

The neural representation of typical and atypical experiences of negative images: comparing fear, disgust and morbid fascination

Suzanne Oosterwijk,¹ Kristen A. Lindquist,^{2,3} Morenikeji Adebayo,^{4,5} and Lisa Feldman Barrett^{4,5}

¹Department of Psychology, University of Amsterdam, Weesperplein 4, 1018 XA, Amsterdam, The Netherlands, ²Department of Psychology, University of North Carolina, 235 E. Cameron Avenue, Chapel Hill, NC 27599, ³Biomedical Research Imaging Center, University of North Carolina, 125 Mason Farm Road, Chapel Hill, NC 27599, ⁴Martinos Center for Biomedical Imaging, 149 Thirteenth Street, Charlestown, MA 02129, and ⁵Department of Psychology, Northeastern University, 125 Nightingale Hall, Boston, MA 02115, USA

Correspondence should be addressed to Suzanne Oosterwijk, Department of Psychology, University of Amsterdam, Amsterdam, The Netherlands. E-mail: s.oosterwijk@uva.nl.

Abstract

Negative stimuli do not only evoke fear or disgust, but can also evoke a state of ‘morbid fascination’ which is an urge to approach and explore a negative stimulus. In the present neuroimaging study, we applied an innovative method to investigate the neural systems involved in typical and atypical conceptualizations of negative images. Participants received false feedback labeling their mental experience as fear, disgust or morbid fascination. This manipulation was successful; participants judged the false feedback correct for 70% of the trials on average. The neuroimaging results demonstrated differential activity within regions in the ‘neural reference space for discrete emotion’ depending on the type of feedback. We found robust differences in the ventrolateral prefrontal cortex, the dorsomedial prefrontal cortex and the lateral orbitofrontal cortex comparing morbid fascination to control feedback. More subtle differences in the dorsomedial prefrontal cortex and the lateral orbitofrontal cortex were also found between morbid fascination feedback and the other emotion feedback conditions. This study is the first to forward evidence about the neural representation of the experimentally unexplored state of morbid fascination. In line with a constructionist framework, our findings suggest that neural resources associated with the process of conceptualization contribute to the neural representation of this state.

Key words: emotion; morbid fascination; false feedback; conceptualization

Introduction

Affective scientists, including neuroscientists, often study emotional processes by presenting participants with negative images and then observing the consequences. So far, the discrete emotions that have dominated the research agenda are typical experiences such as fear and disgust (e.g. Hariri et al., 2002; Wright et al., 2004; Stark et al., 2007; Borg et al., 2013). Nevertheless, people also experience fascination, interest or

curiosity when faced with negative events (Zuckerman and Litle, 1986; Rimé et al., 2005; Turner and Silvia, 2006). These states have received no attention in affective neuroscience so far, even though comparing the neural representation of fascination with those of fear and disgust provides an important opportunity to examine how the brain instantiates different experiences that contain the same affective (i.e. valenced) information. In this article, we present a novel neuroimaging study

Received: 26 June 2014; Revised: 30 June 2015; Accepted: 7 July 2015

© The Author (2015). Published by Oxford University Press. For Permissions, please email: journals.permissions@oup.com

to assess the neural mechanisms that underlie labeling similar negative stimuli as evoking fear, disgust or 'morbid fascination'.

The brain basis of typical and atypical emotional experiences

The Conceptual Act Theory of emotion (CAT; Barrett, 2006) explicitly predicts that the same affective stimulus (e.g. a negative image) can lead to different subjective experiences of emotion (e.g. fear, disgust, fascination) depending on a number of individual and situational factors (Barrett, 2006; 2012; Lindquist and Barrett, 2008). One underlying process that causes diversity in subjective emotional experiences is called conceptualization, and refers to the retrieval of context-relevant conceptual knowledge derived from previous experiences that shapes the way that internal and external sensations are made meaningful (cf. Barrett, 2013). Indeed, the CAT conceives of conceptualization as an essential 'ingredient' in emotional experience, along with sensory information from the world (exteroceptive sensations) and sensory information from the body (interoceptive sensations) (Lindquist and Barrett, 2012).

In this study, we focused on how different conceptualizations of negative stimuli are supported by the 'neural reference space for discrete emotion', a set of regions that is consistently active when people experience discrete emotions (Kober et al., 2008; Lindquist et al., 2012). We specifically predicted involvement of the dorsomedial and ventromedial prefrontal cortex (dmPFC and vmPFC) and the ventrolateral prefrontal cortex (vlPFC), because the CAT argues that these regions support a conceptualization function in emotion (Wilson-Mendenhall et al., 2011; Lindquist and Barrett, 2012; Barrett and Satpute, 2013). Furthermore, there is ample research that demonstrates a role for these regions when people generate subjective emotional experiences (e.g. Wilson-Mendenhall et al., 2011; Oosterwijk et al., 2012) or label/appraise affective or neutral stimuli in emotional terms (e.g. Lieberman et al., 2007; Ochsner et al., 2009; Satpute et al., 2013).

Interestingly, a recent study has further shown that these same regions engage with relatively different strength when individuals generate experiences of atypical as compared with typical experiences of emotion (Wilson-Mendenhall et al., 2014). Wilson-Mendenhall et al. manipulated typicality by shifting the valence of instances of fear, sadness and happiness (e.g. feeling 'disturbing happiness' when making a sarcastic jab at a colleague or 'energizing fear' when performing at an important sports match), and demonstrated that atypical experiences relied more heavily upon regions that support a conceptualization function. In this article we manipulate typicality in a different way. We propose that fear and disgust are likely to be typical ways of conceptualizing one's reaction to a negative image displaying mutilated bodies or crawling insects, whereas morbid fascination, in contrast, is likely to be a more atypical way of conceptualizing one's reaction. In the case of negative stimuli, morbid fascination is atypical because it combines negativity with an intention to approach and explore. Moreover, although morbid fascination may often be expressed in daily life (e.g. 'rubber-necking' on a freeway to view a grisly accident), the term itself may not be often used to describe experiences of negative events. At the very least, morbid fascination is not a 'basic-level' category of emotion that is learned early in childhood and regularly used to describe modal unpleasant emotional experiences.

According to the CAT, conceptualization transforms non-discrete affective experiences into experiences of discrete emotions (such as fear, disgust or fascination) by integrating

internal sensations, external sensations and context-relevant conceptual knowledge (Barrett, 2006, 2012). We hypothesize that the atypical label of morbid fascination will draw upon the process of conceptualization in a different way than the typical labels of fear and disgust. First of all, at an implicit level, fear and disgust are 'efficient' conceptualizations. This means that the brain has a large sample of prior experiences that it can draw from to integrate these labels with the sensory information at hand. The label of morbid fascination, in contrast, requires more elaborative conceptual processing because it is used less often, and therefore has a smaller sample of 'priors' to draw from. In addition, morbid fascination suggests a reconciliation between a tendency to explore the informational content of the stimulus and its negativity. In this context, sensations may take on a more complex meaning, causing individuals to continue to draw on conceptual information to iteratively process the meaning of internal and external information.

The present study

We introduce a novel paradigm that presents false feedback to direct the way participants conceptualize their experience of a negative image. False feedback has been used across multiple psychological contexts, including in social, clinical and neuroscientific studies. Previous applications of false feedback have predominantly focused on false auditory feedback about the participant's heart rate (e.g. Valins, 1966; Wild et al., 2008). For instance, Gray et al. (2007) found that false heart rate feedback influenced participants' affective judgments of neutral faces and led to corresponding changes in the anterior insula and amygdala. In the present experiment we did not provide participants with false feedback about their physiological state. Instead, inspired by recent developments in the decoding of mental states using brain imaging data (e.g. Haynes and Rees, 2006; Shirer et al., 2012) we led participants to believe that we could decode the psychological meaning of their brain states in real time. Subsequently, we gave participants false feedback about the 'content' of their mental experiences when viewing highly arousing, negative images. This allowed us to provide participants with different labels (i.e. fear, disgust, morbid fascination) to prime them to conceptualize their experience in a certain way, without asking them to actively judge the images. In the control condition, we told participants that the acquired neural pattern precluded categorization of their mental state. We chose to incorporate a non-specific label in the control condition so that the presence of feedback was constant across conditions—including the possibility for participants to reflect upon the categorization of their state.

