted for the left amygdala in one of our prior studies (Todorov et al., 2008), the activation was stronger to faces at the extremes of the dimensions than to faces at the middle of the dimension.

The nonlinear responses to face trustworthiness are broadly consistent with the emotion overgeneralization hypothesis (section Emotion Overgeneralization Mechanisms). As our modeling and behavioral findings showed (sections on Computer Modeling of Face Trustworthiness and Face Dominance and Emotion Overgeneralization Mechanisms), variations on the dimension of trustworthiness can be understood in terms of similarity to expressions of happiness on the positive extreme of the dimension and expressions of anger on the negative end. Given that a number of functional neuro-imaging studies have found a stronger amygdala response to happy than to neutral faces (e.g., Breiter et al., 1996; Pessoa et al., 2006; Winston, O'Doherty, & Dolan, 2003; Yang et al., 2002), it should be possible to observe a nonlinear response to face trustworthiness with elevated responses to both extremely trustworthy and untrustworthy faces.

Even if this is the case, one should be able to specify the conditions under which the amygdala's response to face valence is linear and the conditions under which the response is nonlinear. There are at least two hypotheses about these conditions. According to the first hypothesis, the nature of the evaluation—implicit versus explicit—may be critical. In contrast to the study by Engell et al. (2007), participants in Said et al.'s study explicitly evaluated the

faces on trustworthiness, and this may have biased attention to extreme faces. In a recent study, Cunningham et al. (2008) observed similar quadratic responses in the amygdala in a valence evaluation task of famous people. When participants focused on the positivity of the evaluation, the response was enhanced to positive stimuli; when they focused on the negativity, the response was enhanced to negative stimuli.

However, this hypothesis cannot account for all of the data. In Todorov et al. (2008), the task was the same as the task used in Engell et al. According to the second hypothesis, the range of face valence used in a particular study may determine the nature of the amygdala's response. For wider ranges of face valence, the response may be quadratic. For example, we compared the trustworthiness of the faces used in Todorov et al. (2008) and the faces used in Engell et al. (2007) in our computer model of face trustworthiness (Oosterhof & Todorov, 2008). The range of trustworthiness in the former study was from -3.26 to 2.64 in SD units, whereas the range in the latter study was from -1.79 to 1.53 . The range for the faces used in Said et al.'s study, in which participants explicitly evaluated the faces, was from -2.71 to 1.37 . Studies on attractiveness typically use extreme faces (O'Doherty et al., 2003; Winston et al., 2007), and given the high correlation between attractiveness and face valence, this can lead to nonlinear responses in the amygdala as observed by Winston et al. (2007).

Both of these hypotheses, as well as linear and nonlinear responses in the amygdala, are consistent with a common attentional mechanism according to which the amygdala biases attention toward stimuli that are of current motivational significance to the person (Cunningham et al., 2008; LaBar et al., 2001; Vuilleumier, 2005). Interestingly, early studies in social cognition showed that allocation of attention to social stimuli exhibits nonlinear quadratic response to people as a function of their extremeness rather than their valence (Fiske, 1980), and more recent studies have shown that evaluative processes are context-dependent (Ferguson & Bargh, 2004).

As argued in section Emotion Overgeneralization Mechanisms, valence evaluation of faces may be in the service of approach/avoidance decisions. This is consistent with findings that macaque monkeys with bilateral amygdala lesions exhibit uninhibited approach behaviors during social interactions (Emery et al., 2001) and with theories that posit that one of the primary functions of the amygdala is to provide continuous vigilance by evaluating objects and agents prior to interacting with them (Amaral, 2002; Whalen, 1998). Evaluation processes in the amygdala may not only enhance attention and processing of stimuli in perceptual areas (Anderson & Phelps, 2001; Vuilleumier et al., 2004) but may also influence approach/avoidance decisions via interactions with orbital frontal cortex (Baxter et al., 2000).

BEYOND FACIAL APPEARANCE: IMPRESSIONS FROM BEHAVIORS

Before the widespread use of fMRI to study the neural basis of social cognition, Leslie Brothers wrote that "the visual appearance of a face in social cognition is analogous to a stream of speech in linguistic processing: the face stimulus is immediately and obligatorily transformed into the representation of a person (with dispositions and intentions) before having access to consciousness." (Brothers, 1990, p. 35). These prescient insights are consistent with recent behavioral (Todorov & Uleman, 2004) and neuro-imaging studies (Chapter 3; Todorov, Gobbini, Evans, & Haxby, 2007). As Gobbini and her colleagues have shown, faces of significant others activate a network of regions implicated in social cognition such as the medial prefrontal cortex (mPFC) and precuneus (Chapter 3; Gobbini & Haxby, 2007; Gobbini, Leibenluft, Santiago, & Haxby, 2004).

