

Robust learning of affective trait associations with faces when the hippocampus is damaged, but not when the amygdala and temporal pole are damaged

Alexander Todorov,¹ and Ingrid R. Olson²

¹Department of Psychology and Center for the Study of Mind, Brain and Behavior, Princeton University, Princeton, NJ and

²Department of Psychology, Temple University, 1701 N. 13th Street, Philadelphia, PA 19122, USA

People can form evaluative associations with faces after obtaining a small amount of behavioral information. We studied whether patients with medial temporal lobe amnesia can form such associations. Participants were presented with trustworthy- and untrustworthy-looking faces paired with positive or negative descriptions of behaviors. After the learning task, they were asked to rate the same faces on trait dimensions—trustworthiness, likeability and competence—and to make forced-choice judgments between faces. Normal young and older adults judged faces that had been associated with positive behaviors more positively than faces that had been associated with negative behaviors. A patient with hippocampal lesions showed similar learning effects. In contrast, two patients with hippocampal lesions that extended into the left amygdala and temporal pole showed little evidence of learning. All patients judged trustworthy-looking faces more positively than untrustworthy-looking faces. The findings suggest that the hippocampus is not critical for learning affective associations between traits and faces.

Keywords: social cognition; amnesia; face perception; hippocampus; amygdala

People form person impressions from minimal information (Uleman *et al.*, 2005). Such impressions can originate in nonverbal behaviors (Albright *et al.*, 1988; Ambady and Rosenthal, 1992; Ambady *et al.*, 1995), facial appearance (Olson and Marshuetz, 2005; Bar *et al.*, 2006; Willis and Todorov, 2006) or behavioral information (Carlston and Skowronski, 1994; Todorov and Uleman, 2002, 2003, 2004). Here, we focus on how people integrate trait information extracted from facial appearance and behaviors to form person impressions. Both trait inferences (e.g. trustworthy) from facial appearance and behavioral information have been described as occurring rapidly, effortlessly and spontaneously (Todorov and Uleman, 2003; Uleman *et al.*, 2005; Todorov *et al.*, submitted for publication). However, impressions from facial appearance are perceptual, and may not be based on prior person knowledge. For example, 100 ms exposure to a novel face is sufficient for people to form a person impression (Willis and Todorov, 2006). In contrast, impressions from behaviors depend on forming associations between the affective implications of these behaviors with the person performing the behavior (Todorov and Uleman, 2002; Todorov *et al.*, 2007). Thus, it is possible to observe

dissociations between impressions from faces and behaviors when the ability to learn associations is disrupted.

In the current experiment, we used faces that differed in their perceived trustworthiness because of prior research suggesting that faces are spontaneously categorized on this dimension and that the amygdala plays a key role in this categorization (Winston *et al.*, 2002; Engell *et al.*, 2007; Todorov, 2008; Todorov *et al.*, in press). For example, neuroimaging studies show that untrustworthy faces evoke a stronger response in the amygdala than trustworthy faces (Winston *et al.*, 2002; Engell *et al.*, 2007). Moreover, bilateral damage to the human amygdala impairs the ability to discriminate trustworthy- from untrustworthy-looking faces (Adolphs *et al.*, 1998). In contrast, developmental prosopagnosics, who are unable to recognize facial identity, can nevertheless, make normal trustworthiness judgments from faces (Todorov and Duchaine, in press), suggesting that trait judgments from faces are independent of memory for faces.

Although making trait inferences from behaviors and associating these inferences with faces involves more complex processes than making trait inferences from facial appearance, these processes are highly efficient. For example, presenting a description of behavior (e.g. 'Henry turned in someone else's project under his own name'.) for 2 s with an unfamiliar face is sufficient to form an unfavorable impression of Henry (Todorov and Uleman, 2003). This effect persists even if people have no explicit

Received 14 November 2007; Accepted 25 April 2008

Advance Access publication 23 May 2008

We thank Manish Pakrashi, Valerie Loehr, Nick Oosterhof, Marianna Stark and Youssef Ezzyat for help with testing subjects, and our amnesic patients for volunteering their time. This research was supported in part by RO1 MH071615 to I.R.O. and by National Science Foundation Grant BCS-0446846 to A.T.

Correspondence should be addressed to Alexander Todorov, Department of Psychology, Princeton University, Princeton, NJ 08540, USA. E-mail: atodorov@princeton.edu.

recollection of the behavior (Carlston and Skowronski, 1994; Carlston *et al.*, 1995; Todorov and Uleman, 2002). In a highly pertinent study, Johnson *et al.* (1985, Study 2) presented Korsakoff's 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. This was the first study suggesting that amnesiacs can learn affective associations with faces.

The first objective of the current experiment was to test whether the medial temporal lobe memory system—in particular the hippocampus—is critical for learning of affective trait associations with faces. We studied three patients with amnesia due to lesions in the hippocampus and surrounding tissue, as well as younger and older normal participants. For two of the patients, the lesions extended into the left amygdala and temporal pole.

