OXFORD

Neural correlates of cognitive bias modification for interpretation

Kohei Sakaki, [iD][1,2,3] Takayuki Nozawa, [iD][4] Shigeyuki Ikeda,[5] and Ryuta Kawashima[5,6]

[1]Department of Functional Brain Imaging, Graduate School of Medicine, Tohoku University, Sendai 980-8575, Japan, [2]Division for Interdisciplinary Advanced Research and Education, Tohoku University, Sendai 980-8578, Japan, [3]Japan Society for the Promotion of Science, Tokyo 102-0083, Japan, [4]Research Institute for the Earth Inclusive Sensing Empathizing with Silent Voices, Tokyo Institute of Technology, Tokyo 152-8550, Japan, [5]Department of Ubiquitous Sensing, Institute of Development, Aging and Cancer, Tohoku University, Sendai 980-8575, Japan, and [6]Department of Functional Brain Imaging, Institute of Development, Aging and Cancer, Tohoku University, Sendai 980-8575, Japan

Correspondence should be addressed to Kohei Sakaki, Department of Functional Brain Imaging, IDAC, Tohoku University, Seiryo-machi 4-1, Aoba-ku, Sendai 980-8575, Japan. E-mail: sakaki@med.tohoku.ac.jp.

Abstract

The effectiveness of cognitive bias modification for interpretation (CBM-I), a treatment method employed to reduce social anxiety (SA), has been examined. However, the neural correlates of CBM-I remain unclear, and we aimed to elucidate brain activities during intervention and activity changes associated with CBM-I effectiveness in a pre–post intervention comparison. Healthy participants divided into two groups (CBM, control) were scanned before, during and after intervention using functional magnetic resonance imaging. Ambiguous social situations followed by positive outcomes were repeatedly imagined by the CBM group during intervention, while half of the outcomes in the control group were negative. Whole-brain analysis revealed that activation of the somatomotor and somatosensory areas, occipital lobe, fusiform gyrus and thalamus during intervention was significantly greater in the CBM than in the control group. Furthermore, altered activities in the somatomotor and somatosensory areas, occipital lobe and posterior cingulate gyrus during interpreting ambiguous social situations showed a significant group × change in SA interaction. Our result suggests that when facing ambiguous social situations, positive imagery instilled by CBM-I is recalled, and interpretations are modified to contain social reward. These findings may help to suggest an alternative manner of enhancing CBM-I effectiveness from a cognitive-neuroscience perspective.

Key words: cognitive bias; social anxiety; positive thinking; social reward; fMRI

Introduction

Although interpersonal relationships constitute a fundamental human motivation (Baumeister and Leary, 1995), we sometimes experience anxiety and a need to withdraw from social situations. It is not uncommon for people to fear negative evaluations and feel distress or discomfort in social situations (Watson and Friend, 1969). These feelings comprise social anxiety (SA), and it is assumed that SA exists on a wide spectrum, from shyness regarded as a general personality trait to SA disorder (SAD) requiring clinical support (Kashdan, 2007; Henderson *et al.,* 2014).

Adolescents with higher levels of SA reported poorer social functioning, fewer friendships, and less intimacy, companionship, and support in their close friendships (La Greca and Lopez, 1998). Most university students, although comprising a nonclinical sample, experienced symptoms of SA such as sweating and shaking in their daily life (Purdon *et al.,* 2001). Therefore, SA reduction is an issue of great concern to the fields of psychiatry and education.

In the field of psychology, a cognitive model of SA has been investigated in detail (Mathews and MacLeod, 1994). A cognitive process that may mainly contribute to causality and maintain SA symptoms is interpretation bias (Rapee and Heimberg, 1997). Socially anxious individuals tend to interpret social stimuli and events in a negative manner even if the events and stimuli are emotionally ambiguous (Amin *et al.,* 1998; Hirsch and Mathews, 2000; Hirsch and Clark, 2004).

Cognitive bias modification for interpretation (CBM-I) is a training program aimed at reducing SA by modifying negative interpretation bias (Mathews and Mackintosh, 2000; Yiend *et al.,* 2005). In the typical CBM-I paradigm for SA, participants are requested to repeatedly mentally imagine presented series of social scenarios that are initially ambiguous and are followed by a clearly positive (or benign) interpretation outcome (MacLeod and Mathews, 2012). Initial CBM-I studies investigated the effectiveness of positive interpretation training to reduce trait anxiety in healthy populations (Mathews and MacLeod, 2002; Mathews *et al.,* 2007). Recent previous studies have reported that the effect of CBM-I modifying biased interpretation could transfer to reduce SA symptoms (Murphy *et al.,* 2007; Beard and Amir, 2008; Amir and Taylor, 2012; Bowler *et al.,* 2012; Khalili-Torghabeh *et al.,* 2014; Brettschneider *et al.,* 2015; Yeung and Sharpe, 2019). Although CBM-I has promising potential, meta-analyses have indicated that the effect size of CBM-I was small and its effectiveness should be enhanced (Hallion and Ruscio, 2011; Menne-Lothmann *et al.,* 2014; Cristea *et al.,* 2015; Liu *et al.,* 2017; Krebs *et al.,* 2018). Several clinical studies have attempted to overcome this challenge, boosting the effectiveness of CBM-I by combining it with other psychiatric treatments such as cognitive behavior therapy (CBT) (Butler *et al.,* 2015; Stevens *et al.,* 2018) or CBM for attention (Beard *et al.,* 2011; Naim *et al.,* 2018; Yeung and Sharpe, 2019). In addition to such clinical challenges, basic research for greater understanding of the mechanisms underpinning the effectiveness of CBM-I would contribute toward shedding light from different directions and lead to developing an innovative approach to improve CBM-I. Although several experimental studies and clinical trials have been accumulated, the underlying neural correlates of CBM-I remain understudied and unclear.

In general CBT using reappraisal, it is well-known that the dorsal medial and lateral prefrontal cortices play a key role in top–down emotion regulation (Simpson *et al.,* 2000; Ochsner *et al.,* 2002, 2012; Goldin *et al.,* 2008; Ochsner and Gross, 2008; Wager *et al.,* 2008). In contrast, CBM-I was assumed to have relatively implicit and automatic action mechanisms and was reported to work even under conditions when cognitive resources were depleted (Bowler *et al.,* 2012). Mathews (2012) proposed two potential CBM-I action mechanisms; one was semantic priming of a category of affective meanings and the other was transfer of a learned form of processing from training to testing. The affective priming effect is a phenomenon where processing or responses for target stimuli are facilitated by the prior presentation of a prime when they are affectively congruent (Fazio, 2001; Winkielman *et al.,* 2005). Applying this to the CBM-I context, positive outcomes repeatedly presented during intervention act as primes and facilitate the positive interpretation circuit when the participants face novel ambiguous social situations immediately after intervention. In fact, previous CBM-I studies supported that participants trained to make positive (or negative) interpretations by related affective priming homographs were faster to identify positive (or negative) words in a lexical decision task (Grey and Mathews, 2009; Hoppitt *et al.,* 2010b). Here, we hypothesized that the nucleus accumbens (NAcc) would play a key role in the affective priming underlying CBM-I. The NAcc is a core element of the subcortical reward circuitry included in the ventral striatum and activates in response to various reward stimuli such as money, food and drugs (Berridge, 2003). NAcc activation was related to affective priming using happy-face stimuli, which are known to be treated as 'social reward' by the brain (Winkielman *et al.,* 2007; Suslow *et al.,* 2013).

Conversely, only the affective priming effect is not sufficient to explain why modified interpretation could endure over longer periods (Mackintosh *et al.,* 2006) and transfer to novel stimuli that are not similar to those used in the training task (Hoppitt *et al.,* 2010b). Participants may learn a processing rule to select positive meanings following encountering ambiguity through training, and that rule would be unintentionally transferred and applied to subsequent events (Mathews and MacLeod, 2002). Rule learning through CBM-I could be successful under conditions that former parts of stimuli had affective ambiguity (Clarke *et al.,* 2014; Clifton *et al.,* 2016) and subsequent positive interpretations were actively selected (Hoppitt *et al.,* 2010a,b). However, totally self-generation of positive resolutions for ambiguous scenarios was not beneficial to enhance CBM-I effectiveness (Rohrbacher *et al.,* 2014). Additionally, presenting model answers or providing brief feedback would be required to reinforce the association between ambiguous social situations and subsequent positive interpretations (Murphy *et al.,* 2007; Beard and Amir, 2008; Bowler *et al.,* 2012). Based on this rule-learning hypothesis, we expected that brain areas related to memory functions would activate in response to learning rules for positive interpretations during intervention, and their activation would increase after, compared to before, intervention, reflecting retrieval of learned rules. It is well known that the hippocampus is involved in three successive processes comprising encoding, storage and retrieval (Squire, 1992; Burgess *et al.,* 2002). Additionally, the amygdala is associated with emotional learning and memory (Gallagher and Holland, 1994; Phelps, 2004, 2006; Phelps and LeDoux, 2005). The amygdala tends to become more highly active in response to negative, than to positive, emotions, but a previous study found that the amygdala was more active when more optimistic people, who tend to expect better outcomes, imagined future positive, compared to negative, events (Sharot *et al.,* 2007). Furthermore, the hippocampus and amygdala are associated with explicit and implicit memory (Rose *et al.,* 2002; Reber, 2013), associative learning (Killgore *et al.,* 2000) and affective conditioning (Büchel *et al.,* 1999).

Here, brain activities were examined using functional magnetic resonance imaging (fMRI) while participants imagined outcomes of ambiguous social situations before, during and after intervention to examine the neural correlates of CBM-I, which remains understudied. Interpretation bias, which was assessed by subjective ratings of expectation for novel ambiguous social situations, and SA, evaluated by a questionnaire, were assessed at pre- and post-intervention. To examine the neural correlates of CBM-I, we mainly used three types of analysis. First, we analyzed imaging data scanned during intervention to reveal the brain areas that were activated during CBM-I. Second, we analyzed the data derived from the assessment task to locate the brain areas whose activities would be altered by CBM-I. Third, multiple regression analysis using change in assessment scores

was performed to search for brain areas reflecting individual differences in CBM-I effectiveness. Based on the review of the neuroimaging literature mentioned above, our hypotheses were that brain areas associated with social reward and emotional memory, such as the NAcc, hippocampus and amygdala, would be more activated during CBM-I compared to the control intervention, and activities in these areas during interpreting novel ambiguous social situations would increase after intervention. Moreover, participants with greater activity changes in these areas would show greater SA reduction.

