

# The social evaluation of faces: a meta-analysis of functional neuroimaging studies

Peter Mende-Siedlecki,<sup>1</sup> Christopher P. Said,<sup>2</sup> and Alexander Todorov<sup>1,3</sup>

<sup>1</sup>Department of Psychology, Princeton University, Princeton, NJ, USA, <sup>2</sup>Department of Psychology and Center for Neural Science, New York University, New York, NY, USA, and <sup>3</sup>Department of Psychology, Behavioural Science Institute, Radboud University, Nijmegen, The Netherlands

**Neuroscience research on the social evaluation of faces has accumulated over the last decade, yielding divergent results. We used a meta-analytic technique, multi-level kernel density analysis (MKDA), to analyze 29 neuroimaging studies on face evaluation. Across negative face evaluations, we observed the most consistent activations in bilateral amygdala. Across positive face evaluations, we observed the most consistent activations in medial prefrontal cortex, pregenual anterior cingulate cortex (pgACC), medial orbitofrontal cortex (mOFC), left caudate and nucleus accumbens (NAcc). Based on additional analyses comparing linear and non-linear responses, we propose a ventral/dorsal dissociation within the amygdala, wherein separate populations of neurons code for face valence and intensity, respectively. Finally, we argue that some of the differences between studies are attributable to differences in the typicality of face stimuli. Specifically, extremely attractive faces are more likely to elicit responses in NAcc/caudate and mOFC.**

**Keywords:** meta-analysis; fMRI; social cognition; face perception; face evaluation; trait judgments; trustworthiness; attractiveness

## INTRODUCTION

Within a single glance of a face, people automatically appraise face attractiveness and make a host of social attributions (Olson and Marshuetz, 2005; Bar *et al.*, 2006; Willis and Todorov, 2006; Rule *et al.*, 2009; Todorov *et al.*, 2009). For example, 33-ms exposure to a face is sufficient for people to make trustworthiness decisions (Todorov *et al.*, 2009). Additional time exposure simply increases confidence in these decisions (Willis and Todorov, 2006). As one of the founding fathers of modern social psychology, Solomon Asch (1948, p. 258), put it, 'We look at a person and immediately a certain impression of his character forms itself in us. A glance, a few spoken words are sufficient to tell us a story about a highly complex matter. We know that such impressions form with remarkable rapidity and with great ease. Subsequent observations may enrich or upset our view, but we can no more prevent its rapid growth than we can avoid perceiving a given visual object or hearing a melody'.

Recent research confirms Asch's insights (Zebrowitz, 1999; Todorov *et al.*, 2008a,b; Zebrowitz and Montepare, 2008). People rapidly and effortlessly form impressions from facial appearance. Although the validity of such impressions is low (Olivola and Todorov, 2010a), inferences of character and personality have been shown to predict important outcomes

in domains ranging from politics (Todorov *et al.*, 2005; Ballew and Todorov, 2007; Olivola and Todorov, 2010b), law (Zebrowitz and McDonald, 1991; Blair *et al.*, 2004; Eberhardt *et al.*, 2006), mate choice (Olivola *et al.*, unpublished data), business (Hamermesh and Biddle, 1994; Rule and Ambady, 2008) and the military (Mazur *et al.*, 1984).

Despite the importance of first impressions for social interactions, research on their neural basis is in its infancy. Researchers began to use social neuroscience methods to investigate this basis only a decade ago (Adolphs *et al.*, 1998; Nakamura *et al.*, 1998; Aharon *et al.*, 2001; Winston *et al.*, 2002). Although a number of neuroimaging studies have been published on the topic, many of the results have been inconsistent (Todorov *et al.*, 2011). The objective of this article is to provide a quantitative summary of the major findings across studies on face evaluation.

## The neural basis of face evaluation

Neuroimaging research on the social evaluation of faces has usually focused on evaluations along the trait dimensions of trustworthiness and attractiveness. Although these are separable dimensions, psychometric studies of social judgments from faces show that these judgments are highly inter-correlated with each other, with correlations ranging from 0.60 to 0.80 (Oosterhof and Todorov, 2008; Todorov *et al.*, 2008a,b). For example, principal components analyses show that (i) the first component, which indicates general face valence, accounts for >60% of the variance of judgments; and (ii) trustworthiness and attractiveness judgments are highly correlated with this valence component. Given these behavioral data, one would expect to observe

Received 21 January 2011; Accepted 27 November 2011

We thank Hedy Kober and Bryan Denny for assistance with the MKDA toolbox, Tor Wager for making the MKDA toolbox available online, and Jenny Porter and Katie Hansen, who provided the behavioral data for Figure 5B. We also thank Jamil Zaki, Adam Hampshire, and Leigh Nyström for their assistance in obtaining analyses not reported in their original papers. This research was supported by National Science Foundation grant BCS-0823749 (to A.T.), National Science Foundation grant DGE 1148900 (to P.M.-S.), and the Russell Sage Foundation.

Correspondence should be addressed to Peter Mende-Siedlecki. E-mail: pmende@princeton.edu

overlapping regions in neuroimaging studies on attractiveness and trustworthiness.

For the purposes of this meta-analysis, we focus on studies on attractiveness and trustworthiness. Typically, such studies present participants with facial stimuli that vary on the respective dimension—either systematically manipulated via computer modeling, or confirmed by independent behavioral ratings—and subsequently report brain activity that shows a linear relationship with changes in facial appearance along that dimension. For example, some studies have observed increased responses in the amygdala for untrustworthy faces (Winston *et al.*, 2002), while other studies have observed increased responses in the nucleus accumbens (NAcc) and medial orbitofrontal cortex (mOFC) for attractive faces (Aharon *et al.*, 2001; O’Doherty *et al.*, 2003). More recent studies have sought to identify regions that show a quadratic relationship between brain activity and changes in attractiveness or trustworthiness. Researchers have observed non-linear responses in the amygdala for both attractive and unattractive faces (Winston *et al.*, 2007), as well as for both trustworthy and untrustworthy faces (Todorov *et al.*, 2011). While there is convergence between the linear and non-linear approaches, there exists the possibility that these analyses are tapping distinct processes, wherein areas that show a linear pattern of activity are coding for face valence, while areas that show quadratic patterns are coding something more like face intensity.

The first objective of this article is to systematically explore the pattern of observed brain activations across published neuroimaging studies on face evaluation as a function of face valence. The second objective is to examine possible dissociations between linear and non-linear responses. The third and final objective is to explore potential differences between trustworthiness and attractiveness studies.

### Multilevel kernel density analysis

Meta-analysis is a powerful statistical tool that allows researchers to combine the data sets of a collection of similar studies to provide a more accurate, robust estimate of the effect-size of a given phenomena. This approach is widespread within behavioral research, and in recent years, meta-analyses of neuroimaging studies have become more common (Fox *et al.*, 1998; Phan *et al.*, 2002; Wager and Smith, 2003; Wager *et al.*, 2004; Laird *et al.*, 2005; Nielsen *et al.*, 2005). Meta-analyses of neuroimaging data typically compute how frequently studies examining a given psychological phenomenon report activity in a specific brain area (Kober and Wager, 2010). This approach can be used to confirm the prevailing thinking regarding what brain areas are associated with a particular psychological phenomenon or experience. At the same time, meta-analysis can serve a more exploratory purpose—identifying regions that are consistently activated across a large number of studies of the same psychological phenomenon, but that are not typically associated with that phenomenon.

Indeed, a meta-analysis of the social evaluation of faces has been recently published (Bzdok *et al.*, 2011), and in part, motivated the analyses herein. While we ultimately employed slightly different selection criteria in choosing studies to include, we also sought to perform several more targeted analyses, as noted above. Perhaps more importantly, while Bzdok and colleagues conducted an activation likelihood estimation (ALE) meta-analysis, we use a different statistical procedure.

Specifically, we use a Multi-level Kernel Density Analysis (MKDA), which represents an advance in meta-analytic methods for neuroimaging data, because it accounts for the fact that individual activation peaks are nested within contrast maps (maps of particular comparisons within studies), making these maps the unit of analysis, and not the peaks (Wager *et al.*, 2008). Further, MKDA models contrast maps as a random effect, eliminating the possibility of one contrast dominating the meta-analysis.

We conduct several analyses. First, we analyze activations across all contrasts showing (i) stronger brain responses to negative—untrustworthy and unattractive—than positive—trustworthy and attractive—faces; (ii) stronger responses to positive than negative faces; and (iii) stronger responses to positive and negative faces than to neutral faces. Second, within these contrasts, we also explore potential differences between trustworthiness and attractiveness studies.

## METHODS

### Data collection

We searched for neuroimaging studies of the social evaluation of faces using the online databases PsycINFO and PubMed, as well as the scholarly article search engine Google Scholar. We limited our search using combinations of keywords including ‘faces’, ‘social evaluation’, ‘social judgment’, ‘fMRI’, ‘trustworthiness’ and ‘attractiveness’. To be included in our meta-analysis, studies had to involve fMRI or PET investigations of healthy adults,<sup>1</sup> report activations in a standard coordinate system—either Talairach or Montreal Neurological Institute (MNI) coordinates, and explicitly state whether their analyses were performed with fixed or random effects. With respect to in-scanner tasks, we only included studies in which subjects either made explicit judgments regarding the trustworthiness or attractiveness of faces, or were presented with faces that varied on one of these two dimensions during an implicit or a passive viewing task, based upon normative ratings, computer modeling or some other form of categorization. In the case of some studies (Hampshire *et al.*, 2011; Pochon *et al.*, 2008; Zaki *et al.*, 2011), relevant contrasts were not originally reported, but were obtained through personal communication with the respective authors.

We excluded studies that did not report specific coordinates arising from relevant contrasts, but instead referred to

<sup>1</sup>Pinkham *et al.* (2008a) and Baas *et al.* (2008) were neuroimaging studies comparing the social evaluation of faces in patient populations to healthy controls. When composing our MKDA, we only included the coordinates yielded from the analysis of the healthy controls’ data.

