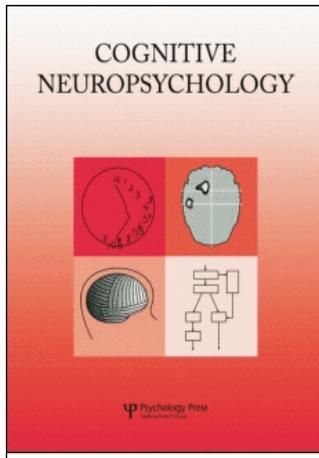


This article was downloaded by:[National Hospital for Neuropsychology]
On: 30 June 2008
Access Details: [subscription number 789757518]
Publisher: Psychology Press
Informa Ltd Registered in England and Wales Registered Number: 1072954
Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Cognitive Neuropsychology

Publication details, including instructions for authors and subscription information:
<http://www.informaworld.com/smpp/title~content=t713659042>

Reading trustworthiness in faces without recognizing faces

Alexander Todorov^a; Bradley Duchaine^b

^a Department of Psychology and Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, NJ, USA

^b Institute of Cognitive Neuroscience, University College London, London, UK

Online Publication Date: 01 May 2008

To cite this Article: Todorov, Alexander and Duchaine, Bradley (2008) 'Reading trustworthiness in faces without recognizing faces', *Cognitive Neuropsychology*, 25:3, 395 — 410

To link to this article: DOI: 10.1080/02643290802044996

URL: <http://dx.doi.org/10.1080/02643290802044996>

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <http://www.informaworld.com/terms-and-conditions-of-access.pdf>

This article maybe used for research, teaching and private study purposes. Any substantial or systematic reproduction, re-distribution, re-selling, loan or sub-licensing, systematic supply or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Reading trustworthiness in faces without recognizing faces

Alexander Todorov

Department of Psychology and Center for the Study of Brain, Mind and Behavior, Princeton University, Princeton, NJ, USA

Bradley Duchaine

Institute of Cognitive Neuroscience, University College London, London, UK

We show that developmental prosopagnosics with severe impairments in both memory for faces and perception of facial identity can make normal trustworthiness judgements of novel faces. We tested four prosopagnosics on three different face sets. The first set consisted of faces that varied on multiple dimensions and that have been used to demonstrate impairments in trustworthiness judgements of patients with bilateral amygdala damage. The second and third sets consisted of standardized faces with direct gaze, neutral expression, and similar age. On all tests, two prosopagnosics made judgements that closely agreed with control judgements while the other two showed weak agreement but within the normal range. The performance of the tests was correlated suggesting that the tests mapped the same underlying judgement irrespective of the specific face stimuli. The normal performance of two of the prosopagnosics suggests that forming person impressions from faces involves mechanisms functionally independent of mechanisms for encoding the identity of faces.

Keywords: Face perception; Social cognition; Prosopagnosia; Trustworthiness; Amygdala.

People can form a person impression after as little as a 100-ms exposure to a face (Bar, Maital, & Linz, 2006; Willis & Todorov, 2006) and can store an extraordinary number of faces (Bahrick, Bahrick, & Wittlinger, 1975). Once formed, person impressions are linked to the identity of the person, and faces of familiar people trigger the spontaneous retrieval of person knowledge (Gobbini & Haxby, 2007; Todorov, Gobbini,

Evans, & Haxby, 2007a; Todorov & Uleman, 2003). Although the processes underlying the encoding of facial identity and the formation of person impressions from the face are closely related, we argue that they are dissociable. In this paper, we show that developmental prosopagnosics who are severely impaired in processing facial identity can make normal trustworthiness judgements from faces.

Correspondence should be addressed to Alexander Todorov, Department of Psychology, Green Hall, Princeton University, Princeton, NJ 08544–1010, USA (E-mail: atodorov@princeton.edu).

We thank Manish Pakrashi, James Chu, Chris Said, and Valerie Loehr for their help in running the experiments, Galit Yovel for providing us with the face part/spacing discrimination task, and Ralph Adolphs for providing the stimuli used to test patients with bilateral amygdala damage and the raw data for our reanalysis. This research was supported by National Science Foundation Grant BCS-0446846.

Making rapid judgements about strangers (e.g., a potential friend or foe) is a task functionally different from the task of tracking the identity of familiar people over time. Thus, it is possible that the mechanisms used for trait judgements are different from mechanisms for representing person identity. This possibility is consistent with findings from functional neuroimaging studies implicating regions in the inferotemporal cortex as critical for perception of the identity of faces (e.g., Kanwisher, McDermott, & Chun, 1997; McCarthy, Puce, Gore, & Allison, 1997; Wada & Yamamoto, 2001) and the amygdala as critical for perception of the trustworthiness of faces (Engell, Haxby, & Todorov, 2007; Todorov, Baron, & Oosterhof, in press; Winston, Strange, O'Doherty, & Dolan, 2002).

We focus on judgements of trustworthiness from facial appearance for two other reasons. First, these judgements are one of the most important decisions in social environments, determining basic approach/avoidance response (Todorov, 2008). Second, judgements of trustworthiness are more efficient than other trait judgements and, in fact, as efficient as judgements of attractiveness (Willis & Todorov, 2006). People start discriminating between trustworthy- and untrustworthy-looking faces after 33 ms of exposure to a face, and the correlations with judgements made in the absence of time constraints do

not improve with exposures longer than 167 ms (Todorov, 2008; Todorov, Pakrashi, Loehr, & Oosterhof, 2007b).

To obtain evidence about the relationship of mechanisms underlying encoding of person identity and forming of person impressions, we investigated person impressions in individuals with developmental prosopagnosia, a condition characterized by severe face recognition deficits due to a failure to develop the required mechanisms (Behrmann & Avidan, 2005; Duchaine & Nakayama, 2006a). Because deficits in face recognition may be due to problems in face memory or to problems in perception of identity, we studied four individuals who had severe impairments in both long-term memory for faces and perception of facial identity (e.g., Table 1). If such individuals were able to make normal trustworthiness judgements from novel faces, this would provide strong evidence about the independence of mechanisms for encoding identity and forming person impressions.

Adolphs, Tranel, and Damasio (1998) conducted the first study suggesting that the amygdala plays a key role in trustworthiness judgements. They showed that patients with bilateral amygdala damage perceived untrustworthy-looking faces as trustworthy. In Experiment 1 below, our prosopagnosics made trustworthiness judgements for the set of faces used with the amygdala patients.

Table 1. Performance of developmental prosopagnosics on tests of memory for faces and perception of facial identity in standard deviation units

	Memory for faces		Perception of identity of faces		
	Famous faces	CFMT	CFPT	Parts	Spacing
J.K.	-11.8 ^a	-3.0	-4.7	-2.6	-4.0
T.U.	-3.1	-2.3	-2.2	-3.4	-4.4
J.P.	-6.1	-2.8	-2.4	-3.8	-2.7
J.L.	-5.8	-4.1	-3.7	-4.1	-4.4
Control mean (<i>SD</i>)	47.3 (6.2)	57.9 (7.9)	36.7 (12.2) ^b	79.7 (6.6)	76.9 (6.2)

Note: The famous faces test and Cambridge face memory test (CFMT) assess long-term memory for faces. The other three tests—the Cambridge face perception test (CFPT) and a discrimination test involving changes to the face parts or the spacing of the face parts—assess the perception of face identity. Performance is presented in standard deviation units.