Material and methods

Participants

A total of 40 healthy right-handed Japanese university students (20 men and 20 women, age range 20–26 years; mean, 22.0 years) participated in the experiment. The sample-size estimation was based on a previous study that investigated the effectiveness of mental imagery CBM-I using group comparisons (Clarke *et al.,* 2014) and was consistent with other relevant studies (Browning *et al.,* 2010; Taylor *et al.,* 2014; Beevers *et al.,* 2015; Britton *et al.,* 2015). Handedness was evaluated using the Edinburgh Handedness Inventory (Oldfield, 1971). No participant had history of neurological or psychiatric illness. The experimental protocol was approved by the ethics committee of the Tohoku University School of Medicine and conformed to the Declaration of Helsinki. Written informed consent was obtained from all participants.

Each of the two groups included the same number of participants (10 men and 10 women) who exhibited similar tendency toward SA. SA tendency was measured with the Japanese version of the fear of negative evaluation scale (FNE; Ishikawa *et al.,* 1992), and the participants were allocated so that they would be roughly matched between the two groups based on their score on the scale.

The data of two participants (one woman in the CBM group and one man in the positive–negative [PN] group) were excluded due to insufficient task performance (response rates <80%).

Intervention task

We constructed an intervention task based on previous studies (Holmes and Mathews, 2005; Holmes *et al.,* 2006; Blackwell and Holmes, 2010; Clarke *et al.,* 2014). One hundred scenarios consisting of three short sentences were created (see Supplementary Data for details). All scenarios had to include some characters and be related to a daily social situation. Participants were instructed to imagine all scenarios as vividly as possible playing the role of the main character. Participants could not determine whether the end of the scenario was positive or negative simply by reading the two initial sentences (e.g. 'I had a presentation at a conference/A professor spoke to me after the presentation'). All scenarios in the CBM group had a positive outcome (e.g. 'The professor praised my presentation as attractive'). Conversely, half the scenarios in the PN group had a negative outcome (e.g. 'The professor castigated my presentation as incomprehensible'). This manner of control intervention had no effect on interpretation bias (Salemink *et al.,* 2009, 2014). Every sentence was presented for 10 s on a screen in an MRI scanner. After three sentences were presented, the participants were asked to rate how vividly they could imagine the scenario on an eight-point Likert scale from 1 (*Not at all vividly*) to 8 (*Very vividly*) within 5 s. Afterward, a small fixation cross was presented for 8–12 s as the intertrial interval (ITI). The experimental design is illustrated in

Figure 1A. One session consisted of 20 scenarios and lasted for ~15 min. The participants rested for 15 min after every session and completed five sessions.

Assessment task for interpretation bias

We also constructed an assessment task for interpretation bias based on a previous study (Holmes *et al.,* 2006). Thirty scenarios consisting of three short sentences were created (see Supplementary Data for details). In the case of the assessment task, only the initial two sentences were presented for 10 s each. The participants were required to freely interpret the initial two sentences, create an outcome for the scenario and imagine it for 10 s. Then, participants rated how positive was the imagined outcome on an eight-point Likert scale from 1 (*It was very negative*) to 8 (*It was very positive*) within 5 s. We defined the interpretation bias as the average scores of the affective ratings in each session (pre and post the intervention sessions) for every participant. Moreover, they rated within 5 s how vividly they imagined the scenario. Afterward, a small fixation cross was presented for 8–12 s as the ITI. The experimental design is illustrated in Figure 1B. One session consisted of 15 scenarios and lasted for ~13 min. The participants performed one session of the assessment task before and after intervention. Two sets of scenarios were prepared and used with the order of the sets counterbalanced among participants.

Assessment of SA

We focused on participant tendency toward SA as the first outcome of intervention. Participants completed the FNE (Watson and Friend, 1969) before and after intervention (see Supplementary Data for further details).

Statistical analysis for behavioral and psychological data

Behavioral and psychological data were analyzed using SPSS (version 23; IBM, Armonk, NY, USA). We performed mixed-ANOVA including time (pre, post) as a within-subjects factor and group (CBM, PN) as a between-subjects factor for the SA and interpretation bias scores.

fMRI data acquisition and preprocessing

Both experimental tasks were conducted in an MRI scanner (3T Philips Achieva, Best, Netherlands). Participants comfortably laid on the bed of the MRI scanner with their heads fixed in the head coil using elastic bands. The visual stimulus was back-projected onto a semi-lucent screen behind the head coil and was viewed via a mirror. The size of the visual stimuli in all tasks subtended a visual angle of less than 5° (see Supplementary Data for further details).

Overview of fMRI data analysis

Both fMRI tasks were modeled using a block design. All imaging data were analyzed using a two-level approach in SPM12. At the first-level, the hemodynamic responses and the time derivatives generated by a participant under the different experimental conditions were assessed at each voxel using a general linear model. We constructed separate canonical regressors corresponding to each phase of the task in the trial, and the mean centered rating values as parametric regressors and confounding factors (head motion and magnetic field drift) were incorporated into
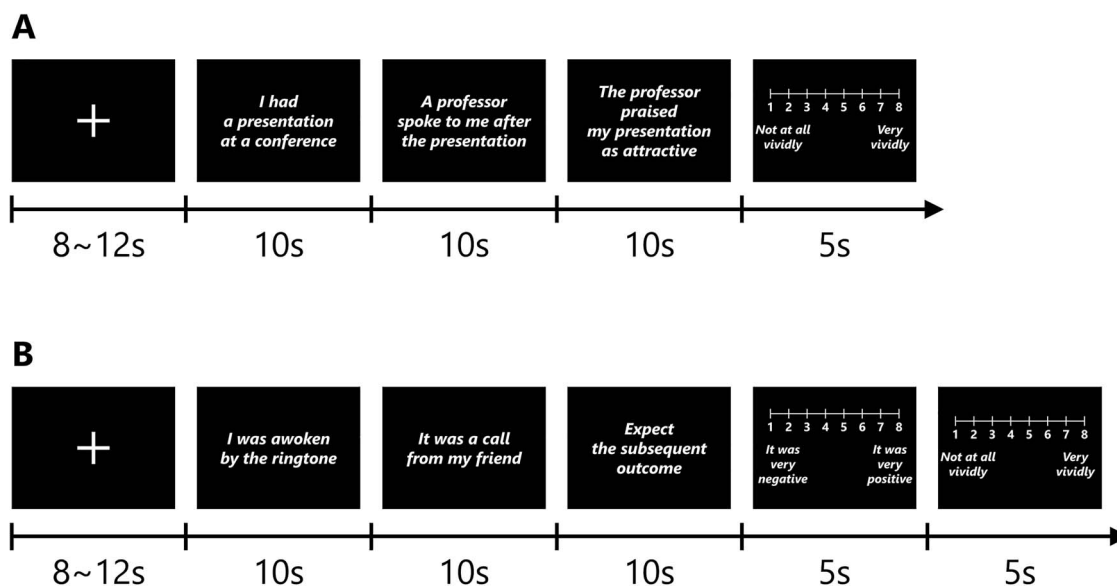
A

B

Fig. 1. Schematic of the sequence of trial events in the intervention task (A) and the assessment task (B). (A) A small fixation cross was presented for 8–12 s as the ITI. Three short sequential sentences composing a single social scenario were presented for 10 s each, and participants imagined the situations. The initial two sentences denoted an ambiguous outcome and the third sentence denoted a positive (or negative) outcome. Participants rated the vividness of their imagery for the present scenario on an eight-point Likert scale from 1 (Not at all vividly) to 8 (Very vividly) within 5 s. (B) A small fixation cross was presented for 8–12 s as the ITI. Two short sequential sentences composing a single social scenario were presented for 10 s each, and participants imagined the situations. The two sentences denoted ambiguous outcome, and participants were required to expect the subsequent outcome for the present scenario and imagine it for 10 s. Participants rated the valence of their expected outcome on an eight-point Likert scale from 1 (It was very negative) to 8 (It was very positive) within 5 s and the vividness of their imagery for the present scenario on an eight-point Likert scale from 1 (Not at all vividly) to 8 (Very vividly) within 5 s.

the model. ITIs were not modeled. For second-level whole-brain analysis, the cluster-forming threshold at the voxel-level was set at $P < 0.001$ (uncorrected) and corrected for multiple comparisons at the cluster-level (family-wise error [FWE], $P < 0.05$, one-tailed).

fMRI data analysis of the intervention task

For second-level analysis, we performed a two-sample t-test using contrast images corresponding to brain activities during the third sentence-imagining phase. From whole-brain analysis, we located the brain areas where activity in the CBM group was greater than that in the PN group. Additionally, based on the a priori hypothesis that brain areas associated with social reward and emotional memory would be more activated during CBM-I than during the control intervention, we applied small volume correction (SVC) using a region of interest (ROI) including the bilateral NAcc, hippocampus and amygdala as defined in the Wake Forest University PickAtlas (Maldjian *et al.,* 2003). For ROI analysis, the statistical threshold was set at $P < 0.05$ FWE-corrected (peak-level, one-tailed).

Furthermore, psychophysiological interaction (PPI) analyses were conducted to examine relations between differences in group conditions and functional connectivity of the bilateral NAcc, hippocampus and amygdala seed regions utilizing the generalized psychophysiological interaction (gPPI) toolbox (McLaren *et al.,* 2012). The cluster-forming threshold at the voxel-level was set at $P < 0.001$ (uncorrected) and corrected for multiple comparisons at the cluster-level (FWE, $P < 0.016$, one-tailed), as three statistical tests were performed for each seed region.

fMRI data analysis of the assessment task

For second-level analysis, we performed a two-sample t-test using contrast images from the difference in brain activities during the interpretation phase between post and pre intervention (post−pre) within subjects and tested for group (CBM, PN) × time (pre, post) interaction on activity. To detect brain areas corresponding to the individual differences in CBM-I effectiveness, i.e. a significant target of this study, we performed voxel-wise multiple regression analysis using the contrast images (post−pre) and the change in SA (ΔSA) as covariates of interest. The groups were included in the same model to examine the group × ΔSA interaction on activity changes (see Supplementary Data for further details). Additionally, based on the a priori hypothesis that activities in brain areas associated with social reward and emotional memory during interpreting novel ambiguous social situations would increase after intervention specifically in the CBM group, we applied SVC using the ROI including the bilateral NAcc, hippocampus and amygdala. Furthermore, gPPI analyses were conducted to examine the group × ΔSA interaction on functional connectivity changes (ΔConnectivity) of the bilateral NAcc, hippocampus and amygdala seed regions.