**Table 1** Breakdown of Studies Included in the Meta-Analysis

Study	Included contrasts	N	Task nature <sup>a</sup>			Study type	Valence			ROI? <sup>c</sup>	Stimulus category
			Implicit	Explicit	Collapsed <sup>b</sup>		Negative	Positive	Non-Linear		
Aharon, et al., 2001	1	6	X			attractiveness		X			extreme
Aron, et al., 2005	1	17	X			attractiveness		X			average
Baas, et al., 2008	4	21	X	X	X	trustworthiness	X	X			average
Blasi, et al., 2009	1	43		X		trustworthiness	X				average
Bray, et al., 2007	2	25	X			attractiveness	X	X			extreme
Chatterjee, et al., 2009	4	13	X	X		attractiveness	X	X			computer generated
Cloutier, et al., 2008	2	48		X		attractiveness	X	X			extreme
Engell, et al., 2007	1	15	X			trustworthiness	X				average
Gordon, et al., 2009	2	6	X			trustworthiness		X	X	X	average
Hampshire, et al., 2011	1	19		X		attractiveness		X			extreme
Iaria, et al., 2008	1	10		X		attractiveness		X			extreme
Kampe, et al., 2001	1	16	X			attractiveness		X			average
Kim, et al., 2007	1	25		X		attractiveness		X			computer generated
Liang, et al., 2010	3	17	X			attractiveness	X	X	X	X	extreme
O'Doherty, et al., 2003	2	25	X			attractiveness	X	X			extreme
Pinkham, et al., 2008	1	12		X		trustworthiness	X			X	average
Pochon, et al., 2008	1	17		X		attractiveness		X			extreme
Said, et al., 2009	3	31		X		trustworthiness	X	X	X		average
Said, et al., 2010	1	24	X			trustworthiness			X		computer generated
Smith, et al., 2010	1	26		X		attractiveness		X			average
Todorov, et al., 2008	2	14	X			trustworthiness	X		X	X	computer generated
Todorov, et al., 2010	2	22		X		trustworthiness	X		X		computer generated
Todorov, et al., 2010	2	22	X			trustworthiness	X		X		computer generated
Tsukiura, et al., 2010a	3	22		X		attractiveness	X	X	X		average
Turk, et al., 2004	1	18		X		attractiveness		X			extreme
VanRijn, et al., 2011	3	18		X		trustworthiness	X	X		X	average
Winston, et al., 2002	2	14			X	trustworthiness	X	X			average
Winston, et al., 2007	2	26			X	attractiveness		X	X		extreme
Zaki, et al., 2011	1	14		X		attractiveness		X			extreme

<sup>a</sup>'Task nature' categorizes only the contrasts included in our meta-analysis. In some cases, like [Winston et al. \(2002\)](#), both explicit and implicit paradigms were employed, but only collapsed analyses were reported.

<sup>b</sup>'Collapsed' analyses refer to analyses in which neural activity was aggregated across both explicit and implicit tasks.

<sup>c</sup>In this column, we note if a given study reported coordinates arising from ROI-based analyses. In some cases, these studies *only* reported such ROI-based analyses (for instance, [Pinkham et al., 2008](#)). As such, these studies have only been able to impact our [supplementary analyses](#), which incorporate ROI-based analyses in addition to whole-brain contrasts.

various ROIs from a functional localizer being more or less active during specific contrasts ([Kranz and Ishai, 2006](#)). In some instances, multiple studies were found which presented analyses of the same data sets ([Todorov and Engell, 2008](#); [Pinkham et al., 2008b](#)). In these cases, we only included one study's reported coordinates, and this choice was made based upon which version of the study ultimately presented the more relevant analyses. Finally, we excluded some studies whose research questions bordered on ours (for instance, aesthetic judgments of paintings of faces, as in [Kawabata and Zeki \(2004\)](#) or neural responses to faces similar to the self varying in trustworthiness, as in [Verosky and Todorov \(2010\)](#) as they ultimately did not report contrasts that were appropriate for inclusion in our analyses. These choices are not trivial, as they represent some of the differences between our meta-analysis and the one conducted by [Bzdok and colleagues \(2011\)](#), in terms of study selection.

This search yielded 28 published papers comprising 29<sup>2</sup> neuroimaging studies on the social evaluation of faces. Seventeen of these studies were on attractiveness evaluations and 12 were on trustworthiness or related evaluations (i.e. 'would you approach or avoid this person'). The latter were included because such approach/avoidance evaluations are highly correlated with trustworthiness evaluations ([Todorov, 2008](#)). This set of studies accounted for 52 separate contrasts ([Table 1](#)). For contrasts to be included in our database, they had to be representative of neural activity that varied parametrically with either facial attractiveness or trustworthiness, and furthermore, the direction and linearity of this relationship had to be clearly stated. We excluded coordinates derived from complex interaction-based

<sup>2</sup>We consider a single study to represent an investigation of the neural responses to a given set of stimuli in the context of one or potentially multiple psychological tasks within the same set of subjects. Thus, [Todorov et al. \(2011\)](#) represents two separate studies, while [Chatterjee et al. \(2009\)](#)—in which the same subjects took part in explicit and implicit tasks in-scanner on separate scanning days—represents one study.

analyses (for instance, stimuli type and gender interactions, as seen in O'Doherty *et al.* (2003), as well as coordinates arising from analyses that were not relevant to our research questions (e.g. effects of face novelty in Kim *et al.* (2007)). Further, overlapping contrasts are often reported in the articles surveyed. For instance, Aharon and colleagues (2001) report separate contrasts detailing neural activity associated with facial attractiveness for male stimuli, female stimuli, and collapsed across both kinds of stimuli. In these cases, we only included the most general reported contrast—for instance, for Aharon *et al.* (2001), we used the collapsed contrast. Studies that report separate results for explicit and implicit paradigms presented a unique problem (see Winston *et al.*, 2002; Baas *et al.*, 2008; Chatterjee *et al.*, 2009). On the one hand, both analyses are relevant to our main research question, and favoring one paradigm over the other in these three cases would bias our results in favor of that task design. On the other hand, the contrasts are undeniably non-independent of each other. Ultimately, we chose to run our analyses using both sets of coordinates for these three studies, which were entered into our database as separate contrasts. To confirm that this approach had no demonstrable impact on our results, we ran complimentary analyses that only included one contrast per study (i.e. only the explicit task contrast from the three studies in question). We observed no practical differences in either the size or localization of consistent activations.

We tabulated the design particulars and parameters of each study, as well as the reported activation points for all relevant contrasts. Specifically, we coded each study in terms of which coordinate system activations were reported in, number of participants, whether a fixed or random effects analysis had been performed, whether activations represented linear or non-linear effects, whether the task was explicit or implicit in nature, and whether the reported activations were the result of a whole-brain or region-of-interest (ROI) analysis. This coding scheme served two purposes. Primarily, this information was fed into the MKDA toolbox and used to determine the proper weighting scheme for the different studies. Secondarily, it served as the basis for contrasting studies against each other on relevant variables. This coding scheme was initially entered by the first author, with subsequent confirmation and complete agreement from the second and third authors. Entered coordinates were checked and re-checked against their original sources numerous times throughout the course of setting up our database.

The studies compiled in our database used a variety of face stimuli. Some studies used computer-generated faces (for instance, Chatterjee *et al.*, 2009), others used standardized photograph sets of volunteer subjects (for instance, Engell *et al.*, 2007), and still others used photographs culled from magazines and newspapers (for instance, O'Doherty *et al.*, 2003). These faces likely differ in terms of their typicality—faces in standardized photographs are more typical

than the extreme faces seen in photographs of models and actresses. Given recent work suggesting that face typicality can partially account for the amygdala's response to the valence of face stimuli (Said *et al.*, 2010), it is possible that different types of face stimuli (e.g. extremely attractive faces that are less typical) could lead to different patterns of neural responses. As such, while we did not exclude studies based on the sources of the face stimuli, we did keep track of the source of each study's stimuli. This allowed for the possibility of comparing the more typical faces (computer-generated and standardized sets) against the more atypical faces (photos of models and actors from print media).

It is important to note that contrasts containing ROI-based analyses pose a problem for inclusion in meta-analyses. On the one hand, including coordinates from ROI-based analyses may bias the results by introducing researchers' a priori predictions about which regions are involved in trustworthiness and attractiveness evaluation. On the other hand, such analyses represent theoretically motivated prior research. Further, because some ROIs like the amygdala and NAcc are relatively small and often difficult to image, excluding ROI-based analyses may miss important findings that are consistent across studies. Given that, we chose to run each of our analyses twice—once limited to whole-brain contrasts, and once with ROI-based contrasts included. In the interest of space, we chose to report the whole-brain analyses in the main text, as well as to note whether or not adding ROI-based contrasts substantially affected the results. (In all cases but one, adding ROI-based contrasts did not have a substantial effect on analyses. For the contrast that produced divergent results, we chose to explicitly note in the text how the two approaches differed.) The specific results for the ROI-based analyses are reported in [supplemental material](#). We note that while some studies reported ROI-based analyses side by side with whole-brain analyses (for instance, Van Rijn *et al.*, 2011), there are a small number of studies that reported only ROI-based analyses (for instance, Pinkham *et al.*, 2008a).

### Data analysis

Our MKDA of 'negative' contrasts comprised all contrasts in which brain activity increases as facial stimuli *decrease* in either trustworthiness or attractiveness. 'Positive' contrasts comprised all contrasts in which brain activity increases as facial stimuli *increase* in either trustworthiness or attractiveness. Non-linear, quadratic contrasts comprised all contrasts in which brain activity increases as facial stimuli increase or decrease in either trustworthiness or attractiveness relative to faces at the middle of the continuum. Not all studies included in our database reported both negative and positive contrasts. Therefore, neither of our primary MKDAs contains contrasts from every study.

When performing these analyses, the peak coordinates from all relevant contrast maps were first separately convolved with a 10-mm spherical kernel, yielding comparison

indicator maps (CIMs). Previous meta-analytic work suggests that this is an appropriate default kernel size (Wager *et al.* 2003; Salimi-Khorshidi *et al.*, 2009). These CIMs were subsequently weighted based upon the number of participants and what type of analysis was performed in each study, following the same parameters used by Kober and colleagues (2008). Specifically, each map was first weighted by the square root of its study's sample size and subsequently multiplied by an adjustment factor accounting for the type of analysis used in the respective study. Random effects studies were multiplied by an adjustment factor of 1; fixed effects studies were multiplied by an adjustment factor of .75. In this fashion, studies received higher weighting if they had large sample sizes and performed random effects analyses.