^a J.K. has a different nationality from that of the other prosopagnosics, and a different test was used to measure her performance (control $M = 53.8$, $SD = 3.1$). ^b CFPT measures number of errors, and, hence, poor performance is demonstrated by scores higher than the control mean.

All four prosopagnosics showed trustworthiness judgements comparable to control judgements. Compared to the performance of the amygdala patients, they also showed normal perceptions of untrustworthy faces.

The faces in the first set varied on a number of dimensions including hair, expression, gaze, and age (Figure 1), and it is possible that prosopagnosics' normal performance on trustworthiness judgements can be accounted for by reliance on expression, gaze, or nonface information. To rule out this explanation, in the second and third studies we used sets of standardized faces. The second set consisted of faces that all had direct gaze, neutral expression, and a similar age. The third set consisted of a subset of faces from the second set. These faces were chosen because they elicited the greatest agreement by controls. The

hair and facial blemishes were removed, and they were presented as greyscale images. Thus, the face sets consisted of increasingly homogeneous face images in order to force reliance on facial structure and complexion. For both the second and third sets of faces, two of the prosopagnosics showed typical trustworthiness judgements that closely agreed with control judgements. The judgements of the other two only weakly agreed with control judgements, particularly on the third set of faces.

We used two measures of performance: agreement in the rating of the faces and mean judgements across faces. These measure different aspects of performance. For example, prosopagnosics and controls can have the same mean judgements but their judgements across faces can be completely uncorrelated, suggesting that prosopagnosics cannot make trustworthiness judgements. In contrast, it is possible that judgements of prosopagnosics agree with control judgements, but that the prosopagnosics show a positivity or negativity bias in their judgements. Such biases can be informative about potential deficits in judgements. For example, the patients with bilateral amygdala damage show a bias to perceive untrustworthy faces as trustworthy (Adolphs et al., 1998), and individuals with Williams syndrome show an overall positivity bias (Bellugi, Adolphs, Cassady, & Chiles, 1999)—that is, faces are rated as more trustworthy on average, although their relative ranking is preserved.

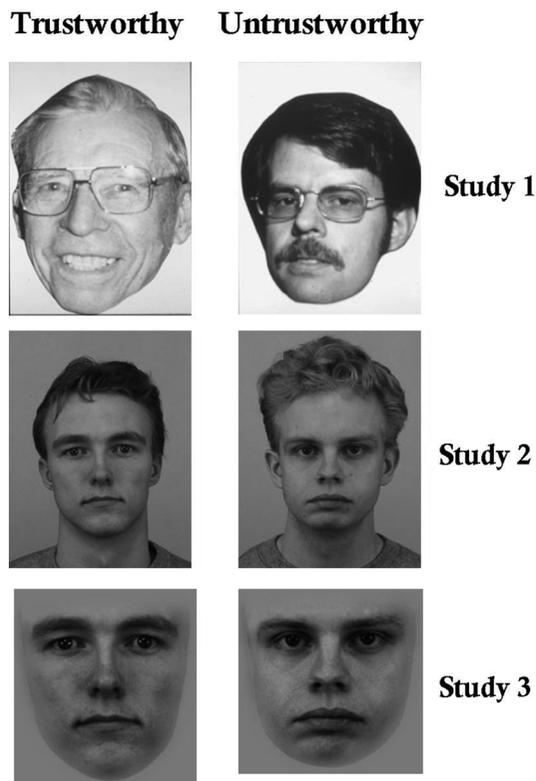


Figure 1. Examples of trustworthy- and untrustworthy-looking faces used in the three studies.

Method

Participants

The control group consisted of 48 undergraduate students (30 males, age range 18 to 32 years, mean age = 20) from Princeton University who participated in the studies in exchange for course credit. There were no significant differences between male and female control participants on both their agreement in the ratings of the faces and their mean trustworthiness judgements. Thus, we collapsed across gender for all analyses. The control group was younger than the developmental prosopagnosics (see Table 2). However,

Table 2. Performance of developmental prosopagnosics on tests of low-level vision in standard deviation units

	Sex	Age	Occupation	Low-level vision tests			
				Birmingham Object Recognition Battery			
				Length match	Size match	Orientation match	Position of gap
J.K.	F	36	Scientist	0.69	0.71	-1.08	0.98
T.U.	M	31	Graphic designer	-0.56	0.71	2.00	0.98
J.P.	F	24	Civil Service	0.06	0.71	-1.08	-0.03
J.L.	F	62	Retired lecturer	-0.56	0.29	0.85	-0.03

Note: Performance is presented in standard deviation units. Age in years.

this difference in age most likely leads to underestimation of the performance of the prosopagnosics, particularly on the sets (second and third sets, see later) with faces of young people.

The four developmental prosopagnosics contacted the Harvard University/University College London Prosopagnosia Research Center (<http://www.faceblind.org>), because they experienced severe face recognition problems in everyday life. None have experienced any brain damage, and all report lifelong face recognition difficulties, so their problems appear to be developmental in origin. All have normal or corrected-to-normal visual acuity. All were tested in person with a battery of face recognition tests (Table 1). The data for three of the patients with bilateral amygdala damage were reported in Adolphs et al. (1998). The raw data for these patients and data collected from a fourth patient were provided for our reanalysis by R. Adolphs.

Face and object recognition tests for participants with prosopagnosia

The prosopagnosics were tested with 11 tests prior to the studies involving trustworthiness judgments. Of these tests, 2 were designed to identify problems in face memory and 2 to identify problems in perception of facial identity. The 5th test was used to demonstrate that the prosopagnosics did not have low-level vision problems. The other 6 tests assessed object and scene recognition, as well as face recognition using the same paradigm.

Tests of face memory

In the Cambridge Face Memory Test (Duchaine & Nakayama, 2006b), participants viewed images of six male target faces in the introduction. Following this stage, they were presented with forced-choice items consisting of three faces, one of which was a target face. The poses and lighting in the images in the test items were different from the study images in the introduction. In the famous face test, closely cropped celebrity faces were presented for 3 seconds. Two versions were used, one adapted for US/Canadian participants and one for UK participants. For all tests, prosopagnosics were compared to age- and nationality-matched controls.

Tests of facial identity perception

Participants in the Cambridge Face Perception Test sorted six faces in frontal view based on their similarity to a three-quarters profile view of a target face (Duchaine, Germine, & Nakayama, 2007a; Duchaine, Yovel, & Nakayama, 2007b). Each frontal view was created by morphing a frontal view of the target face with another face. Images from these morph continua were chosen that contained different percentages of the target face (88%, 76%, 64%, 52%, 40%, and 28%), and these were presented in a random order. Participants had one minute to sort the images. There were eight upright sets to sort. Deviations from the correct order for each item were summed to compute the total number of errors (chance = 94 errors).