Results

Behavioral and psychological data

The summary of the behavioral ratings and scores on the psychological scales is presented in Table 1. A statistically significant interaction was revealed for SA (Figure 2A; $F(1, 36) = 11.1$, $P = 0.002$, $\eta_p^2 = 0.24$) and interpretation bias (Figure 2B; $F(1, 36) = 5.89$, $P = 0.020$, $\eta_p^2 = 0.14$). Additionally, *post-hoc* analysis with Bonferroni correction in the CBM group revealed a significant simple main effect of time (pre, post) for SA ($t(18) = -3.09$, $P = 0.003$, $r = 0.59$) and interpretation bias ($t(18) = 2.08$, $P = 0.038$, $r = 0.44$), which was not significant in the PN group (SA: $t(18) = -1.55$, $P = 0.15$, $r = 0.34$; interpretation bias: $t(18) = 1.33$, $P = 0.21$, $r = 0.30$).

Table 1. Summary of the behavioral ratings and scores on the psychological scales

Measure	CBM group (N = 19)			PN group (N = 19)		
	Pre, mean (s.d.)	Post, mean (s.d.)	Mean difference [95% CI]	Pre, mean (s.d.)	Post, mean (s.d.)	Mean difference [95% CI]
SA	15.42 (8.36)	13.21 (9.46)	−2.21 [−3.59, −0.83]	17.00 (9.15)	18.00 (9.92)	1.00 [−0.38, 2.38]
Interpretation bias	5.28 (0.56)	5.66 (0.91)	0.38 [0.02, 0.73]	5.39 (0.69)	5.16 (0.80)	−0.23 [−0.58, 0.13]
Vividness rating	5.25 (0.85)	5.54 (0.81)	0.29 [−0.06, 0.64]	5.77 (1.07)	5.72 (0.92)	−0.06 [−0.40, 0.29]

Note: SA: the average score of the FNE scale; CBM: cognitive bias modification; PN: positive–negative; interpretation bias: the average score of the positive ratings; vividness rating: the average score of the vividness ratings; 95% CI: 95% confidence intervals (Bonferroni adjusted).
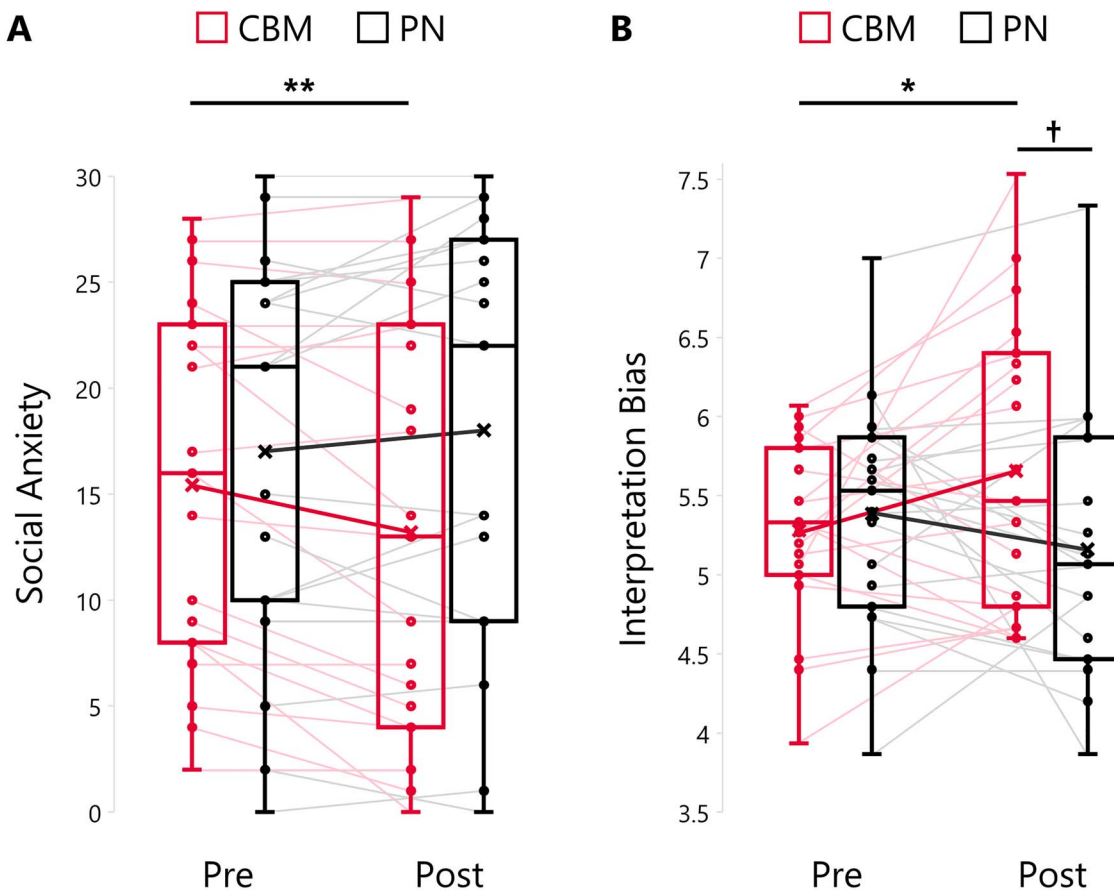

Fig. 2. Distribution of scores of SA (A) and interpretation bias (B). Box plots show the median, interquartile range and minimum/maximum values of the scores for each group and session. (A) SA was assessed by the FNE. A high score on the FNE signifies high SA tendency. (B) Interpretation bias was calculated by the scores extracted from the assessment task. A high score of interpretation bias signifies the tendency of a participant to expect a positive outcome for ambiguous social situations. $^{\dagger}P < 0.1$; $^{*}P < 0.05$; $^{**}P < 0.01$.

fMRI data

Whole-brain analysis for the intervention task indicated that the activity of brain areas including the somatomotor and somatosensory areas, occipital lobe, fusiform gyrus (FuG) and thalamus in the CBM group was significantly greater than that in the PN group (Figure 3, Table 2). ROI analysis found that peak voxel activity in the right hippocampus was significantly greater in the CBM group than in the PN group (Figure 4, Table 3). The gPPI analysis revealed stronger connectivity between the bilateral NAcc and clusters located in the somatomotor and somatosensory areas, FuG and inferior temporal gyrus (ITG) in the CBM group than in the PN group (Figure 5, Table 4). The other seed regions (the hippocampus and amygdala) returned no significant results.

The fMRI data of the assessment task indicated that there were no significant interaction of activity between the group (CBM, PN) and time (pre, post). The multiple regression analysis for the contrast images (post−pre) with the ΔSA as a regressor revealed a significant group × ΔSA interaction on activity changes in brain areas including the somatomotor and somatosensory areas, occipital lobe and posterior cingulate gyrus (PCgG) (Figure 6, Table 5). Part of areas including the somatomotor and somatosensory areas and superior occipital gyrus (SOG) overlapped with the activation area detected in the intervention task (Supplementary Figure S1). Focusing on the parametric regressors, the positive or vividness ratings modulated the activity of no brain area. ROI analysis found a significant group × ΔSA interaction on activity change in a peak
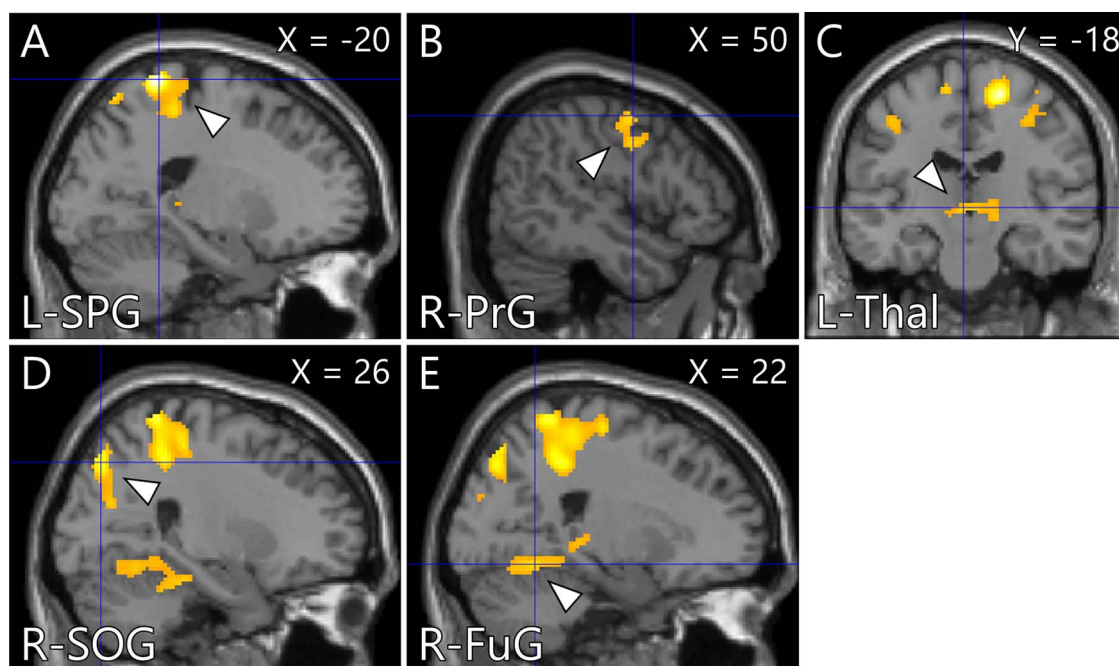
Fig. 3. Brain areas showing greater activation in the CBM group than in the PN group during the intervention task. The cluster-forming threshold at the voxel level was set at $P < 0.001$ (uncorrected); multiple comparisons were corrected for at the cluster level (FWE, $P < 0.05$). (A) Left superior parietal lobule (-20, -44, 72). (B) Right PrG (50, -2, 52). (C) Left thalamus (-2, -18, 0). (D) Right SOG (26, -70, 22). (E) Right fusiform gyrus (22, -56, -12). CBM: cognitive bias modification; PN: positive-negative and FWE: family-wise error.