Second, the weighted CIMs were averaged together, producing a density map. Each voxel of this density map attains a density statistic,  $P$ , which is the weighted proportion of contrasts included in the MKDA that yield activity within 10 mm of that voxel. To identify voxels whose  $P$ -statistic exceeds the frequency expected by chance, a Monte Carlo simulation was conducted. Over 5000 iterations of the Monte Carlo simulation, the observed activation blobs (contiguous regions of activation within the CIMs, holding shape constant) from each CIM were randomly shuffled within a gray-matter mask. Following each iteration, we recorded both the maximum whole brain density statistic ( $P$ , across all studies) and the largest cluster of contiguous voxels. These values were used to create null-hypothesis distributions for the density statistic and the expected size of clusters, respectively.

Third, the weighted  $P$  was subsequently tested against the resulting null-hypothesis  $P_0$  produced by the Monte Carlo simulation. A similar procedure was used to test for the significance of the size of the clusters, allowing us to identify a size threshold at which a certain number of voxels must be activated contiguously for a given cluster to be deemed significant. Hence, we used two types of thresholds—a density height-based threshold and a cluster size threshold, the latter derived from a non-parametric cluster-based thresholding procedure (Nichols and Hayasaka, 2003). For  $P$ , the resulting familywise error rate (FWER)-controlled threshold is the proportion of studies reporting activity within 10 mm of a given voxel that exceeds the maximum  $P$ -statistic across 95% of the resulting Monte Carlo maps. These voxels appear on resulting maps colored in yellow and will be referred to in our results as exceeding the height-based threshold of the MKDA. For the cluster size threshold, the resulting FWER-controlled threshold is the clusters observed at  $P < .001$  and  $P < 0.01$  whose size exceeds the maximum cluster size across 95% of the Monte Carlo maps. These voxels appear on resulting maps in orange and pink, respectively, and will be referred to as exceeding the extent-based threshold. The thresholded maps were overlaid on a canonical MRI image (colin27.img, the single-subject template in SPM5; <http://www.fil.ion.ucl.ac.uk/spm/software/spm5/>), which was co-registered to the MNI brain template.

When reporting areas of consistent activation in our tables, we provide information on whether each area withstood height-based thresholding, extent-based thresholding, or both. Some areas of activation were sizable enough to pass extent-based thresholding but not height-based thresholding. Conversely, other areas were highly consistent across the database and passed height-based thresholding, but were not sufficiently large to pass extent-based thresholding. XYZ-coordinates reported in our tables reflect the peak activation foci which withstand height-based thresholding, or, if activations are less consistent, the center of mass of the cluster at the most stringent level of extent-based thresholding. Further, we report the number of voxels in each cluster which withstood height-based thresholding, or if activations are less consistent, the number of voxels at the most stringent level of extent-based thresholding.

We also performed several smaller, more targeted MKDAs exploring differences between trustworthiness and attractiveness studies and performed several additional exploratory analyses based on stimulus typicality. To perform such analyses, a simple subtraction yields the relative difference in the distribution of peaks between the respective contrasts, which is subsequently thresholded as explained above.

## RESULTS

### Results across negative contrasts

Eleven studies reported 13 negative contrasts—where brain activity increased as attractiveness or trustworthiness decreased—across the whole brain. The MKDA results for these contrasts are summarized in Table 2 and Figure 1. We observed highly consistent activation in right amygdala (withstood height- and extent-thresholding,  $P < 0.001$ ), as well as activation in left amygdala that survived extent- but not height-thresholding ( $P < 0.001$ ). Four studies reported ROI-based coordinates for negative contrasts. When we included these additional coordinates in our analysis as well, we continued to observe highly consistent activation in right amygdala (withstood height- and extent-thresholding,  $P < 0.001$ ), as well as activations which survived extent- but not height-thresholding in left amygdala ( $P < 0.001$ ), right globus pallidus ( $P < 0.01$ ) and a large region of consistent activation encompassing right anterior insula, right inferior frontal gyrus (IFG) and right ventrolateral prefrontal cortex (vlPFC,  $P < 0.01$ , additional results summarized in Supplementary Table S1).

### Results across positive contrasts

Twenty-one studies reported 23 positive contrasts—where brain activity increased as attractiveness or trustworthiness increased—across the whole brain. The MKDA results for these contrasts are summarized in Table 2 and Figure 2. We observed highly consistent activation in left caudate extending into NAcc and mOFC, right thalamus, vmPFC and dACC/pgACC (withstood height- and extent-thresholding,  $P < 0.001$ ), as well as portions of right amygdala right

**Table 2** Areas Consistently Activated During Negative and Positive Face Evaluations

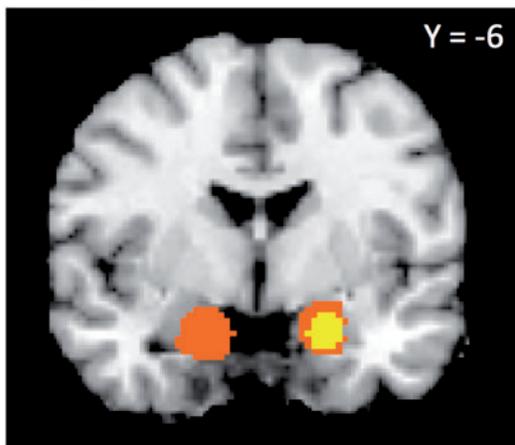
Region	Lat	x	y	z	Vol	%Act	
<i>Increased activity for negative evaluations, collapsed across Untrustworthiness and Unattractiveness, whole-brain contrasts only</i>							
<i>Basal telencephalon</i>							
Amygdala	R	20	-6	-18	1080	0.34	†**
Amygdala	L	-18	-6	-18	1328	0.29	**
<i>Increased activity for positive evaluations, collapsed across Trustworthiness and Attractiveness, whole-brain contrasts only</i>							
<i>Basal telencephalon</i>							
Caudate/nucleus accumbens/medial orbitofrontal cortex	L	-10	10	-4	1888	0.33	†**
Thalamus	R	14	-16	6	384	0.28	†**
Caudate/right amygdala/anterior insula/inferior frontal gyrus	R	4	22	0	76656	0.22	**
Ventral striatum/thalamus/anterior insula/inferior frontal gyrus/ventrolateral prefrontal cortex	R	14	-2	-2	34688	0.20	*
<i>Frontal/insular cortex</i>							
Ventromedial prefrontal cortex	-	0	42	-6	344	0.30	†**
Pregenuar cingulate cortex/dorsal anterior cingulate	-	-2	40	8	96	0.25	†**
Pregenuar cingulate cortex	-	0	36	2	16	0.25	†**

Note. Stereotactic coordinates representing the areas most consistently activated across full database. We report laterality (Right or Left), XYZ coordinates, number of voxels in each cluster (Vol), and weighted percentage of CIMs which activated each cluster (%Act).

†, areas withstanding height-based thresholding.

\*\* , areas withstanding extent-based thresholding ( $p < .001$ ).

\* , areas withstanding extent-based thresholding ( $p < .01$ ).



**Fig. 1** Consistently activated areas across negative evaluations, showing consistent right amygdala activation. Yellow voxels withstood height-based thresholding, orange voxels withstood extent-based thresholding ( $P < 0.001$ ).

anterior insula, right IFG ( $P < 0.001$ ), and bilateral vIPFC that survived extent- but not height-thresholding ( $P < 0.01$ ). Four studies reported ROI-based coordinates for negative contrasts. Including these additional coordinates in our analysis yielded similar results (summarized in Supplementary Table S1).

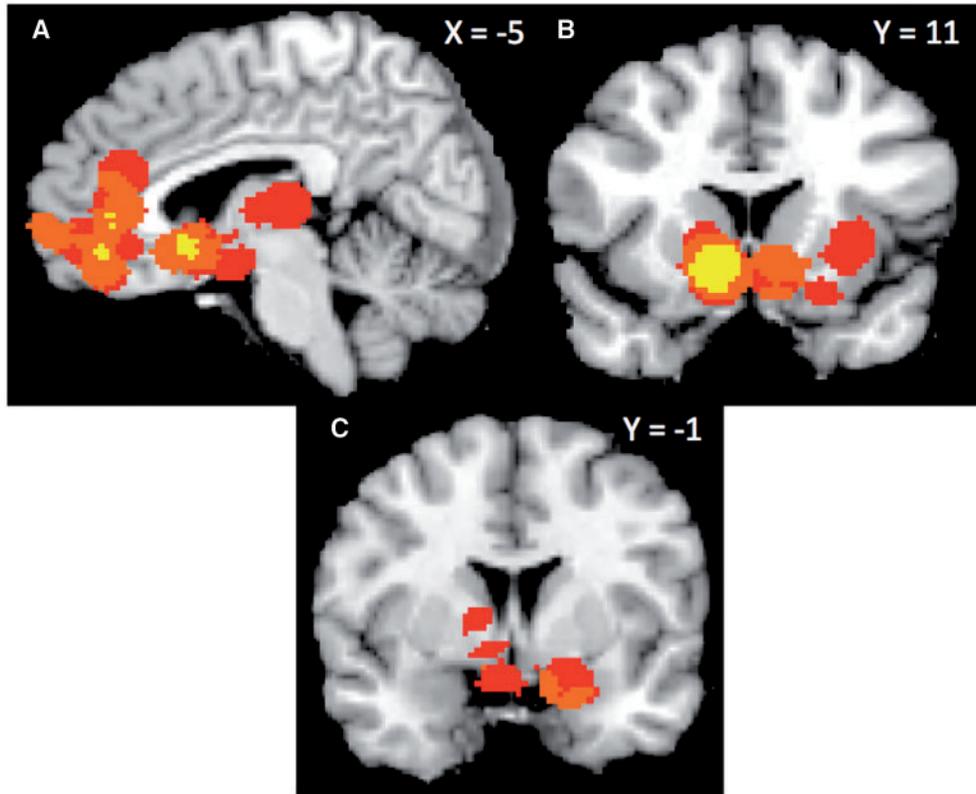
### Non-linear responses

Nine studies within our database conducted non-linear analyses testing for stronger responses to both negative- (unattractive or untrustworthy) and positive-looking (attractive or trustworthy) faces than to faces at the middle of the

continuum. Collapsed across both sets of stimuli, we observed consistent non-linear activation across seven whole-brain contrasts in the right amygdala extending into right putamen (withstood height- and extent-thresholding,  $P < 0.001$ , Table 3). Including two additional ROI-based contrasts in the analysis yielded similar results (Supplementary Table S2). We note that given the relatively small number of contrasts documenting non-linear responses, this analysis is underpowered. Nevertheless, five of the seven whole-brain contrasts reported activity in right amygdala.