In the same–different sequential matching task (Yovel & Duchaine, 2006), two faces were presented for 500 ms each. A total of 40 trials showed the same face twice. The 40 different trials consisted of two types of trial. On spacing trials, the face parts were identical but the spacing between the eyes and between the nose and mouth was varied. On part trials, the face parts were varied but their spacing remained constant. For all tests, prosopagnosics were compared to age- and nationality-matched controls.

Low-level vision tests from the Birmingham Object Recognition Battery

Participants judged whether pairs of stimuli were the same or different. Stimuli varied in line length, circle size, position of gap in a circle, and orientation. The performance of the prosopagnosics was compared to norms from the test manual (Riddoch & Humphreys, 1993).

Object perception tests

The four prosopagnosics were also tested on five tests measuring within-class object and scene recognition (Duchaine et al., 2007a; Duchaine & Nakayama, 2005). The four categories of objects used were cars, guns, tools, and sunglasses, along with images of natural scenes. We also assessed their face recognition using this paradigm to demonstrate that it was sensitive to their face recognition deficits. In the study phase, 10 target images were presented for 3 s each, and each was shown twice. In the test phase, the 10 target images were shown twice, and 30 new images were shown once, and participants decided whether each of these 50 images was a target (old) or nontarget (new) image. A-prime, an unbiased measure of discrimination based on hits and false alarms, was computed to assess performance (Macmillan & Creelman, 1991).

Face stimuli for current studies

For the first study, we used the same 100 faces as those used to test patients with bilateral amygdala damage (Adolphs et al., 1998). These faces varied

on a number of dimensions including hair, eye gaze, expressions, and age. For the second study, we used a set of 68 standardized faces (Lundqvist, Flykt, & Öhman, 1998) with direct gaze and neutral expressions. These were photographs of amateur actors and actresses between 20 and 30 years of age with no facial hair, earrings, eyeglasses, or visible make-up, all wearing grey T-shirts. For the third study, we selected the 30 faces from the second set with the lowest variance in the control judgements. The original colour images were transformed to greyscale, and all hair and facial marks such as moles and blemishes were removed (see Figure 1).

Procedure

The trait judgement studies were conducted on Princeton University's Websurvey platform. Participants were sent an invitation with an individualized login name and password. Once participants had logged in, a detailed set of instructions was provided. For all face sets, participants were told that the study was about first impressions and were encouraged to rely on their "gut feeling". The three sets of faces were presented in three blocks always in the same sequence: the heterogeneous set of 100 faces, the standardized set of 68 faces, and the greyscale set of 30 faces. Within each block, the order of faces was randomized for each participant. Each face was presented at the centre of the screen with a question above the photograph "How trustworthy is this person?" and a response scale below the photograph. For the first set of faces, we used the same response scale as that in Adolphs et al. (1998), as well as the same instructions. Participants were asked to imagine "trusting each person with all your money, or with your life, and then base your rating on this impression". The response scale ranged from -3 (extremely untrustworthy) to 3 (extremely trustworthy). For the second and third sets, the response scale ranged from 1 (extremely untrustworthy) to 9 (extremely trustworthy). Each face was visible until the participant responded.

Results

Performance of prosopagnosics on facial identity tests

The performance of the prosopagnosics on tests of low-level vision (Riddoch & Humphreys, 1993) was normal (Table 2), but they had severe impairments in both long-term memory for faces and perception of facial identity. As shown in Table 1, on five different measures—two measuring face memory problems and three measuring identity perception problems—each prosopagnosic was more than 2 standard deviations below the mean control performance (Duchaine & Nakayama, 2006b; Yovel & Duchaine, 2006). As a group, they were 6.7 standard deviations below the control mean on a test requiring recognition of famous faces and 3.1 standard deviations below the mean on a widely used test requiring learning and then recognizing unfamiliar faces (Bentin, Degutis, D'Esposito, & Robertson, 2007; Duchaine & Nakayama, 2006b). On the Cambridge Face Perception Test (Duchaine et al., 2007a; Duchaine et al., 2007b), participants sorted morphed faces in terms of similarity to a target face, and the prosopagnosic group was 3.3 standard deviations below the mean. On a same/different sequential discrimination involving briefly presented faces, the prosopagnosics averaged 3.5 standard deviations below the mean for discriminations involving the shape of face parts and 3.9 standard deviations below for discriminations involving the spacing of the face parts (Yovel & Duchaine, 2006).

Object and scene recognition

In contrast to their performance on the facial identity tests, most of the prosopagnosics did not appear to have severe problems with object and scene recognition (Table 3). As expected, all four prosopagnosics scored poorly on the face old-new test, but all showed better than average performance on scene recognition. For one of the object categories—sunglasses—all prosopagnosics also showed better than average performance. For two of the object categories—tools and guns—all four showed performance within 2 standard deviations of the control mean. Only J.L.'s performance and J.P.'s performance on car recognition were more than 2 standard deviations below the control performance. These findings suggest that the high-level vision problems of the prosopagnosics primarily affected processing of faces.

Judgements of trustworthiness: Experiment 1

Having established their deficits with facial identity, we asked the prosopagnosics to judge the trustworthiness of three sets of unfamiliar faces (Figure 1). We compared the judgements of the prosopagnosics with the judgements of a control group of 48 participants. The control judgements for all three sets were highly reliable as assessed by the internal consistency of the judgements and correlations with judgements of the same faces by other control groups (Table 4). For the first set of faces, the interrater agreement was high, and the mean control judgements correlated highly with control judgements collected for the

Table 3. Performance of developmental prosopagnosics on tests of high-level object, scene, and face recognition memory

	Faces		Cars		Guns		Sunglasses		Tools		Scenes	
	A^I	SD	A^I	SD	A^I	SD	A^I	SD	A^I	SD	A^I	SD
J.K.	.88	-3.4	.98	1.1	.91	-0.1	.98	1.7	.94	-0.5	1	1.2
T.U.	.83	-5.6	.90	-1.3	.91	-0.1	.95	1.1	.99	1.6	1	1.2
J.P.	.83	-5.8	.81	-3.8	.94	0.7	.95	1.1	.97	0.6	.99	0.9
J.L.	.85	-4.6	.87	-2.2	.85	-1.8	.92	0.4	.95	-0.1	.99	0.9
Control mean (SD)	.96 (.023)		.94 (.037)		.91 (.036)		.91 (.042)		.95 (.025)		.97 (.025)	

Note: Performance is presented in terms of discrimination between old and new stimuli (A^I) and standard deviation units (SD).

Table 4. Measures of reliability of control judgements of trustworthiness for the face sets used in the studies

Face set	Mean IA	Cronbach's α	CGA
1	.43	.97	.88
2	.32	.95	.97
3	.28	.95	.94

Note: IA = interrater agreement (r); CGA = agreement with judgements of other control groups (r). For the first face set, the other control judgements ($n = 47$) were provided by R. Adolphs (Adolphs et al., 1998). For the second face set, the other control judgements ($n = 129$) were provided by A. Todorov (Engell et al., 2007). For the third face set, the other control judgements ($n = 34$) were provided by A. Todorov (unpublished data).

study of the amygdala patients (Adolphs et al., 1998).