Participants were presented with trustworthy- and untrustworthy-looking faces that were paired either with positive or with negative behaviors (Figure 1). After the learning task, participants were asked to rate the faces on

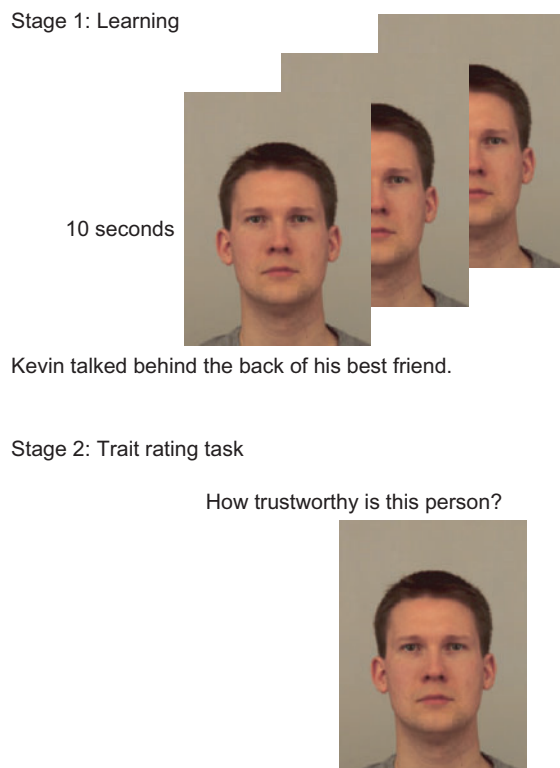


Fig. 1 Experimental paradigm. In the first task, participants were presented with faces and behaviors. Each face was presented on three consecutive trials with three different behaviors. Each trial lasted 10 s. In the second task, participants were presented with the same faces and asked to judge them on trustworthiness, likeability and competence. In the third task, participants were presented with pairs of faces and asked to select the nicer person.

trustworthiness, likeability and competence. After the trait-rating task, participants were presented with pairs of trustworthy- and untrustworthy-looking faces and asked to choose the nicer person. These choices provided additional measures of the effect of learning and the effect of facial appearance on person judgments.

It is well known that hippocampal damage impairs explicit memory but leaves many types of implicit memory intact (Squire, 2004). The expression of learned preferences does not require access to explicit learning (Kunst-Wilson and Zajonc, 1980; Zajonc, 1980; Hill *et al.*, 1989; Esteves *et al.*, 1994; Morris *et al.*, 1998) or memory (Lieberman *et al.*, 2001). Given these findings, we expected that hippocampal damage would spare affective trait-face associative learning.

Although the hippocampus may not be necessary for forming or expressing affective memories, the amygdala may be critical for the consolidation of emotional memories (McGaugh, 2004; Phelps and LeDoux, 2005; Somerville *et al.*, 2006). The temporal pole may also play a role in the formation of associations between affective knowledge and faces because of its involvement in social and emotional semantic memory (Ellis *et al.*, 1989; Olson *et al.*, 2007).

The second objective of the experiment was to test to what extent learning of affective trait associations with faces is preserved in old relative to young adults. As in prior studies with college samples (Carlston and Skowronski, 1994; Carlston *et al.*, 1995; Todorov and Uleman, 2002, 2003, 2004), we expected that younger participants would show robust learning of trait associations with faces. Consistent with prior studies (Johnson *et al.*, 1985; Johnson and Multhaup, 1993), we also expected that this learning would be well preserved in older populations. Although there is abundant research showing the negative impact of aging on explicit memory (Fleischman *et al.*, 2004; Henry *et al.*, 2004), there is mounting evidence that emotional learning and memory are well preserved with age (Mather, 2004). However, it is unclear whether such learning remains completely intact.

To summarize our predictions, we expected that both younger and older normal participants would be able to learn the associations between trait inferences and faces. Specifically, they should rate faces that were associated with positive behaviors more positively than faces that were associated with negative behaviors, and should be more likely to choose the former over the latter in forced-choice judgments. We expected the same pattern of results for the patient with hippocampal lesions but not for the patients with lesions extending into the left amygdala and temporal pole. We also expected that all participants would be affected by the facial appearance of the targets. Specifically, they should rate trustworthy looking faces more positively than untrustworthy looking faces. Although the amygdala has been implicated in judgments of trustworthiness from faces, only patients with bilateral amygdala damage show deficits in these judgments (Adolphs *et al.*, 1998).

METHOD

Participants

Thirty-one normal participants (20 younger controls and 11 older controls) and three patients with lesions in the medial temporal lobe participated in the experiment.

Young controls

Twenty undergraduate students from Princeton University (8 males, 12 females, mean age = 20 years) participated for partial course credit.

Older controls

Eleven older healthy adults (three males, eight females, mean age = 57 years) with an average of 13 years of education participated for \$15 compensation. Average verbal IQ as measured by the Wechsler Adult Intelligence Scale (WAIS), 3rd edition (Wechsler, 1997) was 100. They had normal or corrected-to-normal visual acuity and no evidence of psychiatric or neurological abnormalities.

Patients

Three patients with anterograde amnesia were referred to us by a neurologist at the Hospital of the University of Pennsylvania. Amnesia was verified by (i) presence of bilateral medial temporal lobe damage on MRI scans evaluated by a radiologist and neurologist (Figure 2); (ii) bedside memory testing by a neurologist and more quantitative testing by a staff neuropsychologist and (iii) performance on the Wechsler Memory Scale (WMS), 3rd edition (Wechsler, 1997). Patients had an average of 12 years of education and an average verbal IQ of 92, which is numerically, but not statistically below the mean of 100, as measured by the WAIS (Wechsler, 1997). The mean General Memory score on the WMS, was 60 and the Visual Delayed Memory score was 62, which are both statistically below the mean of 100 and in the

severely impaired range. The only area of lesion overlap in the patients is the anterior hippocampus. All patients have participated in numerous experiments and it has been observed that at all times they exhibit normal affect and show no signs of depression, as assessed by self-report and performance on the Beck Depression Inventory (BDI), or any other affective illnesses. All patients have normal face perception (Ezzyat and Olson, 2008). Detailed information about each patient is listed below and in Table 1; brain scans can be viewed in Figure 2.

