

# A neural mechanism of first impressions

Daniela Schiller<sup>1,2</sup>, Jonathan B Freeman<sup>2,3</sup>, Jason P Mitchell<sup>4</sup>, James S Uleman<sup>2</sup> & Elizabeth A Phelps<sup>1,2</sup>

**Evaluating social others requires processing complex information. Nevertheless, we can rapidly form an opinion of an individual during an initial encounter. Moreover, people can vary in these opinions, even though the same information is provided. We investigated the brain mechanisms that give rise to the impressions that are formed on meeting a new person. Neuroimaging revealed that responses in the amygdala and the posterior cingulate cortex (PCC) were stronger while encoding social information that was consistent, relative to inconsistent, with subsequent evaluations. In addition, these responses scaled parametrically with the strength of evaluations. These findings provide evidence for encoding differences on the basis of subsequent evaluations, suggesting that the amygdala and PCC are important for forming first impressions.**

Making sense of others in a social interaction is not easy, as each person is often a source of ambiguous and complex information. Despite this, when encountering someone for the first time, we are usually quick to judge whether we like that person or not. Indeed, people make relatively accurate and persistent evaluations on the basis of rapid observations of even less than half a minute<sup>1,2</sup>. Here we investigated the brain mechanisms giving rise to the impressions that are rapidly formed on meeting a new person.

When confronted with multifaceted social information, people will usually not be unanimous in their evaluations. For example, a person may be both smart and lazy. Although these traits have a social valence, the former is considered to be a good quality and the latter to be bad, people also assign their own subjective value to these traits on the basis of their personal preferences. Some evaluators might value intelligence more and care less about laziness, generating a positive impression of that person, whereas an opposite valuation would result in a negative impression overall. Thus, individuals vary in their evaluations despite the fact that the same information is provided. Our objective here was to examine the neural encoding of social information reflecting this subjective valuation (that is, the weight ascribed to each bit of social information) and its correlation with subsequent impressions. Neural regions that are specifically involved in the evaluative process of impression formation should show greater responding while encoding information with higher subjective value, which would be consistent with subsequent evaluations, as compared with information that has a lower subjective value.

A similar approach has been used in memory research: neural responses to items that are later remembered are different from items that are forgotten. This difference based on subsequent memory is considered to be evidence for the involvement of a brain region in memory formation and was termed the difference in memory effect (also known as the DM effect)<sup>3–5</sup>. Analogously, in social evaluation processes, information that is used or disregarded might also be

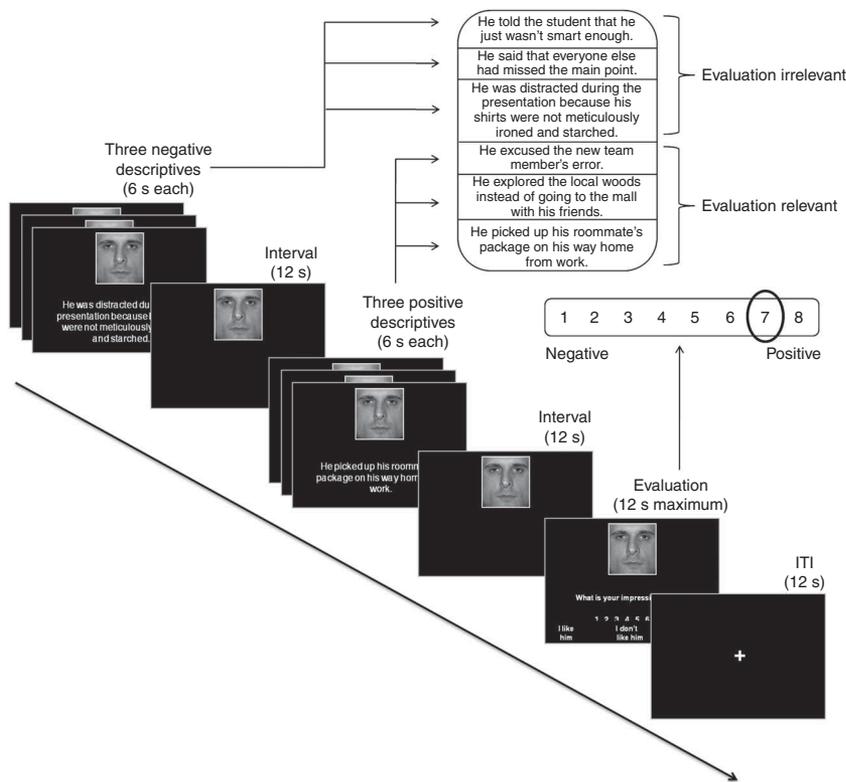
encoded differently and might therefore be predictive of subsequent evaluations. This differential neural response to information that is relevant versus irrelevant to later evaluations will be referred to as the difference in evaluation effect (or, the DE effect).

Previous neuroimaging studies typically compared social impression formation with other types of cognitive processes<sup>6,7</sup>. These studies strongly point to the involvement of the dorsomedial prefrontal cortex (dmPFC). For example, the dmPFC showed increased activation, as measured with functional magnetic resonance imaging (fMRI), when subjects were engaged in impression formation compared with memorizing<sup>6</sup>. Also, in impression formation, the dmPFC was selectively recruited when subjects formed impressions of social others compared with inanimate objects<sup>7</sup>. Beyond this, the mPFC has been extensively implicated in social cognition and the processing of social information during mentalizing, self-knowledge and person perception<sup>8,9</sup>. However, there is still no direct evidence for the involvement of this region in the evaluative process of social impressions.

We were interested in providing such evidence by decomposing the process of impression formation and isolating the evaluation component (that is, an affective judgment about a social other). The dmPFC is a likely candidate to sort through social information and sum it up into an evaluation of another person. However, because impression formation involves an affective judgment, it is also possible that neural systems more broadly involved in emotion and valuation processes, rather than social processing *per se*, are recruited.