To evaluate the performance of the prosopagnosics with each face set, we computed the Pearson correlation between their judgements and the mean control judgements.¹ For the first set of faces, the prosopagnosics' correlations were significantly higher than zero ($p < .0001$), ranging from .50 to .66, indicating that the judgements of the prosopagnosics agreed with control judgements. To construct a distribution of the control judgements, for each control participant we computed the correlation between their judgements and the mean of the judgements of the other control participants. The average correlation was .65 ($SD = .12$). To normalize the distribution of the raw correlations and to avoid the upper bound of 1, we transformed the correlations into Fisher z -scores where $z_r = 0.5 \cdot \ln[(1 + r)/(1 - r)]$. This transformation is also necessary for statistical tests on correlations. The average of the transformed correlations (z_r) was .80 ($SD = .20$). All prosopagnosics scored within 2 standard deviations of the control mean (Figure 2a), and the average correlation for the prosopagnosia group ($z_r = .67$, $SD = .13$) was not significantly different from the correlation for the control group, $t(50) = 1.32$, $p = .19$.

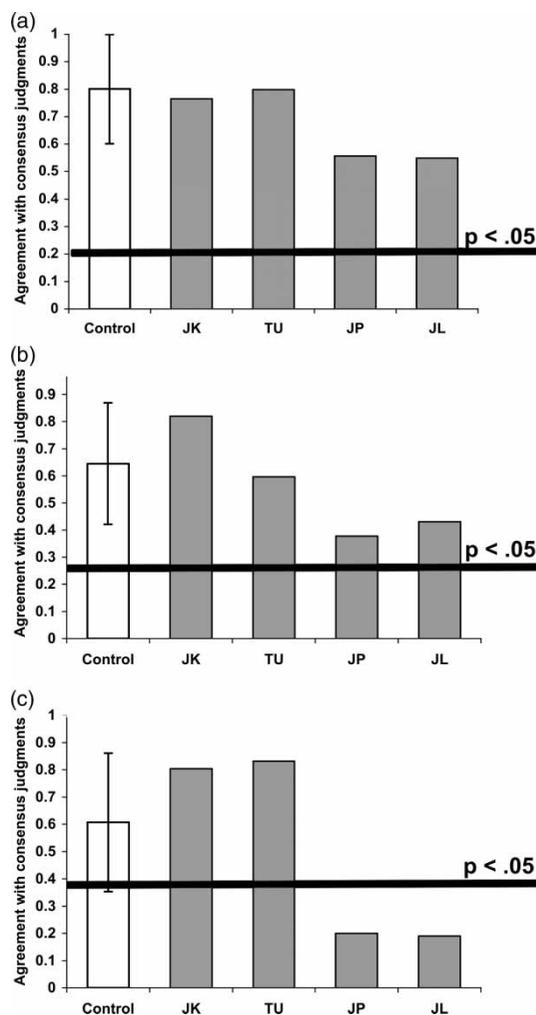


Figure 2. Agreement between prosopagnosics' judgements and consensus judgements of trustworthiness of faces in Experiment 1 (a), Experiment 2 (b), and Experiment 3 (c). The agreement is measured in Fisher z transformations of the raw correlations between individual judgements and the mean judgements of control participants. The error bar shows 1 standard deviation. Scores above the thick black line indicate significantly positive correlations with mean control judgements.

As shown in Table 5, the mean trustworthiness judgements of the prosopagnosics were

¹ We also conducted correlational analyses using the nonparametric Spearman's rho coefficient. The results were practically identical. For example, the Pearson's correlations for the first set of faces were .64, .66, .50, and .50 for J.K., T.U., J.P., and J.L., respectively. The corresponding nonparametric correlations were .63, .66, .50, and .44 ($p < .0001$ for all correlations).

comparable to the mean control judgements. The judgements of J.K., T.U., and J.P. were within 1 standard deviation of the control mean, and the judgements of J.L. were within 2 standard deviations. We also compared the mean judgements for the 50 most trustworthy ($M = 0.77$; $SD = 0.61$) and the 50 least trustworthy faces ($M = -1.04$, $SD = 0.52$), according to the control judgements. Consistent with the correlation analysis, all four prosopagnosics provided more positive ratings for trustworthy than for untrustworthy faces. For each prosopagnosic, the t test comparing their ratings of the 50 trustworthy and 50 untrustworthy faces was highly significant, $t_s(98) > 4.72$, $p < .0001$. The difference between these ratings was also comparable to the difference for the control participants. The average difference for control participants was 1.81 ($SD = 0.55$). The corresponding differences for T.U., J.P., and J.K. were 1.78, 1.70, and 1.28, all within 1 standard deviation of the control difference. The difference for J.L. was 1.12, which was within 2 standard deviations of the control difference.

As described earlier, patients with bilateral amygdala damage had been tested with the first set of faces. Relative to three comparison groups (normal controls, brain-damaged patients with no damage in the amygdala, and patients with unilateral amygdala damage) these patients showed a bias to judge untrustworthy faces as trustworthy

(Adolphs et al., 1998). As shown in Figure 3, for all four amygdala patients, their judgements of untrustworthy faces were more than 2 standard deviations higher than the mean control judgements. In contrast, the judgements of three out of the four prosopagnosics were within 1 standard deviation of the control mean, and the judgements of the fourth were within 2 standard deviations. In terms of the judgements of the trustworthy faces, the judgements of all prosopagnosics and all amygdala patients except for S.M. were within 1 standard deviation of the control mean. S.M.'s judgements were within 2 standard deviations.

At the group level, the judgements of trustworthy faces for both amygdala patients and prosopagnosics were within 1 standard deviation of the control mean. However, whereas the prosopagnosics' judgements of untrustworthy faces were also within 1 standard deviation of the control mean, the bilateral amygdala damage patients' judgements of untrustworthy faces were 2.8 standard deviations above the control mean. Despite the small sample size of the two groups, these judgements were significantly higher than the prosopagnosics' judgements of untrustworthy faces, $t(6) = 5.18$, $p < .002$. This difference was also significant ($p < .037$) when tested with the nonparametric Kolmogorov-Smirnov test.

In terms of correlations of the individual judgements with control judgements, there was large

Table 5. Mean trustworthiness judgements of three different sets of faces for control participants and developmental prosopagnosics

	Trustworthiness								
	Face Set 1			Face Set 2			Face Set 3		
	High	Low	Mean	High	Low	Mean	High	Low	Mean
Control	.77 (.61)	-1.04 (.52)	-.13 (.49)	4.89 (1.13)	3.34 (.92)	4.12 (.96)	5.04 (1.33)	3.61 (.85)	4.32 (1.00)
J.K.	.54	-.74	-.10	5.56	3.79	4.68	5.67	4.53	5.10
T.U.	.94	-.84	.05	5.53	4.32	4.93	6.27	4.27	5.27
J.P.	.86	-.84	.01	5.03	3.97	4.50	4.27	3.20	3.73
J.L.	1.04	-.08	.48	6.06	4.56	5.31	5.67	5.60	5.63

Note: Standard deviations in parentheses. The response scale ranged from -3 (extremely untrustworthy) to 3 (extremely trustworthy) for the first set of faces and from 1 (extremely untrustworthy) to 9 (extremely trustworthy) for the second and third sets of faces. The faces were divided into high and low trustworthiness according to the mean control judgements.