These effects are based on prior person knowledge rather than on facial appearance. There is a long tradition of research in social psychology showing that people form impressions from observing the behaviors of other people (e.g., Gilbert & Malone, 1995; Jones & Davis, 1965; Trope, 1986; for a review, see

Gilbert, 1998). A number of studies on spontaneous trait inferences (STIs) from behaviors have demonstrated that such inferences are associated with the faces that accompanied the behaviors (Carlston & Skowronski, 1994; Carlston, Skowronski, & Sparks, 1995; Goren & Todorov, 2009; Todorov & Uleman, 2002, 2003, 2004). Importantly, in our STI studies, we randomly assigned behaviors to faces to avoid effects of facial appearance on inferences and subsequent judgments.

To study whether rapidly acquired person knowledge affects the neural representation of faces, we conducted an fMRI study (Todorov et al., 2007) modeled after our behavioral paradigm in which faces are presented with single behavioral descriptions for a few seconds. In the first stage of the study, participants familiarized themselves with the faces and behaviors. In the second stage, they were presented with the faces that were associated with behaviors intermixed with novel faces. Although the task was perceptual—deciding whether each face was the same as the preceding one—and did not demand retrieval of person knowledge, the rapidly acquired person knowledge modulated the response to faces in a number of brain regions. Specifically, faces that were associated with behaviors evoked a stronger response in the mPFC and the STS than novel faces. Moreover, the type of behaviors associated with the faces affected the response to the faces. For example, faces associated with disgusting behaviors evoked a stronger response in the anterior insula, a region implicated in the processing of disgust related stimuli (Calder et al., 2000; Phillips et al., 1997), than faces associated with aggressive behaviors. These findings are consistent with Brothers' hypothesis that person knowledge is automatically retrieved in the process of face perception.

From an adaptive point of view, people should be able to rapidly learn about other people and overwrite initial impressions. The robustness of the learning process is demonstrated by findings that person learning (1) occurs after minimal time exposure to faces and behaviors; (2) is relatively independent of

availability of cognitive resources; (3) is independent of explicit goals to form impressions; and (4) subsequent effects on perception and judgments are independent of explicit memory for the behaviors (Todorov & Uleman, 2003).

Findings from studies of patients with brain lesions are consistent with the idea of robust person learning mechanisms (Johnson et al., 1985; Tranel & Damasio, 1993; Todorov & Olson, 2008). In a particularly striking case of brain damage, Tranel and Damasio described a patient (Boswell) with an extensive damage in the medial temporal lobe and orbitofrontal cortex. Boswell had dense amnesia, did not recognize the faces of caregivers, and did not even show increased galvanic skin response to familiar faces, an index of implicit face processing. However, if consistently treated nicely by a caregiver, Boswell had a reliable preference for her face in forced choice preference tasks.

This case is extreme, but research with Korsakoff patients has also showed that they can acquire and preserve affective responses to people's faces despite lack of explicit memory (Johnson et al., 1985). Johnson et al. presented such patients with two pictures and described one of the people as bad (e.g., "... stole a car ... robbed an old man who lived in the neighborhood") and the other as good (e.g., "... joined the Navy ... saved a fellow sailor"). The patients reliably preferred the good person despite lack of memory for the origin of these impressions.

Recently, in a conceptual replication of Johnson et al., we studied how inferences from facial appearance and behavioral descriptions were integrated in person impressions (Todorov & Olson, 2008). Normal participants and three patients with amnesia caused by lesions in the hippocampus were presented with trustworthy- and untrustworthy-looking faces paired with trustworthy and untrustworthy behaviors. After the learning stage of the experiment, participants were asked to judge the faces on a number of trait dimensions. One of the patients with a localized lesion in the hippocampus showed excellent learning just as young and older control participants

did. Faces associated with positive behaviors were judged more positively than faces associated with negative behaviors, and this learning effect was stronger than the effect of facial appearance on judgments. The other two patients, whose lesions extended into the left amygdala and left temporal pole, showed little evidence of learning. At the same time, all patients showed effects of facial appearance on judgments similar to the effects observed for prosopagnosics (Todorov & Duchaine, 2008). These findings suggest that the hippocampus may not be necessary for forming of affective associations with faces. Other structures in the medial temporal lobe like the amygdala may be critical for this process (Somerville, Wig, Whalen, & Kelley, 2006).