The amygdala has been implicated in both domains. There is abundant evidence across species that this region is essential for the formation and expression of affective value (both positive and negative) to neutral objects by way of associative learning<sup>10–12</sup>. In the social domain, the amygdala has been implicated in judgments of trustworthiness<sup>13,14</sup>, in the assessment of emotion from facial expression<sup>15,16</sup> and body movements<sup>17</sup>, and was linked to implicit responses that were

<sup>1</sup>Center for Neural Science, <sup>2</sup>Department of Psychology, New York University, New York, New York, USA. <sup>3</sup>Department of Psychology, Tufts University, Medford, Massachusetts, USA. <sup>4</sup>Department of Psychology, Harvard University, Cambridge, Massachusetts, USA. Correspondence should be addressed to D.S. (schiller@cns.nyu.edu) or E.A.P. (liz.phelps@nyu.edu).



**Figure 1** An example of a person profile. One out of 20 person-profiles presented to each subject. The profile consisted of six person-descriptive sentences arranged from negative to positive (or vice versa). Each profile corresponded with a male face. A 12-s interval with the face alone separated the positive and the negative segments. Subsequently, an evaluation slide instructed subjects to form their impression on an 8-point scale (1, 'I don't like him'; 8, 'I like him'). On the basis of subjects' evaluations, the positive and negative segments were assigned to either the evaluation-relevant or evaluation-irrelevant conditions.

negative versus positive. No valence bias was found (two-tailed  $t$  tests comparing the mean proportion of negative evaluations to 0.5,  $t_{18} = 1.35$ ). Examining a valence bias in individual subjects using binomial tests found a negative bias in 4 ( $P < 0.05$ ) of the 19 subjects.

To confirm that subjects' evaluations were not guided by the order in which positive or negative descriptive sentences were introduced in a profile, we assessed the number of evaluations that were consistent with the valence of the sentences that were presented first (primacy effect) or last (recency effect) in

a profile. No order bias was found (one-sample two-tailed  $t$  tests comparing the mean proportion of evaluations consistent with first segments to 0.5,  $t_{18} = 1.12$ ). Examining an order bias in individual subjects using binomial tests found a primacy effect in 1 ( $P < 0.05$ ) of the 19 subjects.

We also confirmed that none of the faces led to more negative or more positive evaluations. Binomial tests conducted for negative evaluation proportions for each face found a negative bias for 1 face ( $P < 0.05$ ) out of 20.

These results confirm that positive and negative descriptive sentences and sentences presented first versus last were evenly distributed between the evaluation-relevant and evaluation-irrelevant categories. This point is of particular importance when examining the underlying neural responses. Given these results, it is clear that the difference in evaluation effect reflects subjective weighting of person-descriptive information driven by subjects' own interpretation of it rather than by more general effects such as the primacy (order) or negativity (valence) of the stimuli themselves.

It should be noted that order effects have been found using similar sequential presentations<sup>20</sup>. However, these effects were shown to be dependent on procedural variations, such as the sequence of the items, number and spacing of the items and instruction<sup>21</sup>. These effects were also shown to decrease with continued practice<sup>22</sup>, similar to our findings here, as subjects were practiced and familiar with the task beforehand.

To confirm that the difference at encoding on the basis of subsequent evaluation and the difference at encoding on the basis of subsequent memory are separate phenomena, we assessed whether evaluation-relevant descriptive sentences were remembered better. We found no memory difference between evaluation-relevant and irrelevant sentences (paired two-tailed  $t$  test,  $t_{18} = 0.317$ ). Examining this bias in individual subjects using a  $\chi^2$  test found an effect in 2 ( $P < 0.05$ ) of the 19 subjects. One subject remembered the evaluation-relevant

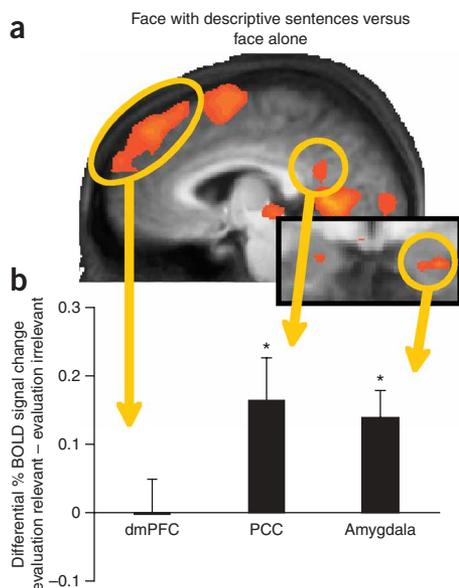
indicative of race bias<sup>18,19</sup>. Thus, the amygdala might participate in the formation of social value assigned to other people.

To test these hypotheses, we developed the difference in evaluation procedure (Fig. 1), allowing us to sort social information encoding trials by subsequent evaluations. More specifically, we measured blood oxygenation level-dependent (BOLD) signals using whole brain fMRI during exposure to different person profiles. Each profile consisted of 6 person-descriptive sentences implying different personality traits. The sentences varied gradually in their positive to negative valence (or vice versa) but evoked equivalent levels of arousal. A 12-s interval with the face alone separated the positive and the negative segments. Subsequently, an evaluation slide instructed subjects to form their impression on an 8-point scale. On the basis of these evaluations, we determined which of the presented descriptive sentences guided evaluations (evaluation relevant) and which did not (evaluation irrelevant). For example, if a subject's evaluation was positive, we assigned the positive segment of the profile to the evaluation-relevant category and the negative segment to the evaluation-irrelevant category. We then identified the brain regions dissociating items from each category (that is, difference in evaluation effect). Notably, we correlated subjects' BOLD signal with their own individual evaluations. This allowed us to identify brain regions that were consistent across subjects in processing evaluation-relevant information regardless of the particular stimuli that they considered. Immediately after the scanning session, subjects underwent a memory-recognition task.

## RESULTS

### Behavioral results

We scanned 19 volunteers while they performed the impression-formation task, which consisted of 20 profile evaluations. To confirm that subjects' impressions were guided by their idiosyncratic evaluations, rather than by the valence of the information that was presented, we assessed the number of profiles that the subjects evaluated as



information better than the evaluation-irrelevant information, whereas the other subject showed an opposite effect. Both evaluation-relevant and evaluation-irrelevant sentences were remembered significantly above chance (0.2; one-sample  $t$  test,  $t_{18} = 5.37$ ,  $P < 0.01$ ,  $t_{18} = 3.91$ ,  $P < 0.01$ , respectively).