Patient H.T. (age 64) has focal bilateral hippocampal damage as evidenced by hyperintensities in the hippocampus on T2 weighted MR scans (left greater than right), as well as a small hyperintensity in the white matter of the left parietal lobe. Damage was caused in the setting of a basilar meningitis and CNS vasculitis. She and her family report that her behavior is unchanged from the past except for a radical decline in her memory. She can no longer read novels or watch television because she cannot follow the story line. She sometimes gets confused when having a conversation, due to an inability to remember the topic of conversation and she cannot navigate to new places because she cannot

Table 1 Demographic and test information for patients with medial temporal lobe damage

Patient	Age	Sex	Education	Lesion	Etiology	Verbal IQ	WMS	BDI
H.T.	64	F	12	H	Basilar meningitis	92	69	11
M.S.	63	F	12	H+	Encephalitis	90	45	9
C.T.	68	M	12	H+	Encephalitis	98	67	—

Notes: 'H' indicates damage limited to the hippocampus; 'H+' indicates damage to the hippocampus and surrounding structures. The WMS measures general memory performance (Wechsler, 1997). For both verbal IQ and WMS, the mean is 100 (s.d. = 15). The BDI is a self-report measure of severity of depression. Scores within the range of 0–13 indicate minimal depression. We did not have data for CT on the BDI.

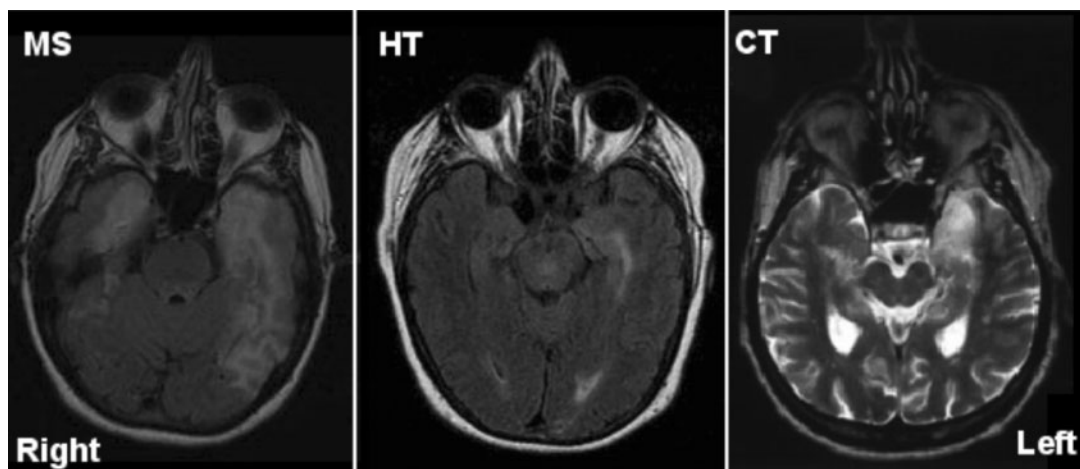


Fig. 2 Axial MRI scans from the three patients with MTL damage—M.S., H.T. and C.T.—shown in radiological convention. Images for M.S. and H.T. are fluid-attenuated inversion recovery (FLAIR) and for C.T. are T2 weighted. Patient H.T. had hippocampal damage, without amygdala damage, while M.S. and C.T. had both hippocampus and amygdala damage, greater on the left than on the right.

remember where she was a few minutes ago. She exhibits no other neuropsychological problems. H.T. exhibits no damage to the amygdala.

Patient M.S. (age 62) has bilateral medial temporal lobe damage as a result of herpes encephalitis in 1999. Damage on the left extends into the temporal pole, amygdala, perirhinal and hippocampal cortex, while on the right it extends into entorhinal and hippocampal cortex, as assessed by MRI. Damage on the left also extends into middle-inferior temporal regions. M.S. is densely amnesic and also has some anomia, most likely due to left lateral temporal and polar damage (Lezak, 1995). This problem has steadily decreased since her neurological insult. Because she has amygdala damage, additional tests were administered. In line with one prior study of patients with amygdala damage, when shown Heider and Simmel (1944) 'animacy' displays, she does not see any social meaning in the moving geometric objects (Heberlein and Adolphs, 2004).

Patient C.T. (age 68) has bilateral medial temporal lobe (MTL) damage as a result of encephalitis in 2001. The damage extends to the left anterior hippocampus, entorhinal cortex, amygdala and temporal pole. Damage on the right is restricted to anterior portions of the hippocampus. His only cognitive complaint is dense amnesia. He no longer watches TV, reads or drives due to his memory impairment. No other cognitive problems were observed or reported.

Stimulus material

Stimuli consisted of frontal head-shot photographs with neutral expressions and a direct gaze from the Karolinska Directed Emotional Faces set (Lundqvist *et al.*, 1998). These were photographs of amateur actors between 20 and 30 years of age with no beards, mustaches, earrings, eyeglasses or visible make-up, all wearing gray T-shirts. All of the faces were previously rated on trustworthiness in our laboratory (Engell *et al.*, 2007). The ratings were collected from a large sample of raters ($n=129$) and the trustworthiness judgments were highly reliable (Cronbach's $\alpha=0.98$). Based on the mean trustworthiness judgments, we selected 12 trustworthy looking faces (six males and six females; $M=5.22$, $s.d.=0.67$, on a 9-point scale) and 12 untrustworthy looking faces (six males and six females; $M=4.40$, $s.d.=0.71$) that were matched on facial attractiveness.