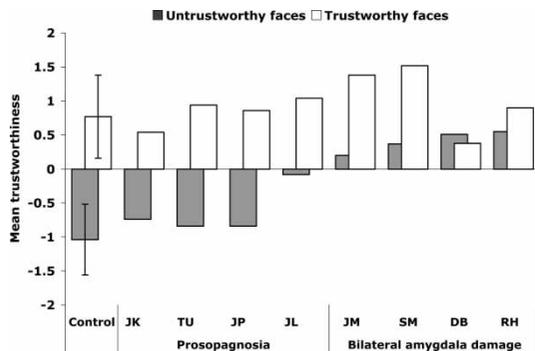


Figure 3. Mean perceived trustworthiness of faces categorized as trustworthy and untrustworthy for four patients with bilateral amygdala damage and four developmental prosopagnosics. The trustworthiness was measured on a scale from -3 (untrustworthy) to $+3$ (trustworthy). The error bar shows 1 standard deviation of the mean control judgements.

variation in the performance of the patients with bilateral amygdala damage. Two showed correlations in the normal range, one was 3 standard deviations below the control mean, and one was more than 4 standard deviations below the control mean. Similarly, in terms of the differences in the judgements of the 50 trustworthy and 50 untrustworthy faces, the judgements of two of the patients (J.M. and S.M.) were within 2 standard deviations of the control difference, the difference for R.H.'s judgements was 2.6 standard deviations smaller than the control difference, and D.B. did not discriminate between trustworthy and untrustworthy faces at all.

Judgements of trustworthiness: Experiment 2

For the second set of faces, the average interrater correlation was lower than that for the first set of faces (Table 4), reflecting the fact that the faces in the second set were more homogeneous (Hönekopp, 2006) and did not vary on cues that strongly influence trustworthiness judgements such as emotional expression (Winston et al., 2002; Zebrowitz, Voinescu, & Collins, 1996). The mean judgements correlated highly with the judgements of a different control group of 129 participants (Engell et al., 2007).

The average correlation between individual control judgements and the mean of the other control participants was .55 ($SD = .15$). The average Fisher z -score was 0.64 ($SD = 0.22$). As in the first set, all four correlations between the prosopagnosics' judgements and the mean control judgements were significantly higher than zero ($p < .003$), ranging from .36 to .68. All prosopagnosics scored within 2 standard deviations of the control mean (Figure 2b). The average correlation for the prosopagnosia group ($r = .56$, $SD = .20$) was not significantly different from the correlation for the control group, $t(50) < 1$, $p = .44$.

The mean trustworthiness judgements of the prosopagnosics were comparable to the mean control judgements (Table 3). The judgements of J.K., T.U., and J.P. were within 1 standard deviation of the control mean, and the judgements of J.L. were within 2 standard deviations. Consistent with the correlation analysis, all four prosopagnosics provided more positive ratings for the 34 trustworthy than for the 34 untrustworthy faces. For each prosopagnosic, the t test comparing their ratings of trustworthy and untrustworthy faces was significant, $t_s(66) > 4.00$, $p < .0001$, for T.U., J.P., and J.K., and $t(66) = 2.09$, $p < .040$, for J.P., respectively. The difference between these ratings was also comparable to the difference for the control participants. The average difference for control participants was 1.55 ($SD = 0.75$). The corresponding differences for J.K., J.L., T.U., and J.P. were 1.77, 1.50, 1.21, and 1.06, all within 1 standard deviation of the control difference.

Judgements of trustworthiness: Experiment 3

It is clear from Figures 2a and 2b (see also Figure 4a), as well as Table 5, that the judgements of the prosopagnosics are within the normal range and that the judgements of two prosopagnosics (J.K. and T.U.) are typical. However, it is possible that the prosopagnosics agreed with the controls, because both groups were influenced by extrafacial cues such as hairstyle and skin blemishes. To rule out this explanation, we used greyscale faces without hair or facial marks (Figure 1). As with

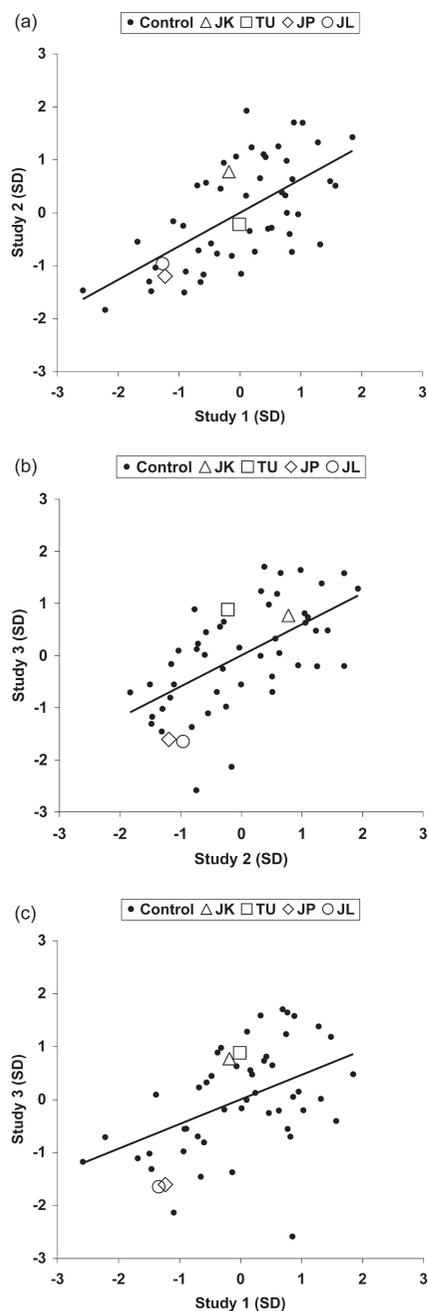


Figure 4. Scatter plots of performance of control participants and prosopagnosics on three sets of faces judged on trustworthiness. Performance is shown in standard deviation units. (a) Performance on the first two sets of faces. (b) Performance on the second and third sets of faces. (c) Performance on the first and third sets of faces.

the other two sets, the control judgements were highly reliable (Table 4) and correlated highly with judgements of a different control group of 34 participants.

For this set of faces, the average correlation between individual control judgements and the mean of the other control participants was .52 ($SD = .19$). The average Fisher z -score was 0.61 ($SD = 0.25$). As in the previous sets, the correlations for the prosopagnosics were positive, ranging from .19 ($p = .32$) to .68 ($p < .0001$). Although the correlations for two of the prosopagnosics (J.P. and J.L.) were substantially reduced and did not reach significance, the correlations for the other two (J.K. and T.U.) were higher than the mean control correlation (Figure 2c). The average correlation for the prosopagnosia group ($r = .51$, $SD = .36$) was not significantly different from the correlation for the control group, $t(50) < 1$, $p = .46$.