We conducted an additional study on a separate group of subjects without the fMRI scan to examine memory immediately after each profile presentation. This task was similar to the one used in the scanner, with the exception that, after the evaluation slide of each profile, subjects were presented with the face along with six descriptive sentences (five new and one old). The subjects had to identify which of the sentences was previously presented with the face. Again, we found no difference in memory for evaluation-relevant and irrelevant descriptive sentences and no valence or order biases in memory (see **Supplementary Results** and **Supplementary Fig. 1** online). Together, these results eliminate differential memory as an alternative explanation of the findings in our procedure. Thus, evaluations were not driven by episodic memory for item information, confirming that the difference in evaluation effect is different from the difference in memory effect.

### Neuroimaging results

In our neuroimaging analysis, we examined which regions showed the difference in evaluation effect out of regions that were broadly engaged in the impression-formation task. Functional regions of interest (ROIs) were identified contrasting faces with person-descriptive information and face-alone presentations (false discovery rate  $< 0.05$ ; **Fig. 2** and **Table 1**). To examine whether the regions revealed by this contrast (**Table 1**) show the difference in evaluation effect, we extracted the BOLD response from each of these ROIs and compared the mean percentage BOLD signal change during the presentation of evaluation-relevant versus evaluation-irrelevant person-descriptive sentences (two-tailed  $t$  tests,  $P < 0.05$ ; **Table 1**). The only regions showing significantly greater BOLD responses to evaluation-relevant sentences were the amygdala and the PCC (**Fig. 2b**), and the thalamus ( $P < 0.05$ ; **Table 1**). There were no regions showing the opposite effect.

For comparison, we examined an extensive region of dmPFC revealed by the face with descriptive sentences versus face alone contrast (**Fig. 2a**). However, the analysis separating dmPFC responses into evaluation-relevant versus evaluation-irrelevant person-descriptive sentences revealed no difference in evaluation effect (**Fig. 2b**).

**Figure 2** Brain regions demonstrating the difference in evaluation effect out of regions broadly engaged in the impression-formation task. **(a)** Functional ROIs were identified by contrasting faces with person-descriptive sentences versus face-alone presentations (false discovery rate  $< 0.05$ ). The dmPFC ( $x = -7$ ,  $y = 24$ ,  $z = 53$ ; Brodmann area 8/9), PCC ( $x = 0$ ,  $y = -51$ ,  $z = 23$ ; Brodmann area 23) and left amygdala ( $x = -23$ ,  $y = -8$ ,  $z = -16$ ) are denoted by yellow circle on the statistical activation map. The full list of regions revealed by this contrast is detailed in **Table 1**. **(b)** To examine whether these ROIs show the DE effect, we extracted the BOLD response from each of these regions (dmPFC, 6,561-mm<sup>3</sup> voxels; PCC, 760-mm<sup>3</sup> voxels; amygdala, 111-mm<sup>3</sup> voxels) and compared the mean percentage BOLD signal change during the presentation of evaluation-relevant versus evaluation-irrelevant person-descriptive sentences. The differential score was calculated by subtracting evaluation-irrelevant from evaluation-relevant responses, so positive scores correspond to stronger responses to the evaluation-relevant information. A significant differential responding (two-tailed  $t$  tests) was shown by the PCC and the amygdala, but not by the dmPFC ( $P < 0.05$ ). Error bars indicate s.e.

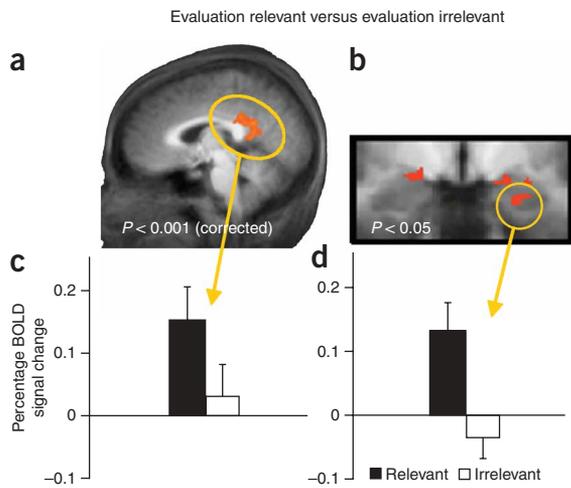
To confirm these findings, we conducted an additional analysis using a contrast directly comparing BOLD responses during the presentation of evaluation-relevant versus evaluation-irrelevant person-descriptive sentences. As expected, this contrast revealed only the PCC ( $P < 0.001$  corrected; **Fig. 3a**), the amygdala (**Fig. 3b**) and the thalamus (see **Supplementary Results**). For the latter two regions, we used a more liberal threshold ( $P < 0.05$  and  $P < 0.005$  uncorrected,

**Table 1** Talairach coordinates of regions extracted from the face with person-descriptive information  $>$  face alone contrast (false discovery rate  $< 0.05$ )

Region	Side	Coordinates				Volume (mm <sup>3</sup> )
		x	y	z	BA	
PCC*	M	0	-51	23	23	760
Amygdala*	L	-23	-8	-16	-	111
Thalamus*	L	-7	-13	12	-	333
Thalamus*	R	9	-13	10	-	291
Caudate	R	9	5	10	-	170
Caudate	L	-9	3	12	-	181
Hippocampus	R	27	-11	-13	-	125
dmPFC	L	-7	52	37	9	2,930
dmPFC	L	-9	26	55	8	2,997
dmPFC	R	11	32	52	8	1,584
Superior frontal gyrus	L	-4	-1	61	6	2,128
Ventromedial prefrontal cortex	M	0	48	-8	10	2,931
Inferior frontal gyrus	L	-43	24	-1	47	2,426
Inferior frontal gyrus	R	44	26	-6	47	2,917
Middle frontal gyrus	L	-46	10	27	8	4,478
Middle frontal gyrus	R	40	5	30	6	1,184
Superior temporal gyrus	L	-46	14	-14	38	4,291
Superior temporal gyrus	R	52	12	-13	38	3,833
Superior temporal gyrus (posterior)	L	-50	-57	16	21/22	2,615
Middle temporal gyrus (anterior)	R	53	-2	-18	21	4,712
Middle temporal gyrus (anterior)	L	-52	-2	-18	21	5,330
Middle temporal gyrus (posterior)	L	-54	-29	1	21	4,293
Middle temporal gyrus (posterior)	R	52	-37	4	21	2,537
Precuneus	L	-45	-7	47	4	5,404
Fusiform gyrus	L	-38	-48	-15	20/37	6,371
Lingual gyrus	R	21	-58	-1	37/19	2,414
Lingual gyrus	L	-17	-58	-1	19	2,585
Cerebellum	M	2	-64	-25	-	4,631

BA, Brodmann area; L, left; M, middle; R, right.