We generated 36 descriptions of positive and 36 descriptions of negative behaviors. The behaviors were unambiguously positive (e.g. 'Tony volunteered his time as a big brother to a fatherless child'.) or negative (e.g. 'Ryan knowingly engaged in unprotected sex after testing positive for HIV'.). To validate that the sentences had the intended affective meaning, we asked 10 participants to rate each behavioral statement on a seven-point scale ranging from -3 (extremely negative) to 3 (extremely positive). Every participant rated the positive behaviors as positive ($M=1.59$, $s.d.=0.58$) and the negative behaviors as negative

($M=-1.78$, $s.d.=0.42$), $t(9)=10.98$, $P<0.001$. The behaviors are available on request from the authors.

For each type of behavior, we randomly divided the behaviors into groups of 3 (12×3 positive behaviors and 12×3 negative behaviors). The groups of behaviors were randomly assigned to the 24 faces with the constraint that half of the positive behaviors were assigned to untrustworthy faces (equal number of male and female faces) and half of the negative behaviors were assigned to trustworthy faces (equal number of male and female faces). This created a 2 (Behavior: Positive vs negative) $\times 2$ (Face: Untrustworthy vs trustworthy) design.

Procedure

The experiment consisted of three tasks: (i) person learning task; (ii) person trait rating task and (iii) person preference task. Participants were tested individually on either a laptop computer or a desktop computer. The experiment was programmed in Eprime for PC.

Participants were told that this was a study on face impressions. For the first task, they were presented with faces and behaviors and instructed to carefully read the behavioral information and try to imagine the person performing the behavior. Each trial consisted of a face-behavior pair, with the behavior positioned below the face on the screen, presented for 10 s (Figure 1). Each face was presented on three consecutive trials with three different behaviors with the same valence. The order of the faces was randomized for each participant. Participants were presented with 72 trials ($24 \text{ faces} \times 3 \text{ behaviors}$).

For the second task, participants were shown the same set of faces, without any information about their behaviors. The task was to look at each face and judge it on three different trait dimensions: trustworthiness, competence and likeability. The judgments were made on a seven-point scale, the anchors of which were tailored for each judgment. For example, for the judgments of likeability ('How likeable is this person?'), the scale ranged from 1 (least likeable) to 7 (most likeable). Each trial consisted of a trait judgment question presented above the face and a response scale presented below the face. Participants responded by pressing the corresponding number key on the keyboard. Each face was presented until a response was entered into the keyboard. The next trial was presented after a 1 s delay. The order of the trials was randomized for each participant. Each participant was presented with 72 trials ($24 \text{ faces} \times 3 \text{ judgments}$).

For the third task, participants were presented with pairs of faces, again drawn from the same set used in the first task and asked to choose which person was nicer, based on their gut instinct. Each trial consisted of one trustworthy and one untrustworthy face (based on their appearance) of the same gender. We created 72 face pairs (all 36 pair-wise comparisons for the 6 trustworthy and 6 untrustworthy males, and all 36 pair-wise comparisons for 6 trustworthy

and 6 untrustworthy females). Participants responded by pressing the 'F' key to select the face on the left and the 'J' key to select the face on the right. The position of the faces was counterbalanced across participants. The inter-stimulus interval was 1 s. The order of trials was randomized for each participant.

RESULTS

Person trait rating task

The data for the normal controls were submitted to a 2 (Behavior: positive vs negative) \times 2 (Face: trustworthy vs untrustworthy) \times 3 (Trait judgment: trustworthiness vs likeability vs competence) \times 2 (Group: younger vs older control) mixed-subjects ANOVA. As expected, there was a strong learning effect of behavior. Participants rated faces associated with positive behaviors more positively ($M = 4.56$, $s.d. = 0.84$) than faces associated with negative behaviors ($M = 3.13$, $s.d. = 0.84$), $F(1, 29) = 42.89$, $P < 0.001$, $\eta^2 = 0.60$. The pattern of means in Table 2 suggests that the effect of behavior was weaker for older than younger controls, but the interaction between behavior and group did not reach significance, $F(1, 29) = 2.35$, $P = .14$, $\eta^2 = 0.08$ ($F < 1$ for the main effect of group). To demonstrate that the learning effect was robust for older controls, we conducted an analysis on their data only. The effect of behavior remained highly significant, $F(1, 10) = 14.88$, $P < 0.003$, $\eta^2 = 0.60$.

The overall analysis also revealed a main effect of judgment, $F(1, 29) = 6.22$, $P < 0.004$, $\eta^2 = 0.18$, showing that participants' judgments of competence ($M = 3.98$,

$s.d. = 0.55$) were more positive than their judgments of trustworthiness ($M = 3.79$, $s.d. = 0.47$) and likeability ($M = 3.78$, $s.d. = 0.44$). More important, this effect was qualified by an interaction with behavior, $F(2, 58) = 3.25$, $P < 0.046$, showing that the learning effect was stronger for the more relevant judgments of trustworthiness and liking than for judgments of competence (Figure 3). After all, a person who commits negative behaviors, although dislikeable, can be highly competent.

The only other significant effect was the effect of facial appearance. Participants rated trustworthy looking faces ($M = 3.96$, $s.d. = 0.47$) more positively than untrustworthy looking faces ($M = 3.73$, $s.d. = 0.54$), $F(1, 29) = 6.06$, $P < 0.020$, $\eta^2 = 0.17$, although the effect of facial appearance was much weaker than the learning effect. In sum, normal participants were affected by both the behavior information and the facial appearance of the targets.

Patient H.T. showed excellent learning (Table 2 and Figure 3). On all three judgments—trustworthiness, likeability and competence—she showed typical performance. The learning effect as measured by the difference between her ratings of faces associated with positive behaviors and faces associated with negative behaviors was larger than the learning effect for the controls with similar age (0.83 s.d. above the mean older performance). In contrast to H.T., patients M.S. and C.T. showed less evidence of learning (Table 2 and Figure 3). Patient M.S.' performance on the trustworthiness judgments was within the typical range for the older controls (Figure 3a), although below their mean performance. Across the three judgments, her performance was 0.91 s.d. below the older controls' performance, and we cannot rule out the possibility of some learning. Patient C.T. did not show any evidence of learning. His performance was invariably at the bottom of the distribution (Figure 3). Across the judgments, his performance was 2.14 s.d. below the older controls' performance.