Table 2. Brain areas showing greater activation in the CBM group than in the PN group during the intervention task

Area		MNI peak coordinates (mm)				
		x	y	z	t value	k
Superior parietal lobule	L	−20	−44	72	6.49	4742
Supplementary motor area	R	18	−18	62	5.97	
	L	−12	−50	72	5.3	
PrG	R	50	−2	52	3.98	415
		48	−2	36	3.83	
PoG	R	56	0	42	3.67	
Thalamus	L	−2	−18	0	4.07	547
	R	14	−36	−4	3.98	
	L	−14	−30	2	3.96	
SOG	R	26	−76	44	4.79	322
		14	−86	34	4.00	
		26	−70	22	3.8	
Fusiform gyrus	R	22	−56	−12	4.00	405
		28	−34	−20	3.79	
		28	−48	−12	3.59	

Note: For each area, the coordinates (x, y, z) of the activation peak in MNI space, peak t value, and size of the activated cluster in a number (k) of voxels ($2 \times 2 \times 2$ mm^3) are presented. The cluster-forming threshold at the voxel level was set at $P < 0.001$ (uncorrected) and corrected for multiple comparisons at the cluster level (FWE, $P < 0.05$). CBM: cognitive bias modification; PN: positive–negative; MNI: Montreal Neurological Institute and FWE: family-wise error.
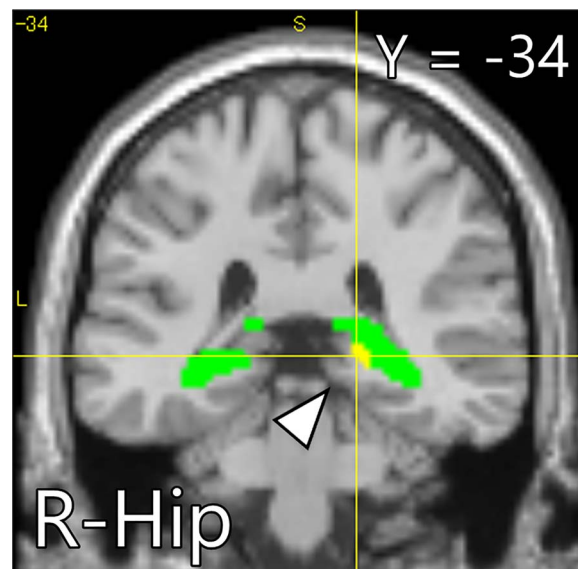

Fig. 4. Brain areas showing greater activation (yellow) within the ROI (green) including the bilateral NAcc, hippocampus and amygdala in the CBM group than in the PN group during the intervention task. The statistical threshold at the voxel level was set at $P < 0.001$ (uncorrected) and corrected for SVC within the structural ROI at the peak level (FWE, $P < 0.05$). ROI: region on interest; CBM: cognitive bias modification; PN: positive-negative and FWE: family-wise error.

voxel in the right amygdala (Figure 7, Table 6). The gPPI analysis revealed a significant group × ∆SA interaction on ∆Connectivity between the NAcc and clusters located in the inferior parietal lobule (IPL), PCgG and superior frontal gyrus (SFG) (Figure 8, Table 7). The other seed regions returned no significant results.

Discussion

To our knowledge, this study is the first to investigate neural correlates associated with CBM-I. We found greater brain activation in the somatomotor and somatosensory areas, occipital lobe, FuG and thalamus during CBM-I than during the control intervention. Stronger functional connectivity of the somatomotor
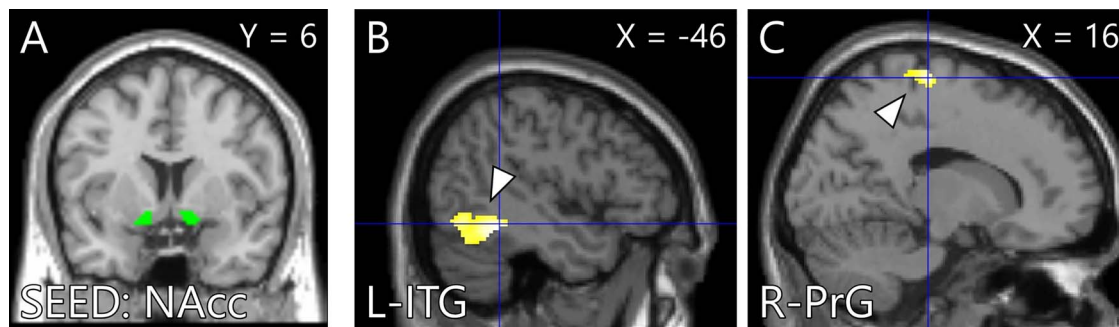
Fig. 5. Brain areas showing greater functional connectivity with the bilateral NAcc seed region (A) in the CBM group than in the PN group during the intervention task. The cluster-forming threshold at the voxel level was set at $P < 0.001$ (uncorrected); multiple comparisons were corrected for at the cluster level (FWE, $P < 0.016$). (B) Left ITG (-46, -44, -12). (C) Right PrG (16, -24, 70). NAcc: nucleus accumbens; CBM: cognitive bias modification; PN: positive–negative and FWE: family-wise error.

Table 3. A peak voxel in the ROI showing greater activation in the CBM group than in the PN group during the intervention task

Area		MNI peak coordinates (mm)
		x	y	z	t value
Hippocampus	R	18	-34	0	3.93

Note: For each voxel, the coordinates (x, y, z) in MNI space and t value are presented. The statistical threshold at the voxel level was set at $P < 0.001$ (uncorrected) and corrected for SVC within the structural ROI at the peak level (FWE, $P < 0.05$). ROI: region on interest; CBM: cognitive bias modification; PN: positive–negative; MNI: Montreal Neurological Institute and FWE: family-wise error.

Table 4. Brain areas showing greater functional connectivity with the bilateral NAcc seed region in the CBM group than in the PN group during the intervention task

Area		x	y	z	t value	k
ITG	L	-46	-44	-12	4.90	668
Fusiform gyrus	L	-18	-88	-10	4.69	
Inferior occipital gyrus	L	-34	-82	-10	4.52	
PrG	R	16	-24	70	4.77	673
PoG	L	-14	-38	72	4.68	
Paracentral lobule		0	-36	70	4.37	

Note: For each area, the coordinates (x, y, z) of the activation peak in MNI space, peak t value, and size of the activated cluster in a number (k) of voxels ($2 \times 2 \times 2$ mm^3) are presented. The cluster-forming threshold at the voxel level was set at $P < 0.001$ (uncorrected) and corrected for multiple comparisons at the cluster level (FWE, $P < 0.05$). NAcc: nucleus accumbens; CBM: cognitive bias modification; PN: positive–negative; MNI: Montreal Neurological Institute and FWE: family-wise error.

Table 5. Brain areas in which the change in the amount of activity in the assessment task was associated with an interaction effect between the group and the change in SA

Area		x	y	z	t value	k
PoG	R	14	-32	76	6.35	787
		26	-36	72	5.84	
PrG	R	30	-24	66	4.64	
Lingual gyrus	R	10	-74	-10	5.02	1100
	L	-12	-76	-14	4.22	
Cerebellum	R	2	-72	-32	3.94	
PCgG	R	14	-44	34	4.97	359
Cuneus	R	10	-84	8	4.71	365
	L	-10	-76	14	4.61	
	R	4	-72	16	3.50	
SOG	R	32	-68	48	4.62	746
		24	-62	34	4.32	
		26	-76	18	3.92	

Note: For each area, the coordinates (x, y, z) of the activation peak in MNI space, peak t value, and size of the activated cluster in a number (k) of voxels ($2 \times 2 \times 2$ mm^3) are presented. The cluster-forming threshold at the voxel level was set at $P < 0.001$ (uncorrected) and corrected for multiple comparisons at the cluster level (FWE, $P < 0.05$). MNI: Montreal Neurological Institute and FWE: family-wise error.

Table 6. A peak voxel in the ROI in which the change in the amount of activity in the assessment task was associated with an interaction effect between the group and the change in SA

Area		x	y	z	t value
Amygdala	R	24	4	-20	3.74

Note: For each area, the coordinates (x, y, z) of the activation peak in MNI space, peak t value, and size of the activated cluster in a number (k) of voxels ($2 \times 2 \times 2$ mm^3) are presented. The statistical threshold at the voxel level was set at $P < 0.001$ (uncorrected) and corrected for multiple comparisons at the cluster level (FWE, $P < 0.05$). ROI: region of interest; MNI: Montreal Neurological Institute and FWE: family-wise error.

and somatosensory areas, FuG and ITG with the NAcc was also found. Regarding brain activity changes, we found a group × ΔSA interaction on activity changes in the somatomotor and somatosensory areas, occipital lobe and PCgG during interpreting ambiguous social situations. Interaction on functional connectivity of the IPL, PCgG and SFG with the NAcc was also found.

Neural correlates of CBM-I

Here, the activity in the somatomotor and somatosensory areas, occipital lobe, FuG and thalamus during the intervention task was significantly greater in the CBM than in the PN group. The FuG is known as the 'face area' (Rossion *et al.*, 2003; Kanwisher