With this paradigm we tested the following hypotheses. First, we tested whether regions in the neural reference space for emotion (e.g. dmPFC, vmPFC, vlPFC) that support the process of conceptualization engaged more strongly after participants received emotion feedback (i.e. fear, disgust, morbid fascination) compared with the control condition. Second, based on the previous findings regarding the role of conceptualization in atypical emotional experiences (Wilson-Mendenhall et al., 2014), we tested whether these regions engaged more strongly when people received morbid fascination feedback (i.e. an atypical state) as compared with fear or disgust feedback (i.e. typical states).

Materials and methods

Participants

Participants were recruited from the Greater Boston Area through an online advertisement. Potential participants were

excluded if they had a history of psychiatric illness, were using psychoactive or systemic medications or if MRI contraindications were present. Twenty-seven participants gave informed consent according to the Partners Health Care Institutional Review Board. Two participants did not complete the scanning session because they felt uncomfortable in the scanner, and one participant was excluded because of movement artifacts (>3 mm motion across multiple time points across multiple runs). Analyses were performed on the remaining twenty-four participants (12 females, $M_{\text{age}} = 24.7$, $SD_{\text{age}} = 3.8$). Participants were paid \$150 for their participation.

Manipulation

We manipulated experiential labels by telling participants (falsely) that we could 'read-out' their mental experience while they were viewing negative images in the scanner and that we would give them feedback on this categorization. Three experimental conditions combined negative images with disgust feedback, fear feedback and morbid fascination feedback to direct conceptualization with a typical label (as 'fear' or 'disgust') or with an atypical label (as 'morbid fascination'). The 'could not be calculated' condition served as a control condition in that it kept the presence of a feedback cue constant, without providing a label to direct conceptualization.

Our instruction proceeded as follows. First, we explained to participants that we had 'identified three specific brain states that reliably predicted whether someone was experiencing fear, disgust or morbid fascination' and that the purpose of the study was to validate these profiles. Then, we explained that the experiment would start with three 'collection runs' to 'capture your brain state while you view the pictures', followed by three 'feedback runs' that would give feedback about 'what you were experiencing when you saw the image for the first time'. We further mentioned that when our algorithm could not find a match, participants would receive 'could not be calculated' feedback. Furthermore, all participants were instructed to 'prolong the state that you initially experienced when viewing the picture' (e.g. during the collection runs), 'during the full duration of the second viewing' (e.g. during the feedback runs). The full instruction can be found in the Supplementary Materials. Because the manipulation incorporates deception, we used a thorough debriefing procedure.

Procedure

The complete experiment consisted of two sessions ~5 days apart. In the first session the participants viewed and rated all images presented in the study. In the second session the participants performed the experimental task in the scanner, viewed and rated all images again and filled out questionnaires and an exit interview. The rating task was programmed in MATLAB and presented each image in combination with the following five rating dimensions: threat, repulsion, interest, negativity and emotional intensity. Ratings were made on a continuous sliding scale with labeled endpoints (e.g., 'not at all threatening' and 'extremely threatening'). To avoid response demand characteristics, we deliberately chose words that indexed 'world-focused' experiences of each emotion (Lindquist and Barrett, 2008), that were different from the terms used in the feedback manipulation.

The scanning session consisted of six runs. In three passive viewing runs (or 'collection runs') participants viewed 96 negative images, 24 positive images and 24 neutral images passively.

Positive and neutral images were intermixed with negative images to prevent habituation to the highly aversive images. Each trial presented a fixation cross for 3 s, followed by an image for 3 s. Several 3–12 s jitter periods presenting a black screen were intermixed with the trials. Jitter length and the order of stimulus events was optimized using optseq (<http://surfer.nmr.mgh.harvard.edu/optseq/>).

In three 'feedback runs', participants viewed 96 negative images with feedback, and 24 positive and 24 neutral images without feedback. Each trial started with a cue presented for 3 s. Negative images were combined with a feedback cue that combined a label (i.e. fear, disgust, morbid fascination or control) with an artificially created image of brain activation (see section on Stimulus Materials). Positive and neutral images were preceded by a 'no feedback' cue. Cues were followed by an image presented for 3 s and 3–12 s jitter periods were intermixed with the trials. In order to separately model the image-viewing period from the feedback cue period, we included 24 catch trials that presented feedback cues without a following image. See Figure 1 for a visual overview of the paradigm.

To ensure that participants kept attending to the images, participants were asked to press a button when the image on the screen was surrounded by a red box. The red box appeared twice per run around positive and neutral images. All runs were programmed and presented in Eprime (Psychology Software Tools).

The scanning session was followed by a second world-focused emotion rating task. We reasoned that participants who received morbid fascination feedback, for instance, would be more likely to later rate those images as 'interesting' as compared with 'threatening' or 'repulsive'. We did not include trial-by-trial ratings of world-focused emotions or feedback accuracy, as this would have confounded the feedback manipulation with an explicit judgement of the stimulus, or an explicit evaluation of the label. Moreover, such an explicit evaluation may have undermined our cover story. In addition to including post-scanner ratings of world-focused emotions, we concluded the session with a thorough exit interview to evaluate whether participants believed that the feedback correctly identified their mental state. We asked participants to 'judge the percentage of trials for which the feedback correctly identified your state' on an 11-point scale ranging from 0 to 100% for each of the experimental feedback categories. Finally, the experimenter asked the participant about the goal of the study. Only six participants showed suspicion about the falsity of the feedback manipulation. As we indicate below, our results did not differ as a result of participants' suspicion.

Stimulus materials

Images with a valence score <4.5 (i.e. negative) and an arousal score >4.5 (i.e. high arousal) were selected from the International Affective Pictures System database (Lang et al., 2008). In order to exclude the possibility that differences in strength and content between conditions could influence our effects, we organized the images into four different lists that were counterbalanced across the fear, disgust, morbid fascination and control feedback conditions. Lists did not differ in mean valence, $F(3, 95) = 0.70$; $P = 0.55$, ($M_{\text{list1}} = 2.8$, $SE = 0.16$; $M_{\text{list2}} = 2.7$, $SE = 0.16$; $M_{\text{list3}} = 2.8$, $SE = 0.16$; $M_{\text{list4}} = 3.0$, $SE = 0.14$), nor in mean arousal, $F(3, 95) = 0.72$; $P = 0.54$, ($M_{\text{list1}} = 5.9$, $SE = 0.11$; $M_{\text{list2}} = 6.0$, $SE = 0.12$; $M_{\text{list3}} = 5.8$, $SE = 0.12$; $M_{\text{list4}} = 5.7$, $SE = 0.16$). Furthermore, lists were fully matched in terms of specific content. Each list included an equal number of images

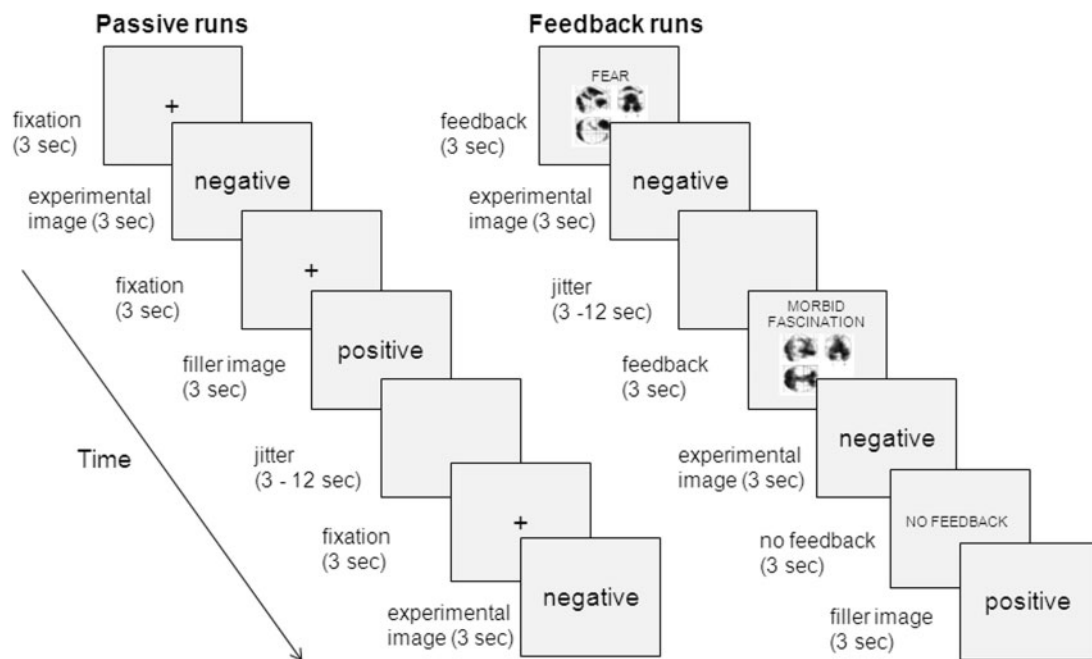


Fig. 1. Overview of the false feedback design. During the passive viewing runs, participants passively viewed negative, positive and neutral images. During the feedback runs, participants received false feedback (i.e. 'fear', 'disgust', 'morbid fascination' or 'could not be calculated') preceding each negative image. Positive and neutral images were combined with a preceding 'no feedback' cue.

portraying hand or head mutilations, burns, dead bodies, conflict scenes, skulls, attacking animals, weapons, car crashes, etc.