Although the three patients exhibited different degrees of learning, ranging from normal learning to no learning, they all showed effects of facial appearance on their trait judgments. All patients rated trustworthy faces more positively than untrustworthy faces (Table 2 and Figure 4). Compared to normal participants, this effect was especially pronounced for patients H.T. and M.S. (Table 2 and Figure 4). The difference between their ratings of trustworthy- and untrustworthy-looking faces was greater than the mean difference for both younger and older controls (2.50 and 1.41 s.d. above the older controls' performance for H.T. and M.S., respectively). Patient C.T.'s difference in ratings was equivalent in size to the mean difference for older controls (-0.07 s.d.).

Table 2 Mean trait judgments (standard deviations) for normal participants and amnesiac patients as a function of valence of behavior, perceived trustworthiness of faces and trait judgment

	Positive behaviors		Negative behaviors	
	Trustworthy-looking face	Untrustworthy-looking face	Trustworthy-looking face	Untrustworthy-looking face
Trustworthiness				
Younger controls	4.71 (0.99)	4.70 (1.17)	3.20 (0.66)	2.75 (0.62)
Older controls	4.20 (0.68)	4.15 (0.66)	3.30 (0.63)	3.12 (1.00)
H.T.	5.33	4.67	3.17	2.50
M.S.	5.00	3.83	4.50	3.17
C.T.	4.00	3.17	4.00	4.67
Likeability				
Younger controls	4.73 (0.92)	4.53 (1.21)	3.05 (0.52)	2.77 (0.53)
Older controls	4.62 (0.65)	4.17 (0.73)	3.23 (0.82)	3.23 (0.70)
H.T.	5.83	4.00	4.00	2.67
M.S.	5.00	4.17	4.83	3.83
C.T.	3.50	3.00	4.83	4.33
Competence				
Younger controls	4.75 (0.87)	4.66 (1.15)	3.39 (0.74)	3.03 (0.74)
Older controls	4.52 (0.95)	4.39 (0.55)	3.77 (1.04)	3.26 (0.78)
H.T.	5.17	4.67	3.83	2.00
M.S.	3.83	3.83	4.67	3.67
C.T.	3.50	3.33	3.83	3.33

Forced choice judgments

After the trait judgments task, participants were presented with 72 pairs of trustworthy- and untrustworthy-looking faces and asked to select the nicer person. Normal