and Yovel, 2006) and becomes active in response to perception of happy facial expressions (Fusar-Poli *et al.*, 2009). The following structures are supplementally involved in response to happy facial stimuli: the superior parietal lobule, which plays a key role in various sensory and cognitive processes, including attention shifts (Corbetta *et al.*, 1995), sensorimotor integration (Wolpert *et al.*, 1998), visuomotor control (Culham *et al.*, 2006), mental
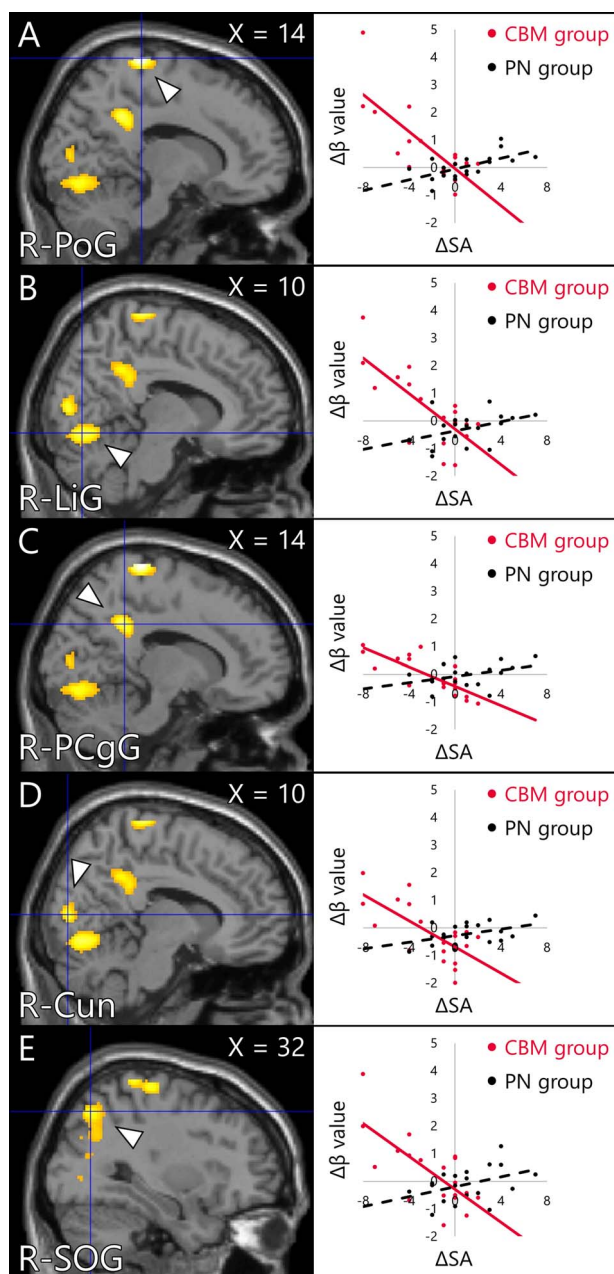
Fig. 6. Brain areas in which the change in the amount of activity in the assessment task was associated with an interaction effect between the group and the change in SA. The cluster-forming threshold at the voxel level was set at P < 0.001 (uncorrected); multiple comparisons were corrected for at the cluster level (FWE, P < 0.05). (A) Right PoG (14, −32, 76). (B) Right lingual gyrus (10, −74, −10). (C) Right PCgG (14, −44, 34). (D) Right cuneus (10, −84, 8). (E) Right SOG (32, −68, 48). The $\Delta\beta$ values described in scatter plots were calculated by the difference of the peak voxel activity between pre and post intervention (post−pre) for each group (MIN = −1.99, MAX = 4.89). The ΔSAs were calculated by the difference of the FNE score between pre and post intervention (post−pre) for each group (MIN = −8, MAX = 7). FWE: family-wise error; SA: social anxiety and FNE: fear of negative evaluation.

rotation (Murphy, 2003) and working memory (Koenigs *et al.*, 2009); the thalamus that relays information between different subcortical areas and the cerebral cortex (Sherman and Guillery, 1996), and the SOG (Morris *et al.*, 1998; Lee *et al.*, 2002; Sato *et al.*, 2004). The precentral gyrus (PrG) and postcentral gyrus (PoG), which include the primary somatosensory cortex (Penfield and

Table 7. Brain areas in which the change in the amount of connectivity with the bilateral NAcc seed region in the assessment task was associated with an interaction effect between the group and the change in SA

Area		MNI peak coordinates (mm)				
		x	*y*	*z*	*t* value	*k*
IPL	L	−50	−54	44	6.06	1106
Angular gyrus	L	−44	−68	36	5.05	
Supramarginal gyrus	L	−56	−50	28	3.92	
PCgG	L	−2	−24	32	5.80	516
		−14	−20	32	4.49	
SFG	L	−24	62	2	5.65	481
Middle frontal gyrus	L	−40	56	0	5.02	
SFG	L	−12	70	8	4.17	

Note: For each area, the coordinates (*x*, *y*, *z*) of the activation peak in MNI space, peak *t* value, and size of the activated cluster in a number (*k*) of voxels (2 × 2 × 2 mm³) are presented. The cluster-forming threshold at the voxel level was set at P < 0.001 (uncorrected) and corrected for multiple comparisons at the cluster level (FWE, P < 0.05). NAcc: nucleus accumbens; MNI: Montreal Neurological Institute and FWE: family-wise error.

Rasmussen, 1950), activate in response to words and pictures bearing positive meaning (Kensinger and Schacter, 2006). The participants in the CBM group repeatedly imagined social situations ending consistently positively during the intervention task. The greater activation in the somatomotor and somatosensory areas, occipital lobe, FuG and thalamus possibly reflects the greater number of positive feelings induced by the happy faces and actions of the characters in the social situations that the participants repeatedly imagined during the task.

ROI analysis found that the hippocampus showed greater activation in the CBM group. The hippocampus was reported to be more active during memory encoding of pleasant stimuli than of neutral stimuli (Hamann *et al.*, 1999; Bulganin and Wittmann, 2015). Contrary to our initial hypothesis, the greater activation in the NAcc and amygdala was not significant. Moreover, gPPI analysis found stronger functional connectivity between the NAcc and the somatomotor and somatosensory areas, FuG and ITG in the CBM group. Numerous neuropsychological studies have shown that the NAcc is associated with various social reward such as happy and attractive faces (Cloutier *et al.*, 2008; Spreckelmeyer *et al.*, 2009; Mende-Siedlecki *et al.*, 2013), social approval and avoidance of social punishment (Kohls *et al.*, 2013), acquiring a positive reputation of oneself (Izuma *et al.*, 2008), social approach (Radke *et al.*, 2016), and prosocial behavior (van der Meulen *et al.*, 2016). All these types of social reward were possibly evoked during the intervention task. The stronger connectivity of the NAcc with the somatomotor and somatosensory areas and FuG also possibly supported the notion that the greater number of imageries in the CBM group was processed as social reward. Here, significant activity or functional connectivity differences between groups in the intervention task were not noted in the dorsal medial or lateral prefrontal cortex in line with previous findings.

Brain activity changes induced by CBM-I

Here, the whole-brain analysis revealed a significant group × ΔSA interaction on activity changes in several brain areas including the somatomotor and somatosensory areas, occipital lobe and PCgG, observed while participants interpreted novel social
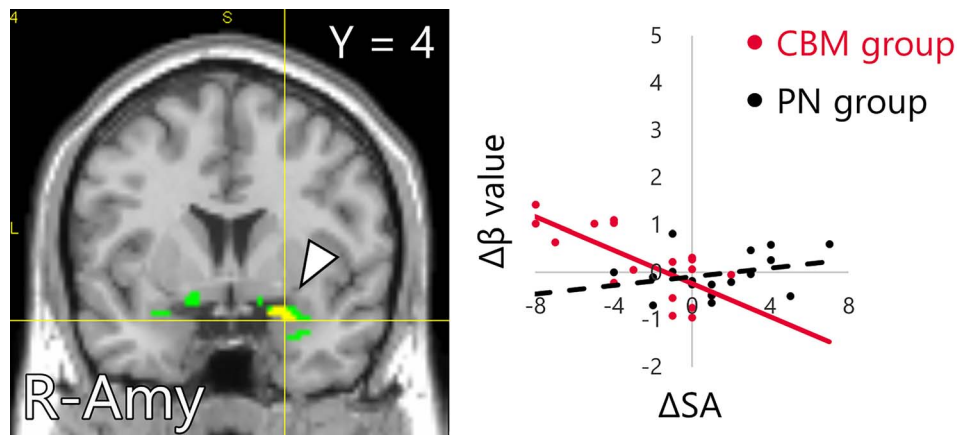
Fig. 7. Brain areas in which the change in the amount of activity in the assessment task was associated with an interaction effect between the group and the change in SA (yellow) within the ROI including the bilateral NAcc, hippocampus and amygdala (green). The $\Delta\beta$ values described in scatter plots were calculated by the difference of the peak voxel activity between pre and post intervention (post−pre) for each group (MIN = −0.94, MAX = 1.44). The ΔSAs were calculated by the difference of the FNE score between pre and post intervention (post−pre) for each group (MIN = −8, MAX = 7). ROI: region of interest; SA: social anxiety and FNE: fear of negative evaluation.
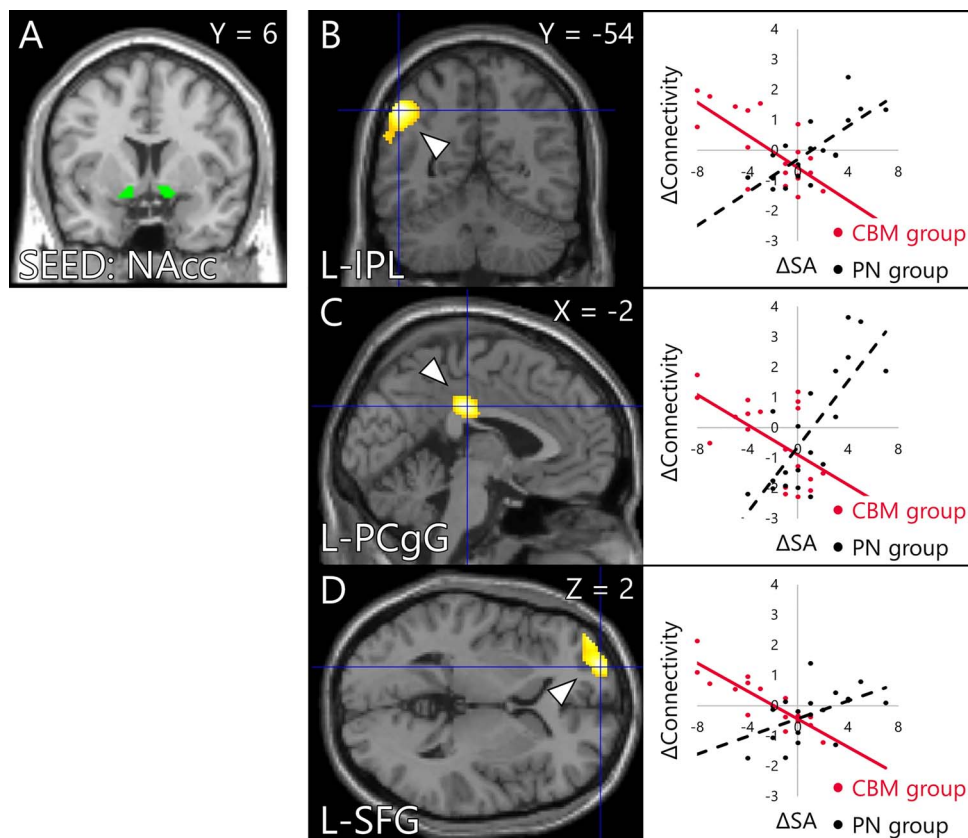

Fig. 8. Brain areas in which the change in the amount of connectivity with the bilateral NAcc seed region (A) in the assessment task was associated with an interaction effect between the group and the change in SA. The cluster-forming threshold at the voxel level was set at $P < 0.001$ (uncorrected); multiple comparisons were corrected for at the cluster level (FWE, $P < 0.016$). (B) Left IPL (−50, −54, 44). (C) Left PCgG (−2, −24, 32). (D) Left SFG (−24, 62, 2). The connectivity changes (ΔConnectivity) described in scatter plots were calculated by the difference of the peak voxel connectivity between pre and post intervention (post−pre) for each group (MIN = −2.28, MAX = 3.65). The ΔSAs were calculated by the difference of the FNE score between pre and post intervention (post − pre) for each group (MIN = −8, MAX = 7). NAcc: nucleus accumbens; FWE: family-wise error; SA: social anxiety; and FNE: fear of negative evaluation.

situations in the assessment tasks. Furthermore, ROI analysis also revealed a group × ΔSA interaction on the activity change in the right amygdala. To wit, participants whose activity in these brain areas were further increased showed greater SA reduction after the CBM-I intervention. Moreover, gPPI analysis found a

significant group × ΔSA interaction on ΔConnectivity between the NAcc and the IPL, PCgG and SFG.