The findings show that learning can overwrite initial impressions based on facial appearance. However, at present, we lack models specifying how person representations are dynamically updated in the brain. One of the most important tasks for future research is to specify models of how different sources of person information are integrated in coherent person representations.

CONCLUSIONS AND OUTSTANDING QUESTIONS

People rapidly form impressions of other people based on minimal information. In this chapter, I focused on the processes underlying evaluation of faces. Although people evaluate faces on multiple trait dimensions, these evaluations are highly correlated with each other. Findings from data-driven methods suggest that these evaluations can be represented within a two-dimensional space defined by valence and dominance evaluation of faces (section A Dimensional Model of Evaluation of Emotionally Neutral Faces). Computer modeling findings suggest that whereas valence evaluation is based on facial cues resembling emotional expressions signaling approach/avoidance behavior, dominance evaluation is based on cues signaling physical strength (section Computer Modeling of Face Trustworthiness and Face Dominance). Additional behavioral and computer modeling

studies provide convergent evidence that evaluation of emotionally neutral faces is rooted in adaptive mechanisms for inferring emotional states and corresponding behavioral intentions (section Emotion Overgeneralization Mechanisms). The two-dimensional model provides a unifying framework for the study of face evaluation. In light of this framework, a re-analysis of functional neuro-imaging data from an implicit face evaluation paradigm has showed that the amygdala (1) is engaged in general valence evaluation rather than in specific trait evaluations of faces and (2) modulates the activity in perceptual areas in temporal and occipital cortices (section The Role of the Amygdala in Evaluation of Emotionally Neutral Faces). Finally, I reviewed initial evidence about how person representations are updated in light of new knowledge (section Beyond Facial Appearance: Impressions From Behaviors).

Although we have made progress in identifying some of the key phenomena in face evaluation and the key brain regions involved in this evaluation, there are a number of outstanding questions. First, although people do agree in making judgments from facial appearance, there are also individual differences in these judgments (Engell et al., 2007; Hönokopp, 2006). Understanding these differences may be critical for understanding perceptual learning and top-down effects on social perception. Two likely sources of idiosyncratic contributions to judgments of novel faces are self-resemblance (e.g., DeBruine, 2002, 2005) and resemblance to faces of familiar people. This latter source is directly related to how person knowledge can affect evaluation of novel faces. At present, there are no compelling tests of this hypothesis. Similarly, we do not know what neural regions subserve idiosyncratic contributions to person judgments. To begin addressing these questions, we need statistical models that can estimate consensus and idiosyncratic contributions to these judgments.

Second, the dimensional model of face evaluation was designed as a general model of implicit face evaluation and may be most applicable to situations where no specific evaluative context is provided (e.g., Engell et al., 2007). Specific

judgmental dimensions (e.g., variance in competence judgments not shared with the valence component) may be prominent in contexts that make these dimensions relevant. For example, in electoral decisions, voters believe that competence is the most important attribute for a politician and evaluations of competence but not trustworthiness predict electoral success (Todorov et al., 2005). Similarly, in mating decisions, physical attractiveness could trump evaluations on other dimensions including trustworthiness (DeBruine, 2005). Such decisional contexts may change the nature of brain responses and recruit specific sets of brain regions. As argued in the section The Role of the Amygdala in Evaluation of Emotionally Neutral Faces, the nature of the evaluation task may be critical for the shape of the amygdala's response to faces.

Third, according to the emotion overgeneralization hypothesis, the same neural systems should underlie face evaluation on trait dimensions and perception of emotional expressions. The direct tests of this hypothesis remain to be conducted. Multi-voxel pattern analysis methods (Norman, Polyn, Detre, & Haxby, 2007) and methods designed to measure adaptation effects in fMRI designs (Aguirre, 2007) would be particularly useful to test these hypotheses. For example, if overgeneralization underlies evaluation of emotionally neutral faces, pattern classifiers trained to detect the pattern of neural activation associated with the perception of expressions (Said, Moore, Engell, Todorov, & Haxby, 2010) should be able to detect similar patterns of activation during the perception of neutral faces that vary on trait dimensions associated with the particular expressions (e.g., dominance and anger).