Feedback images were created in Photoshop. 'Neural activity' associated with the experimental feedback manipulations was mimicked in three different patterns on a glass brain image (see the Supplementary Materials for examples).

Scan parameters

Data were collected with a 3 T Siemens Magnetom trio MR scanner with a 12-channel matrix head coil. We acquired a structural T1-weighted multi-echo MPRAGE image (van der Kouwe et al., 2008) (TR = 2530 ms, TE1 = 1.74 ms, TE2 = 3.6 ms, TE3 = 5.46 ms, TE4 = 7.32 ms, flip angle = 7°, 1 mm³ voxels) for structural registration and 816 T2* weighted functional images (TR = 3000 ms, TE = 30 ms, flip angle = 90°, FOV = 220 mm, 3.4 × 3.4 × 2 mm voxels) across six different runs.

Data analysis

The structural data were preprocessed using the standard Freesurfer (<http://surfer.nmr.mgh.harvard.edu/>) protocol (Dale et al., 1999; Fischl et al., 1999a,b). Functional pre-processing included motion correction, slice time correction, spatial smoothing (5-mm full-width/half-max), high-pass filtering (1/128 Hz) and registration of the functional images onto the anatomical scan. Hemodynamic responses were modeled with a gamma function; the design matrix included motion parameters as nuisance regressors. Time-points with movement exceeding 2 mm were excluded from the analysis. Through the inclusion of catch trials in the feedback runs, we were able to separately model the feedback phase and the image phase within each trial; analyses only focused on the neural responses when participants were viewing negative images. A random effects analysis on the group level was performed to calculate contrasts.

Hypotheses testing

To test our hypotheses we first created a mask representing 'the neural reference space for discrete emotion' (cf. Lindquist et al., 2012) that includes regions consistently active during emotion experience. This mask combined the following surface-based labels from the Desikan-Killiany cortical atlas (Desikan et al., 2006): superior frontal, medial orbitofrontal, rostral and caudal anterior cingulate, pars orbitalis, pars opercularis, pars triangularis, lateral orbitofrontal and insula. In the first step, we identified significant clusters within the mask that reflected general task activity. We selected clusters that survived a threshold provided by a Monte Carlo simulation within the contrast that compared all conditions against a baseline consisting of the blank screen jitter periods. In the second step, we extracted percent signal change from each of these functionally defined regions of interest (ROIs) for each of the feedback conditions separately.

For cluster selection we implemented a so-called leave-one-out procedure (Esterman, Tamber-Rosenau et al., 2010) to refrain from introducing circularity in our ROI analysis (Kriegeskorte et al., 2009). More precisely, we created 24 'n-1' datasets that included all participants minus one. Then, for each n-1 dataset, we performed a Monte Carlo simulation to select functional clusters that reflected general task activity. All 24 datasets revealed significant clusters in the left dmPFC/supplementary motor area (SMA) and left vlPFC (vertex-wise $P < 0.005$; cluster-wise $P < 0.01$), and in the left IOFC (vertex-wise $P < 0.005$; cluster-wise $P < 0.05$). In the next step, for each participant that was left out, we extracted percent signal change for each condition separately from the functional ROIs selected from the corresponding n-1 dataset (i.e. the dataset in which the participant was left out). This data were then subjected to three separate repeated measures ANOVAs (Bonferroni-corrected) that tested whether the extracted percent signal change was significantly different between feedback conditions. Because of the novel nature of our paradigm, we report uncorrected P-values for the

follow-up paired comparisons and note when these comparisons reached a Bonferroni-corrected level of significance.

Because previous research has shown decreased activation of the amygdala when people label emotional stimuli (Hariri et al., 2002; Lieberman et al., 2007), we also performed a functional ROI analysis focused on the amygdala. This analysis did not demonstrate any significant effects (see Supplementary Materials). Furthermore, we performed exploratory whole brain analyses for the contrasts comparing the feedback conditions among each other, and each feedback condition with its corresponding passive viewing trials. To balance type I and type II error (Lieberman and Cunningham, 2009), we applied a cluster-search threshold that identified clusters with a vertex-wise threshold of $P < 0.005$ and a minimum cluster-size of 50 mm².

Results

Accuracy judgments

In a first manipulation check, we analyzed the judged percentage of trials where participants believed the emotion feedback correctly identified their mental state. This analysis demonstrated that judged accuracy was equally high for fear ($M = 67.9\%$; $SE = 4.7$), disgust ($M = 73.8\%$; $SE = 3.6$) and morbid fascination ($M = 67.1\%$; $SE = 4.8$) feedback, $F(2, 46) = 1.18$; $P = 0.32$. This finding not only shows that our manipulation was successful, but also that participants, on average, had few reservations in applying different labels to very similar evocative images.

Subjective ratings

In a second manipulation check, we analyzed the ratings of threat, repulsion and interest collected at Time 1 and 2 to test whether feedback influenced ratings on the corresponding 'world-focused emotion' dimension (e.g. higher interest ratings for images paired with morbid fascination feedback). A 2 (time) \times 3 (rating) \times 4 (condition) repeated measures ANOVA did not reveal any significant effects (all P 's > 0.077), although the means did show that ratings decreased from Time 1 to Time 2 in absolute terms. This is likely the result of habituation because subjects viewed the images for the first time at Time 1 (outside the scanner), and for the fourth time at Time 2 (after viewing the images twice in the scanner during collection and feedback runs).

Because habituation may have obscured the effect of the feedback manipulation in our overall analysis, we also performed a 3 (rating) \times 4 (condition) repeated measures ANOVA separately on the ratings at Time 1 and Time 2. Importantly, there was neither a main effect of rating, $F(2, 44) = 1.96$; $P = 0.15$ nor an interaction between rating category and condition at Time 1, $F < 1$. These findings suggest that on average, participants endorsed threat ($M = 50.4$; $SE = 4.6$), repulsion ($M = 56.7$; $SE = 3.8$) and interest ratings ($M = 54.2$; $SE = 4.6$) equally at the start of the experiment. Critically, we did find a significant effect of the false feedback manipulation at Time 2, indicated by an interaction between rating category and condition, $F(6, 132) = 2.62$; $P = 0.02$; $\eta^2 = 0.11$. As shown in Figure 2, images combined with fear feedback produced the strongest threat ratings; images combined with disgust feedback produced the strongest repulsion ratings and images combined with morbid fascination feedback produced the strongest interest ratings. Paired comparisons showed significant differences between fear and morbid fascination and fear and control for threat, and between morbid fascination and control for interest ($P < 0.05$, uncorrected).

ROI analyses

We tested the hypothesis that a negative image combined with morbid fascination feedback would be associated with stronger activation in the neural reference space for emotion than a negative image combined with fear, disgust or control feedback. First, based on an iterated leave-one-out Monte Carlo corrected cluster search, we identified significant clusters in the dmPFC/SMA, IOFC and vlPFC that reflected general task activity. Figure 3 contains a visual representation of the functionally defined ROIs by presenting the conjunction of clusters produced by the iterated leave-one-out cluster search. Table 1 contains the peak coordinates and size of the clusters within this conjunction.