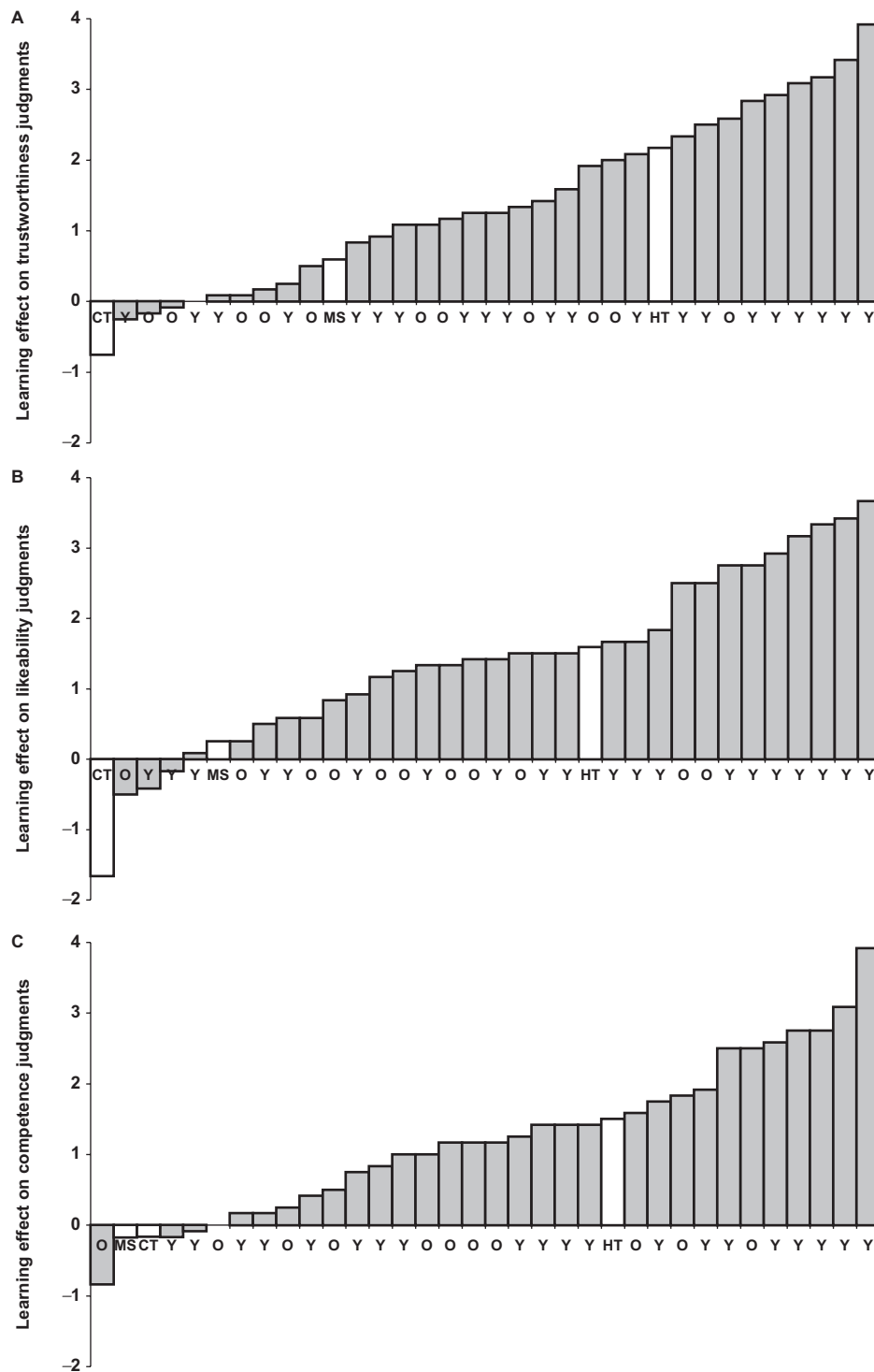


Fig. 3 Learning effect of behavior on trait judgments for younger (Y) and older (O) controls, patients H.T., M.S. and C.T. The bars show the difference between trait judgments of faces that were associated with positive behaviors and trait judgments of faces that were associated with negative behaviors for each participant. (A) Trustworthiness judgments. (B) Likeability judgments. (C) Competence judgments.

participants chose the trustworthy over the untrustworthy face on 56% of the trials (s.d. = 15), and this proportion was significantly higher than the chance level of 50%, $t(30) = 2.79$, $P < 0.009$. This effect was equivalent for younger and older controls, $t < 1$ for the difference.

Despite the overall bias to select trustworthy faces, normal participants' choices were strongly affected by the learned trait associations with the faces. When the behavior information was congruent with facial appearance—trustworthy faces associated with positive behaviors and untrustworthy

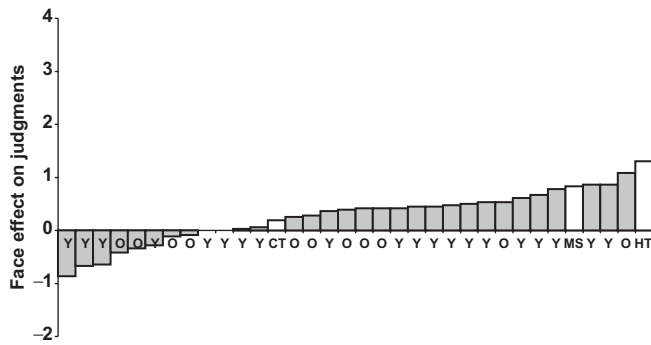


Fig. 4 Effect of facial appearance on trait judgments for younger (Y) and older (O) controls, patients H.T., M.S. and C.T. The bars show the difference between trait judgments of trustworthy- and untrustworthy-looking faces for each participant.

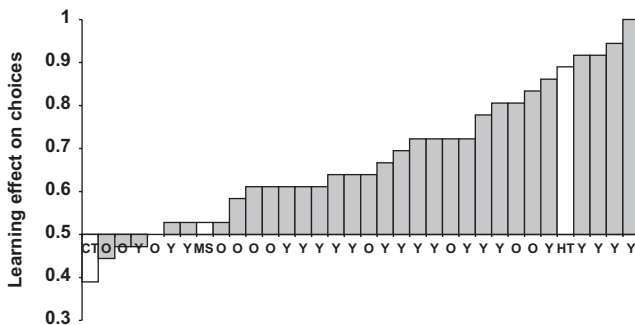


Fig. 5 Learning effect of behavior on choices for younger (Y) and older (O) controls, patients H.T., M.S. and C.T. The bars show the preference for faces associated with positive behaviors over faces associated with negative behaviors.

faces associated with negative behaviors—participants chose the trustworthy face over the untrustworthy face on 74% of the trials ($s.d. = 15$). In contrast, when the behavior information was incongruent with facial appearance—trustworthy faces associated with negative behaviors and untrustworthy faces associated with positive behaviors—participants' choices reversed and the majority chose the untrustworthy face over the trustworthy face ($M = 62\%$, $s.d. = 22\%$).

Although it seems that the learning effect on choices was stronger for congruent than incongruent face-behavior associations, correcting for the bias to prefer trustworthy over untrustworthy faces (6%) produced equivalent learning effects—68% preference for faces associated with positive behaviors over faces associated with negative behaviors. The overall learning effect on choices was stronger for younger than older controls, $t(29) = 1.97$, $P < 0.058$ (Figure 5). However, older controls remembered who was nicer at better-than chance rates, $t(10) = 2.91$, $P < 0.015$.

As in the trait-rating task, patient H.T. showed excellent learning (Figure 5). She outperformed all of the older controls and most of the younger controls on the forced choice task, showing a significant preference for faces associated with positive behaviors over faces associated with negative behaviors, $P < 0.001$ from a binomial test. In contrast to patient H.T., patients M.S. and C.T. showed little

evidence of learning. Patient M.S. showed a slight preference, which was indistinguishable from chance ($P = 0.87$), for faces associated with positive behaviors over faces associated with negative behaviors. Patient C.T. showed a preference for faces associated with negative behaviors over faces associated with positive behaviors, but this preference was not significantly different from chance ($P = 0.24$).

DISCUSSION

In this study, we investigated how people integrate information from facial appearance and behaviors in forming person impressions. On two judgment tasks, both younger and older normal participants showed robust effects of learning on judgments. Participants rated faces associated with positive behaviors more positively than faces associated with negative behaviors. Similarly to the findings of the trait-rating task, participants were more likely to choose faces associated with positive behaviors than faces associated with negative behaviors in a forced choice task. Although the learning effect appeared to be weaker for older participants, this effect was highly reliable. This finding suggests that learning of affective associations with faces or, at least, the implicit components of this learning may be relatively unaffected by aging.