As in the other face sets, the mean trustworthiness judgements of the prosopagnosics were comparable to the mean control judgements (Table 5). The judgements of J.K., T.U., and J.P. were within 1 standard deviation of the control mean, and the judgements of J.L. were within 2 standard deviations. All four prosopagnosics provided more positive ratings for the 15 trustworthy than for the 15 untrustworthy faces, although this difference was only significant for J.K., $t(28) = 2.70$, $p < .012$, and T.U., $t(28) = 3.37$, $p < .002$. The average difference between these ratings for control participants was 1.43 ($SD = 1.00$). The corresponding differences for J.K., J.P., and T.U. were 1.14, 1.07, and 2.00, all within 1 standard deviation of the control difference. The difference for J.L. was very small (0.06), $t < 1$.

One explanation for the good performance of J.K. and T.U. on the third set of faces is that they relied on their memory for the judgements in the second set. However, this explanation is quite unlikely, given the prosopagnosics' severe impairments in face memory. Furthermore, test-retest correlations between judgements of the same set of faces are generally low even for normal controls. For example, for the faces used in the second set, the average test-retest

correlation between judgements of the identical face images was .49 ($SD = .14$) in Engell et al., 2007.² This correlation should be even lower for judgements of faces presented in different image formats. Although the faces in the third set were a subset of the second set, the images in the third set were greyscale, and hair and blemishes were removed from the faces (Figure 1). In fact, for the control group, the average correlation between the judgements of the faces used in both sets was .38 ($SD = .19$). The corresponding correlations were .44 for J.K. and .17 for T.U. J.P.'s judgements of the faces were completely uncorrelated ($r = .00$), and J.L.'s judgements were negatively correlated ($r = -.11$).

Performance on the three trustworthiness tests

The performance on the three sets of faces was highly correlated in both controls and prosopagnosics (Figure 4). Participants who performed well on one of the face sets also performed well with the other sets, indicating that the tests mapped the same underlying judgement irrespective of the specific face stimuli. Although the scores for two of the prosopagnosics were at the bottom of the distribution, the scores were within the normal range of 2 standard deviations. More importantly, the scores for the other two were either at the mean or above the mean of the distribution for all three tests, indicating typical trustworthiness judgements.

We examined whether the prosopagnosics with lower correlations (J.P. and J.L.) were more severely impaired with the facial identity tests than were the prosopagnosics with clearly normal correlations (J.K. and T.U.), but there was no difference (see Table 1). J.K. and T.U. were 5.0 standard deviations below the control mean on the face memory tests, whereas J.P. and J.L. were 4.7 standard deviations below the control mean. Similarly, J.K. and T.U. were 3.54 standard

deviations below the control mean on the perception of face identity tests, whereas J.P. and J.L. were 3.50 standard deviations below the control mean.

Discussion

With three different sets of faces, four developmental prosopagnosics with severe impairments for facial identity perception and memory made normal trustworthiness judgements. The face sets consisted of increasingly homogeneous images making it unlikely that their performance can be accounted for by reliance on information other than facial structure, shape, and complexion. The performance of two of the four prosopagnosics was consistently in the middle of the distribution of normal performance. The performance of the other two was within the normal range although their performance was poor for the third set of faces. On this set, although the correlations of their judgements with the mean control judgements were within 2 standard deviations of the control performance, these correlations failed to reach significance.³

The current findings suggest that forming person impressions from faces involves mechanisms functionally independent of mechanisms for encoding the identity of faces. This possibility is consistent with functional imaging studies implicating the amygdala as one of the key regions in the spontaneous encoding of faces on the dimension of trustworthiness (Engell et al., 2007; Todorov et al., in press; Winston et al., 2002) and studies implicating the inferotemporal cortex in encoding of face identity (e.g., Kanwisher et al., 1997; McCarthy et al., 1997; Yovel & Kanwisher, 2005). Given that there are patients with bilateral amygdala damage who process the identity of faces normally (Adolphs, Tranel, Damasio, & Damasio, 1995), the results from

² Note that although within-subjects correlations are not very high, they are generally higher than between-subjects (interrater agreement) correlations. For the faces and the sample used in Engell et al. (2007), the average between-subjects correlation was .28 ($SD = .17$).

³ However, it should be noted that the third face set had the smallest number of faces, and, correspondingly, the significance threshold was more stringent because of the fewer degrees of freedom.

amygdala patients and prosopagnosic participants suggest a possible double dissociation between encoding identity and forming trustworthiness judgements. In addition, because our participants' face recognition deficits resulted from developmental difficulties, their dissociations indicate that mechanisms used for trustworthiness judgements are constructed by different developmental processes from mechanisms used for identity processing.

Navigating the social world requires tracking the identity of people over time, reading their expressions to understand the meaning of the immediate situation, and making quick trait decisions about strangers (e.g., a potential friend or foe?). The first two functional tasks have been incorporated in models of face perception (Bruce & Young, 1986; Haxby, Hoffman, & Gobbini, 2000; Perrett, Hietanen, Oram, & Benson, 1992). These models posit different neural pathways for processing of relatively invariant facial features essential for recognizing identity and gender and changeable facial features essential for recognizing emotions and gaze direction (but see Calder & Young, 2005, for an alternative view). Trait judgements from faces do not fit neatly in this categorization. These judgements are based on relatively invariant features and concern stable characteristics of people, but our results show that they can be dissociated from computations involving facial identity.

Recent computer modelling work suggests that trait judgements from facial appearance, in particular judgements of trustworthiness, are grounded in the similarity of facial features to emotional expressions signalling approach and avoidance behaviours. Using a statistically driven model of face representation, in which faces are represented as points in a multidimensional space, Oosterhof and Todorov (2008; see also Todorov, 2008) built a dimension optimal in changing face trustworthiness. Specifically, trustworthiness judgements of computer-generated faces were used to find a trustworthiness vector in the multidimensional face space based on the correlations of these judgements with the dimensions (50 orthogonal principal components) of

the face space. Trustworthiness judgements of faces that varied on the trustworthiness dimension tracked the model-predicted trustworthiness. More importantly, exaggerating the features of the faces in the positive direction of the trustworthiness dimension (e.g., 8 standard deviations) produced happy faces, and exaggerating the features in the negative direction produced angry faces (e.g., -8 standard deviations).

Emotional expressions often signal the behavioural intentions of the person displaying the emotion (Fridlund, 1994). Specifically, expressions of anger communicate that the person should be avoided (Adams, Ambady, Macrae, & Kleck, 2006; Marsh, Ambady, & Kleck, 2005), and expressions of happiness communicate that the person can be approached. That is, trustworthiness judgements from emotionally neutral faces can be construed as an attempt to infer behavioural intentions in the absence of emotional cues (Todorov, 2008). These judgements seem to be derived from facial features that resemble emotional expressions signalling approach/avoidance behaviours. The computer modelling findings suggest that one of the factors that may determine whether prosopagnosics are able to make trustworthiness judgements is their ability to recognize emotional expressions of anger and happiness. The findings also suggest that trustworthiness judgements may be used as a subtle test of preserved emotional processing of faces.