In the second step, we extracted percent signal change from each of the functionally defined ROIs for each of the feedback conditions separately. Figure 3 presents the means and standard errors for each feedback condition. As predicted, we found that the left dmPFC/SMA was significantly modulated by the feedback manipulation, $F(3, 69) = 3.92$; $P = 0.036$; $\eta^2 = 0.15$. Consistent with our hypothesis, the dmPFC/SMA was more strongly engaged after morbid fascination feedback than fear feedback ($P = 0.033$), disgust feedback ($P = 0.036$) and control feedback ($P = 0.003$). Disgust and fear feedback did not differ significantly from the control condition (both P 's > 0.21). A similar pattern was present in the left IOFC, $F(3, 69) = 4.52$; $P = 0.018$; $\eta^2 = 0.16$, with stronger activation after morbid fascination than after fear feedback ($P = 0.022$) and control feedback ($P = 0.001$). The comparison between morbid fascination feedback and disgust feedback was marginally significant ($P = 0.074$). Disgust and fear feedback did not differ significantly from the control condition (both P 's > 0.12). The modulation of the left vlPFC by the feedback manipulation was marginally significant, $F(3, 69) = 3.53$; $P = 0.057$; $\eta^2 = 0.13$. Paired comparisons demonstrated that the vlPFC engaged significantly more after morbid fascination feedback ($P = 0.004$), fear feedback ($P = 0.014$) and disgust feedback ($P = 0.051$) than after control feedback. Disgust and fear feedback did not differ significantly from morbid fascination feedback (both P 's > 0.61). For each of the above functional ROIs, the comparison between morbid fascination and control feedback also passed a Bonferroni corrected level of significance.

Interaction with accuracy judgments

A possible alternative explanation for our findings concerns the individual variability in how people judged the accuracy of the

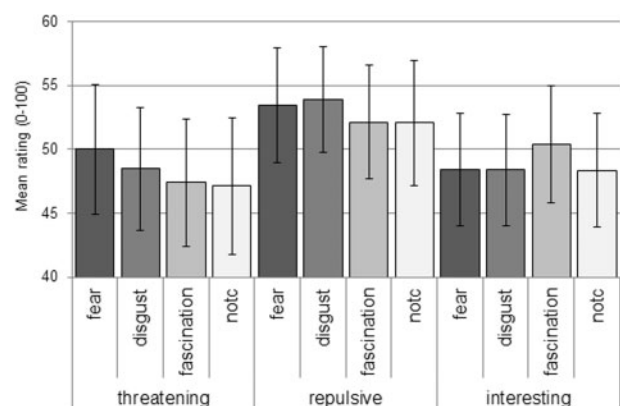


Fig. 2. Interaction between feedback condition (x-axis, vertical labels) and rating category (x-axis, horizontal labels). Error bars represent standard errors.

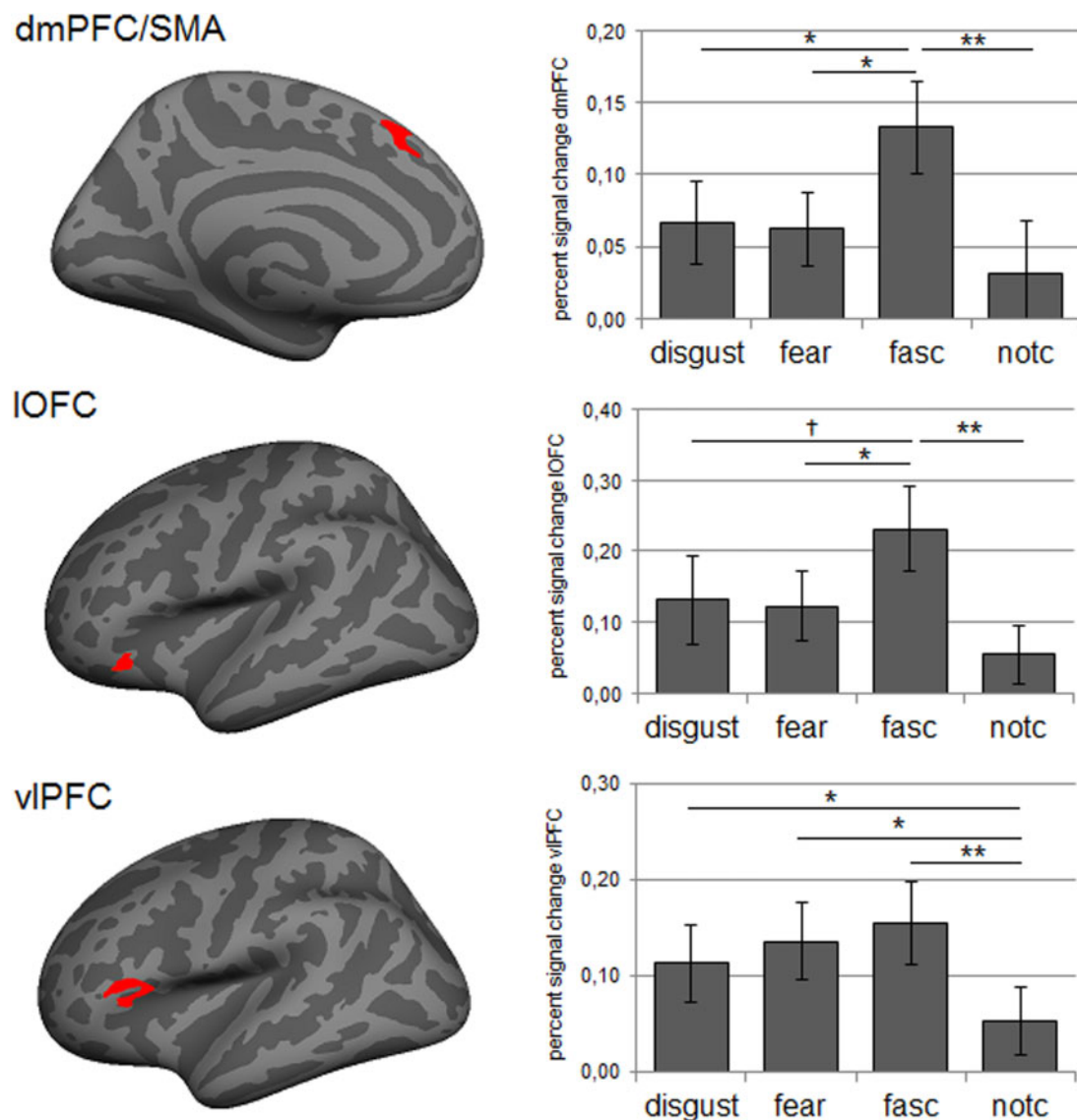


Fig. 3. Left: visual representation of functional ROIs (conjunction) in dmPFC/SMA, IOFC and vIPFC activated by the task. Right: mean percent signal change (error bars represent standard errors) for images paired with disgust, fear, morbid fascination (fasc) and could not be calculated (notc) feedback, extracted from functional ROIs produced by the iterated leave-one-out cluster search. [†] $P < 0.1$; * $P \leq 0.05$; ** $P < 0.01$.

Table 1. Representation of functional ROIs within neural reference space for emotion

Region	hemi	max	Size mm ²	x	y	z	k
dmPFC/SMA	lh	2.81	140	-7.3	28.5	42.4	241
IOFC	lh	3.48	105	-42.5	26.2	-11.5	216
vIPFC	lh	3.43	286	-51.6	27.6	5.2	520

Note: Clusters reflect a conjunction across clusters produced by the iterated leave-one-out Monte Carlo corrected cluster search. For a visual representation of the clusters see Figure 3. k is size in vertices; coordinates are in Talairach space.

feedback. The patterns of neural activity that we found could have been driven solely by the participants with low accuracy reports, possibly because they were confused by the feedback. In order to test this, we split our sample into two groups. One group included participants that judged that 'all' feedback was correct in 70% or more of the trials ($n = 11$). The other group

included participants that thought that 'one or more' of the feedback types was correct in 60% or less of the trials ($n = 13$). It is important to note that this latter group included all participants that showed any suspicion about the feedback manipulation. We then re-ran all ROI analyses with group membership as a between-subjects factor. None of analyses showed a significant interaction effect (all $P > 0.23$, uncorrected).

Whole brain contrasts feedback

In addition to the ROI analyses, we also calculated whole brain contrasts comparing the different categories of feedback directly (see Table 2 for an overview). Consistent with the ROI analyses, we found that morbid fascination feedback activated the vIPFC, dmPFC and IOFC more than control feedback. Furthermore, the contrast between disgust feedback and control feedback revealed a cluster in the left vIPFC. Comparing the control feedback to the emotion feedback conditions, we consistently found clusters in the right temporoparietal junction (TPJ)

and the frontal pole. The contrasts that compared emotion feedback conditions with one another demonstrated clusters in the superior parietal lobule and fusiform gyrus for morbid fascination vs fear feedback; clusters in the precuneus for morbid fascination vs disgust feedback; and clusters in the precentral gyrus for fear vs disgust feedback.