\*Regions showing the difference in evaluation effect (two-tailed  $t$  tests,  $P < 0.05$ ).



**Figure 3** Brain regions revealed by the evaluation-relevant versus evaluation-irrelevant contrast. **(a)** PCC BOLD responses ( $x = 0, y = -51, z = 24; P < 0.001$  corrected). **(b)** Amygdala BOLD responses ( $x = -23, y = -8, z = -16; P < 0.05$  uncorrected). **(c,d)** Mean percentage BOLD signal change extracted from all PCC ( $2,187 \text{ mm}^3$ ) and amygdala ( $108 \text{ mm}^3$ ) voxels. The mean BOLD response for each trial type is shown: evaluation relevant (black bars) and evaluation irrelevant (orange bars). Error bars indicate s.e.

respectively; minimal cluster size  $> 100 \text{ mm}^3$ ). No other regions of activation were observed, even at the most liberal threshold ( $P < 0.05$  uncorrected). We determined the mean and variance of the PCC and amygdala BOLD responses (**Fig. 3c,d**). Finally, no regions were revealed by the opposite contrast (evaluation irrelevant versus evaluation relevant), even at a more liberal threshold ( $P < 0.05$  uncorrected).

The categorical comparison between the evaluation-relevant and evaluation-irrelevant person-descriptive information is predicated on the idea that the overall evaluation involves either the positive or negative information alone. However, it might also reflect the weighted sum of both positive and negative information. To capture this, we conducted a parametric analysis, correlating the level of subjects' evaluations (ranging from 1 to 8) with PCC and amygdala mean BOLD responses (**Fig. 4**). We found that the mean BOLD response was stronger during the encoding of negative information for subjects giving negative evaluations (**Fig. 4a,c**) and was stronger during the encoding of positive information for subjects giving positive evaluations (**Fig. 4b,d**). These data suggest that PCC and amygdala responses reflect the integration of positive and negative information into an overall impression. Consistent with our behavioral findings, in which no valence bias was observed, these BOLD responses also do not map

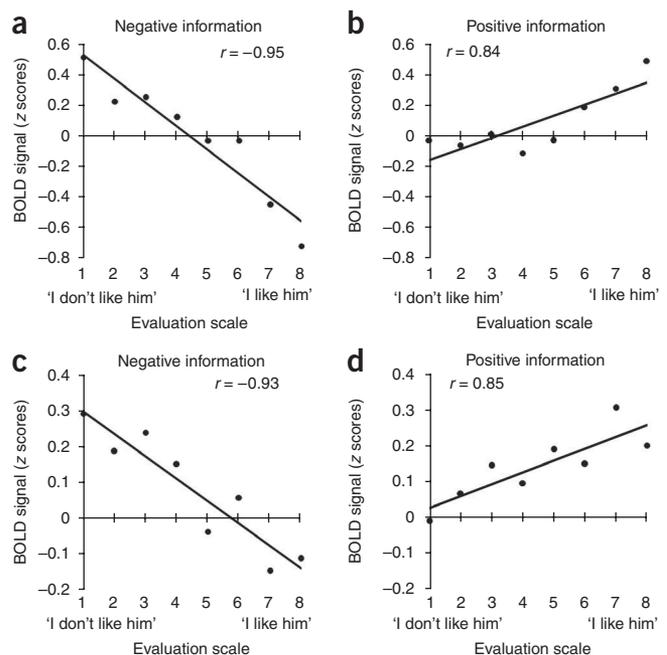
**Figure 4** BOLD signal in PCC and amygdala correlates with strength of evaluation. **(a–d)** Correlations between mean normalized PCC **(a,b)** and amygdala **(c,d)** BOLD signal (extracted from the evaluation relevant versus evaluation irrelevant contrast) and level of evaluation ranging from 1 ('I don't like him') to 8 ('I like him') are presented. For both regions, a more negative evaluation led to a stronger mean BOLD response during the encoding of negative person-descriptive information **(a,c)** and a more positive evaluation led to a stronger mean BOLD response during the encoding of positive person-descriptive information **(b,d)**. This was supported by a significant correlation between the mean BOLD response and level of evaluation. Because lower scores on the evaluation scale represent more negative evaluations, the correlation was negative (PCC,  $r = -0.95$ , **a**; amygdala,  $r = -0.93$ , **c**;  $P < 0.01$ ) during negative information trials and positive during positive information trials (PCC,  $r = 0.84$ , **b**; amygdala,  $r = 0.85$ , **d**;  $P < 0.01$ ).

onto the valence of the information *per se*. Instead, these regions encode social information that is subjectively meaningful and more heavily weighted in later evaluations. Thus, both positive and negative person-descriptive information receives more processing as long as it is relevant for subsequent evaluations (see **Supplementary Fig. 2** and **Supplementary Tables 1** and **2**).

## DISCUSSION

We provide new evidence for the manner in which social information is encoded in the brain to form impressions of others. When required to rapidly judge others, we appear to be efficient evaluators. We sift through available social information, weighting what matters to us. Sorting information by level of importance is a matter of subjective preference. Our results show where this sorting occurs in the brain, suggesting a neural mechanism by which impressions are formed. Subjects regarded different segments of person-descriptive information as being relevant or irrelevant for their subsequent evaluations. The idiosyncratic basis for this differential relevance might be factors such as personal experience, social values and pet peeves among others, which could make particular items more salient in the eyes of the beholder. First impressions, therefore, are tightly connected with the enduring biases subjects bring along. Such biases shape how subjects weight different types of information and which information is selected for additional processing. Even though these factors may vary widely between subjects, across subjects, the same brain regions, the PCC, the amygdala and the thalamus, dissociated these two types of information. The extent to which these regions were recruited during encoding of person-descriptive information correlated with how subjects valued it, as was evident in their subsequent evaluation scores.