Among the three patients with amnesia, H.T. showed excellent learning. This was the most important finding of the study. Although she was older than the participants in the older control group, she outperformed them on both tasks. In fact, her trait judgments were as affected by the behavioral information as the judgments of younger controls. Her forced choice judgments were comparable to the judgments of the best young participants. H.T.'s MTL lesions were restricted to the hippocampus. It is well known that the hippocampus has a critical role in episodic memory, as well as some forms of semantic memory and working memory (Squire, 2004). Interestingly, our findings, as well as the findings of Johnson *et al.* (1985), suggest that the hippocampus is not necessary for learning and memory of affective associations with faces.

In contrast to patient H.T., patients M.S. and C.T. showed little evidence of learning. In particular, patient C.T. was invariably at the bottom of the performance distributions (Figures 3 and 5). The lesions of M.S. and C.T. extended outside the anterior hippocampus and encompassed other MTL regions including the left amygdala and temporal pole. Because the lesions were not restricted to the latter two regions, we cannot conclude that these two regions are necessary for learning of affective associations with faces. However, previous findings strongly suggest that the amygdala plays a critical role in consolidation of emotional memories (McGaugh, 2004). The amygdala appears to be critical for the boost that emotion gives to memories. Patients with bilateral amygdala damage fail to accurately remember emotional pictures or words (Markowitsch *et al.*, 1994), and fail to show the normal enhancement of both

short-term (Hamann, 2001) and long-term memory by emotion (Adolphs *et al.*, 1997; Cahill *et al.*, 1995). There is also good evidence that the temporal poles are involved in binding of emotional information to high-level perceptual representations and that they can play a specific role in social memory (Ellis *et al.*, 1989; Olson *et al.*, 2007). In sum, in light of previous findings, the current findings suggest that the amygdala and temporal poles may be part of a network involved in updating of person representations. Yet, given the more extensive lesions of patients M.S. and C.T., we cannot exclude other possibilities. These patients were also older than the older controls and this difference in age may have contributed to their poor performance.

Although younger and older normal participants seemed to differ in the amount of learning, they showed equivalent effects of facial appearance on person judgments. They rated trustworthy looking faces more positively than untrustworthy looking faces and were more likely to choose the former over the latter in forced choice judgments. These effects were substantially smaller than the learning effects, suggesting that behavioral information can overwrite initial face impressions. However, it should be noted that this information was constructed to have extreme evaluative implications. In situations, where the behavioral information has ambiguous implications, the effect of facial appearance may be stronger (Trope, 1986; Hassin and Trope, 2000).

The trait judgments of all three patients were affected by the facial appearance of the targets. As the control participants, they rated trustworthy looking faces more positively than untrustworthy looking faces. In fact, relative to controls, this effect was particularly strong for H.T. and M.S. (Figure 4). An interesting possibility is that amnesiacs who cannot retrieve episodic memories about people may be overly influenced by real-time affective information available from facial appearance. This may create situations in which learning does not offset the effect of facial appearance. The finding that the patients showed normal effects of facial appearance on trait judgments despite their poor memory for faces (Ezzyat and Olson, 2008; Olson *et al.*, 2006) is consistent with Todorov and Duchaine's finding (in press) that prosopagnosics can make normal trustworthiness judgments from faces. This is further evidence for the functional independence of encoding of facial identity and trait judgments from faces.

CONCLUSIONS

People make rapid trait judgments from facial appearance. However, when given more reliable information about the person, they update their judgments accordingly. We showed here that these processes are relatively independent of age. Both young and old normal participants showed robust learning effects. We also showed that a patient with hippocampal lesions showed similar learning effects, suggesting that the hippocampus is not necessary for learning of affective associations with faces. In contrast to this patient,

two other patients with lesions that extended outside the hippocampus and included the left amygdala and temporal pole failed to learn these associations, suggesting that these regions may be critical for updating of person representations.

REFERENCES

- Adolphs, R., Cahill, L., Schul, R., Babinsky, R. (1997). Impaired declarative memory for emotional material following bilateral amygdala damage in humans. *Learning & Memory*, 4, 291–300.
- Adolphs, R., Tranel, D., Damasio, A.R. (1998). The human amygdala in social judgment. *Nature*, 393, 470–4.
- Albright, L., Kenny, D.A., Malloy, T.E. (1988). Consensus in personality judgments at zero acquaintance. *Journal of Personality and Social Psychology*, 55, 387–95.
- Ambady, N., Hallahan, M., Rosenthal, R. (1995). On judging and being judged accurately in zero-acquaintance situations. *Journal of Personality and Social Psychology*, 69, 518–29.
- Ambady, N., Rosenthal, R. (1992). Thin slices of expressive behavior as predictors of interpersonal consequences: a meta-analysis. *Psychological Bulletin*, 111, 256–74.
- Bar, M., Maital, N., Linz, H. (2006). Very first impressions. *Emotion*, 6, 269–78.
- Cahill, L., Babinsky, R., Markowitsch, H.J., McGaugh, J.L. (1995). The amygdala and emotional memory. *Nature*, 377, 295–6.
- 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, 840–56.
- 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, 420–36.
- Ellis, A.W., Young, A.W., Critchley, E.M.R. (1989). Loss of memory for people following temporal lobe damage. *Brain*, 112, 1469–83.
- 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, 1508–19.
- Esteves, F., Dimberg, U., & Öhman, A. (1994). Automatically elicited fear-conditioned skin conductance responses to masked facial expressions. *Cognition & Emotion*, 8, 393–413.
- Ezzyat, Y., Olson, I.R. (2008). The medial temporal lobe and visual working memory: comparisons across tasks, delays, and visual similarity. *Cognitive, Affective, and Behavioral Neuroscience*, 8, 32–40.
- Fleischman, D.A., Wilson, R.S., Gabrieli, J.D., Bienias, J.L., Bennett, D.A. (2004). A longitudinal study of implicit and explicit memory in old persons. *Psychology and Aging*, 19, 617–25.
- Hamann, S. (2001). Cognitive and neural mechanisms of emotional memory. *Trends in Cognitive Sciences*, 5, 394–400.
- Hassin, R., Trope, Y. (2000). Facing faces: studies on the cognitive aspects of physiognomy. *Journal of Personality and Social Psychology*, 78, 837–52.
- Heberlein, A.S., Adolphs, R. (2004). Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proceedings of the National Academy of the USA*, 111, 7487–91.
- Heider, F., Simmel, M. (1944). An experimental study of apparent behavior. *American Journal of Psychology*, 57, 243–59.
- Henry, J.D., MacLeod, M.S., Phillips, L.H., Crawford, J.R. (2004). A meta-analytic review of prospective memory and aging. *Psychology and Aging*, 19, 27–39.
- Hill, T., Lewicki, P., Czyzewska, M., Boss, A. (1989). Self-perpetuating development of encoding biases in person perception. *Journal of Personality & Social Psychology*, 57, 373–87.
- 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.