Whole brain contrasts with passive viewing

As a final examination of the effect of feedback, we compared activation patterns after each feedback condition with activation patterns while people 'passively' viewed the same set of images (i.e. during the 'collection' runs) (see Table 3 for an overview). Consistent with the ROI analyses, we found that both disgust feedback and morbid fascination feedback showed increased activation within the left vlPFC, compared with passively viewing the same

images. Morbid fascination feedback further demonstrated clusters in the left dmPFC, dlPFC, IOFC, TPJ and precuneus (see Figure 4). When comparing control feedback with passive viewing, we found clusters of activation in the TPJ and precuneus. For passive viewing, we consistently found clusters in bilateral occipital lobe (e.g. cuneus, pericalcarine and lingual gyrus) and superior parietal lobule across all comparisons. This may reflect stronger visual processing and spatial orientation when viewing images for the first time in the scanner. In addition, for all conditions except for morbid fascination, we found that passive viewing was associated with increased activation in the right IOFC.

Discussion

This study documents novel evidence for the neural mechanism that supports morbid fascination in reaction to negative

Table 2. Contrasts comparing emotion and control feedback

Contrast	hemi	max	Size mm ²	x	Y	z	k	Region
fear feedback vs control feedback	lh	no clusters						
	rh	2.94	64	4.5	-35.3	30.5	145	Precentral
disgust feedback vs control feedback	lh	3.59	78	-49.6	0.7	39.1	135	Precentral
		3.48	89	-36.2	31.2	-5.2	198	vlPFC
	rh	3.28	87	57.8	-18.5	4.1	167	Superior temporal
		2.99	52	46.1	30.6	-3.5	83	vlPFC
Fascination feedback vs control feedback	lh	5.25	55	-6.9	25.3	48.9	106	dmPC
		4.08	59	-36.4	30	-5.8	112	vlPFC
		3.98	176	-34.9	-48.5	56.2	443	Superior parietal
		3.59	102	-44.9	30.1	7.8	116	vlPFC
		3.04	85	-47.8	21.8	5.7	150	vlPFC
		2.73	155	-36.2	36.6	-10.3	301	IOFC
	rh	3.5	80	31.8	-6.1	43.9	180	Precentral
		3.35	169	26.9	-50.5	46.9	376	Superior parietal
		2.58	50	37.6	5	-3	94	Insula
	lh	no clusters						
Control feedback vs. fear feedback	rh	-4.2	60	4.5	-35.3	30.5	145	Posterior cingulate
		-3.08	55	56	-43.5	-7.4	87	Middle temporal
		-3.07	76	23.8	56.4	3.8	94	Frontal pole
		-2.94	295	35	-67.8	41.7	554	IPL/TPJ
		-2.62	59	48.5	-49.9	30.9	117	IPL/TPJ
Control feedback vs disgust feedback	lh	no clusters						
	rh	-3.89	72	49.1	-49.8	31.1	145	IPL/TPJ
		-3.4	105	56.8	-42.6	-7	160	Middle temporal
		-2.95	54	12.1	-53.8	20.6	109	Precuneus
		-2.75	50	23.5	56.7	3.1	61	Frontal pole
control feedback vs fascination feedback	lh	no clusters						
	rh	-3.21	61	24	57	5.1	70	Frontal pole
fear feedback vs disgust feedback	lh	-2.77	86	37.3	-62.9	46.5	164	IPL/TPJ
	rh	3.41	68	-35.2	-19.1	47.2	153	Precentral
fear feedback vs fascination feedback	lh	no clusters						
	rh	no clusters						
Disgust feedback vs fear feedback	lh	no clusters						
	rh	no clusters						
Disgust feedback vs. fascination feedback	lh	no clusters						
	rh	no clusters						
Fascination feedback vs fear feedback	lh	no clusters						
	rh	-4.31	97	16.7	-62.2	51.8	168	Superior parietal
		-2.97	71	28.3	-58.9	-7.6	132	Fusiform
Fascination feedback vs disgust feedback	lh	no clusters						
	rh	-2.38	52	8.7	-55.8	21.4	85	Precuneus

Note: Clusters are significant at $P < 0.005$ with a minimum cluster size of 50 mm²; k is size in vertices; coordinates are in Talairach space.

IOFC, lateral orbitofrontal cortex; dmPFC, dorsomedial prefrontal cortex; vlPFC, ventrolateral prefrontal cortex; TPJ, temporoparietal junction; IPL, inferior parietal lobe.

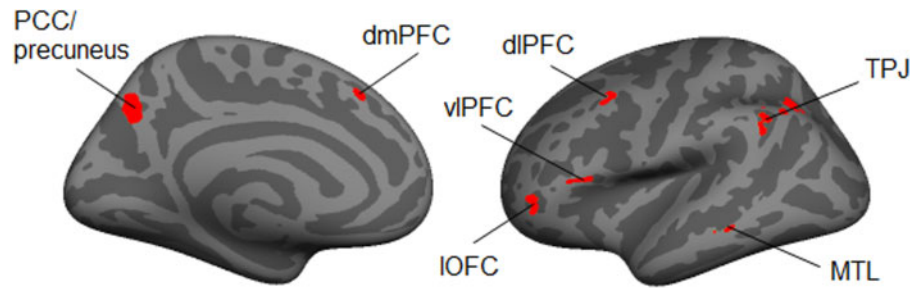


Fig. 4. Significant clusters ($P < 0.005$; size $> 50 \text{ mm}^3$) when comparing viewing after fascination feedback with passively viewing the same images. PCC, posterior cingulate cortex; dmPFC, dorsomedial prefrontal cortex; dlPFC, dorsolateral prefrontal cortex; vlPFC, ventrolateral prefrontal cortex; IOFC, lateral orbitofrontal cortex; TPJ, temporoparietal junction; MTL, middle temporal lobe.

images. To manipulate the way participants conceptualized a negative image, we used a novel false feedback paradigm in which we claimed to 'decode' participants' psychological state from their brain activity. In reality, we used the feedback as a form of suggestion to direct how participants 'might' conceptualize their experience of negative images in the moment. This method allowed us to directly compare the relative involvement of neural regions following morbid fascination, fear and disgust feedback.

Behavioral results suggest that our manipulation was successful: participants reported that the feedback was correct on the majority of trials. Moreover, an analysis of ratings after scanning indicated that false feedback influenced participants' subjective experience of the negative stimuli. This difference was subtle, but given that these ratings occurred well after the manipulation, they may be a conservative estimate of the shift in participants' experience caused by the feedback manipulation. It should be noted that we deliberately decided not to include trial-by-trial ratings of emotional experience or feedback accuracy during our task, as this would have confounded the feedback manipulation with the explicit evaluation of a stimulus or a label. These limitations notwithstanding, we believe that the subtle change in participants' world-focused emotional experiences, along with participants' general agreement about the accuracy of the feedback, are suggestive that our false feedback influenced participants' emotional reactions, just as priming emotion concepts shapes emotional experiences in the behavioral literature (Schachter and Singer, 1962; Lindquist and Barrett, 2008; Oosterwijk et al., 2010; Crum et al., 2013; Kassam and Mendes, 2013).

Our results reflected different patterns of activation in the neural reference space for discrete emotions, notably within regions associated with conceptualization (i.e. the dmPFC/SMA, IOFC and vlPFC). Consistent with our first hypothesis, the vlPFC showed a trend towards increased activation following emotion label feedback (i.e. fear, disgust, morbid fascination) as compared with control feedback. This finding is consistent with research observing vlPFC activation during labeling of emotional feelings (i.e. 'affect labeling'; Lieberman et al., 2007; Satpute et al., 2013), semantic retrieval more generally (Wagner et al., 2001) and with constructionist predictions that language plays a role in emotion production (Barrett, 2006; Lindquist et al., 2015). Consistent with our second hypothesis, the dmPFC/SMA showed increased activation following morbid fascination feedback, as compared with all other conditions. The dmPFC has been implicated when individuals show increased attention to their internal state (Satpute et al., 2013) and is also part of a distributed network of brain regions involved in representing semantic knowledge (Binder et al., 2009). We thus interpret this finding in line with the proposal of the CAT (Barrett, 2006, 2009)

that 'atypical' emotions prompt a relatively greater reliance on conceptualization (Wilson-Mendenhall et al., 2014).