It is important to note that the differential encoding on the basis of subsequent evaluation effect that we report here is dissociable from the previously reported effect of differential encoding on the basis of subsequent memory. Behaviorally, evaluation-relevant items were not remembered better than irrelevant ones. Consistent with this, encoding of evaluation-relevant information did not selectively engage memory related areas such as the hippocampus or the dorsolateral prefrontal



cortex. It appears that the relevance of particular information for subsequent evaluations does not confer a mnemonic advantage (as would be shown by a DM effect). One might expect information that is valued more to be better remembered. However, these results are not surprising given previous research showing that individuals with amnesia develop impressions of social others<sup>23</sup> and show attitude change<sup>24</sup> without having recollection of the information on which these affective judgments were based. Normal subjects have long been known to show such dissociation in rapid on-line impression formation<sup>25</sup>, as well as in judgments made a few months after person-descriptive information was introduced<sup>26</sup>. Taken together, these data suggest that episodic memory is probably not a mediating factor in the use of information for subsequent evaluations.

Understanding the neural substrates of social cognition has been one of the core motivations driving the burgeoning field of social neuroscience. A number of studies have highlighted the dmPFC in the processing of social information<sup>6–9</sup>. Our results provide further evidence that the dmPFC is recruited to process person-descriptive information during impression formation. However, BOLD responses in this region do not dissociate evaluation-relevant from evaluation-irrelevant information, suggesting that the dmPFC is not essential for the evaluative component of impression formation. In fact, social evaluation recruits brain regions that are not socially specialized<sup>27</sup> but are more generally involved in valuation and emotional processes.

Valuation and emotional processes, as a substantial amount of research has shown, are characteristic of the amygdala. In particular, the amygdala is considered to be a crucial region in learning about motivationally important stimuli<sup>10–12</sup>. It is also implicated in social inferences that are based on facial and bodily expressions<sup>15–17</sup>, in inferences of trustworthiness<sup>13,14</sup> and in the capacity to infer social attributes<sup>28</sup>. Moreover, the involvement of amygdala in social inferences might be independent of awareness or explicit memory. For example, increased amygdala responses were correlated with implicit, but not explicit, measures of the race bias<sup>19</sup>, as well as with presentation of faces previously presented in an emotional, but not neutral, context, regardless of whether subjects could explicitly retrieve this information<sup>26</sup>. Here we provide evidence linking the two domains of affective learning and social processing by showing that the amygdala is engaged in the formation of subjective value assigned to another person in a social encounter.

Although the amygdala is typically implicated in the processing of negative affect and negative stimuli have been shown to modulate it more than positive stimuli<sup>29</sup>, we found that the amygdala processed both positive and negative evaluation-relevant information, suggesting that amygdala activity is driven by factors other than mere valence, such as the motivational importance or salience of the stimuli. This result is consistent with recent findings<sup>30,31</sup> showing enhanced amygdala responses for both positive and negative stimuli as a function of motivational importance.

Evidence related to the PCC has been more diverse. There have been reports in the social domain, such as involvement in theory of mind<sup>32</sup> and self-referential outward-focused thought<sup>33</sup>, in memory related processes such as autobiographical memory of family and friends<sup>34</sup>, and in emotional modulation of memory<sup>35</sup> and attention<sup>36</sup>. More recently, the PCC has been linked with economic decision making, the assignment of subjective value to rewards<sup>37</sup> under risk and uncertainty<sup>38</sup>, and credit assignment in a social exchange<sup>39</sup>. A common denominator of these studies might be that all involved either a social or an outward-directed valuation component. Our task also encompasses these features, extending

the role of the PCC to value assignment to social information guiding our first impressions of others.

The amygdala and the PCC are both interconnected with the thalamus as part of a larger circuitry that is implicated in emotion, arousal and learning<sup>40</sup>. Beyond the known role of the amygdala and the PCC in social-information processing and value representation, our results suggest a neural mechanism underlying the online formation of first impressions. When encoding everyday social information during a social encounter, these regions sort information on the basis of its personal and subjective importance and summarize it into an ultimate score, a first impression. Other regions, such as the ventromedial PFC, the striatum and the insula, have also been implicated in valuation processes<sup>41–45</sup>. However, these regions did not emerge in our difference in evaluation effect analysis. This might suggest a possible dissociation in the valuation network between regions engaged in the formation of value and its subsequent representation and updating. The latter regions would not be engaged during encoding and therefore would not show a difference in evaluation effect but would instead have an effect once the evaluation is formed. The amygdala and the PCC probably participate in both value formation and its representation. The difference in evaluation procedure may provide a useful tool for disentangling the different components of the valuation system and their specific contributions to social versus nonsocial evaluations.

In sum, the complexity of social evaluation in forming a first impression is evident in the recruitment of multiple brain systems that are involved in social-information processing, emotion and valuation. Although it has been suggested that some neural systems are specialized for social-information processing, it seems that when it comes to evaluating others, these systems are not enough. Additional regions specialized in affective processing are recruited to evaluate other people in our initial encounters with them, providing a neural signature of first impressions.

## METHODS

**Subjects.** We recruited 19 right-handed normal volunteers (12 males) between 18 and 31 years of age (mean = 22.68, s.d. = 4.57) for the fMRI evaluation task. The experiment was approved by the New York University Committee on Activities Involving Human Subjects. All subjects gave informed consent and were paid for their participation.

**Stimuli.** We constructed 20 person profiles using 120 person-descriptive sentences implying different personality traits (for example, considerate: “He promised not to smoke in his apartment since his roommate was trying to quit.”)<sup>6</sup>. These sentences were pretested ( $n = 30$ ) for valence (1 = very negative and 8 = very positive) and arousal (1 = not arousing and 8 = very arousing). Each profile consisted of six unique sentences, arranged sequentially from either negative to positive or positive to negative. The valence transition was gradual according to the mean valence ratings from the pretest in the following way (for negative to positive profiles): the first was very negative ( $1.00 < \text{mean} < 2.74$ ), the second and third were moderately negative ( $2.75 < M < 4.49$ ), the fourth and fifth were moderately positive ( $4.50 < M < 6.24$ ), and the sixth was very positive ( $6.25 < M < 8.00$ ). This order was reversed for positive-to-negative profiles. Two-tailed  $t$  tests were conducted between mean arousal ratings of positive and negative sentences in each profile to ensure that there were no significant differences ( $P > 0.1$  for all profiles). We paired each profile with a picture of a monochrome male face of neutral expression, all of which were taken with identical lighting source and camera angle (Extended Yale Database B).