We also compared non-linear responses against linear responses, though these comparisons are, by virtue of the smaller number of non-linear contrasts, unavoidably unbalanced. Contrasting negative linear contrasts (13 contrasts) against non-linear contrasts (seven contrasts), we observed a ventral portion of the right amygdala that was more consistently active in negative linear contrasts (withstood height- but not extent-thresholding, Supplementary Table S3), while a more dorsal portion of the right amygdala was more consistently active in non-linear contrasts. Including ROI-based contrasts in the analysis yielded similar results (Supplementary Table S4). (As this contrast is unbalanced, we have provided information regarding the frequency of activation at the peak voxels of those areas that withstood height-thresholding, Supplementary Table S5A)

Contrasting positive linear contrasts (23 contrasts) against non-linear contrasts (seven contrasts), we observed a set of regions that were more consistently active in positive linear contrasts, including bilateral caudate, vmPFC/OFC, dACC/pgACC (withstood height- and extent-thresholding,  $P < 0.001$ ), right thalamus (withstood height-thresholding,



**Fig. 2** Consistently activated areas across positive evaluations, including pgACC, vmPFC (A), left caudate/NAcc extending into mOFC (A and B), and right amygdala (C). Yellow voxels withstood height-based thresholding, orange voxels withstood extent-based thresholding ( $P < 0.001$ ), and pink voxels withstood extent-based thresholding ( $P < 0.01$ ).

**Table 3** Consistently Activated Areas Displaying Non-linear Response Profiles

Region	Lat	x	y	z	Vol	%Act
<i>Non-Linear responses, collapsed across Attractiveness and Trustworthiness studies</i>						
<i>Basal telencephalon</i>						
Amygdala	R	20	-2	-10	1040	0.63 †

*Note.* Stereotactic coordinates representing the areas most consistently activated across full database. We report laterality (Right or Left), XYZ coordinates, number of voxels in each cluster (Vol), and weighted percentage of CIMs which activated each cluster (%Act).

†, areas withstanding height-based thresholding.

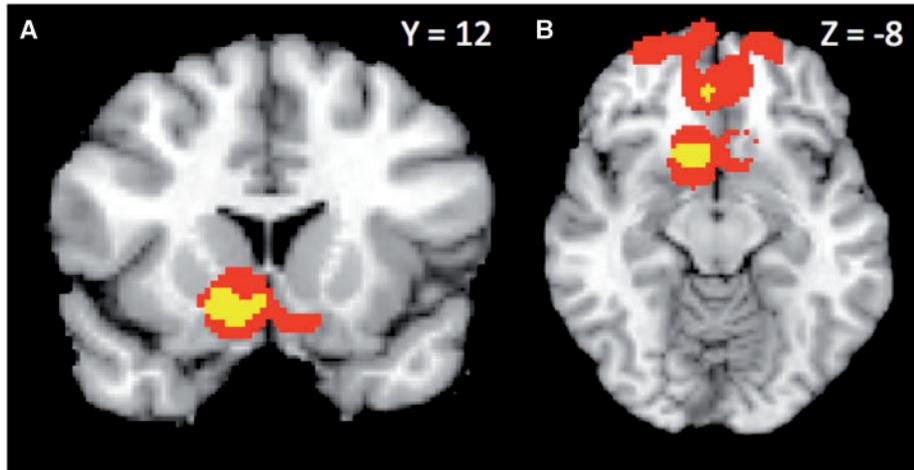
$P < 0.001$ ), sgACC and bilateral vPFC ( $P < 0.01$ , [Supplementary Table S3](#)). Conversely, we observed a region of right dorsal amygdala extending into right putamen (withstood height- but not extent-thresholding, results summarized in [Supplementary Table S3](#)) that was more consistently active in non-linear contrasts. Including ROI coordinates in the analysis yielded similar results (noted in [Supplementary Table S4](#)). (Information regarding the frequency of activation at the peak voxels of those areas that withstood height-thresholding is provided in [Supplementary Table S5B](#).)

### Negative linear responses in attractiveness and trustworthiness studies

We contrasted negative linear responses in attractiveness (six contrasts) and trustworthiness studies (seven contrasts), observing one activation in the right amygdala that was more consistently active for negative linear responses to trustworthiness than attractiveness (withstood height- and extent-thresholding,  $P < 0.001$ ). We observed no regions that were consistently more active for negative linear responses to attractiveness than trustworthiness. (Results are summarized in [Supplementary Table S5](#).) Including ROI coordinates (from one unattractiveness study and three untrustworthiness studies) in the analysis yielded similar results (noted in [Supplementary Table S6](#)).

### Positive linear responses in attractiveness and trustworthiness studies

We contrasted positive linear responses in attractiveness (18 contrasts) and trustworthiness (5 contrasts). We observed activations in left caudate extending into NAcc, vmPFC/OFC and pgACC extending dorsally into dACC (withstood height- and extent-thresholding,  $P < 0.01$ ) that were more consistent for positive linear responses to attractiveness than trustworthiness. We observed no regions that were



**Fig. 3** Consistently activated areas in the contrast 'Extreme' Attractiveness > Trustworthiness, (A) NAcc extending into mOFC, (B) pgACC and vmPFC. (The Average Attractiveness > Trustworthiness contrast produces no regions of consistent activation.) Yellow voxels withstood height-based thresholding and pink voxels withstood extent-based thresholding ( $P < 0.01$ ).

consistently more active for positive linear responses to trustworthiness than attractiveness. (Results are summarized in [Supplementary Table S5](#).) Including ROI coordinates (from one attractiveness study and three trustworthiness studies) in the analysis yielded similar results (noted in [Supplementary Table S6](#)). (Information regarding the frequency of activation at the peak voxels of those areas that withstood height-thresholding is provided in [Supplementary Table S5C](#).)

### Separating attractiveness studies by stimulus type

The differences between trustworthiness and attractiveness studies are interesting but also puzzling given that evaluations on these two dimensions are highly correlated. There were no obvious differences between these two sets of studies (for instance, they were well-balanced between implicit and explicit tasks) except for the nature of the face stimuli used in the studies. Whereas eleven of the attractiveness studies used atypical, extremely attractive faces (culled from magazines and print media, often of models), none of the trustworthiness studies used such faces (typically, these were standardized sets of faces or computer-generated faces).

If the differences between attractiveness and trustworthiness studies are partly due to differences in stimuli, then the regions that differentiate these studies should also appear in contrasts involving the extremeness of faces. We can test this proposition by splitting attractiveness studies into two groups—those that used extremely attractive stimuli and those that used average or computer-generated stimuli. Comparing extreme attractiveness studies to the set of trustworthiness studies should yield areas of consistent activation in NAcc/caudate and mOFC, for example, while there should be fewer differences between average attractiveness and trustworthiness studies.

Indeed, when contrasting extreme attractiveness (11 contrasts) against trustworthiness (8 contrasts), we observed consistent activation in left caudate and NAcc, extending into mOFC, pgACC, and vmPFC (withstood height- and extent-thresholding,  $p < .01$ ; see [Figure 3](#), results summarized in [Supplemental Table 7](#)). Further, we observed a consistent pattern of activation centered in pgACC and extending broadly into both vmPFC and vlPFC that withstood extent-thresholding ( $p < .01$ ) but not height-thresholding. Contrasting average attractiveness against trustworthiness produced no areas of consistent activation.

Similarly within the set of attractiveness studies, when contrasting studies that used extremely attractive faces (11 contrasts) against studies that used more typical faces (7 contrasts), we observed consistent activation in left caudate, vmPFC/mOFC, and pgACC/dACC (withstood height- and extent-thresholding,  $p < .01$ ), while a larger activation extending broadly through mOFC, vmPFC and vlPFC withstood extent-thresholding ( $p < .01$ ), but not height-thresholding. The reverse contrast produced no areas of consistent activation.

Further, differences due to face stimuli should be apparent in studies that used implicit evaluation paradigms. Because no evaluative dimension is specified in such paradigms, stimulus properties should drive the neural responses. Contrasting implicit paradigm studies that used extremely attractive faces with implicit paradigm studies that used more typical faces produced consistent activation in right amygdala, left caudate extending into NAcc and right inferior frontal gyrus (withstood height- but not extent thresholding; results are summarized in [Supplemental Table S7](#)).

### DISCUSSION

Using multi-level kernel density analysis, a statistically rigorous method of meta-analysis that treats contrasts as the unit

of analysis instead of individual activation peaks, we performed a meta-analysis on 29 neuroimaging studies of the social evaluation of faces. We split these studies by valence into two MKDAs, one focusing on brain responses to negative evaluations like unattractiveness and untrustworthiness, and the other focusing on brain responses to positive evaluations like attractiveness and trustworthiness.

Our negative MKDA revealed the most consistent activation in right amygdala. Less consistent areas of activation were observed in left amygdala, right anterior insula, right IFG, right vIPFC and right globus pallidus. These results are remarkably consistent with previous findings regarding the neural responses to angry faces (Morris *et al.*, 1998; Whalen *et al.*, 2001; Monk *et al.*, 2006; Dannlowski *et al.*, 2007). Amygdala responses to angry faces have been widely observed and characterized (Morris *et al.*, 1998; Whalen *et al.*, 2001; Nomura *et al.*, 2004; Taylor *et al.*, 2006; Dannlowski *et al.*, 2007; Monk *et al.*, 2008; Vrticka *et al.*, 2008). Furthermore, a functional connectivity between the amygdala and vIPFC has been proposed and demonstrated (Nomura *et al.*, 2004; Taylor *et al.*, 2006; Monk *et al.*, 2006, 2008), suggesting that in response to angry faces, the vIPFC may serve to modulate amygdala reactivity, effectively regulating emotional responses. Right IFG (Dannlowski *et al.*, 2007), right insula (Dannlowski *et al.*, 2007; Vrticka *et al.*, 2008) and right globus pallidus (Jackson *et al.*, 2008) have also all been implicated in the neural response to angry faces.