- Johnson, M.K., Multhaup, K.S. (1993). Emotion and MEM. In: Christianson, S.A., editor. *The Handbook of Emotion and Memory: Current Research and Theory*. Hillsdale, NJ: Erlbaum Associates, pp. 33–66.
- Kunst-Wilson, W.R., Zajonc, R.B. (1980). Affective discrimination of stimuli that cannot be recognized. *Science*, 207, 557–8.
- Lezak, M.D. (1995). *Neuropsychological Assessment*. New York: Oxford University Press.
- Lieberman, M.D., Ochsner, K.N., Gilbert, D.T., Schacter, D.L. (2001). Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychological Science*, 12, 135–40.
- Lundqvist, D., Flykt, A., Ohman, A. (1998). *The Karolinska Directed Emotional Faces* [Database of standardized facial images]. S-171 76 Stockholm, Sweden: Psychology Section, Department of Clinical Neuroscience, Karolinska Hospital.
- Markowitsch, H.J., Calabrese, P., Wierker, M., et al. (1994). The amygdala's contribution to memory: a study on two patients with Urbach–Wiethe disease. *NeuroReport*, 5, 1349–52.
- Mather, M. (2004). Aging and emotional memory. In: Reisberg, D., Hertel, P., editors. *Memory and Emotion*. NY: Oxford University Press, pp. 272–307.
- McGaugh, J.L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience*, 27, 1–28.
- Morris, J.S., Ohman, A., Dolan, R.J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393, 467–70.
- Olson, I.R., Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, 5, 498–502.
- Olson, I.R., Moore, K., Stark, M., Chatterjee, A. (2006). Visual working memory is impaired when the medial temporal lobe is damaged. *Journal of Cognitive Neuroscience*, 18, 1087–97.
- Olson, I.R., Plotzer, A., Ezzyat, Y. (2007). The enigmatic temporal pole: a review of findings on social and emotional processing. *Brain*, 1320, 1718–31.
- Phelps, E.A., LeDoux, J.E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, 48, 175–87.
- Squire, L. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of Learning and Memory*, 82, 171–7.
- 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, 1253–65.
- Todorov, A. (2008). Evaluating faces on trustworthiness: An extension of systems for recognition of emotions signaling approach/avoidance behaviors. In Kingstone, A., Miller, M., editors. *The Year in Cognitive Neuroscience 2008*, *Annals of the New York Academy of Sciences*, 1124, 208–24.
- Todorov, A., Baron, S.G., Oosterhof, N.N. (In press). Evaluating face trustworthiness: a model based approach. *Social, Cognitive, & Affective Neuroscience*. [Epub ahead of print; doi:10.1093/scan/nsn009].
- Todorov, A., Duchaine, B. (In press). Reading trustworthiness in faces without recognizing faces. *Cognitive Neuropsychology*.
- Todorov, A., Gobbini, M.I., Evans, K.K., Haxby, J.V. (2007). Spontaneous retrieval of affective person knowledge in face perception. *Neuropsychologia*, 45, 163–73.
- 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, 1051–65.
- Todorov, A., Uleman, J.S. (2003). The efficiency of binding spontaneous trait inferences to actor's faces. *Journal of Experimental Social Psychology*, 39, 549–62.
- Todorov, A., Uleman, J.S. (2004). The person reference process in spontaneous trait inferences. *Journal of Personality and Social Psychology*, 87, 482–93.
- Trope, Y. (1986). Identification and inferential processes in dispositional attribution. *Psychological Review*, 93, 239–57.
- Uleman, J.S., Blader, S., Todorov, A. (2005). Implicit impressions. In: Hassin, R., Uleman, J.S., Bargh, J.A., editors. *The New Unconscious*. New York: Oxford University Press, pp. 362–392.
- Wechsler, D. (1997). *The Wechsler Memory Scale*, 3rd edn. San Antonio Texas: The Psychological Corporation.
- Willis, J., Todorov, A. (2006). First impressions: making up your mind after 100 ms exposure to a face. *Psychological Science*, 17, 592–8.
- Winston, J., Strange, B., O'Doherty, J., Dolan, R. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of face. *Nature Neuroscience*, 5, 277–83.
- Zajonc, R.B. (1980). Feeling and thinking: preferences need no inferences. *American Psychologist*, 35, 151–75.