**Procedure.** Participants were told that they would see information about different people and would be asked to give their impressions of them. We presented 20 person profiles in one of four orders, counterbalanced across subjects. These four sets were designed to control for order of profile presentation, order of sentences in each profile and assignment of faces to

profiles. None of the sentences or faces repeated in each order. At the onset of each profile, the first, second and third sentences were presented with the face for 6 s each (18 s total). Following these was a 12-s interstimulus interval displaying only the face. Next, the fourth, fifth and sixth sentences were presented with the face for 6 s each (18 s total). Another face-only interstimulus interval ensued for 12 s. Subsequently, a feedback slide with the face still being apparent asked subjects to answer the question, 'What is your impression of him?' on a Likert-type scale between 1 ('I don't like him') and 8 ('I like him'). After subjects made a button response, they were requested to confirm their choice. Following confirmation, a 'Thank you' slide was displayed for the remainder of the 12 s from the onset of the feedback slide (that is, feedback slide and confirmation slide and 'Thank you' slide = 12 s). Finally, an intertrial interval (ITI) with fixation point was displayed for 12 s, after which the next profile was presented.

After the fMRI session, subjects participated in a memory task that consisted of 40 trials in one of two orders. In each trial, five faces, which were previously encountered during the fMRI task, were coupled with one sentence presented with one of these five faces during the fMRI task. Subjects had to select which of the five faces had been described with this sentence. These were sampled from all 20 profiles, with equal representation from the positive, negative, first and last segments.

**Evaluation analysis.** To verify that subjects' evaluations were guided by their individual subjective preferences rather than by any other effect, we probed the data for valence bias (more positive than negative evaluations or vice versa), order bias (primacy or recency effects; that is, more evaluations with the same valence as the segment presented first (first, second and third sentences) compared with last (fourth, fifth and sixth sentences) or vice versa). We also examined a face bias (that is, whether particular faces led to more negative or positive evaluations).

Subjects' responses on the 8-point scale were coded as being negative (1 to 4) or positive (5 to 8). To test whether subjects were biased by valence in their evaluations, we compared the proportion of negative evaluations to the proportion that would be measured if there were zero bias (that is, 0.5) using a one-sample two-tailed *t* test for the group analysis and binomial tests for individual subject analyses. Similarly, to test whether subjects were biased by order, we compared the proportion of evaluations with the same valence as the segment presented first to the proportion that would be measured if there were zero bias (that is, 0.5) using a one-sample two-tailed *t* test for the group analysis and binomial tests for individual subject analyses. To test whether a particular face led to more negative or more positive evaluations, binomial tests comparing the proportion of negative evaluations to the proportion that would be measured if there were zero bias (that is, 0.5) were conducted for each face. An alpha level of 0.05 was set for all statistical comparisons.

**Memory analysis.** To verify that the different types of information did not affect memory performance, we examined the memory data for valence bias (negative sentences were remembered better than positive or vice versa), order bias (sentences presented in the first segment (first, second or third presentations) were remembered better than those presented in the last segment (fourth, fifth or sixth presentations) in profile (primacy effect) or vice versa (recency effect)) and evaluation-relevance bias (evaluation-relevant sentences (having the same valence as the subsequent evaluation) were remembered better than evaluation-irrelevant sentences (with valences different from the subsequent evaluation)).

To test for these memory biases, we used a recognition accuracy measure that was defined as correct recognition responses divided by the total number of sentences in category. For group analyses, we compared mean recognition accuracy in the two corresponding categories of each bias test using paired two-tailed *t* tests. For individual subject analyses, numerical counts corresponding with the two categories were compared using  $\chi^2$  tests. Finally, overall recognition success was tested using a one-sample *t* test comparing the proportion of correct recognition responses to chance level. An alpha level of 0.05 was set for all statistical comparisons.

**fMRI acquisition.** A 3T Siemens Allegra head-only scanner and Siemens standard head coil were used for data acquisition. Anatomical images were

acquired using a T1-weighted protocol (256 × 256 matrix, 176 1-mm sagittal slices). Functional images were acquired using a single-shot gradient echo EPI sequence (repetition time, 2.0 s; echo time, 25 ms; field of view, 192 cm, flip angle = 75°, bandwidth = 4,340 Hz px<sup>-1</sup> and echo spacing = 0.29 ms). We obtained 39 contiguous oblique-axial slices (3 × 3 × 3-mm voxels) parallel to the anterior commissure–posterior commissure line. Analysis of the imaging data was conducted using BrainVoyager QX software package (Brain Innovation). Functional imaging data preprocessing included motion correction, slice scan time correction (using sinc interpolation), spatial smoothing using a three-dimensional Gaussian filter (4-mm full width at half maximum), and voxel-wise linear detrending and high-pass filtering of frequencies (above three cycles per time course). The structural and functional data of each participant were transformed to standard Talairach stereotaxic space<sup>46</sup>.

**fMRI analysis.** A random-effects general linear model analysis was conducted on the fMRI signal during the evaluation task with the following predictors: evaluation-relevant person-descriptive information, evaluation-irrelevant person-descriptive information, face-alone after first information segment, face-alone after second information segment and evaluative response. The predictors were convolved with a standard canonical hemodynamic response function.

**Contrast analyses.** The primary contrast of interest was faces with person descriptive sentences versus face alone presentations. On the basis of this contrast, we defined the ROIs on which to examine the difference in evaluation effect (see below). Regions on the statistical map showing a significant response (clusters of at least 100-mm<sup>3</sup> contiguous voxels whose false discovery rate was < 0.05) are detailed in **Table 1**.