Our positive MKDA revealed highly consistent activations in left caudate extending into NAcc/mOFC, vmPFC, dACC/pgACC, right thalamus as well as less consistent activations in right amygdala, insula, IFG and vIPFC. Once again, this pattern of activations bears a strong resemblance to that associated with a different emotional expression—happiness. Happy faces have been observed to elicit responses in parts of the striatum (Morris *et al.*, 1996; Phillips *et al.*, 1998; Fu *et al.*, 2007; Vrticka *et al.*, 2008), mPFC (Phillips *et al.*, 1998; Kesler-West *et al.*, 2001), ACC (Dolan *et al.*, 1996; Phillips *et al.*, 1998; Kesler-West *et al.*, 2001; Vrticka *et al.*, 2008), thalamus (Dolan *et al.*, 1996), IFG (Dolan *et al.*, 1996) and insula (Lee *et al.*, 2002). Furthermore, a number of studies have observed amygdala responses to happy faces (Breiter *et al.*, 1996; Canli *et al.*, 2002; Pessoa *et al.*, 2002; Yang *et al.*, 2002; Somerville *et al.*, 2004).

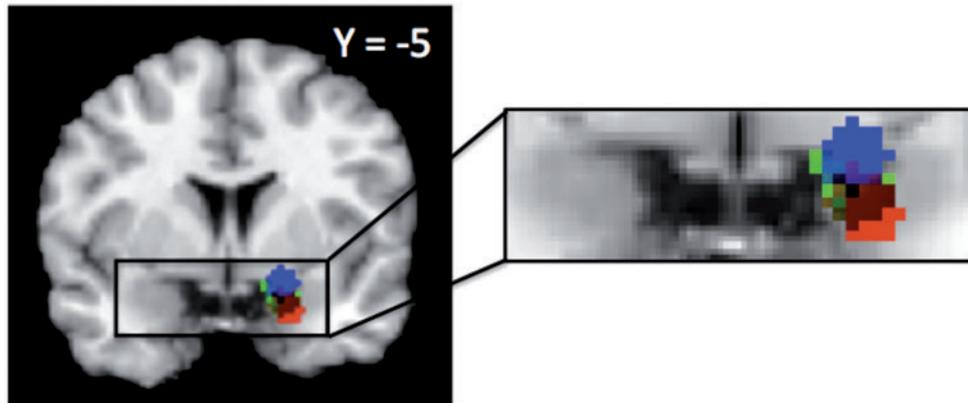
The consistent activation in the left caudate nucleus, extending broadly into the nucleus accumbens, suggests that positive evaluation of faces may depend, in part, on the recruitment of structures implicated in reward-processing (Knutson *et al.*, 2001a,b; Haruno *et al.*, 2004). However, we note that consistent activation in this area was almost entirely driven by attractiveness contrasts, and, therefore, may not be part of a general network for face evaluation. Nonetheless, the highly consistent presence of these areas in our meta-analysis suggests that under certain task and stimulus conditions, attractive faces modulate activity in reward-related regions of the brain.

The similarities between the neural correlates of negatively and positively evaluated faces and angry and happy faces, respectively, parallels perceptual similarities between these types of faces. In computer models of facial trustworthiness, extreme untrustworthiness resembles anger and extreme trustworthiness resembles happiness (Oosterhof and Todorov, 2008, 2009; Todorov *et al.*, 2008a,b). Further, behavioral adaptation studies suggest common neural underpinnings for evaluations of trustworthiness and anger/happiness (Engell *et al.*, 2010). These observations are consistent with the emotion overgeneralization hypothesis (Montepare and Dobish, 2003; Todorov *et al.*, 2008a,b; Zebrowitz and Montepare, 2008; Said *et al.*, 2009a,b), according to which evaluative judgments of faces are based on configurations of facial features resembling emotional expressions. In the context of positive and negative evaluation, these configurations signal approach and avoidance behaviors, respectively (Todorov, 2008). Our meta-analysis findings are also consistent with the hypothesis that novel faces are automatically evaluated with respect to their approach/avoidance value.

### The role of the amygdala in face evaluation

The amygdala is critical for adaptive social behavior (Adolphs, 2010; Sander *et al.*, 2003) and, possibly, for normal face perception and evaluation (Todorov, 2011). Large meta-analyses of PET and fMRI studies on emotional processing show that faces are one of the most potent stimuli for eliciting responses in the amygdala (Costafreda *et al.*, 2008; Sergerie *et al.*, 2008). The role of the amygdala in face evaluation is also consistent with neurophysiology findings of face selective responses in the amygdala (Nakamura *et al.*, 1992; Rolls, 2000a,b; Gothard *et al.*, 2007). The amygdala receives input from the inferior temporal (IT) cortex and projects back not only to IT cortex but also to extrastriate and striate visual areas (Amaral *et al.*, 1992). The amygdala also has strong interconnections with rACC, OFC, mPFC, basal ganglia and anterior insula. This anatomical position of the amygdala allows for it to serve as an affective hub of information. The current findings, together with the findings of a recent ALE-based meta-analysis of a smaller and only partially overlapping set of 16 studies on face evaluation (Bzdok *et al.*, 2011), further buttress the importance of the amygdala in face perception and evaluation.

Importantly, the amygdala responded not only to negatively evaluated faces but also to positively evaluated faces, consistent with meta-analyses of its responses to the valence of emotional expressions (Sergerie *et al.*, 2008). Interestingly, we observed different loci of activation within the amygdala for linear and non-linear responses (Figure 4). Whereas a ventral portion responded more consistently to negative faces only, a dorsal portion of the amygdala responded more consistently to both negative and positive faces than to neutral faces. This dissociation of linear and non-linear responses in the human amygdala parallels the findings of a



**Fig. 4** Linear and non-linear response patterns in right amygdala. Blue indicates voxels more consistently active across non-linear contrasts, red indicates voxels more consistently active across negative evaluations, and green indicates voxels consistently active across positive evaluations. Blue and red clusters withstood height-based thresholding, while the green cluster withstood extent-based thresholding ( $P < 0.001$ ).

high-resolution fMRI study on non-human primates (Hoffman et al., 2007). Hoffman and colleagues observed a linear response in ventral portions of the amygdala (comprising the basolateral amygdala)—specifically, stronger responses to threatening faces and progressively weaker responses to neutral and appeasing faces. However, in a more dorsal portion (comprising the central nucleus and the bed nucleus of the stria terminalis), they observed a non-linear response—stronger responses to both threatening and appeasing faces than to neutral faces. This ventral/dorsal distinction also parallels a distinction made by Whalen and his colleagues (Whalen et al., 2001; Kim et al., 2003, Somerville, et al., 2006; Davis, et al., 2010). They have argued that while the ventral portion of the amygdala is involved in processing valence, the dorsal portion of the amygdala is recruited in determining the value of ambiguous information (e.g. expressions of surprise) in a given context. These authors suggest further that given the dorsal amygdala's response to surprised (Kim et al., 2003), fearful (Whalen et al., 1998, 2001) and happy faces (Breiter et al., 1996; Whalen et al., 1998), it may be tracking the salience of these faces, more generally. This hypothesis is consistent with the current findings.

These findings open the door to future work along those lines. One possibility is that there exist separate populations of neurons within the amygdala that code for stimulus valence and stimulus salience, respectively. Ultimately, the findings are in line with previous work proposing a shift away from conceptualizing the amygdala as simply a fear or threat module and instead toward an account of the amygdala as also tracking stimulus intensity (Anderson et al., 2003; Small et al., 2003) or motivational salience (Sander et al., 2003; Cunningham et al., 2008; Adolphs, 2010; Todorov, 2011). These findings also serve as an excellent reminder that one of the additional benefits of the meta-analytic method is the possibility of generating new, testable hypotheses for future research.

Faces that are tagged as affectively significant in the amygdala can be further processed in prefrontal regions, which, in

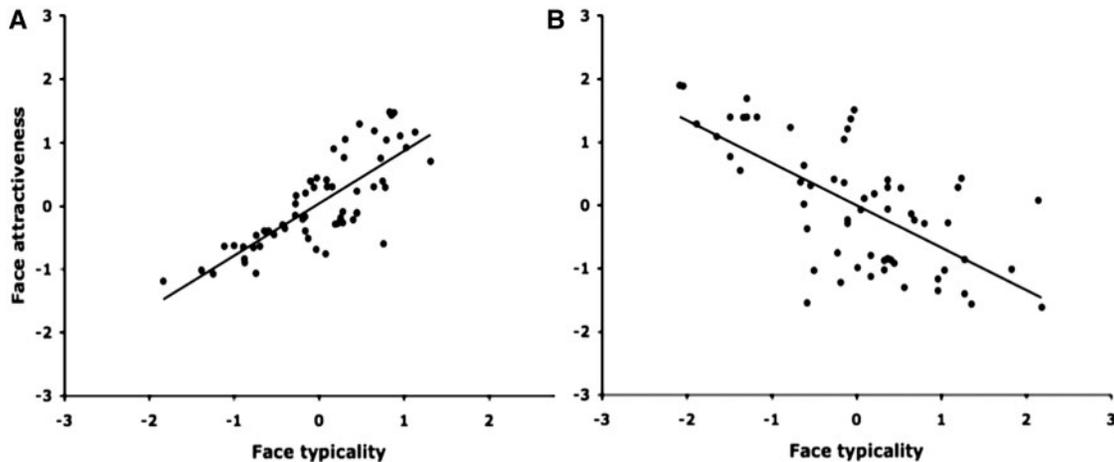
turn, can serve to modulate amygdala activity. Prefrontal-amygdala connections have been explored in the vmPFC (Quirk et al., 2003, Heinz et al., 2004), as well as the pgACC (Pezawas et al., 2005; Stein et al., 2007; Zink et al., 2010), both of which were observed as consistently activated across our set of positive contrasts.