Fourth, although several studies have implicated the amygdala as one of the key regions in face evaluation (section The Role of the Amygdala in Evaluation of Emotionally Neutral Faces), we know very little about the interactions between the regions involved in this evaluation. The amygdala is in the perfect anatomical position—with connections to orbitofrontal, temporal, and occipital cortices—to serve as an interface between perceptual and decision making regions in the brain. We need better mechanistic models of the interactions

between these regions based on both anatomical evidence and causal modeling of their dynamic interactions.

Addressing all these questions will require well-defined behavioral models that constrain and guide research on the neural basis of social cognition. We have tried to follow this research strategy in our lab.

REFERENCES

- Aguirre, G. K. (2007). Continuous carry-over designs for fMRI. *Neuroimage*, *35*, 1480–1494.
- Adams, R. B., Ambady, N., Macrae, N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation & Emotion*, *30*, 179–188.
- Adolphs, R. (2002). Neural systems for recognizing emotion. *Current Opinion in Neurobiology*, *12*, 169–177.
- Adolphs, R. (2003). Cognitive neuroscience of human social behavior. *Nature Reviews Neuroscience*, *4*, 165–178.
- Adolphs, R., Tranel, D., & Damasio, A. R. (1998). The human amygdala in social judgment. *Nature*, *393*, 470–474.
- Adolphs, R., Tranel, D., Damasio, H., & Damasio, A. R. (1995). Fear and the human amygdala. *The Journal of Neuroscience*, *15*, 5879–5891.
- Albright, L., Kenny, D. A., & Malloy, T. E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, *55*, 387–395.
- Allison, T., Puce, A., & McCarthy, G. (2000). Social perception from visual cues: role of the STS region. *Trends in Cognitive Sciences*, *4*, 267–278.
- Amaral, D. G. (2002). The primate amygdala and the neurobiology of social behavior: implications for understanding social anxiety. *Biological Psychiatry*, *51*, 11–17.
- Amaral, D. G., Behniea, H., & Kelly, J. L. (2003). Topographic organization of projections from the amygdala to the visual cortex in the macaque monkey. *Neuroscience*, *118*, 1099–1120.
- Anderson, A. K., & Phelps, E. A. (2001). Lesions of the human amygdala impair enhanced perception of emotionally salient events. *Nature*, *411*, 305–309.
- Ambady, N., Hallahan, M., & Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology*, *69*, 518–529.

- Ambady, N., & Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychological Bulletin*, *111*, 256–274.
- Asch, S. E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, *41*, 258–290.
- Ballew, C. C., & Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of the USA*, *104*(46), 17,948–17,953.
- Bar, M., Neta, M., & Linz, H. (2006). Very first impressions. *Emotion*, *6*, 269–278.
- Baxter, M. G., Parker, A., Lindner, C. C. C., Izquierdo, A. D., & Murray, E. A. (2000). Control of response selection by reinforcer value requires interaction of amygdala and orbital prefrontal cortex. *Journal of Neuroscience*, *20*, 4311–4319.
- Bentin, S., Degutis, J. M., D'Esposito, M., & Robertson, L. C. (2007). Too many trees to see the forest: Performance, event-related potential, and functional magnetic resonance imaging manifestations of integrative congenital prosopagnosia. *Journal of Cognitive Neuroscience*, *19*, 132–146.
- Blair, I. V., Judd, C. M., & Chapleau, K. M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, *15*, 674–679.
- Blanz, V., & Vetter, T. (1999). A morphable model for the synthesis of 3D faces. In *Proceedings of the 26th annual conference on Computer graphics and interactive techniques*, 187–194.
- Blanz, V., & Vetter, T. (2003). Face recognition based on fitting a 3D morphable model. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, *25*, 1063–1074.
- Breiter, H. C., Etcoff, N. L., Whalen, P. J., et al. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron*, *17*, 875–887.
- Brothers, L. (1990). The social brain: a project for integrating primate behavior and neurophysiology in a new domain. *Concepts in Neuroscience*, *1*, 27–51.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
- Bruner, J. S., & Tagiuri, R. (1954). The perception of people. In G. Lindzey (Ed.), *Handbook of Social Psychology*, Vol. 2. Cambridge, MA: Addison-Wesley.
- Calder, A. J., Keane, J., Manes, F., Antoun, N., & Young, A. W. (2000). Impaired recognition and experience of disgust following brain injury. *Nature Neuroscience*, *3*, 1077–1078.
- Calder, A. J., Lawrence, A. D., & Young, A. W. (2001). Neuropsychology of fear and loathing. *Nature Reviews Neuroscience*, *2*, 352–363.
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641–651.
- Carlston, D. E., & Skowronski, J. J. (1994). Saving in the relearning of trait information as evidence for spontaneous inference generation. *Journal of Personality and Social Psychology*, *66*, 840–856.
- Carlston, D. E., Skowronski, J. J., & Sparks, C. (1995). Savings in relearning: II. On the formation of behavior-based trait associations and inferences. *Journal of Personality and Social Psychology*, *69*, 420–436.
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, *25*, 215–224.
- Cronbach, L. J. (1955). Processes affecting scores on understanding others and assuming "similarity." *Psychological Bulletin*, *52*, 177–193.
- Cunningham, W. A., Van Bavel, J. J., & Johnsen, I. R. (2008). Affective flexibility: evaluative processing goals shape amygdala activity. *Psychological Science*, *19*, 152–160.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, *13*, 89–109.
- Davis, M., & Whalen, P. J. (2001). The amygdala: vigilance and emotion. *Molecular Psychiatry*, *6*, 13–34.
- DeBruine, L. M. (2002). Facial resemblance enhances trust. *Proceedings of the Royal Society B*, *269*, 1307–1312.
- DeBruine, L. M. (2005). Trustworthy but not trustworthy: context-specific effects of facial resemblance. *Proceedings of the Royal Society B*, *272*, 919–922.
- Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal emotion recognition in a developmental prosopagnosic. *Perception*, *32*, 827–838.
- Eberhardt, J. L., Davies, P. G., Purdie-Vaughns, V. J., & Johnson, S. L. (2006). Looking deathworthy: perceived stereotypicality of Black defendants