To validate the primary analysis, the secondary contrast of interest was evaluation-relevant versus evaluation-irrelevant person-descriptive information, which directly explored the difference in evaluation effect (that is, regions showing neural differentiation of encoding on the basis of subsequent evaluations). We expected to reveal similar regions in these two independent analyses. Calculation of significant values in the activation map revealed by this contrast (**Fig. 3a**) was based on the individual voxel significance ( $P < 0.001$  corrected) and on the minimum cluster size of 890-mm<sup>3</sup> voxels. The probability of a false positive was determined from the frequency count of cluster sizes in the entire brain using a Monte Carlo simulation<sup>47</sup>. Because this analysis revealed only the region of the PCC, we used a gradually more liberal threshold to reveal additional regions (minimal cluster size of 100-mm<sup>3</sup> voxels,  $P < 0.005$  uncorrected for thalamus; **Supplementary Fig. 1**;  $P < 0.05$  uncorrected for amygdala; **Fig. 3b**). We extracted the mean percentage BOLD signal change at peak activation from the PCC (**Fig. 3c**) and the amygdala (**Fig. 3d**) for evaluation-relevant versus evaluation-irrelevant person-descriptive sentences.

**Difference in evaluation effect analysis.** A comparison between responses to evaluation-relevant and evaluation-irrelevant person-descriptive information was conducted on BOLD responses in ROIs revealed by the face with descriptive information versus face contrast (false discovery rate < 0.05; **Table 1**). For each ROI, we used paired two-tailed *t* tests comparing the mean percentage BOLD signal change from baseline (12-s ITI where a fixation point was presented between profile presentations) of the evaluation-relevant versus the evaluation-irrelevant trials. Each trial consisted of the mean of the nine BOLD measurements (repetition time = 2 s), during which the descriptive information was presented, and which were averaged across 20 trials in each condition and across 19 subjects. The alpha level for statistical comparisons was set at 0.05.

**Correlation analysis.** Correlations were computed between mean normalized BOLD signals and levels of evaluation ranging from 1 to 8. fMRI responses were extracted from the evaluation-relevant versus evaluation-irrelevant contrast for PCC ( $x = 0, y = -51, z = 24$ ;  $P < 0.001$  corrected, cluster size of 2,187-mm<sup>3</sup> voxels) and amygdala ( $x = -23, y = -8, z = -16$ ;  $P < 0.05$  uncorrected, cluster size of 108-mm<sup>3</sup> voxels). The fMRI signal was averaged across 19 subjects at each level of evaluation. The correlations were calculated separately for fMRI responses during the presentation of negative and positive information.

*Note: Supplementary information is available on the Nature Neuroscience website.*

## ACKNOWLEDGMENTS

We thank I. Levy and D. Amodio for fruitful discussions and comments, C. Raio for assistance with data collection, and K. Sanzenbach and the Center for Brain Imaging at New York University for technical assistance. This study was funded by a Seaver Foundation grant to the Center for Brain Imaging, a James S. McDonnell Foundation grant to E.A.P. and a Fulbright award to D.S.

## AUTHOR CONTRIBUTIONS

D.S. designed the experiments, collected and analyzed data, interpreted the data, and wrote the first draft of the manuscript. J.B.F. contributed to data collection, analysis and interpretation, and the final version of the manuscript. J.P.M., J.S.U. and E.A.P. contributed to experimental design, data interpretation and the final version of the manuscript.

Published online at <http://www.nature.com/natureneuroscience/>

Reprints and permissions information is available online at <http://www.nature.com/reprintsandpermissions/>