Identity recognition and expression recognition have been shown to dissociate in developmental prosopagnosia (Duchaine, Parker, & Nakayama, 2003; Humphreys, Avidan, & Behrmann, 2007), and the possibility that trait judgements are a by-product of expression recognition mechanisms makes a straightforward prediction for prosopagnosics. Prosopagnosics with normal expression recognition should perform normally with trait judgements whereas those with expression deficits should also have trait judgement deficits. We do not have enough data to determine whether or not the prosopagnosics we have tested conform to this prediction, but we plan to examine this question in both developmental and acquired prosopagnosics.

As a preliminary test of the hypothesis that prosopagnosics' normal judgements of trustworthiness can be explained by their reliance on facial features resembling happy and angry expressions, we asked 21 normal participants to judge the second set of faces on expressiveness of anger and happiness.⁴ Both judgements of anger and judgements of happiness were highly correlated with the mean control judgements of trustworthiness, $-.62$ and $.61$ ($p < .0001$), respectively. Analyses at the level of individual participants showed that the emotion judgements significantly predicted the trustworthiness judgements of all prosopagnosics with the exception of judgements of anger for J.L. (Figure 5a). As shown in Figure 5a, relative to control participants, J.K., T.U., and J.P. showed typical correlations between their trustworthiness judgements and the controls' emotion judgements. However, for three of the four prosopagnosics, these correlations were not sufficient to account for their trustworthiness judgements. The trustworthiness judgements of J.K., T.U., and J.L. correlated with the control trustworthiness judgements even after the shared variance of the latter with emotion judgements was removed. These correlations were also in the typical range (Figure 5b). In contrast, after controlling for the contribution of emotion to trustworthiness judgements, J.P.'s trustworthiness judgements did not correlate significantly with control judgements, and the correlation was more than 2 standard deviations below the mean control correlation. These preliminary findings suggest that whereas similarity of facial features of emotionally neutral faces to expressions of anger and happiness contribute to judgements of trustworthiness, they are not sufficient to account for the performance of the prosopagnosics. However, future studies need to directly control for the presence of emotion recognition deficits in prosopagnosics.

Two other hypotheses that may account for the prosopagnosics' normal trustworthiness

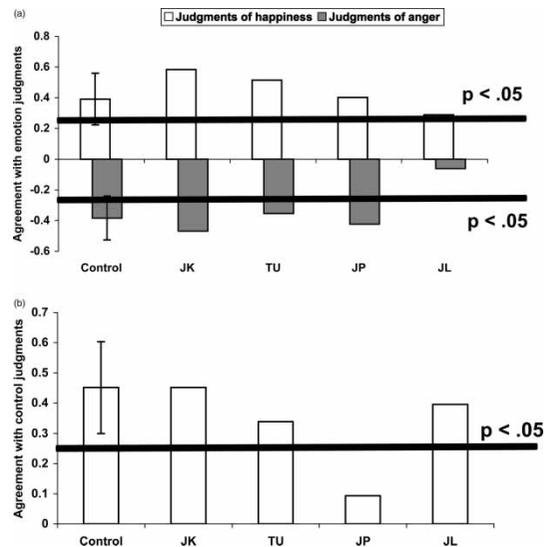


Figure 5. Agreement between prosopagnosics' judgements of trustworthiness and control judgements of anger and happiness for the second set of faces (a). The agreement is measured in Fisher z transformations of the raw correlations between individual trustworthiness judgements and the mean emotion judgements of control participants. (b) Agreement between prosopagnosics' judgements of trustworthiness and control trustworthiness judgements after removing the shared variance of the latter with judgements of anger and happiness. The agreement is measured in Fisher z transformations of the raw correlations between individual judgements and the residuals of mean control trustworthiness judgements after regressing them on emotion judgements. The error bars shows 1 standard deviation. Scores above (or below for judgements of anger) the thick black line indicate significant correlations.

judgements need to be tested. The first hypothesis is based on prior research suggesting that face recognition relies on configural or holistic processing (e.g., Farah, Wilson, Drain, & Tanaka, 1998; Young, Hellawell, & Hay, 1987). Preliminary evidence suggests that holistic processing is not necessarily required for trait judgements from faces. Todorov et al. (2007b) showed that trustworthiness judgements from facial parts (e.g., eyes) are correlated with judgements from the

⁴ Participants were told that although the faces were emotionally neutral, they could show subtle emotional information and were asked to rate how angry or happy the person is on a 9-point scale, ranging from 1 (not at all) to 9 (moderately). Both judgements were reliable, Cronbach's $\alpha = .92$ for judgements of happiness and $\alpha = .84$ for judgements of anger.

whole face, suggesting that incomplete face information is often sufficient for people to make a trait judgement. In fact, after short time exposures (e.g., 33 ms) facial parts carried as much information for the trait judgement as the whole face. If prosopagnosics are able to process information from facial parts but are unable to integrate this information (e.g., Barton, Cherkasova, Press, Intriligator, & O'Connor, 2003), they can demonstrate both trait judgements within the normal range and impaired face recognition. However, all four prosopagnosics in our study and many in another study (Yovel & Duchaine, 2006) demonstrated deficits with face part discrimination, so the relationship between part processing, trait judgements, and prosopagnosia is not straightforward.

The second hypothesis is that trait judgements and encoding of identity are sensitive to different bands of spatial frequency information. Bar et al. (2006) argued that trait judgements from faces are derived from low-spatial-frequency information. Similar arguments have been made for expressions of emotions (Schyns & Oliva, 1999), and functional imaging studies suggest that the amygdala is particularly sensitive to low-spatial-frequency information in the discrimination of fearful from neutral faces (Vuilleumier, Armony, Driver, & Dolan, 2003). In contrast, there have been several accounts that face recognition is more sensitive to middle-range spatial-frequency information, which provides more details for face identification (e.g., Costen, Parker, & Craw, 1996; Nasanen, 1999). Thus, if prosopagnosics have impairments in the processing of middle-range spatial frequency from faces but not in low-spatial-frequency information, this would account for abnormal performance on face recognition tasks and normal performance on trait judgements tasks. Finally, it is possible that identity and trait judgements depend on identical information but are computed in different neural systems. These hypotheses remain to be tested, and it would be particularly important to extend the current findings to acquired prosopagnosics with well-defined brain lesions.

Prosopagnosia provides a unique window for the study of specialized face processes.

Dissociations within face processing have been critical for the development of models of face perception (Bruce & Young, 1986; Ellis & Lewis, 2001; Haxby et al., 2000). Studies with prosopagnosics have found evidence for normal processing of emotional expressions (Bentin et al., 2007; Damasio, Tranel, & Damasio, 1990; Duchaine et al., 2003; Humphreys, Avidan, & Behrmann, 2007; Tranel, Damasio, & Damasio, 1988) and intact categorical inferences about age and gender (Tranel et al., 1988). The current studies find another islet of dissociable face processing—forming person impressions.