- predicts capital-sentencing outcomes. *Psychological Science*, 17, 383–386.
- Emery, N. J., Capitanio, J. P., Mason, W. A., Machado, C. J., Mendoza, S. P., & Amaral, D. G. (2001). The effects of bilateral lesions of the amygdala on dyadic social interactions in rhesus monkeys (*Macaca mulatta*). *Behavioral Neuroscience*, 115, 515–544.
- Engell, A. D., Haxby, J. V., & Todorov, A. (2007). Implicit trustworthiness decisions: automatic coding of face properties in human amygdala. *Journal of Cognitive Neuroscience*, 19, 1508–1519.
- Engell, A., Todorov, A., & Haxby, J. (In press). Common neural mechanisms for the evaluation of facial trustworthiness and emotional expressions as revealed by behavioral adaptation. *Perception*.
- Fiske, S. T. (1980). Attention and weight in person perception: the impact of negative and extreme behavior. *Journal of Personality and Social Psychology*, 38, 889–906.
- Fiske, S. T., Cuddy, A. J. C., & Glick, P. (2007). Universal dimensions of social cognition: warmth and competence. *Trends in Cognitive Sciences*, 11, 77–83.
- Ferguson, M. J., & Bargh, J. A. (2004). Liking is for doing: the effects of goal pursuit on automatic evaluation. *Journal of Personality and Social Psychology*, 87, 557–572.
- Fox, C. J., & Barton, J. J. (2007). What is adapted in face adaptation? The neural representations of expression in the human visual system. *Brain Research*, 1127, 80–89.
- Fridlund, A. J. (1994). *Human Facial Expression: An Evolutionary View*. San Diego, CA: Academic Press.
- Gilbert, D. T. (1998). Ordinary personology. In D. T. Gilbert, S. T. Fiske, & G. Lindzey (Eds.), *The Handbook of Social Psychology*, Vol. 2 (pp. 89–150). New York: McGraw-Hill.
- Gilbert, D. T., & Malone, P. S. (1995). The correspondence bias. *Psychological Bulletin*, 117, 21–38.
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, 45, 32–41.
- Gobbini, M. I., Leibenluft, E., Santiago, N., & Haxby, J. V. (2004). Social and emotional attachment in the neural representation of faces. *NeuroImage*, 22, 1628–1635.
- Goren, A., & Todorov, A. (2009). Two faces are better than one: eliminating false trait associations with faces. *Social Cognition*, 27, 222–248.
- Hall, C. C., Goren, A., Chaiken, S., & Todorov, A. (2009). Shallow cues with deep effects: trait judgments from faces and voting decisions. In E. Borgida, J. L. Sullivan, & C. M. Federico (Eds.), *The Political Psychology of Democratic Citizenship* (pp. 73–99). London: Oxford University Press.
- Hamermesh, D., & Biddle, J. (1994). Beauty and the labor market. *The American Economic Review*, 84, 1174–1194.
- Hassin, R., & Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, 78, 837–852.
- Haxby, J. V., Gobbini, M. I., Furey, M. L., Ishai, A., Schouten, J. L., & Pietrini, P. (2001). Distributed and overlapping representations of faces and objects in ventral temporal cortex. *Science*, 293, 2425–2430.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2000). The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4, 223–233.
- Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. (2002). Human neural systems for face recognition and social communication. *Biological Psychiatry*, 51, 59–67.
- Haxby, J. V., Ungerleider, L. G., Clark, V. P., Schouten, J. L., Hoffman, E. A., & Martin, A. (1999). The effect of face inversion on activity in human neural systems for face and object perception. *Neuron*, 22, 189–199.
- Hess, U., Blairy, S., & Kleck, R. E. (2000). The influence of facial emotion displays, gender, and ethnicity on judgments of dominance and affiliation. *Journal of Nonverbal Behavior*, 24, 265–283.
- Hoffman, E., & Haxby, J. (2000). Distinct representations of eye gaze and identity in the distributed human neural system for face perception. *Nature Neuroscience*, 3, 80–84.
- Hönekopp, J. (2006). Once more: is beauty in the eye of the beholder? Relative contributions of private and shared taste to judgments of facial attractiveness. *Journal of Experimental Psychology: Human Perception and Performance*, 32, 199–209.
- Humphreys, K., Avidan, G., & Behrmann, M. (2007). A detailed investigation of facial expression processing in congenital prosopagnosia as compared to acquired prosopagnosia. *Experimental Brain Research*, 176, 356–373.