As a potential limitation, we should note that only the comparison between morbid fascination feedback and control feedback was significant at the most stringent level in our ROI analysis. Whole brain contrasts comparing morbid fascination feedback to control feedback and to passive viewing forwarded highly similar results. This suggests that our strongest effects result from comparing an atypical emotion label with the control condition, which did not explicitly require participants to draw on knowledge about emotion categories. Nevertheless, we chose to present and discuss all findings since we worked with a novel paradigm that likely produced a subtle effect, both behaviorally and neurally. Participants were in no way instructed to apply any feedback label or judge its accuracy during scanning, which may have resulted in less robust patterns of activation than a more explicit instruction. These points notwithstanding, the brain activity we observed in the functional ROI analyses largely replicated the brain areas that are consistently observed in emotion meta-analyses (Vytal and Hamann, 2010; Lindquist et al., 2012) and in studies on affect labeling (Lieberman et al., 2007; Satpute et al., 2013), suggesting that our findings are reliable.

Consistent with our interpretation that morbid fascination, as an atypical emotional state, involved increased conceptualization, the dmPFC/SMA, vlPFC and IOFC clusters that demonstrated strong involvement for morbid fascination feedback fell within the boundaries of a so-called 'default mode network' (DMN) (i.e. Yeo et al., 2011; see also Oosterwijk et al., 2012). The DMN is thought to support conceptualization by using representations of prior experience to make meaning of sensations in the moment (Bar, 2009; Lindquist et al., 2012; Oosterwijk et al., 2015). In line with this hypothesis, the DMN has been shown to have increased task-related functional activity across a number of different psychological processes, such as self-referential processing (e.g. Kelley et al., 2001; Gusnard et al., 2001), semantic judgments (e.g. Binder et al., 2009), and emotional experience (Wilson-Mendenhall et al., 2011; Lindquist et al., 2012; Oosterwijk et al., 2012). The suggestion that the default network may be important in the construction of complex and atypical states such as morbid fascination, is consistent with work that demonstrates DMN activation for other types of information that reflect atypicality, such as ambiguous or ambivalent stimuli (Jenkins and Mitchell, 2010; Nohlen et al., 2013). The possible link between the DMN and the experience of fascination in general is a topic of investigation that we aim to address in future research.

Although we interpret our findings in terms of an increased role for conceptualization in the experience of morbid

Table 3. Contrasts comparing activation after feedback with activation during passive viewing

Contrast	hemi	max	Size mm ²	x	y	z	k	Region
fear feedback vs passive viewing	lh	No clusters						
	rh	3.17	54	35.8	17.9	31	115	Caudal middle frontal
		2.54	52	47.6	-53.7	44.1	110	Inferior parietal
passive viewing vs fear feedback	lh	-6.36	2973	-26.6	-62.8	7.2	5233	Cuneus/pericalcarine/lingual
		-4.72	143	-19.9	-60.3	53.7	322	Superior parietal
		-2.73	66	-24.1	-43.7	-3.5	151	Lingual
		-2.31	57	-29.2	-65	25.5	140	Inferior parietal
	rh	-7.62	2835	20.1	-69.8	9.2	4815	Pericalcarine
		-3.34	97	44.8	32.5	4.7	166	Pars triangularis
		-3.31	77	20.7	-56.1	49.8	151	Superior parietal
		-3.21	66	30.9	25.8	-16.6	112	IOFC
		-3.06	56	22.6	7.9	56.9	102	Superior frontal
		-2.79	70	26.7	-62	29.3	119	Superior parietal
	lh	2.92	63	-39	38	2	86	dIPFC/vIPFC
disgust feedback vs passive viewing	rh	3.55	109	12	-62	35.9	217	Precuneus
		2.84	58	8.4	39.3	1.8	107	Rostral anterior cingulate
passive viewing vs disgust feedback	lh	-7.73	3765	-3.2	-72.4	17.2	6475	Cuneus/pericalcarine/lingual
		-3.54	182	-17	-61.9	56.8	409	Superior parietal
		-2.99	148	-33.1	-74	24	221	Inferior parietal
		-2.53	90	-19.7	-59.8	37.7	189	Superior parietal
	rh	-6.57	3575	18.7	-34.1	-6.1	6141	Parahippocampal
		-4.21	384	25.9	-52.3	46.4	932	Superior parietal
		-3.04	106	34.5	-75	21.8	169	Inferior parietal
		-3.01	80	32.8	23.9	-17.4	140	IOFC
		-2.98	86	13	-75.6	-0.5	95	Lingual
		-2.87	203	43.2	-70.8	19.8	318	Inferior parietal
		-2.85	132	21.4	-80.1	24.7	195	Superior parietal
		-2.75	102	37	-36.8	-9.3	215	Parahippocampal
	lh	-2.64	97	32.5	-64.7	26.7	183	Inferior parietal
		4.71	249	-5.7	-64.4	40.4	488	Precuneus
		4.23	251	-31.6	-55.9	37.4	531	IPL/TPJ
fascination feedback vs. passive viewing		3.5	209	-48.4	-47.5	33.5	479	TPJ
		3.5	178	-40.1	42.7	-8.2	240	IOFC
		3.35	61	-61.1	-35.9	-6.9	106	Middle temporal
		3.26	60	-46.9	-58	41.7	151	IPL/TPJ
		3.23	109	-51.6	27.6	5.2	186	vIPFC
		3.2	128	-38.4	14.9	46.6	201	dIPFC
		3.15	74	-6.7	26.8	39.9	123	dmPFC
	rh	4.7	249	7.2	-65.8	38.5	510	Precuneus
		3.75	200	39.8	-53.4	40.8	443	Inferior parietal
		2.42	63	14	-90.3	16.2	80	Lateral occipital
	lh	-7.37	3124	-26.4	-60.1	7.5	5542	Precuneus
		-3.64	142	-22.2	-59	52	320	Superior parietal
	rh	-7.84	2722	16.7	-65.1	12.2	4494	Pericalcarine
		-2.53	83	21.5	-77.6	41.7	123	Superior parietal
notc feedback vs. passive viewing	lh	3.8	142	-37.8	-63.8	45.6	293	IPL/TPJ
		3.59	294	-48.1	-55.5	40.1	692	IPL/TPJ
		3.2	66	-40.6	13.7	38.6	86	dIPFC
		3.17	116	-46.4	-54.6	26.7	257	IPL/TPJ
		3.06	70	-11.3	-49.3	37.7	134	Precuneus
		2.85	129	-5.7	-63.3	39.2	256	Precuneus
	rh	3.93	731	42.7	-59.1	41.5	1524	Inferior parietal
		3.46	60	38.6	23	33.1	120	Caudal middle frontal
		3.45	82	9.1	-66.4	39.3	170	Precuneus
passive viewing vs. notc feedback		3.41	73	17.1	-96.5	13.4	98	Lateral occipital
	lh	-7.61	5305	-14.3	-78.4	12.7	9298	Pericalcarine
		-6.05	139	-30.7	-47.3	49.4	330	Superior parietal
		-5.16	62	-44.3	-0.8	-18.3	155	Superior temporal/ATL
		-4.13	58	-17	-58.8	59.6	135	Superior parietal
		-3.93	63	-32.6	-50.3	-5.8	131	Fusiform
		-3.39	83	-36.4	36	-10.1	145	IOFC
		-2.97	84	-43.8	-69.4	14.2	158	Inferior parietal

(continued)

Table 3. Continued

Contrast	hemi	max	Size mm ²	x	y	z	k	Region
		–2.77	91	–46.5	–58.4	3.8	160	Middle temporal
		–2.66	59	–16.8	–35.3	48.4	153	Paracentral
	rh	–6.54	5531	25.8	–71.5	26	9247	Superior parietal
		–4.03	138	30	–40.2	–13.4	300	Fusiform
		–3.85	72	42.6	–54.4	14.5	173	Inferior parietal
		–3.02	56	7.2	7.5	48.7	110	Superior frontal
		–2.56	61	41.1	27.8	–12.6	83	IOFC

Note: Clusters are significant at $P < 0.005$ with a minimum cluster size of 50 mm²; k is size in vertices; coordinates are in Talairach space.