These activity changes in the somatomotor and somatosensory areas including the PoG and PrG, which were also observed in the intervention task (Supplementary Figure S1A and B),

possibly reflect increase in the number of happy faces and actions imagined in the novel social situations during the assessment task. The occipital lobe, which was also observed in the intervention task (Supplementary Figure S1C), including the SOG and cuneus forming the visual cortex, showed greater activity in response to emotional stimuli (Lang *et al.,* 1998; Park *et al.,* 2010). The greater activity increase in the visual cortex possibly reflects that participants could imagine more pleasant interpretations, which resulted in more positive emotional valence. In the PN group, this activity increase was not correlated with SA reduction because the emotional valence of social scenarios presented during the intervention task was not consistent. The NAcc and DLPFC are involved in reward prediction and motivation for goal-directed behavior (Goto and Grace, 2005; Knutson and Cooper, 2005; Watanabe and Sakagami, 2007; Ballard *et al.,* 2011). The activity increase in the visual cortex and reinforced functional connectivity between the NAcc and DLPFC that correlated with SA reduction also possibly supported the notion that the participants' imagery was modified to be processed as higher social reward.

The IPL is associated with memory retrieval (Wagner *et al.,* 2005; Vilberg and Rugg, 2008) and working memory (LaBar *et al.,* 1999). The PCgG is also implicated in memory (Maddock *et al.,* 2001) and self-referential thinking (Johnson *et al.,* 2002; Kelley *et al.,* 2002; Northoff and Bermpohl, 2004; Northoff *et al.,* 2006). The IPL and PCgG are involved in the retrieval process particularly for self-referential episodic memory (Lou *et al.,* 2004). The activity increase in these memory-related areas and reinforced functional connectivity with the NAcc that correlated with SA reduction possibly reflects that participants attempted to self-referentially retrieve and recall the memory of positive interpretations for ambiguous social situations, learned during the intervention task. Furthermore, the activation increase in the amygdala might suggest that the participants who were more affected by intervention and experienced further reduction in SA became more optimistic as also shown by a previous study (Sharot *et al.,* 2007).

Study implications

Although not directly supported by the present findings, we could provide some suggestions to improve the CBM-I protocol. To maximize the effect of CBM-I, experimenters should attempt to create stimuli that could be easily imagined vividly to further activate brain areas related to imagining, such as the face area, occipital lobe and somatomotor and somatosensory areas and induce highly positive emotions to further activate the NAcc and be strongly impressed in the memory of participants to activate brain areas involved in memory functions such as the hippocampus, amygdala, IPL and PCgG. Moreover, the present findings provided from the field of cognitive neuroscience could propose an alternative manner of enhancing effectiveness of CBM-I as a future direction. For example, neurofeedback, which also has an automatic and implicit action mechanism, could be well combined with CBM-I. Neurofeedback is a technique providing feedback on the brain activity of any ROI in real time as sensory input and self-regulating brain activity by reference to the feedback with or without explicit instructions (Heinrich *et al.,* 2007; Sulzer *et al.,* 2013; Stoeckel *et al.,* 2014; Scharnowski and Weiskopf, 2015). Based on the present findings, if brain activity (e.g. in the NAcc, hippocampus and amygdala) could be presented as feedback to participants during intervention, they could attempt to imagine social situations that could result in further activation in the relevant brain area. Thus, effectiveness

of CBM-I could possibly be enhanced. In this manner, the present findings could both elucidate the neural correlates of CBM-I and lead to future studies aiming to improve and develop the intervention protocol for CBM-I.

Limitations

University students that did not comprise a clinical sample were recruited. Further investigations are required to confirm whether the present results could be generalized to clinical populations or younger/older participants. For example, 12 20-min CBM-I intervention sessions delivered over 6 weeks were conducted in a previous randomized control trial that treated a clinical sample with SAD (Amir and Taylor, 2012). In another previous study, children with a diagnosis of SAD completed three 30-min sessions within 2 weeks (Orchard *et al.,* 2017). In this case, we constructed an intervention protocol that required participants to complete five 15-min sessions in a day with the aim to investigate acute effect of CBM-I. However, we would have to revise the intervention protocol following previous trials to reduce the burden of treatment on participants when especially patients or children would participate in the experiments. Moreover, in terms of neural correlates, previous neuroimaging studies have suggested that there were differences in widespread brain activities, connectivity and structures of patients with SAD compared to healthy controls (Bruhl *et al.,* 2014). Additionally, we should conduct follow-up assessments in a future study to confirm the long-term and far-transfer effects of treatment.

We expected that the NAcc, hippocampus and amygdala would be specifically associated with the CBM-I intervention and activities in these areas would increase after intervention. However, activities in these areas were no longer found to be significant in the whole-brain analysis for both the intervention and assessment tasks, and even the ROI analysis only partially returned significant results. This is the first study to address the neural correlates of CBM-I, and we focused on the acute effect of short-term intervention. It would be expected that longer-term intervention would show that the psychological and neurological acute effects verified here could accumulate and lead to more significant results. Furthermore, future research needs to investigate in detail the association of brain areas clarified here with specific potential action mechanisms of CBM-I such as affective priming and rule-learning.

The present experimental design could not completely exclude several confounding factors regarding the intervention task that potentially offer alternative explanations for the present results. For example, differences between the groups regarding the degrees of arousal, familiarity or movement and changes in mood or emotional states with the SA reduction may have influenced the differences in brain activities. Such potential confounding factors within the experimental and control conditions should be carefully removed to the extent possible in a future study (see Supplementary Data for further details).

Conclusion

We identified brain areas associated with CBM-I and the effectiveness of CBM-I. Brain activation during CBM-I was greater in the somatomotor and somatosensory areas, occipital cortex, FuG and thalamus compared with the control intervention. Moreover, functional connectivity of the somatomotor and somatosensory areas, FuG and ITG with the NAcc was also stronger. This finding suggested that CBM-I activated the brain

areas related to social reward perception and imagery of happy faces and actions via consistently and repeatedly imagining positive interpretations for ambiguous social situations. Regarding individual differences in intervention effectiveness, activity increase in the somatomotor and somatosensory areas, occipital lobe and PCgG during interpreting ambiguous social situations was associated with SA reduction specifically in the CBM group. Reinforced functional connectivity of the IPL, PCgG and SFG with the NAcc was also significantly correlated with SA reduction. This finding suggested that CBM-I effectiveness for SA reduction possibly reflected that memories of positive interpretations for ambiguous social situations, which were repeatedly instilled during the CBM-I intervention, was self-referentially retrieved and recalled, and imagery was modified to contain more social reward.

Supplementary data

Supplementary data are available at *SCAN* online.

Acknowledgements

Funding

Conflict of interest

The authors declare that they have no conflict of interest.

References

Amin, N., Foa, E.B., Coles, M.E. (1998). Negative interpretation bias in social phobia. *Behaviour Research and Therapy*, **36**, 945–57.

Amir, N., Taylor, C.T. (2012). Interpretation training in individuals with generalized social anxiety disorder: a randomized controlled trial. *Journal of Consulting and Clinical Psychology*, **80**, 497–511.

Ballard, I.C., Murty, V.P., Carter, R.M., MacInnes, J.J., Huettel, S.A., Adcock, R.A. (2011). Dorsolateral prefrontal cortex drives mesolimbic dopaminergic regions to initiate motivated behavior. *Journal of Neuroscience*, **31**, 10340–6.

Baumeister, R.F., Leary, M.R. (1995). The need to belong: desire for interpersonal attachments as a fundamental human motivation. *Psychological Bulletin*, **117**, 497–529.

Beard, C., Amir, N. (2008). A multi-session interpretation modification program: changes in interpretation and social anxiety symptoms. *Behaviour Research and Therapy*, **46**, 1135–41.

Beard, C., Weisberg, R.B., Amir, N. (2011). Combined cognitive bias modification treatment for social anxiety disorder: a pilot trial. *Depression and Anxiety*, **28**, 981–8.

Beevers, C.G., Clasen, P.C., Enock, P.M., Schnyer, D.M. (2015). Attention bias modification for major depressive disorder: effects on attention bias, resting state connectivity, and symptom change. *Journal of Abnormal Psychology*, **124**, 463–75.

Berridge, K.C. (2003). Pleasures of the brain. *Brain and Cognition*, **52**, 106–28.

Blackwell, S.E., Holmes, E.A. (2010). Modifying interpretation and imagination in clinical depression: a single case series using cognitive bias modification. *Applied Cognitive Psychology*, **24**, 338–50.

Bowler, J.O., Mackintosh, B., Dunn, B.D., Mathews, A., Dalgleish, T., Hoppitt, L. (2012). A comparison of cognitive bias modification for interpretation and computerized cognitive behavior therapy: effects on anxiety, depression, attentional control, and interpretive bias. *Journal of Consulting and Clinical Psychology*, **80**, 1021–33.

Brettschneider, M., Neumann, P., Berger, T., Renneberg, B., Boettcher, J. (2015). Internet-based interpretation bias modification for social anxiety: a pilot study. *Journal of Behavior Therapy and Experimental Psychiatry*, **49**, 21–9.