- Jeffrey, L., Rhodes, G., & Busey, T. (2006). View-specific coding of face shape. *Psychological Science, 17*, 501–505.
- Johnson, M. K., Kim, J. K., & Risse, G. (1985). Do alcoholic Korsakoff's Syndrome patients acquire affective reactions? *Journal of Experimental Psychology: Learning, Memory, and Cognition, 11*, 22–36.
- Jones, E. E., & Davis, K. E. (1965). From acts to dispositions: the attribution process in person perception. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 2, pp. 219–266). New York: Academic Press.
- Kanwisher, N., McDermott, J., & Chun, M. M. (1997). The fusiform face area: A module in human extrastriate cortex specialized for face perception. *Journal of Neuroscience, 17*, 4302–4311.
- Kim, M. P., & Rosenberg, S. (1980). Comparison of two structural models of implicit personality theory. *Journal of Personality and Social Psychology, 38*, 375–389.
- Knutson, B. (1996). Facial expressions of emotion influence interpersonal trait inferences. *Journal of Nonverbal Behavior, 20*, 165–181.
- Krumhuber, E., Manstead, A. S. R., Kappas, A., Cosker, D., Marshall, D., & Rosin, P. L. (2007). Facial dynamics as indicators of trustworthiness and cooperative behavior. *Emotion, 7*, 730–735.
- LaBar, K. S., Gitelman, D. R., Parrish, T. B., Kim, Y.-H., Nobre, A. C., & Mesulam, M.-M. (2001). Hunger selectively modulates corticolimbic activation to food stimuli in humans. *Behavioral Neuroscience, 115*, 493–500.
- Langlois, J. H., Kalakanis, L., Rubenstein, A. J., Larson, A., Hallam, M., & Smoot, M. (2000). Maxims or myths of beauty? A meta-analytic and theoretical review. *Psychological Bulletin, 126*, 390–423.
- 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.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska Directed Emotional Faces*. Psychology section, Department of Clinical Neuroscience, Karolinska Institute, Stockholm, Sweden.
- Macrae, C. N., Quinn, K. A., Mason, M. F., & Quadflieg, S. (2005). Understanding others: the face and person construal. *Journal of Personality and Social Psychology, 89*, 686–695.
- Marsh, A. A., Ambady, N., & Kleck, R. E. (2005). The effects of fear and anger facial expressions on approach- and avoidance-related behaviors. *Emotion, 5*, 119–124.
- Mazur, A., Mazur, J., & Keating, C. (1984). Military rank attainment of a West Point class: effects of cadets' physical features. *American Journal of Sociology, 90*, 125–150.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neurosciences, 9*, 605–610.
- Montepare, J. M., & Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior, 27*, 237–254.
- Montepare, J. M., & Zebrowitz, L. A. (1998). Person perception comes of age: the salience and significance of age in social judgments. *Advances in Experimental Social Psychology, 30*, 93–161.
- Morris, J. S., Frith, C. D., Perrett, D. I., Rowland, D., Young, A.W., Calder, A. J., & Dolan, R. J. (1996). A differential neural response in the human amygdala to fearful and happy expressions. *Nature, 383*, 812–815.
- Mueller, U., & Mazur, A. (1996). Facial dominance of West Point cadets as a predictor of later military rank. *Social Forces, 74*, 823–850.
- Norman, K. A., Polyn, S. M., Detre, G. J., & Haxby, J. V. (2007). Beyond mind-reading: multi-voxel pattern analysis of fMRI data. *Trends in Cognitive Sciences, 10*, 424–430.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D. M., & Dolan, R. J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia, 41*, 147–155.
- Olivola, C. Y., & Todorov, A. (2010). Fooled by first impressions? Re-examining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology, 46*, 315–324.
- Olson, I. R., & Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion, 5*, 498–502.
- Oosterhof, N. N., & Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of the USA, 105*, 11087–11092.
- Oosterhof, N. N., & Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion, 9*, 128–133.