IOFC, lateral orbitofrontal cortex; dlPFC, dorsolateral prefrontal cortex; dmPFC, dorsomedial prefrontal cortex; vlPFC, ventrolateral prefrontal cortex; ATL, Anterior temporal lobe; TPJ, temporoparietal junction; IPL, inferior parietal lobe.

fascination, our findings are also consistent with an interpretation in terms of an increased role for ‘top-down’ interpretative processes (e.g. reappraisal). For example, previous research has observed patterns of activation in the dmPFC/SMA, IOFC and vlPFC when individuals regulate their emotional state by reappraising the meaning of a stimulus (Wager et al., 2008; Diekhof et al., 2011; Buhle et al., 2013) or interpret neutral stimuli as negative (Ochsner et al., 2009). Thus, the relatively stronger activation of the dmPFC/SMA and IOFC for morbid fascination feedback may suggest a link between the experience of fascination and reappraisal specifically. This link may have been emphasized by our instruction, which mentioned that morbid fascination could be interpreted as a state in which people ‘get drawn to a stimulus to examine the details’. Although this systematic way of viewing the stimulus is similar to the instructions that people receive in a reappraisal paradigm, it is also a characteristic of the mental state of fascination itself. Note, furthermore, that our instruction did not ask people to reappraise the negativity of the stimulus, nor to down-regulate their emotions.

It is important to note that conceptualization and reappraisal account are not mutually exclusive. In fact, conceptualization has been offered as a mechanism for achieving reappraisal (Barrett et al., 2014). We argue that paradigms in which participants are instructed to reinterpret a stimulus ‘top-down’ are in fact evoking a re-conceptualization of the sensory information present in that particular situation. Moreover, regions involved in emotion regulation (Buhle et al., 2013; Diekhof et al., 2011) are also commonly involved during emotion experience (Vytal and Hamann, 2010; Lindquist et al., 2012). Thus, the patterns of activation found in the present study may also reflect the production of emotional reactions afresh. Ultimately, there may not be a strict division between the mechanisms involved in creating an emotion in the first place or modifying it after the fact (cf., Gross and Barrett, 2011; Lindquist et al., 2012; Ochsner et al., 2012).

Limitations

As the first study to use false mental state feedback, our study has several limitations. A first possible limitation concerns our choice of control condition. Although we deliberately chose to include a condition with a non-specific label, this could have led participants to give trials with this type of feedback reflective scrutiny (e.g. ‘What am I actually feeling?’), resulting in increased activity in regions associated with conceptualization. Thus, the comparison between emotion feedback and control feedback is likely conservative, which may explain why we only found a significant difference between fear, disgust and control feedback in the vlPFC. Nevertheless, morbid fascination

feedback could also have resulted in self-reflective thoughts (‘Why am I fascinated by this image?’), thus matching morbid fascination and control feedback in this regard. In short, although there may be psychological implications of the control condition, these differences cannot explain the different patterns of neural activation following control feedback and morbid fascination feedback.

Second, there may be possible psychological consequences of giving people (false) morbid fascination feedback that could have influenced the patterns of neural activation in this study (e.g. by causing doubt, surprise). In this context, it is important to note that additional analyses suggested that our results did not differ between participants who reported that the labels were accurate descriptions of their feelings and those who reported low accuracy of the feedback. Although null findings are hard to interpret, this provides some support for the assumption that the patterns forwarded by our study were not merely driven by participants who were confused or in doubt about the veracity of the feedback labels. Nevertheless, as a future direction, it is important to replicate the current findings by targeting the non-manipulated experience of morbid fascination.

Finally, although our experimental design presented fear, disgust and morbid fascination as mutually exclusive states, it is likely that these experiences often co-occur. The subjective ratings collected at Time 1 indeed demonstrated that the negative images were rated equally in terms of threat, repulsion and interest, suggesting that experiencing a stimulus as disgusting does not preclude experiencing that stimulus as interesting. The integration of multiple experiences towards negative images, and an understanding of how these experiences develop over time is an important avenue for future research. Another important side note is that the present findings do not speak to the possible differences or similarities in how the brain represents fascination for negative stimuli, and for positive or non-emotional stimuli. Incorporating such a comparison was beyond the scope of the present project, but we believe that this is another important topic for future research.

Conclusion

To summarize, using an innovative paradigm, this study documents the neural mechanisms underlying atypical and typical experiences towards negative images. Our findings provide an important starting point for future research into the experience of morbid fascination and emphasize that it is relevant to take this state into consideration when investigating reactions to negative events.

Acknowledgements

This research was supported by a National Institutes of Health Director's Pioneer Award to Lisa Feldman Barrett (DP1OD003312), a Marie Curie International Outgoing Fellowship (275214) awarded by the European Commission's Seventh Framework Programme and Netherlands Organization for Scientific Research VENI grant (451-13-008) to Suzanne Oosterwijk and a Harvard University Mind/Brain/Behavior Postdoctoral Fellowship to Kristen Lindquist. The authors would like to thank Evangeline Barnard, Mara Koerner and Jamie Nichols for their assistance with stimulus selection and development.

Supplementary data

Supplementary data are available at SCAN online.

Conflict of interest. None declared.

References

- Bar, M. (2009). The proactive brain: memory for predictions. *Philosophical Transactions of the Royal Society B*, **364**, 1235–43.
- Barrett, L.F. (2006). Solving the emotion paradox: categorization and the experience of emotion. *Personality and Social Psychology Review*, **10**, 20–46.
- Barrett, L.F. (2009). Variety is the spice of life: a psychological constructionist approach to understanding variability in emotion. *Cognition and Emotion*, **23**, 1284–306.
- Barrett, L.F. (2012). Emotions are real. *Emotion*, **12**, 413–29.
- Barrett, L.F. (2013). Psychological construction: a Darwinian approach to the science of emotion. *Emotion Review*, **5**, 379–89.
- Barrett, L.F., Satpute, A.B. (2013). Large-scale brain networks in affective and social neuroscience: towards and integrative functional architecture of the brain. *Current Opinion in Neurobiology*, **23**, 1–12.
- Barrett, L.F., Wilson-Mendenhall, C.D., Barsalou, L.W. (2014). A psychological construction account of emotion regulation and dysregulation: the role of situated conceptualizations. In: Gross, J.J., editor. *The Handbook of Emotion Regulation*, 2nd edn, pp. 447–465, New York: Guilford.
- Barsalou, L.W. (1985). Ideals, central tendency, and frequency of instantiation as determinants of graded structure in categories. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, **11**, 629–54.
- Binder, J.R., Desai, R.H., Graves, W.W., Conant, L.L. (2009). Where is the semantic system? A critical review and meta-analysis of 120 functional neuroimaging studies. *Cerebral Cortex*, **19**(12), 2767–96.
- Borg, C., de Jong, P.J., Renken, R., Georgiadis, J.R. (2013). Disgust trait modulates frontal-posterior coupling as a function of disgust domain. *Social Cognitive and Affective Neuroscience*, **8**, 351–8.
- Buhle, J.T., Silvers, J.A., Wager, T.D., et al. (2013). Cognitive reappraisal of emotion: a meta-analysis of human neuroimaging studies. *Cerebral Cortex*, **24**(11), 2981–90.
- Crum, A., Salovey, P., Achor, S. (2013). Rethinking stress: the role of mindsets in determining the stress response. *Journal of Personality and Social Psychology*, **104**, 716–33.
- Dale, A.M., Fischl, B., Sereno, M.I., (1999). Cortical surface-based analysis. I. Segmentation and surface reconstruction. *Neuroimage*, **9**, 179–94.
- Desikan, R.S., Segonne, F., Fischl, B., et al. (2006). An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. *Neuroimage*, **31**, 968–80.
- Diekhof, E.K., Geier, K., Falkai, P., Gruber, O. (2011). Fear is only as deep as the mind allows: a coordinate-based meta-analysis of neuroimaging studies on the regulation of negative affect. *Neuroimage*, **58**, 275–85.
- Esterman, M., Tamber-Rosenau, B.J., Chiu, Y., Yantis, S. (2010). Avoiding non-independence in fMRI data analysis: Leave one subject out. *Neuroimage*, **50**, 572–6.
- Fischl, B., Sereno, M.I., Dale, A.M., (1999a). Cortical surface-based analysis. II: inflation, flattening, and a surface-based coordinate system. *Neuroimage*, **9**, 195–207.
- Fischl, B., Sereno, M.I., Tootell, R.B., Dale, A.M., (1999b). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, **8**, 272–84.
- Gray, M.A., Harrison, N.A., Wiens, S., Critchley, H.D. (2007). Modulation of emotional appraisal by false physiological feedback during fMRI. *PLoS One* **2**(6): e546.
- Gross, J.J., Barrett, L.F. (2011). Emotion generation and emotion regulation: one or two depends on your point of view. *Emotion Review*, **3**, 8–16.
- Gusnard, D.A., Akbudak, E., Shulman, G.L., Raichle, M.E. (2001). Medial prefrontal cortex and self-referential mental activity: relation to a default mode of brain function. *Proceedings of the National Academy of Sciences*, **98**, 4259–64.
- Hariri, A.R., Tessitore, A., Mattay, V.S., Fera, F., Weinberger, D.R. (2002). The amygdala response to emotional stimuli: a comparison of faces and scenes. *Neuroimage*, **17**, 317–23.
- Haynes, J., Rees, G. (2006). Decoding mental states from brain activity in humans. *Nature Reviews Neuroscience* **7**, 523–34.
- Jenkins, A., Mitchell, J.P. (2010). Mentalizing under uncertainty: dissociated neural responses to ambiguous and unambiguous mental state inferences. *Cerebral Cortex*, **20**, 404–10.
- Kassam, K.S., Mendes, W.B. (2013). The effects of measuring emotion: physiological reactions to emotional situations depend on whether someone is asking. *PLoS One*, **8**(7), e64959.
- Kelley, W.M., Macrae, C.N., Wyland, C.L., Caglar, S., Inati, S., Heatherton, T.F. (2002). Finding the self? an event-related fMRI study. *Journal of Cognitive Neuroscience*, **14**, 785–94.
- Kober, H., Barrett, L.F., Joseph, J., Bliss-Moreau, E., Lindquist, K.A., Wager, T.D. (2008). Functional networks and cortical-subcortical interactions in emotion: a meta-analysis of neuroimaging studies. *Neuroimage*, **42**, 998–1031.
- Kriegeskorte, N., Simmons, W.K., Bellgowan, P.S.F., Baker, C. (2009). Circular analysis in systems neuroscience – the dangers of double dipping. *Nature Neuroscience*, **12**, 535–40.
- Lang, P.J., Bradley, M.M., Cuthbert, B.N. (2008). *International Affective Pictures System (IAPS): Technical manual and affective ratings* (Tech. Rep. A-8). Gainesville: University of Florida, Research Center for Research in Psychophysiology.
- Lieberman, M.D., Cunningham, W.A. (2009). Type I and Type II error concerns in fMRI research: re-balancing the scale. *Social Cognitive and Affective Neuroscience*, **4**, 423–8.
- Lieberman, M.D., Eisenberger, N.I., Crockett, M.J., Tom, S.M., Pfeifer, J.H., Way, B.M. (2007). Putting feelings into words: affect labeling disrupts amygdala activity in response to affective stimuli. *Psychological Science*, **18**, 421–8.
- Lindquist, K.A. (2013). Emotions emerge from more basic psychological ingredients: a modern psychological constructionist approach. *Emotion Review*, **5**, 356–68.
- Lindquist, K., Barrett, L.F. (2008). Constructing emotion: the experience of fear as a conceptual act. *Psychological Science*, **19**, 898–903.