#### Stimulus effects on neuroimaging findings

We also performed several smaller MKDAs to compare between study type, within negative and positive linear responses. These more targeted MKDAs offered evidence that our negative and positive analyses were driven by untrustworthiness and attractiveness, respectively. These two sets of studies were associated with different loci of activations: the right amygdala was more consistently active as facial trustworthiness decreased, while the NAcc/caudate and vmPFC/pgACC were all more consistently active as facial attractiveness increased. This distinction mirrors the results we observed in our primary analyses. In contrast, no brain regions were consistently activated across contrasts where facial attractiveness decreased or facial trustworthiness increased, respectively, even when including ROI coordinates in the analyses. We should note that these contrasts—especially the comparison between positive linear responses in attractiveness and trustworthiness studies—are certainly unbalanced, rendering the results more suggestive than confirmative.

As noted in the introduction, given that attractiveness and trustworthiness judgments from faces are highly correlated, this pattern is puzzling. These differences between trustworthiness and attractiveness studies cannot be explained by researchers' a priori focus on different regions because our results hold even for whole brain analyses that did not include ROI coordinates from individual studies.

The apparent differences in the neural bases of attractiveness and trustworthiness are also puzzling in the context of studies that used the same set of faces to examine responses to facial attractiveness and trustworthiness (Todorov and



**Fig. 5** The relationship between perceptions of face typicality and face attractiveness for (A) a standardized set of faces (the Karolinska set) and (B) faces sampled to be extremely attractive from websites of models. The judgments are shown in standardized units. Each point of the plots represents a face.

Engell, 2008). Specifically, Todorov and Engell re-analyzed the data from Engell *et al.* (2007), using 14 different social judgments of the same set of faces. Most of the brain responses were accounted by a general valence dimension rather than by specific dimensions such as attractiveness and trustworthiness (both of these were highly correlated with this dimension).

What could be driving the differences in neuroimaging studies on attractiveness and trustworthiness? One possibility is that the type of faces used in these studies may lead to different responses. Specifically, a third variable that is correlated with both trustworthiness and attractiveness but could vary across sets of faces may account for such differences. One candidate is face typicality. Recently, Said and colleagues (2010) showed that coding face typicality is a more parsimonious explanation of prior findings of the involvement of the amygdala in face evaluation than coding face valence.

Face typicality could vary across data sets and lead to different results. For example, in many standardized data sets of natural faces, typicality is positively correlated with both attractiveness and trustworthiness judgments (Figure 5A). In studies using these stimuli, the amygdala shows stronger responses to more atypical faces that happened also to be more negative (Todorov and Engell, 2008). In studies using artificial stimuli created by a statistical model, the most atypical faces are faces at the extremes of the dimension. In such studies, the amygdala responds to more atypical faces that happened to be more positive or more negative (Said *et al.*, 2010). The important distinction here is not between real and artificial faces. Judgments of artificial faces that have not been manipulated to exaggerate differences along social dimensions are linearly correlated with their perceived typicality. Finally, in attractiveness studies that use extremely attractive faces (e.g. Aharon *et al.*, 2001), the most attractive faces may be the least typical

(Figure 5B). In such studies, the amygdala may respond to both extremely attractive and extremely unattractive faces as observed in Winston *et al.* (2007).

The typicality hypothesis predicts that faces that systematically differ in their perceived typicality may lead to different neural responses. In fact, in the contrast of attractiveness studies using extreme faces and studies using more typical faces, focusing specifically on studies employing implicit paradigms, we observed consistent patterns of activation in right amygdala and NAcc/caudate. This suggests that when task demands are controlled for, the driving force behind NAcc and caudate activations observed in these studies was the usage of extremely attractive, atypical faces. This result also lends additional support to the suggestion that extreme, atypical faces will drive amygdala activity, regardless of their trustworthiness. Further, the consistent activations in vmPFC and pgACC that were observed across attractiveness contrasts but not trustworthiness contrasts can also be accounted for by face typicality. Contrasting extreme attractiveness and trustworthiness continued to produce consistent activation in these regions, while contrasting more typical attractiveness against trustworthiness did not.

It may be the case that in the context of face evaluation, some of the regions implicated in reward processing are only activated upon the presence of real and extremely attractive faces, or the goal to evaluate face attractiveness, or some combination of stimulus features and task demands. Unfortunately, we do not have a sufficient number of studies to test for more specific effects.

### Recommendations for studies using face stimuli

Our findings suggest that the type of face stimuli selected for a particular study matters a great deal. For example, using more 'extreme' faces resulted in more consistently observed activation in the NAcc. Given that stimuli are often selected

in an *ad hoc* fashion and rarely shared among research groups, this complicates comparisons across studies. Moreover, it undermines the generalizability of results.

To overcome these problems, researchers need to use a shared set of stimuli, not necessarily the same stimuli but stimuli sampled from a common pool. One approach is to use parametrically manipulated faces generated by an explicitly specified statistical model (e.g. Oosterhof and Todorov, 2008; Todorov and Oosterhof, 2011). This approach has the benefit of providing researchers with a full spectrum of faces—one that is not biased towards one portion of a given dimension. Our laboratory has made a number of such databases available for academic research (<http://web.script.princeton.edu/~tlab/databases/>). However, artificial faces may not be the best stimuli for many investigators. In this case, it would be best to create a common bank of stimuli that are shared with other research groups. These stimuli could be validated on a number of important variables such as typicality, and these variables could be further used to facilitate comparisons across studies.

## SUPPLEMENTARY DATA

Supplementary data are available at SCAN online.

## Conflict of Interest

None declared.

## REFERENCES

- Adolphs, R. (2010). What does the amygdala contribute to social cognition? *Annals of the New York Academy of Sciences*, 1191, 42–61.
- Adolphs, R., Tranel, D., Damasio, A.R. (1998). The human amygdala in social judgment. *Nature*, 393, 470–4.
- Aharon, I., Etcoff, N., Ariely, D., Chabris, C.F., O'Connor, E., Breiter, H.C. (2001). Beautiful faces have variable reward value: fMRI and behavioral evidence. *Neuron*, 32, 537–51.
- Amaral, D.G., Price, J.L., Pitkänen, A., Carmichael, S.T. (1992). Anatomical organization of the primate amygdaloid complex. In: Aggleton, J.P., editor. *The Amygdala: Neurobiological Aspects of Emotion, Memory, and Mental Dysfunction*. New York: Wiley-Liss, pp. 1–66.
- Anderson, A.K., Christoff, K., Stappen, I., et al. (2003). Dissociated neural representations of intensity and valence in human olfaction. *Nature Neuroscience*, 6, 196–202.
- Aron, A., Fisher, H., Mashek, D.J., Strong, G., Li, H., Brown, L.L. (2005). Reward, motivation, and emotion systems associated with early-stage intense romantic love. *Journal of Neurophysiology*, 94, 327–37.
- Asch, S.E. (1946). Forming impressions of personality. *Journal of Abnormal and Social Psychology*, 41, 258–90.
- Baas, D., Aleman, A., Vink, M., Ramsey, N.F., de Haan, E.H.F., Kahn, R.S. (2008). Evidence of altered cortical and amygdala activation during social decision-making in schizophrenia. *NeuroImage*, 40, 719–27.
- Bar, M., Neta, M., Linz, H. (2006). Very first impressions. *Emotion*, 6, 269–78.
- Ballew, C.C., Todorov, A. (2007). Predicting political elections from rapid and unreflective face judgments. *Proceedings of the National Academy of Sciences of USA*, 104, 17948–53.
- Blair, I.V., Judd, C.M., Chapleau, K.M. (2004). The influence of Afrocentric facial features in criminal sentencing. *Psychological Science*, 15, 674–9.
- Blasi, G., Hariri, A.R., Alce, G., et al. (2009). Preferential amygdala reactivity to the negative assessment of neutral faces. *Biological Psychiatry*, 66, 847–53.
- Bray, S., O'Doherty, J. (2007). Neural coding of reward-prediction error signals during classical conditioning with attractive faces. *Journal of Neurophysiology*, 97, 3036–45.
- Breiter, H.C., Etcoff, N.L., Whalen, P.J., et al. (1996). Response and habituation of the human amygdala during visual processing of facial expression. *Neuron*, 17, 875–87.
- Bzdok, D., Langner, R., Caspers, S., et al. (2011). ALE meta-analysis on facial judgments of trustworthiness and attractiveness. *Brain Structure and Function*, 215, 209–23.
- Canli, T., Sivers, H., Whitfield, S.L., Gotlib, I.H., Gabrieli, J.D. (2002). Amygdala response to happy faces as a function of extraversion. *Science*, 296, 2191.
- Chatterjee, A., Thomas, A., Smith, S.E., Aguirre, G.K. (2009). The neural response to facial attractiveness. *Neuropsychology*, 23, 135–43.
- Cloutier, J., Heatherton, T.F., Whalen, P.J., Kelley, W.M. (2008). Are attractive people rewarding? Sex differences in the neural substrates of facial attractiveness. *Journal of Cognitive Neuroscience*, 20, 941–51.
- Costafreda, S.G., Brammer, M.J., David, A.S., Fu, C.H.Y. (2008). Predictors of amygdala activation during the processing of emotional stimuli: A meta-analysis of 385 PET and fMRI studies. *Brain Research Reviews*, 58, 57–70.
- Cunningham, W.A., Van Bavel, J.J., Johnsen, I.R. (2008). Affective flexibility: evaluative processing goals shape amygdala activity. *Psychological Science*, 19, 152–60.
- Dannlowski, U., Ohrmann, P., Bauer, J., et al. (2007). Amygdala reactivity predicts automatic negative evaluations for facial emotions. *Psychiatry Research: Neuroimaging*, 154, 13–20.
- Davis, F.C., Johnstone, T., Mazzulla, E.C., Oler, J.A., Whalen, P.J. (2010). Regional response differences across the human amygdaloid complex during social conditioning. *Cerebral Cortex*, 20, 612–21.
- Dolan, R.J., Fletcher, P., Morris, J., Kapur, N., Deakin, J.F.W., Frith, C.D. (1996). Neural activation during covert processing of positive emotional facial expressions. *NeuroImage*, 4, 194–200.
- Eberhardt, J.L., Davies, P.G., Purdie-Vaughns, V.J., Johnson, S.L. (2006). Looking deathworthy: Perceived stereotypicality of Black defendants predicts capital-sentencing outcomes. *Psychological Science*, 17, 383–6.
- 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.
- Engell, A.D., Todorov, A., Haxby, J.V. (2010). Common neural mechanisms for the evaluation of facial trustworthiness and emotional expressions as revealed by behavioral adaptation. *Perception*, 39, 931–41.
- Fox, P.T., Parsons, L.M., Lancaster, J.L. (1998). Beyond the single study: function/location meta-analysis in cognitive neuroimaging. *Current Opinions in Neurobiology*, 8, 178–87.
- Fu, C.H.Y., Williams, S.C.R., Brammer, M.J., et al. (2007). Neural responses to happy facial expressions in major depression following antidepressant treatment. *American Journal of Psychiatry*, 164, 599–607.
- Gordon, D.S., Platek, S.M. (2009). Trustworthy? The brain knows: Implicit neural responses to faces that vary in Dark Triad Personality characteristics and trustworthiness. *Journal of Social, Evolutionary, and Cultural Psychology*, 3, 182–200.
- Gothard, K.M., Battaglia, F.P., Erickson, C.A., Spitzer, K.M., Amaral, D.G. (2007). Neural responses to facial expression and face identity in the monkey amygdala. *Journal of Neurophysiology*, 97, 1671–83.
- Hamermesh, D., Biddle, J. (1994). Beauty and the labor market. *The American Economic Review*, 84, 1174–94.
- Hampshire, A., Chudhry, A.M., Owen, A.M., Roberts, A.C. (2011). Dissociable roles for lateral orbitofrontal cortex and lateral prefrontal cortex during preference driven reversal learning. *Neuroimage*, [Epub 29 October, 2011, doi: 10.1016/j.neuroimage.2011.10.072].