- Ambady, N. & Rosenthal, R. Half a minute: predicting teacher evaluations from thin slices of behavior and physical attractiveness. *J. Pers. Soc. Psychol.* **64**, 431–441 (1993).
- Uleman, J.S., Blader, S. & Todorov, A. Implicit impressions. in *The New Unconscious* (eds. Hassin, R. Uleman, J.S. & Bargh, J.A.) 362–392 (Oxford University Press, New York, 2005).
- Paller, K.A., Kutas, M. & Mayes, A.R. Neural correlates of encoding in an incidental learning paradigm. *Electroencephalogr. Clin. Neurophysiol.* **67**, 360–371 (1987).
- Brewer, J.B. *et al.* Making memories: brain activity that predicts how well visual experience will be remembered. *Science* **281**, 1185–1187 (1998).
- Wagner, A.D. *et al.* Building memories: remembering and forgetting of verbal experiences as predicted by brain activity. *Science* **281**, 1188–1191 (1998).
- Mitchell, J.P., Macrae, C.N. & Banaji, M.R. Encoding-specific effects of social cognition on the neural correlates of subsequent memory. *J. Neurosci.* **24**, 4912–4917 (2004).
- Mitchell, J.P., Heatherton, T.F. & Macrae, C.N. Distinct neural systems subserve person and object knowledge. *Proc. Natl. Acad. Sci. USA* **99**, 15238–15243 (2002).
- Amodio, D.M. & Frith, C.D. Meeting of minds: the medial frontal cortex and social cognition. *Nat. Rev. Neurosci.* **7**, 268–277 (2006).
- Mitchell, J.P., Banaji, M.R. & Macrae, C.N. The link between social cognition and self-referential thought in the medial prefrontal cortex. *J. Cogn. Neurosci.* **17**, 1306–1315 (2005).
- Phelps, E.A. & LeDoux, J.E. Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron* **48**, 175–187 (2005).
- Everitt, B.J. & Robbins, T.W. Amygdala-ventral striatal interactions and reward-related processes. in *The Amygdala: Neurobiological Aspects of Emotion, Memory and Mental Dysfunction* (ed. Aggleton J.P.) 401–430 (New York, Wiley-Liss, 1992).
- LeDoux, J.E. Emotion circuits in the brain. *Annu. Rev. Neurosci.* **23**, 155–184 (2000).
- Engell, A.D., Haxby, J.V. & Todorov, A. Implicit trustworthiness decisions: automatic coding of face properties in the human amygdala. *J. Cogn. Neurosci.* **19**, 1508–1519 (2007).
- Winston, J.S., Strange, B.A., O'Doherty, J. & Dolan, R.J. Automatic and intentional brain responses during evaluation of trustworthiness of faces. *Nat. Neurosci.* **5**, 277–283 (2002).
- Adolphs, R., Tranel, D., Damasio, H. & Damasio, A. Impaired recognition of emotion in facial expressions following bilateral damage to the human amygdala. *Nature* **372**, 669–672 (1994).
- Adolphs, R. Recognizing emotion from facial expressions: psychological and neurological mechanisms. *Behav. Cogn. Neurosci. Rev.* **1**, 21–62 (2002).
- Hadjikhani, N. & de Gelder, B. Seeing fearful body expressions activates the fusiform cortex and amygdala. *Curr. Biol.* **13**, 2201–2205 (2003).
- Cunningham, W.A. *et al.* Separable neural components in the processing of black and white faces. *Psychol. Sci.* **15**, 806–813 (2004).
- Phelps, E.A. *et al.* Performance on indirect measures of race evaluation predicts amygdala activation. *J. Cogn. Neurosci.* **12**, 729–738 (2000).
- Asch, S.E. Forming impression of personality. *J. Abnorm. Soc. Psychol.* **41**, 258–290 (1946).
- Jones, E.E. & Goethals, G.R. Order effects in impression formation. in *Attribution: Perceiving the Causes of Behavior*. (eds. Jones, E.E., Kanouse, D.E., Kelley, H.H., Nisbett, R.E., Valins, S., & Weiner, B.) 27–46 (New Jersey, Lawrence Erlbaum Associates, 1987).
- Anderson, N.H. & Barrios, A.A. Primacy effects in personality impression formation. *J. Abnorm. Soc. Psychol.* **63**, 346–350 (1961).
- Johnson, M.K., Kim, J.K. & Risse, G. Do alcoholic Korsakoffs syndrome patients acquire affective reactions? *J. Exp. Psychol. Learn. Mem. Cogn.* **11**, 22–36 (1985).
- Lieberman, M.D., Ochsner, K.N., Gilbert, D.T. & Schacter, D.L. Do amnesics exhibit cognitive dissonance reduction? The role of explicit memory and attention in attitude change. *Psychol. Sci.* **12**, 135–140 (2001).
- Hastie, R. & Park, B. The relationship between memory and judgment depends on whether the judgment task is memory-based or on-line. *Psychol. Rev.* **93**, 258–268 (1986).
- Somerville, L.H., Wig, G.S., Whalen, P.J. & Kelley, W.M. Dissociable medial temporal lobe contributions to social memory. *J. Cogn. Neurosci.* **18**, 1253–1265 (2006).
- Frith, C.D. The social brain? *Philos. Trans. R. Soc. Lond. B Biol. Sci.* **362**, 671–678 (2007).
- Heberlein, A.S. & Adolphs, R. Impaired spontaneous anthropomorphizing despite intact perception and social knowledge. *Proc. Natl. Acad. Sci. USA* **101**, 7487–7491 (2004).
- Kim, H. *et al.* Contextual modulation of amygdala responsivity to surprised faces. *J. Cogn. Neurosci.* **16**, 1730–1745 (2004).
- Cunningham, W.A., Van Bavel, J.J. & Johnsen, I.R. Affective flexibility: evaluative processing goals shape amygdala activity. *Psychol. Sci.* **19**, 152–160 (2008).
- Todorov, A., Baron, S.G. & Oosterhof, N.N. Evaluating face trustworthiness: a model-based approach. *Soc. Cogn. Affect. Neurosci.* **3**, 119–127 (2008).
- Fletcher, P.C. *et al.* Other minds in the brain: a functional imaging study of “theory of mind” in story comprehension. *Cognition* **57**, 109–128 (1995).
- Johnson, M.K. *et al.* Dissociating medial frontal and posterior cingulate activity during self-reflection. *Soc. Cogn. Affect. Neurosci.* **1**, 56–64 (2006).
- Maddock, R.J., Garrett, A.S. & Buonocore, M.H. Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience* **104**, 667–676 (2001).
- Maddock, R.J., Garrett, A.S. & Buonocore, M.H. Posterior cingulate cortex activation by emotional words: fMRI evidence from a valence decision task. *Hum. Brain Mapp.* **18**, 30–41 (2003).
- Pessoa, L. & Padmala, S. Quantitative prediction of perceptual decisions during near-threshold fear detection. *Proc. Natl. Acad. Sci. USA* **102**, 5612–5617 (2005).
- Kable, J.W. & Glimcher, P.W. The neural correlates of subjective value during intertemporal choice. *Nat. Neurosci.* **10**, 1625–1633 (2007).
- McCoy, A.N. & Platt, M.L. Risk-sensitive neurons in the macaque cingulate cortex. *Nat. Neurosci.* **8**, 1220–1227 (2005).
- Tomlin, D. *et al.* Agent-specific responses in the cingulate cortex during economic exchanges. *Science* **312**, 1047–1050 (2006).
- Taber, K.H., Wen, C., Khan, A. & Hurlley, R.A. The limbic thalamus. *J. Neuropsychiatry Clin. Neurosci.* **16**, 127–132 (2004).
- O'Doherty, J.P. Reward representations and reward-related learning in the human brain: insights from neuroimaging. *Curr. Opin. Neurobiol.* **14**, 769–776 (2004).
- Schiller, D., Levy, I., Niv, Y., LeDoux, J.E. & Phelps, E.A. From fear to safety and back—reversal of fear in the human brain. *J. Neurosci.* **28**, 11517–11525 (2008).
- Delgado, M.R., Li, J., Schiller, D. & Phelps, E.A. The role of striatum in aversive learning and aversive prediction errors. *Phil. Trans. R. Soc. Lond. B* **363**, 3787–3800 (2008).
- Seymour, B. & McClure, S.M. Anchors, scales and the relative coding of value in the brain. *Curr. Opin. Neurobiol.* **18**, 173–178 (2008).
- Knutson, B., Rick, S., Wimmer, G.E., Prelec, D. & Loewenstein, G. Neural predictors of purchases. *Neuron* **53**, 147–156 (2007).
- Talairach, J. & Tournoux, P. *Co-planar Stereotaxic Atlas of the Human Brain: an Approach to Medical Cerebral Imaging* (Thieme, New York, 1988).
- Forman, S.D. *et al.* Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* **33**, 636–647 (1995).