- Lindquist, K.A., Barrett, L.F. (2012). A functional architecture of the human brain: Insights from emotion. *Trends in Cognitive Sciences*, **16**, 533–54.
- Lindquist, K.A., Satpute, A.B., Gendron, M. (2015). Does language do more than communicate emotion? *Current Directions in Psychological Science*, **24**(2), 99–108.
- Lindquist, K.A., Wager, T.D., Kober, H., Bliss-Moreau, E., Barrett, L.F. (2012). The brain basis of emotion: a meta-analytic review. *Behavioral Brain Sciences*, **35**, 121–43.
- Nohlen, H.U., van Harreveld, F., Rotteveel, M., Lelieveld, G., Crone, E.A. (2013). Evaluating ambivalence: social-cognitive and affective brain regions associated with ambivalent decision-making. *Social Cognitive and Affective Neuroscience*, **9**(7), 924–31.
- Ochsner, K.N., Ray, R.R., Hughes, B., et al. (2009). Bottom-up and top-down processes in emotion generation: common and distinct neural mechanisms. *Psychological Science*, **20**, 1322–31.
- Ochsner, K.N., Silvers, J.A., Buhle, J.T. (2012). Functional imaging studies of emotion regulation: a synthetic review and evolving model of the cognitive control of emotion. *Annals of the New York Academy of Sciences*, **1251**, E1–24.
- Oosterwijk, S., Lindquist, K.A., Anderson, E., Dautoff, R., Moriguchi, Y., Barrett, L.F. (2012). States of mind: emotions, body feelings and thoughts share distributed neural networks. *Neuroimage*, **62**, 2110–28.
- Oosterwijk, S., Topper, M., Rotteveel, M., Fischer, A. H. (2010). When the mind forms fear: embodied fear knowledge potentiates bodily reactions to fearful stimuli. *Social Psychological and Personality Science*, **1**, 65–72.
- Oosterwijk, S., Touroutoglou, A., Lindquist, K.A. (2015). The neuroscience of construction: what neuroimaging approaches can tell us about how the brain creates the mind. In: Barrett, L.F., Russell, J.A., editors. *The Psychological Construction of Emotion*. New York: Guilford.
- Rimé, B., Delfosse, C., Corsini, S. (2005). Emotional fascination: responses to viewing pictures of September 11 attacks. *Cognition and Emotion*, **19**, 923–32.
- Satpute, A.B., Shu, J., Weber, J., Roy, M., Ochsner, K.N. (2013). The functional neural architecture of self-reports of affective experience. *Biological Psychiatry*, **73**, 631–8.
- Schachter, S., Singer, J. (1962). Cognitive, social, and physiological determinants of emotional state. *Psychological Review*, **69**, 379–99.
- Shirer, W.R., Ryali, S., Rykhlevskaia, E., Menon, V., Greicius, M.D. (2012). Decoding subject-driven cognitive states with whole-brain connectivity patterns. *Cerebral Cortex*, **22**(1), 158–65.
- Stark, R., Zimmermann, M., Kagerer, S., et al. (2007). Hemodynamic brain correlates of disgust and fear ratings. *Neuroimage*, **37**, 663–73.
- Touroutoglou, A., Lindquist, K.A., Dickerson, B.C., Barrett, L.F. (2015). Intrinsic connectivity in the human brain does not reveal networks for “basic” emotions. *Frontiers in Human Neuroscience*, **10**, 1257–65.
- Turner, S.A., Silvia, P.J. (2006). Must interesting things be pleasant? A test of competing appraisal structures. *Emotion*, **6**, 670–4.
- Valins, S. (1966). Cognitive effects of false heart-rate feedback. *Journal of Personality and Social psychology*, **4**, 400–8.
- Van der Kouwe, A.J., Benner, T., Salat, D.H., Fischl, B. (2008). Brain morphometry with multiecho MPRAGE. *Neuroimage*, **40**, 559–69.
- Vytal, K., Hamann, S. (2010). Neuroimaging support for discrete neural correlates of basic emotions: a voxel-based meta-analysis. *Journal of Cognitive Neuroscience*, **22**, 2864–85.
- Wager, T.D., Davidson, M.L., Hughes, B.L., Lindquist, M.A., Ochsner, K.N. (2008). Prefrontal-subcortical pathways mediating successful emotion regulation. *Neuron*, **59**, 1037–50.
- Wagner, A.D., Paré-Blagoev, E.J., Clark, J., Poldrack, R.A. (2001). Recovering meaning: left prefrontal cortex guides controlled semantic retrieval. *Neuron*, **31**(2), 329–38.
- Wild, J., Clark, D.M., Ehlers, A., McManus, F. (2008). Perception of arousal in social anxiety: effects of false feedback during a social interaction. *Journal of Behavioral Therapy and Experimental Psychiatry*, **39**, 102–16.
- Wilson-Mendenhall, C.D., Barrett, L.F., Barsalou, L.W. (2014). Variety in emotional life: within-category typicality of emotional experiences is associated with neural activity in large-scale brain networks. *Social Cognitive and Affective Neuroscience*, **10**(1), 67–71.
- Wilson-Mendenhall, C.D., Barrett, L.F., Simmons, W.K., Barsalou, L.W. (2011). Grounding emotion in situated conceptualization. *Neuropsychologia*, **49**, 1105–27.
- Wright, P., He, G., Shapira, N.A., Goodman, W.K., Liu, Y. (2004). Disgust and the insula: fMRI responses to pictures of mutilation and contamination. *NeuroReport*, **15**, 2347–51.
- Yeo, B.T.T., Krienen, F.M., Sepulcre, J., et al. (2011). The organization of the human cerebral cortex estimated by functional connectivity. *Journal of Neurophysiology*, **106**, 1125–65.
- Zuckerman, M., Litle, P. (1986). Personality and curiosity about morbid and sexual events. *Personality and Individual Differences*, **7**, 49–56.