- Haruno, M., Kuroda, T., Doya, K., et al. (2004). A neural correlate of reward-based behavioral learning in caudate nucleus: a functional magnetic resonance imaging study of a stochastic decision task. *Journal of Neuroscience*, 24, 1660–5.
- Heinz, A., Braus, D.F., Smolka, M.N., et al. (2004). Amygdala-prefrontal coupling depends on a genetic variation of the serotonin transporter. *Nature Neuroscience*, 8, 20–1.
- Hoffman, K.L., Gothard, K.M., Schmid, M.C., Logothetis, N.K. (2007). Facial-expression and gaze-selective responses in the monkey amygdala. *Current Biology*, 17, 766–72.
- Iaria, G., Fox, C.J., Waite, C.T., Aharon, I., Barton, J.J.S. (2008). The contribution of the fusiform gyrus and superior temporal sulcus in processing facial attractiveness: Neuropsychological and neuroimaging evidence. *Journal of Neuroscience*, 155, 409–22.
- Jackson, M.C., Wolf, C., Johnston, S.J., Raymond, J.E., Linden, D.E.J. (2008). Neural correlates of enhanced visual short-term memory for angry faces: an fMRI study. *PLoS One*, 3, e3536.
- Kampe, K.K., Frith, C.D., Dolan, R.J., Frith, U. (2001). Reward value of attractiveness and gaze. *Nature*, 413, 589.
- Kawabata, H., Zeki, S. (2004). Neural correlates of beauty. *Journal of Neurophysiology*, 91, 1699–05.
- Kesler-West, M.L., Andersen, A.H., Smith, C.D., Avison, M.J., Davis, C.E., Kryscio, R.J., Blonder, L.X. (2001). Neural substrates of facial emotion processing using fMRI. *Cognitive Brain Research*, 11, 213–26.
- Kim, H., Somerville, L.H., McLean, A.A., Johnstone, T., Shin, L.M., Whalen, P.J. (2003). Functional MRI response of the human dorsal amygdala/substantia innominata region to facial expressions of emotion. *Annals of the New York Academy of Sciences*, 985, 533–5.
- Kim, H., Adolphs, R., O'Doherty, J.P., Shimojo, S. (2007). Temporal isolation of neural processes underlying face preference decisions. *Proceedings of the National Academy of Sciences of USA*, 104, 18253–8.
- Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K., Wager, T.D. (2008). Functional grouping and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *NeuroImage*, 42, 998–1031.
- Kober, H., Wager, T.D. (2010). Meta-analyses of neuroimaging data. *Wiley Interdisciplinary Reviews*, 1, 293–300.
- Knutson, B., Adams, C.M., Fong, G.W., Hommer, D. (2001a). Anticipation of increasing monetary reward selectively recruits nucleus accumbens. *Journal of Neuroscience*, 21, 1–5.
- Knutson, B., Fong, G.W., Adams, C.M., Varner, J.L., Hommer, D. (2001b). Dissociation of reward anticipation and outcome with event-related fMRI. *NeuroReport*, 12, 3683–7.
- Kranz, F., Ishai, A. (2006). Face perception is modulated by sexual preference. *Current Biology*, 16, 63–8.
- Laird, A.R., Fox, P.M., Price, C.J., et al. (2005). ALE meta-analysis: controlling the false discovery rate and performing statistical contrasts. *Human Brain Mapping*, 25, 155–64.
- Lee, T.M.C., Liu, H., Hoosain, R., et al. (2002). Gender differences in neural correlates of recognition of happy and sad faces in humans assessed by functional magnetic resonance imaging. *Neuroscience Letters*, 1, 13–16.
- Liang, X., Zebrowitz, L.A., Zhang, Y. (2010). Neural activation in the 'reward circuit' shows a nonlinear response to facial attractiveness. *Social Neuroscience*, 5, 320–34.
- Mazur, A., Mazur, J., Keating, C. (1984). Military rank attainment of a West Point class: effects of cadets' physical features. *American Journal of Sociology*, 90, 125–50.
- Monk, C.S., Nelson, E.E., McClure, E.B., et al. (2006). Ventrolateral prefrontal cortex activation and attentional bias in response to angry faces in adolescents with generalized anxiety disorder. *American Journal of Psychiatry*, 163, 1091–7.
- Monk, C.S., Telzer, E.H., Mogg, K., et al. (2008). Amygdala and ventrolateral prefrontal cortex activation to masked angry faces in children and adolescents with generalized anxiety disorder. *Archives of General Psychiatry*, 65, 568–76.
- Montepare, J.M., Dobish, H. (2003). The contribution of emotion perceptions and their overgeneralizations to trait impressions. *Journal of Nonverbal Behavior*, 27, 236–54.
- Morris, J.S., Frith, C.D., Perrett, D.I., et al. (1996). A differential neural response in the human amygdala to fearful and happy expressions. *Nature*, 383, 812–5.
- Morris, J.S., Ohman, A., Dolan, R.J. (1998). Conscious and unconscious emotional learning in the human amygdala. *Nature*, 393, 467–70.
- Nakamura, K., Kawashima, R., Nagumo, S., et al. (1998). Neuro-anatomical correlates of the assessment of facial attractiveness. *Neuroreport*, 9, 753–7.
- Nakamura, K., Mikami, A., Kubota, K. (1992). Activity of single neurons in the monkey amygdala during performance of a visual discrimination task. *Journal of Neurophysiology*, 67, 1447–63.
- Nichols, T., Hayasaka, S. (2003). Controlling the familywise error rate in functional neuroimaging: a comparative review. *Statistical Methods in Medical Research*, 12, 419–46.
- Nielsen, F.A., Copenhagen, D., Lyngby, D. (2005). Mass meta-analysis in Talairach space. *Advances in Neural Information Processing Systems*, 17, 985–92.
- Nomura, M., Ohira, H., Haneda, K., et al. (2004). Functional association of the amygdala and ventral prefrontal cortex during cognitive evaluation of facial expressions primed by masked angry faces: an event-related fMRI study. *NeuroImage*, 21, 352–63.
- Pessoa, L., McKenna, M., Gutierrez, E., Ungerleider, L.G. (2002). Neural processing of emotional faces requires attention. *Proceedings of the National Academy of Sciences of USA*, 99, 11458–63.
- Pezawas, L., Meyer-Lindenberg, A., Drabant, E.M., et al. (2005). 5-HTTLPR polymorphism impacts human cingulate-amygdala interactions: a genetic susceptibility mechanism for depression. *Nature Neuroscience*, 8, 828–34.
- Phan, K.L., Wager, T.D., Taylor, S.F., Liberzon, I. (2002). Functional neuro-anatomy of emotion: a meta-analysis of emotion activation studies in PET and fMRI. *NeuroImage*, 16, 331–48.
- Phillips, M.L., Bullmore, E.T., Howard, R., et al. (1998). Investigation of facial recognition memory and happy and sad facial expression perception: an fMRI study. *Psychiatry Research: Neuroimaging Section*, 83, 127–38.
- Pinkham, A.E., Hopfinger, J.B., Pelphrey, K.A., Piven, J., Penn, D.L. (2008a). Neural bases for impaired social cognition in schizophrenia and autism spectrum disorders. *Schizophrenia Research*, 99, 164–75.
- Pinkham, A.E., Hopfinger, J.B., Ruparel, K., Penn, D.L. (2008b). An investigation of the relationship between activation of a social cognitive neural network and social functioning. *Schizophrenia Bulletin*, 34, 688–97.
- Pochon, J.-B., Riis, J., Sanfey, A.G., Nystrom, L.E., Cohen, J.D. (2008). Function imaging of decision conflict. *Journal of Neuroscience*, 28, 2468–3473.
- O'Doherty, J., Winston, J., Critchley, H., Perrett, D., Burt, D.M., Dolan, R.J. (2003). Beauty in a smile: the role of medial orbitofrontal cortex in facial attractiveness. *Neuropsychologia*, 41, 147–55.
- Olivola, C.Y., Todorov, A. (2010a). Elected in 100 milliseconds: appearance-based trait inferences and voting. *Journal of Nonverbal Behavior*, 34, 83–110.
- Olivola, C.Y., Todorov, A. (2010b). Fooled by first impressions? Re-examining the diagnostic value of appearance-based inferences. *Journal of Experimental Social Psychology*, 46, 315–24.
- Olson, I.R., Marshuetz, C. (2005). Facial attractiveness is appraised in a glance. *Emotion*, 5, 498–502.
- Oosterhof, N.N., Todorov, A. (2008). The functional basis of face evaluation. *Proceedings of the National Academy of Sciences of USA*, 105, 11087–92.
- Oosterhof, N.N., Todorov, A. (2009). Shared perceptual basis of emotional expressions and trustworthiness impressions from faces. *Emotion*, 9, 128–33.