- Osgood, C. E., Suci, G. I., & Tennenbaum, P. H. (1957). *The Measurement of Meaning*. Urbana: University of Illinois Press.
- Pessoa, L., Japee, S., Sturman, D., & Underleider, L. G. (2006). Target visibility and visual awareness modulate amygdala responses to fearful faces. *Cerebral Cortex*, *16*, 366–375.
- Phan, K. L., Wager, T., Taylor, S. F., & Liberzon, I. (2002). Functional neuroanatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *Neuroimage*, *16*, 331–348.
- Phelps, E. A., & LeDoux, J. E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, *48*, 175–187.
- Phillips, M. L., Young, A. W., Senior, C., et al. (1997). A specific neural substrate for perceiving facial expressions of disgust. *Nature*, *389*, 495–498.
- Puce, A., Allison, T., Benit, S., Gore, J. C., & McCarthy, G. (1998). Temporal cortex activation of humans viewing eye and mouth movements. *Journal of Neuroscience*, *18*, 2188–2199.
- Rhodes, G., Jeffery, L., Watson, T. L., Clifford, C. W. G., & Nakayama, K. (2003). Fitting the mind to the world: face adaptation and attractiveness aftereffects. *Psychological Science*, *14*, 558–566.
- Rosenberg, S., Nelson, C., & Vivekananthan, P. S. (1968). A multidimensional approach to the structure of personality impressions. *Journal of Personality and Social Psychology*, *9*, 283–294.
- Said, C. P., Baron, S., & Todorov, A. (2009). Nonlinear amygdala response to face trustworthiness: Contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience*, *21*, 519–528.
- Said, C. P., Moore, C. D., Engell, A. D., Todorov, A., & Haxby, J. V. (2010). Distributed representations of dynamic facial expressions in the superior temporal sulcus. *Journal of Vision*, *10*(5), 1–12.
- Said, C., Sebe, N., & Todorov, A. (2009). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, *9*, 260–264.
- Schneider, D. J. (1973). Implicit personality theory: a review. *Psychological Bulletin*, *79*, 294–309.
- Singular Inversions (2006). *FaceGen 3.1 Full SDK Documentation*. <http://facegen.com>. Date last accessed July 30, 2008.
- Somerville, L. H., Wig, G. S., Whalen, P. J., & Kelley, W. M. (2006). Dissociable medial temporal lobe contributions to social memory. *Journal of Cognitive Neuroscience*, *18*, 1253–1265.
- Todorov, A. (2008). Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. In A. Kingstone & M. Miller (Eds.), *The Year in Cognitive Neuroscience 2008, Annals of the New York Academy of Sciences*, *1124*, 208–224.
- Todorov, A., Baron, S., & Oosterhof, N. N. (2008). Evaluating face trustworthiness: a model based approach. *Social, Cognitive, & Affective Neuroscience*, *3*, 119–127.
- Todorov, A., & Duchaine, B. (2008). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology*, *25*, 395–410.
- Todorov, A., & Engell, A. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social, Cognitive, & Affective Neuroscience*, *3*, 303–312.
- Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*, 163–173.
- Todorov, A., & Olson, I. (2008). Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged. *Social, Cognitive, & Affective Neuroscience*, *3*, 195–203.
- Todorov, A., Mandisodza, A. N., Goren, A., & Hall, C. C. (2005). Inferences of competence from faces predict election outcomes. *Science*, *308*, 1623–1626.
- Todorov, A., Pakrashi, M., & Oosterhof, N. N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, *27*, 813–833.
- Todorov, A., & Uleman, J. S. (2002). Spontaneous trait inferences are bound to actor's faces: evidence from a false recognition paradigm. *Journal of Personality and Social Psychology*, *83*, 1051–1065.
- Todorov, A., & Uleman, J. S. (2003). The efficiency of binding spontaneous trait inferences to actor's faces. *Journal of Experimental Social Psychology*, *39*, 549–562.
- Todorov, A., & Uleman, J. S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology*, *87*, 482–493.