Britton, J.C., Suway, J.G., Clementi, M.A., Fox, N.A., Pine, D.S., Bar-Haim, Y. (2015). Neural changes with attention bias modification for anxiety: a randomized trial. *Social Cognitive and Affective Neuroscience*, **10**, 913–20.

Browning, M., Holmes, E.A., Murphy, S.E., Goodwin, G.M., Harmer, C.J. (2010). Lateral prefrontal cortex mediates the cognitive modification of attentional bias. *Biological Psychiatry*, **67**, 919–25.

Bruhl, A.B., Delsignore, A., Komossa, K., Weidt, S. (2014). Neuroimaging in social anxiety disorder-a meta-analytic review resulting in a new neurofunctional model. *Neuroscience and Biobehavioral Reviews*, **47**, 260–80.

Büchel, C., Dolan, R.J., Armony, J.L., Friston, K.J. (1999). Amygdala–hippocampal involvement in human aversive trace conditioning revealed through event-related functional magnetic resonance imaging. *The Journal of Neuroscience*, **19**, 10869–76.

Bulganin, L., Wittmann, B.C. (2015). Reward and novelty enhance imagination of future events in a motivational-episodic network. *PLoS One*, **10**, e0143477.

Burgess, N., Maguire, E.A., O'Keefe, J. (2002). The human hippocampus and spatial and episodic memory. *Neuron*, **35**, 625–41.

Butler, E., Mobini, S., Rapee, R.M., Mackintosh, B., Reynolds, S.A. (2015). Enhanced effects of combined cognitive bias modification and computerised cognitive behaviour therapy on social anxiety. *Cogent Psychology*, **2**, 1011905.

Clarke, P.J., Nanthakumar, S., Notebaert, L., Holmes, E.A., Blackwell, S.E., Macleod, C. (2014). Simply imagining sunshine, lollipops and rainbows will not budge the bias: the role of ambiguity in interpretive bias modification. *Cognitive Therapy and Research*, **38**, 120–31.

Clifton, J.L., Hedley, S., Mountier, E., Tiszai, B., Grimshaw, G.M. (2016). Practice makes perfect: training the interpretation of emotional ambiguity. *Cognition & Emotion*, **30**, 654–68.

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.

Corbetta, M., Shulman, G.L., Miezin, F.M., Petersen, S.E. (1995). Superior parietal cortex activation during spatial attention shifts and visual feature conjunction. *Science*, **270**, 802–5.

Cristea, I.A., Kok, R.N., Cuijpers, P. (2015). Efficacy of cognitive bias modification interventions in anxiety and depression: meta-analysis. *British Journal of Psychiatry*, **206**, 7–16.

Culham, J.C., Cavina-Pratesi, C., Singhal, A. (2006). The role of parietal cortex in visuomotor control: what have we learned from neuroimaging? *Neuropsychologia*, **44**, 2668–84.

Fazio, R.H. (2001). On the automatic activation of associated evaluations: an overview. *Cognition & Emotion*, **15**, 115–41.

Fusar-Poli, P., Placentino, A., Carletti, F., *et al.* (2009). Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of Psychiatry and Neuroscience*, **34**, 418–32.

Gallagher, M., Holland, P.C. (1994). The amygdala complex: multiple roles in associative learning and attention. *Proceedings of the National Academy of Sciences of the United States of America*, **91**, 11771–6.

Goldin, P.R., McRae, K., Ramel, W., Gross, J.J. (2008). The neural bases of emotion regulation: reappraisal and suppression of negative emotion. *Biological Psychiatry*, **63**, 57786.

Goto, Y., Grace, A.A. (2005). Dopaminergic modulation of limbic and cortical drive of nucleus accumbens in goal-directed behavior. *Nature Neuroscience*, **8**, 805–12.

Grey, S.J., Mathews, A.M. (2009). Cognitive bias modification—priming with an ambiguous homograph is necessary to detect an interpretation training effect. *Journal of Behavior Therapy and Experimental Psychiatry*, **40**, 338–43.

Hallion, L.S., Ruscio, A.M. (2011). A meta-analysis of the effect of cognitive bias modification on anxiety and depression. *Psychological Bulletin*, **137**, 940–58.

Hamann, S.B., Ely, T.D., Grafton, S.T., Kilts, C.D. (1999). Amygdala activity related to enhanced memory for pleasant and aversive stimuli. *Nature Neuroscience*, **2**, 289–93.

Heinrich, H., Gevensleben, H., Strehl, U. (2007). Annotation: neurofeedback—train your brain to train behaviour. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, **48**, 3–16.

Henderson, L., Gilbert, P., Zimbardo, P. (2014). Shyness, social anxiety, and social phobia. In: Hofmann, S., DiBartolo, P., editors. *Social Anxiety*, 3rd edn, Cambridge: Academic Press, pp. 95–115.

Hirsch, C.R., Clark, D.M. (2004). Information-processing bias in social phobia. *Clinical Psychology Review*, **24**, 799–825.

Hirsch, C.R., Mathews, A. (2000). Impaired positive inferential bias in social phobia. *Journal of Abnormal Psychology*, **109**, 705–12.

Holmes, E.A., Mathews, A. (2005). Mental imagery and emotion: a special relationship? *Emotion*, **5**, 489–97.

Holmes, E.A., Mathews, A., Dalgleish, T., Mackintosh, B. (2006). Positive interpretation training: effects of mental imagery versus verbal training on positive mood. *Behavior Therapy*, **37**, 237–47.

Hoppitt, L., Mathews, A., Yiend, J., Mackintosh, B. (2010a). Cognitive bias modification: the critical role of active training in modifying emotional responses. *Behavior Therapy*, **41**, 73–81.

Hoppitt, L., Mathews, A., Yiend, J., Mackintosh, B. (2010b). Cognitive mechanisms underlying the emotional effects of bias modification. *Applied Cognitive Psychology*, **24**, 312–25.

Ishikawa, R., Sasaki, K., Fukui, I. (1992). Standardization of Japanese version of FNE and SADS. *Koudou Ryouhou Kenkyu*, **18**, 10–7.

Izuma, K., Saito, D.N., Sadato, N. (2008). Processing of social and monetary rewards in the human striatum. *Neuron*, **58**, 284–94.

Johnson, S.C., Baxter, L.C., Wilder, L.S., Pipe, J.G., Heiserman, J.E., Prigatano, G.P. (2002). Neural correlates of self-reflection. *Brain*, **125**, 1808–14.

Kanwisher, N., Yovel, G. (2006). The fusiform face area: a cortical region specialized for the perception of faces. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, **361**, 2109–28.

Kashdan, T.B. (2007). Social anxiety spectrum and diminished positive experiences: theoretical synthesis and meta-analysis. *Clinical Psychology Review*, **27**, 348–65.

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.

Kensinger, E.A., Schacter, D.L. (2006). Processing emotional pictures and words: effects of valence and arousal. *Cognitive, Affective, & Behavioral Neuroscience*, **6**, 110–26.

Khalili-Torghabeh, S., Fadardi, J.S., Mackintosh, B., Reynolds, S., Mobini, S. (2014). Effects of a multi-session cognitive bias modification program on interpretative biases and social anxiety symptoms in a sample of Iranian socially-anxious students. *Journal of Experimental Psychopathology*, **5**, 514–27.

Killgore, W.D., Casasanto, D.J., Yurgelun-Todd, D.A., Maldjian, J.A., Detre, J.A. (2000). Functional activation of the left amygdala and hippocampus during associative encoding. *Neuroreport*, **11**, 2259–63.

Knutson, B., Cooper, J.C. (2005). Functional magnetic resonance imaging of reward prediction. *Current Opinion in Neurology*, **18**, 411–7.

Koenigs, M., Barbey, A.K., Postle, B.R., Grafman, J. (2009). Superior parietal cortex is critical for the manipulation of information in working memory. *Journal of Neuroscience*, **29**, 14980–6.

Kohls, G., Perino, M.T., Taylor, J.M., *et al.* (2013). The nucleus accumbens is involved in both the pursuit of social reward and the avoidance of social punishment. *Neuropsychologia*, **51**, 2062–9.

Krebs, G., Pile, V., Grant, S., Degli Esposti, M., Montgomery, P., Lau, J.Y.F. (2018). Research review: cognitive bias modification of interpretations in youth and its effect on anxiety: a meta-analysis. *Journal of Child Psychology and Psychiatry, and Allied Disciplines*, **59**, 831–44.

La Greca, A.M., Lopez, N. (1998). Social anxiety among adolescents: linkages with peer relations and friendships. *Journal of Abnormal Child Psychology*, **26**, 83–94.

LaBar, K.S., Gitelman, D.R., Parrish, T.B., Mesulam, M.M. (1999). Neuroanatomic overlap of working memory and spatial attention networks: a functional MRI comparison within subjects. *NeuroImage*, **10**, 695–704.

Lang, P.J., Bradley, M.M., Fitzsimmons, J.R., *et al.* (1998). Emotional arousal and activation of the visual cortex: an fMRI analysis. *Psychophysiology*, **35**, 199–210.

Lee, T.M.C., Liu, H.-L., 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*, **333**, 13–6.

Liu, H., Li, X., Han, B., Liu, X. (2017). Effects of cognitive bias modification on social anxiety: a meta-analysis. *PLoS One*, **12**, e0175107.

Lou, H.C., Luber, B., Crupain, M., *et al.* (2004). Parietal cortex and representation of the mental self. *Proceedings of the National Academy of Sciences of the United States of America*, **101**, 6827–32.

Mackintosh, B., Mathews, A., Yiend, J., Ridgeway, V., Cook, E. (2006). Induced biases in emotional interpretation influence stress vulnerability and endure despite changes in context. *Behavior Therapy*, **37**, 209–22.

MacLeod, C., Mathews, A. (2012). Cognitive bias modification approaches to anxiety. *Annual Review of Clinical Psychology*, **8**, 189–217.

Maddock, R.J., Garrett, A.S., Buonocore, M.H. (2001). Remembering familiar people: the posterior cingulate cortex and autobiographical memory retrieval. *Neuroscience*, **104**, 667–76.

Maldjian, J.A., Laurienti, P.J., Kraft, R.A., Burdette, J.H. (2003). An automated method for neuroanatomic and cytoarchitectonic atlas-based interrogation of fMRI data sets. *NeuroImage*, **19**, 1233–9.

Mathews, A. (2012). Effects of modifying the interpretation of emotional ambiguity. *Journal of Cognitive Psychology*, **24**, 92–105.