- Quirk, G.J., Likhtik, E., Pelletier, J.G., Paré, D. (2003). Stimulation of medial prefrontal cortex decreases the responsiveness of central amygdala output neurons. *Journal of Neuroscience*, 23, 8800–7.
- Rolls, E.T. (2000a). The orbitofrontal cortex and reward. *Cerebral Cortex*, 10, 284–94.
- Rolls, E.T. (2000b). Neurophysiology and function of the primate amygdala, and neural basis of emotion. In: Aggleton, J.P., editor. *The Amygdala: A Functional Analysis*. USA: Oxford University Press, pp. 447–78.
- Rule, N.O., Ambady, N. (2008). Brief exposures: male sexual orientation is accurately perceived at 50 ms. *Journal of Experimental Social Psychology*, 44, 1100–5.
- Rule, N.O., Ambady, N., Adams, R.B.Jr (2009). Personality in perspective: judgmental consistency across orientations of the face. *Perception*, 38, 1688–99.
- Said, C.P., Baron, S.G., Todorov, A. (2009a). Nonlinear amygdala response to face trustworthiness: contributions of high and low spatial frequency information. *Journal of Cognitive Neuroscience*, 21, 519–28.
- Said, C.P., Dotsch, R., Todorov, A. (2010). The amygdala and FFA track both social and non-social face dimensions. *Neuropsychologia*, 48, 3596–605.
- Said, C.P., Haxby, J.V., Todorov, A. (2011). Brain systems for the assessment of the affective value of faces. *Philosophical Transactions of the Royal Society, B*, 336, 1660–70.
- Said, C.P., Sebe, N., Todorov, A. (2009b). Structural resemblance to emotional expressions predicts evaluation of emotionally neutral faces. *Emotion*, 9, 260–4.
- Sander, D., Grafman, J., Zalla, T. (2003). The human amygdala: an evolved system for relevance detection. *Reviews in the Neurosciences*, 14(4), 303–16.
- Salimi-Khorshidi, G., Smith, S.M., Keltner, J.R., Wager, T.D., Nichols, T.E. (2009). Meta-analysis of neuroimaging data: a comparison of image-based and coordinate-based pooling of studies. *NeuroImage*, 45, 810–23.
- Sergerie, K., Chochol, C., Armony, J.L. (2008). The role of the amygdala in emotional processing: a quantitative meta-analysis of functional neuroimaging studies. *Neuroscience and Biobehavioral Reviews*, 32, 811–30.
- Small, D.M., Gregory, M.D., Mak, Y.E., Gitelman, D., Mesulam, M.M., Parrish, T. (2003). Dissociation of neural representation of intensity and affective valuation in human gestation. *Neuron*, 39, 701–11.
- Smith, D.V., Hayden, B.Y., Truong, T., Song, A.W., Platt, M.L., Huettel, S.A. (2010). Distinct value signals in anterior and posterior ventromedial prefrontal cortex. *Journal of Neuroscience*, 30, 2490–95.
- Somerville, L.H., Kim, H., Johnstone, T., Alexander, A.L., Whalen, P.J. (2004). Human amygdala responses during presentation of happy and neutral faces: correlations with state anxiety. *Biological Psychiatry*, 55, 897–903.
- 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.
- Stein, J.L., Wiedholz, L.M., Bassett, D.S., et al. (2007). A validated network of effective amygdala connectivity. *NeuroImage*, 36, 736–45.
- Taylor, S.E., Eisenberger, N.I., Saxbe, D., Lehman, B.J., Lieberman, M.D. (2006). Neural responses to emotional stimuli are associated with childhood family stress. *Biological Psychiatry*, 60, 296–301.
- Todorov, A. (2008). Evaluating faces on trustworthiness: an extension of systems for recognition of emotions signaling approach/avoidance behaviors. *Annals of the New York Academy of Sciences*, 1124, 208–24.
- Todorov, A. (2011). The role of the amygdala in face perception and evaluation. *Motivation and Emotion*. [Published online August 2, 2011; doi:10.1007/s11031-011-9238-5].
- Todorov, A., Baron, S., Oosterhof, N.N. (2008a). Evaluating face trustworthiness: a model based approach. *Social, Cognitive, and Affective Neuroscience*, 3, 119–27.
- Todorov, A., Engell, A. (2008). The role of the amygdala in implicit evaluation of emotionally neutral faces. *Social, Cognitive, and Affective Neuroscience*, 3, 303–12.
- Todorov, A., Mandisodza, A.N., Goren, A., Hall, C.C. (2005). Inferences of competence from faces predict election outcomes. *Science*, 308, 1623–6.
- Todorov, A., Oosterhof, N.N. (2011). Modeling social perception of faces. *IEEE Signal Processing Magazine*, 28, 117–22.
- Todorov, A., Pakrashi, M., Oosterhof, N.N. (2009). Evaluating faces on trustworthiness after minimal time exposure. *Social Cognition*, 27, 813–33.
- Todorov, A., Said, C.P., Engell, A.D., Oosterhof, N.N. (2008b). Understanding evaluation of faces on social dimensions. *Trends in Cognitive Sciences*, 12, 455–60.
- Todorov, A., Said, C.P., Oosterhof, N.N., Engell, A.D. (2011). Task-invariant brain responses to the social value of faces. *Journal of Cognitive Neuroscience*. [Epub ahead of print; doi:10.1162/jocn.2011.21616].
- Todorov, A., Said, C.P., Verosky, S.C. (2011). Personality impressions from facial appearance. In: Calder, A., Haxby, J.V., Johnson, M., Rhodes, G., editors. *Handbook of Face Perception*. USA: Oxford University Press, pp. 631–52.
- Tsukiura, T., Cabeza, R. (2010). Shared brain activity for aesthetic and moral judgments: implications for the Beauty-is-Good stereotype. *Social Cognitive and Affective Neuroscience*. [Published online March 15, 2010; doi: 10.1093/scan/nsq025].
- Turk, D.J., Banfield, J.F., Walling, B.R., Heatherton, T.F., Grafton, S.T., Handy, T.C., et al. (2004). From facial cue to dinner for two: the neural substrates of personal choice. *NeuroImage*, 22, 1281–90.
- Van Rijn, S., Swaab, H., Baas, D., de Haan, E., Kahn, R.S., Aleman, A. (2011). Neural systems for social cognition in Klinefelter syndrome (47,XXY): evidence from fMRI. *Social Cognitive and Affective Neuroscience*. [Published online July 6, 2011; doi: 10.1093/scan/nsr041].
- Verosky, S.C., Todorov, A. (2010). Differential neural responses to faces physically similar to the self as a function of their valence. *NeuroImage*, 49, 1690–8.
- Vrticka, P., Andersson, F., Grandjean, D., Sander, D., Vuilleumier, P. (2008). Individual attachment style modulates human amygdala and striatum activation during social appraisal. *PLoS ONE*, 3, e2868.
- Wager, T.D., Barrett, L.F., Bliss-Moreau, E., et al. (2008). The neuroimaging of emotion. In: Lewis, M., Haviland-Jones, J.M., Barrett, L.F., editors. *Handbook of Emotion* 3rd edn. New York: Guilford, pp. 249–71.
- Wager, T.D., Smith, E.E. (2003). Neuroimaging studies of working memory: a meta-analysis. *Cognitive Affective Behavioural Neuroscience*, 3, 255–74.
- Wager, T.D., Reading, S., Jonides, J. (2004). Neuroimaging studies of shifting attention: a meta-analysis. *NeuroImage*, 22, 1679–93.
- Whalen, P.J., Rauch, S.L., Etcoff, N.L., McInerney, S.C., Lee, M.B., Jenike, M.A. (1998). Masked presentations of emotional facial expressions modulate amygdala activity without explicit knowledge. *Journal of Neuroscience*, 18, 411–8.
- Whalen, P.J., Shin, L.M., McInerney, S.C., Fischer, H., Wright, C.I., Rauch, S.L. (2001). A function MRI study of human amygdala responses to facial expressions of fear versus anger. *Emotion*, 1, 70–83.
- 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.S., O’Doherty, J., Kilner, J.M., Perrett, D.L., Dolan, R.J. (2007). Brain systems for assessing facial attractiveness. *Neuropsychologia*, 45, 195–206.
- Winston, J.A., Strange, B., O’Doherty, J., Dolan, R.J. (2002). Automatic and intentional brain responses during evaluation of trustworthiness of face. *Nature Neuroscience*, 5, 277–83.
- Yang, T.T., Menon, C., Eliez, S., et al. (2002). Amygdalar activation associated with positive and negative facial expressions. *Neuroreport*, 13, 1737–41.

- Zaki, J., Schirmer, J., Mitchell, J.P. (2011). Social influence modulates the neural computation of value. *Psychological Science*, 22, 894–900.
- Zebrowitz, L.A. (1999). Reading faces: window to the soul?. Boulder, CO: Westview Press.
- Zebrowitz, L.A., McDonald, S.M. (1991). The impact of litigants' babyfacedness and attractiveness on adjudications in small claims courts. *Law and Human Behavior*, 15, 603–23.
- Zebrowitz, L.A., Montepare, J.M. (2008). Social psychological face perception: why appearance matters. *Social and Personality Psychology Compass*, 2, 1497–1517.
- Zink, C.F., Stein, J.L., Kempf, L., Hakimi, S., Meyer-Lindenberg, A. (2010). Vasopressin modulates medial prefrontal cortex-amygdala circuitry during emotion processing in humans. *Journal of Neuroscience*, 30, 7017–22.