- Tranel, D., & Damasio, A. R. (1993). The covert learning of affective valence does not require structures in hippocampal system or amygdala. *Journal of Cognitive Neuroscience*, 5, 79–88.
- Tranel, D., Damasio, A. R., & Damasio, H. (1988). Intact recognition of facial expression, gender, and age in patients with impaired recognition of face identity. *Neurology*, 38, 690–696.
- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, 93, 239–257.
- Uleman, J. S., Blader, S., & Todorov, A. (2005). Implicit impressions. In R. Hassin, J. S. Uleman, & J. A. Bargh (Eds.), *The New Unconscious* (pp. 362–392). New York: Oxford University Press.
- Uleman, J. S., Newman, L. S., & Moskowitz, G. B. (1996). People as flexible interpreters: evidence and issues from spontaneous trait inference. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 28, pp. 211–279). San Diego, CA: Academic Press.
- Vuilleumier, P. (2005). How brains beware: neural mechanisms of emotional attention. *Trends in Cognitive Sciences*, 9, 585–594.
- Vuilleumier, P., Richardson, M., Armony, J. L., Driver, J., & Dolan, R. J. (2005). Distant influences of amygdala lesion on visual cortical activation during emotional face processing. *Nature Neuroscience*, 7, 1271–1278.
- Webster, M. A., Kaping, D., Mizokami, Y., & Duhamel, P. (2004). Adaptation to natural facial categories. *Nature*, 428(6982), 557–561.
- Webster, M. A., & MacLin, O. H. (1999). Figural aftereffects in the perception of faces. *Psychonomic Bulletin & Review*, 6(4), 647–653.
- Whalen, P. J. (1998). Fear, vigilance, and ambiguity: initial neuroimaging studies of the human amygdala. *Current Directions in Psychological Science*, 7, 177–188.
- Wiggins, J. S. (1979). A psychological taxonomy of trait descriptive terms: the interpersonal domain. *Journal of Personality and Social Psychology*, 37, 395–412.
- Wiggins, J. S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behavior: evidence concerning some untested assumptions underlying diagnostic classification. *Journal of Personality and Social Psychology*, 56, 296–305.
- Willis, J., & Todorov, A. (2006). First impressions: making up your mind after 100 ms exposure to a face. *Psychological Science*, 17, 592–598.
- Winston, J., O'Doherty, J., & Dolan, R. J. (2003). Common and distinct neural responses during direct and incidental processing of multiple facial emotions. *Neuroimage*, 20, 84–97.
- Winston, J., O'Doherty, J., Kilner, J. M., Perrett, D. I., & Dolan, R. J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, 45, 195–206.
- Winston, J., Strange, B., O'Doherty, J., & Dolan, R. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of face. *Nature Neuroscience*, 5, 277–283.
- Wiggins, J. S., Phillips, N., & Trapnell, P. (1989). Circular reasoning about interpersonal behavior: evidence concerning some untested assumptions underlying diagnostic classification. *Journal of Personality and Social Psychology*, 56, 296–305.
- Yang, T. T., Menon, V., Eliez, S., et al. (2002). Amygdalar activation associated with positive and negative facial expressions. *NeuroReport*, 13, 1737–1741.
- Zebrowitz, L. A. (1999). *Reading Faces: Window to the Soul?* Boulder, CO: Westview Press.
- Zebrowitz, L. A. (2004). The origins of first impressions. *Journal of Cultural and Evolutionary Psychology*, 2, 93–108.
- Zebrowitz, L. A., & McDonald, S. M. (1991). The impact of litigants' babyfacedness and attractiveness on adjudications in small claims courts. *Law and Behavior*, 15, 603–623.
- Zebrowitz, L. A. & Montepare, J. M. (2006). The ecological approach to person perception: evolutionary roots and contemporary offshoots. In M. Schaller, J. A. Simpson, & D. T. Kenrick. *Evolution and Social Psychology*, New York: Psychology Press.
- Zebrowitz, L. A. & Montepare, J. M. (2008). Social psychological face perception: why appearance matters. *Social and Personality Psychology Compass*, 2, 1497–1517.