Mathews, A., Mackintosh, B. (2000). Induced emotional interpretation bias and anxiety. *Journal of Abnormal Psychology*, **109**, 602–15.

Mathews, A., MacLeod, C. (1994). Cognitive approaches to emotion and emotional disorders. *Annual Review of Psychology*, **45**, 25–50.

Mathews, A., MacLeod, C. (2002). Induced processing biases have causal effects on anxiety. *Cognition & Emotion*, **16**, 331–54.

Mathews, A., Ridgeway, V., Cook, E., Yiend, J. (2007). Inducing a benign interpretational bias reduces trait anxiety. *Journal of Behavior Therapy and Experimental Psychiatry*, **38**, 225–36.

McLaren, D.G., Ries, M.L., Xu, G., Johnson, S.C. (2012). A generalized form of context-dependent psychophysiological interactions (gPPI): a comparison to standard approaches. *NeuroImage*, **61**, 1277–86.

Mende-Siedlecki, P., Said, C.P., Todorov, A. (2013). The social evaluation of faces: a meta-analysis of functional neuroimaging studies. *Social Cognitive and Affective Neuroscience*, **8**, 285–99.

Menne-Lothmann, C., Viechtbauer, W., Hohn, P., *et al.* (2014). How to boost positive interpretations? A meta-analysis of the effectiveness of cognitive bias modification for interpretation. *PLoS One*, **9**, e100925.

van der Meulen, M., van, I.M.H., Crone, E.A. (2016). Neural correlates of prosocial behavior: compensating social exclusion in a four-player cyberball game. *PLoS One*, **11**, e0159045.

Morris, J.S., Friston, K.J., Buchel, C., *et al.* (1998). A neuromodulatory role for the human amygdala in processing emotional facial expressions. *Brain*, **121**, 47–57.

Murphy, G.L. (2003). The downside of categories. *Trends in Cognitive Sciences*, **7**, 513–4.

Murphy, R., Hirsch, C.R., Mathews, A., Smith, K., Clark, D.M. (2007). Facilitating a benign interpretation bias in a high socially anxious population. *Behaviour Research and Therapy*, **45**, 1517–29.

Naim, R., Kivity, Y., Bar-Haim, Y., Huppert, J.D. (2018). Attention and interpretation bias modification treatment for social anxiety disorder: a randomized clinical trial of efficacy and synergy. *Journal of Behavior Therapy and Experimental Psychiatry*, **59**, 19–30.

Northoff, G., Bermpohl, F. (2004). Cortical midline structures and the self. *Trends in Cognitive Sciences*, **8**, 102–7.

Northoff, G., Heinzel, A., de Greck, M., Bermpohl, F., Dobrowolny, H., Panksepp, J. (2006). Self-referential processing in our brain—a meta-analysis of imaging studies on the self. *NeuroImage*, **31**, 440–57.

Ochsner, K.N., Gross, J.J. (2008). Cognitive emotion regulation: insights from social cognitive and affective neuroscience. *Current Directions in Psychological Science*, **17**, 153–8.

Ochsner, K.N., Bunge, S.A., Gross, J.J., Gabrieli, J.D. (2002). Rethinking feelings: an FMRI study of the cognitive regulation of emotion. *Journal of Cognitive Neuroscience*, **14**, 1215–29.

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.

Oldfield, R.C. (1971). The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia*, **9**, 97–113.

Orchard, F., Apetroaia, A., Clarke, K., Creswell, C. (2017). Cognitive bias modification of interpretation in children with social anxiety disorder. *Journal of Anxiety Disorders*, **45**, 1–8.

Park, J.Y., Gu, B.M., Kang, D.H., *et al.* (2010). Integration of cross-modal emotional information in the human brain: an fMRI study. *Cortex*, **46**, 161–9.

Penfield, W., Rasmussen, T. (1950). *The Cerebral Cortex of Man: A Clinical Study of Localization of Function*, Oxford: Macmillan.

Phelps, E.A. (2004). Human emotion and memory: interactions of the amygdala and hippocampal complex. *Current Opinion in Neurobiology*, **14**, 198–202.

Phelps, E.A. (2006). Emotion and cognition: insights from studies of the human amygdala. *Annual Review of Psychology*, **57**, 27–53.

Phelps, E.A., LeDoux, J.E. (2005). Contributions of the amygdala to emotion processing: from animal models to human behavior. *Neuron*, **48**, 175–87.

Purdon, C., Antony, M., Monteiro, S., Swinson, R.P. (2001). Social anxiety in college students. *Journal of Anxiety Disorders*, **15**, 203–15.

Radke, S., Seidel, E.M., Eickhoff, S.B., *et al.* (2016). When opportunity meets motivation: neural engagement during social approach is linked to high approach motivation. *NeuroImage*, **127**, 267–76.

Rapee, R.M., Heimberg, R.G. (1997). A cognitive-behavioral model of anxiety in social phobia. *Behaviour Research and Therapy*, **35**, 741–56.

Reber, P.J. (2013). The neural basis of implicit learning and memory: a review of neuropsychological and neuroimaging research. *Neuropsychologia*, **51**, 2026–42.

Rohrbacher, H., Blackwell, S.E., Holmes, E.A., Reinecke, A. (2014). Optimizing the ingredients for imagery-based interpretation bias modification for depressed mood: is self-generation more effective than imagination alone? *Journal of Affective Disorders*, **152–154**, 212–8.

Rose, M., Haider, H., Weiller, C., Buchel, C. (2002). The role of medial temporal lobe structures in implicit learning: an event-related FMRI study. *Neuron*, **36**, 1221–31.

Rossion, B., Caldara, R., Seghier, M., Schuller, A.M., Lazeyras, F., Mayer, E. (2003). A network of occipito-temporal face-sensitive areas besides the right middle fusiform gyrus is necessary for normal face processing. *Brain*, **126**, 2381–95.

Salemink, E., van den Hout, M., Kindt, M. (2009). Effects of positive interpretive bias modification in highly anxious individuals. *Journal of Anxiety Disorders*, **23**, 676–83.

Salemink, E., Kindt, M., Rienties, H., van den Hout, M. (2014). Internet-based cognitive bias modification of interpretations in patients with anxiety disorders: a randomised controlled trial. *Journal of Behavior Therapy and Experimental Psychiatry*, **45**, 186–95.

Sato, W., Kochiyama, T., Yoshikawa, S., Naito, E., Matsumura, M. (2004). Enhanced neural activity in response to dynamic facial expressions of emotion: an fMRI study. *Brain Research. Cognitive Brain Research*, **20**, 81–91.

Scharnowski, F., Weiskopf, N. (2015). Cognitive enhancement through real-time fMRI neurofeedback. *Current Opinion in Behavioral Sciences*, **4**, 122–7.

Sharot, T., Riccardi, A.M., Raio, C.M., Phelps, E.A. (2007). Neural mechanisms mediating optimism bias. *Nature*, **450**, 102–5.

Sherman, S.M., Guillery, R.W. (1996). Functional organization of thalamocortical relays. *Journal of Neurophysiology*, **76**, 1367–95.

Simpson, J.R., Ongur, D., Akbudak, E., *et al.* (2000). The emotional modulation of cognitive processing: an fMRI study. *Journal of Cognitive Neuroscience*, **12**(Suppl 2), 157–70.

Spreckelmeyer, K.N., Krach, S., Kohls, G., *et al.* (2009). Anticipation of monetary and social reward differently activates mesolimbic brain structures in men and women. *Social Cognitive and Affective Neuroscience*, **4**, 158–65.

Squire, L.R. (1992). Memory and the hippocampus: a synthesis from findings with rats, monkeys, and humans. *Psychological Review*, **99**, 195–231.

Stevens, E.S., Behar, E., Jendrusina, A.A. (2018). Enhancing the efficacy of cognitive bias modification for social anxiety. *Behavior Therapy*, **49**, 995–1007.

Stoeckel, L.E., Garrison, K.A., Ghosh, S., *et al*. (2014). Optimizing real time fMRI neurofeedback for therapeutic discovery and development. *Neuroimage Clinical*, **5**, 245–55.

Sulzer, J., Haller, S., Scharnowski, F., *et al*. (2013). Real-time fMRI neurofeedback: progress and challenges. *NeuroImage*, **76**, 386–99.

Suslow, T., Kugel, H., Ohrmann, P., *et al*. (2013). Neural correlates of affective priming effects based on masked facial emotion: an fMRI study. *Psychiatry Research*, **211**, 239–45.

Taylor, C.T., Aupperle, R.L., Flagan, T., *et al*. (2014). Neural correlates of a computerized attention modification program in anxious subjects. *Social Cognitive and Affective Neuroscience*, **9**, 1379–87.

Vilberg, K.L., Rugg, M.D. (2008). Memory retrieval and the parietal cortex: a review of evidence from a dual-process perspective. *Neuropsychologia*, **46**, 1787–99.

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., Shannon, B.J., Kahn, I., Buckner, R.L. (2005). Parietal lobe contributions to episodic memory retrieval. *Trends in Cognitive Sciences*, **9**, 445–53.

Watanabe, M., Sakagami, M. (2007). Integration of cognitive and motivational context information in the primate prefrontal cortex. *Cerebral Cortex*, **17**(Suppl 1), i101–9.

Watson, D., Friend, R. (1969). Measurement of social-evaluative anxiety. *Journal of Consulting and Clinical Psychology*, **33**, 448–57.

Winkielman, P., Berridge, K.C., Wilbarger, J.L. (2005). Unconscious affective reactions to masked happy versus angry faces influence consumption behavior and judgments of value. *Personality and Social Psychology Bulletin*, **31**, 121–35.

Winkielman, P., Knutson, B., Paulus, M., Trujillo, J.L. (2007). Affective influence on judgments and decisions: moving towards core mechanisms. *Review of General Psychology*, **11**, 179–92.

Wolpert, D.M., Goodbody, S.J., Husain, M. (1998). Maintaining internal representations: the role of the human superior parietal lobe. *Nature Neuroscience*, **1**, 529–33.

Yeung, E.S., Sharpe, L. (2019). Cognitive bias modification for social anxiety: the differential impact of modifying attentional and/or interpretation bias. *Cognitive Therapy and Research*. doi: 10.1007/s10608-019-10012-3.

Yiend, J., Mackintosh, B., Mathews, A. (2005). Enduring consequences of experimentally induced biases in interpretation. *Behaviour Research and Therapy*, **43**, 779–97.