Manuscript received 11 July 2007

Revised manuscript received 16 November 2007

Revised manuscript accepted 10 March 2008

First published online 18 May 2008

REFERENCES

- Adams, R. B., Ambady, N., Macrae, N., & Kleck, R. E. (2006). Emotional expressions forecast approach-avoidance behavior. *Motivation & Emotion, 30*, 179–188.
- 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.
- Bahrnick, H. P., Bahrnick, P. O., & Wittlinger, R. (1975). Fifty years of memory for names and faces: A cross-sectional approach. *Journal of Experimental Psychology: General, 104*, 54–75.
- Bar, M., Maital, N., & Linz, H. (2006). Very first impressions. *Emotion, 6*, 269–278.
- Barton, J. J. S., Cherkasova, M., Press, D. Z., Intriligator, J., & O'Connor, M. (2003). Developmental prosopagnosia: A study of three patients. *Brain and Cognition, 51*, 12–30.
- Behrmann, M., & Avidan, G. (2005). Congenital prosopagnosia: Face blind from birth. *Trends in Cognitive Sciences, 9*, 180–187.
- Bellugi, U., Adolphs, R., Cassady, C., & Chiles, M. (1999). Towards the neural basis for hypersociability in a genetic syndrome. *Neuroreport, 10*, 1653–1657.

- 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.
- Bruce, V., & Young, A. (1986). Understanding face recognition. *British Journal of Psychology*, *77*, 305–327.
- Calder, A. J., & Young, A. W. (2005). Understanding the recognition of facial identity and facial expression. *Nature Reviews Neuroscience*, *6*, 641–651.
- Costen, N. P., Parker, D. M., & Craw, I. (1996). Effects of high-pass and low-pass spatial filtering on face identification. *Perception & Psychophysics*, *58*, 602–612.
- Damasio, A. R., Tranel, D., & Damasio, H. (1990). Face agnosia and the neural substrates of memory. *Annual Review of Neuroscience*, *13*, 89–109.
- Duchaine, B., Germine, L., & Nakayama, K. (2007a). Family resemblance: Ten family members with prosopagnosia and within-class object agnosia. *Cognitive Neuropsychology*, *24*, 419–430.
- Duchaine, B., & Nakayama, K. (2005). Dissociations of face and object recognition in developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *17*, 249–261.
- Duchaine, B., & Nakayama, K. (2006a). Developmental prosopagnosia: A window to content-specific face processing. *Current Opinion in Neurobiology*, *16*, 166–173.
- Duchaine, B., & Nakayama, K. (2006b). The Cambridge Face Memory Test: Results for neurologically intact individuals and an investigation of its validity using inverted face stimuli and prosopagnosic participants. *Neuropsychologia*, *44*, 576–585.
- Duchaine, B. C., Parker, H., & Nakayama, K. (2003). Normal emotion recognition in a developmental prosopagnosic. *Perception*, *32*, 827–838.
- Duchaine, B., Yovel, G., & Nakayama, K. (2007b). No global processing deficit in the Navon task in 14 developmental prosopagnosics. *Social, Cognitive, and Affective Neuroscience*, *2*, 104–113.
- Ellis, H. D., & Lewis, M. B. (2001). Capgras delusion: A window on face recognition. *Trends in Cognitive Sciences*, *5*, 149–156.
- 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.
- Farah, M. J., Wilson, K. D., Drain, M., & Tanaka, J. N. (1998). What is “special” about face perception? *Psychological Review*, *105*, 482–498.
- Fridlund, A. J. (1994). *Human facial expression: An evolutionary view*. San Diego, CA: Academic Press.
- Gobbini, M. I., & Haxby, J. V. (2007). Neural systems for recognition of familiar faces. *Neuropsychologia*, *45*, 32–41.
- 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.
- 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.
- 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.
- Lundqvist, D., Flykt, A., & Öhman, A. (1998). *The Karolinska directed emotional faces*. Stockholm: Karolinska Institute, Department of Clinical Neuroscience.
- Macmillan, N. A., & Creelman, C. D. (1991). *Detection theory: A user's guide*. New York: Cambridge University Press.
- 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.
- McCarthy, G., Puce, A., Gore, J. C., & Allison, T. (1997). Face-specific processing in the human fusiform gyrus. *Journal of Cognitive Neuroscience*, *9*, 605–610.
- Nasanen, R. (1999). Spatial frequency bandwidth used in the recognition of facial images. *Vision Research*, *39*, 3824–3833.
- Oosterhof, N. N., & Todorov, A. (2008). *The functional basis of face evaluation*. Manuscript submitted for publication.
- Perrett, D. I., Hietanen, J. K., Oram, M. W., & Benson, P. J. (1992). Organization and functions of cells responsive to faces in the temporal cortex. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *335*, 23–30.

- Riddoch, M. J., & Humphreys, G. W. (1993). *Birmingham Object Recognition Battery*. Hove, UK: Lawrence Erlbaum Associates.
- Schyns, P. G., & Oliva, A. (1999). Dr. Angry and Mr. Smile: When categorization flexibly modifies the perception of faces in rapid visual presentations. *Cognition*, *69*, 243–265.
- Todorov, A. (2006). Unpublished raw data. Princeton University, NJ.
- Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. In M. Miller & A. Kingstone (Eds.), *The year in cognitive neuroscience 2008*, *1*, 208–224.
- Todorov, A., Baron, S., & Oosterhof, N. N. (in press). Evaluating face trustworthiness: A model based approach. *Social, Cognitive, & Affective Neuroscience*.
- Todorov, A., Gobbini, M. I., Evans, K. K., & Haxby, J. V. (2007a). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, *45*, 163–173.
- Todorov, A., Pakrashi, M., Loehr, V. R., & Oosterhof, N. (2007b). *Evaluating faces on trustworthiness: Automatic assessment of face valence*. Manuscript submitted for publication.
- 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.
- 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.
- Vuilleumier, P., Armony, J. L., Driver, J., & Dolan, R. J. (2003). Distinct spatial frequency sensitivities for processing faces and emotional expressions. *Nature Neuroscience*, *6*, 624–631.
- Wada, Y., & Yamamoto, T. (2001). Selective impairment of facial recognition due to a haematoma restricted to the right fusiform and lateral occipital region. *Journal of Neurology, Neurosurgery, and Psychiatry*, *71*, 254–257.
- 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., Strange, B., O'Doherty, J., & Dolan, R. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of face. *Nature Neuroscience*, *5*, 277–283.
- Young, A. W., Hellawell, D., & Hay, D. C. (1987). Configurational information in face perception. *Perception*, *16*, 747–759.
- Yovel, G., & Duchaine, B. (2006). Specialized face perception mechanisms extract both part and spacing information: Evidence from developmental prosopagnosia. *Journal of Cognitive Neuroscience*, *18*, 580–593.
- Yovel, G., & Kanwisher, N. (2005). The neural basis of the behavioral face-inversion effect. *Current Biology*, *15*, 2256–2262.
- Zebrowitz, L. A., Voinescu, L., & Collins, M. A. (1996). “Wide-eyed” and “crooked-faced”: Determinants of perceived and real honesty across the life span. *Personality and Social Psychology Bulletin*, *22*, 1258–1269.