Available online at www.sciencedirect.com
ScienceDirect
Biological Psychology 79 (2008) 80–90
BIOLOGICAL PSYCHOLOGY
www.elsevier.com/locate/biopsycho

The neural correlates of verbal encoding and retrieval in monozygotic twins at low or high risk for depression and anxiety

Saskia P.A. Wolfensberger [a,b,*], D.J. Veltman [a,b], W.J.G. Hoogendijk [a,b], M.B. De Ruiter [a], D.I. Boomsma [b,c], E.J.C. de Geus [b,c]

[a] Department of Psychiatry, Vrije Universiteit Medical Center, Amsterdam, The Netherlands
[b] Centre for Neurogenomics and Cognitive Research (CNCR), Amsterdam, The Netherlands
[c] Department of Biological Psychology,Vrije Universiteit, Amsterdam, The Netherlands

Received 15 November 2007; accepted 9 January 2008
Available online 18 January 2008

Abstract

Emotional processing and brain activation were examined during an encoding and recognition paradigm using emotionally salient words in a sample of monozygotic twin pairs at low or high risk for anxiety and depression. Discordant twin pairs were used to chart the effects of environmental risk factors and concordant twin pairs were used to chart the effects of genetic risk factors on performance and brain activation.

Performance data did not support the existence of a negative response bias in subjects at high risk. At the neural level, however, increased left inferior frontal gyrus (LIFG) activation by negative words was found in high-risk subjects, most prominently during recognition. Increased LIFG activity was found in subjects at high risk through either genetic or environmental risk factors. These results suggest that fMRI activation of the LIFG in a verbal emotional memory task may be a useful vulnerability marker for anxiety and depression.
© 2008 Elsevier B.V. All rights reserved.

Keywords: fMRI; Encoding; Retrieval; Twins; Anxiety; Depression

1. Introduction

Psychological theories of major depression have emphasized the role of negative biases in information processing in the etiology and the maintenance of the disorder (Beck, 1967). Such biases have been reported both for the interpretation and storage of emotional information (Bradley et al., 1995; Leppanen, 2006; Murphy et al., 1999; Phillips et al., 2003). For example, in facial expression recognition tasks, depressed patients show a bias away from positive towards negative emotions, i.e. reduced recognition of positive expressions and increased perception of negative expressions (Gur et al., 1992; Bouhuys et al., 1995; Surguladze et al., 2004). Negative perceptual and memory biases have also been found in healthy volunteers following negative mood induction (Bouhuys et al., 1995; Teasdale and Russell, 1983). Furthermore, there is evidence that abnormal emotional processing may persist in the non-depressed state (Bhagwagar et al., 2004; Hayward et al., 2005; Leppanen, 2006) and is associated with poor outcome (Beevers et al., 2007). These observations raise the possibility that emotional biases might pre-date the onset of clinical depression and thereby represent a risk factor for the subsequent development of illness in predisposed individuals.

To investigate emotional bias prior to the onset of depression it is necessary to study people who are at risk for anxiety and depression but who are not clinically depressed. Until now, studies combining brain imaging and neuropsychological testing in groups at high risk for depression and anxiety, but not yet affected, have been scarce. Moreover, these studies have focused primarily on executive functioning, whereas functional imaging studies investigating emotion processing in a memory paradigm have been lacking. Here we examined emotion processing, memory performance, and brain activation during an encoding and recognition paradigm using emotionally salient words in a sample of longitudinally followed monozygotic (MZ) twin pairs selected to be at low or high risk for anxiety disorder and major depression based on their ratings on neuroticism, anxiety, and depression in longitudinal surveys. A

* Corresponding author at: Department of Psychiatry, Nuclear Medicine and PET Research, Vrije Universiteit Medical Center, P.O. Box 7057, 1007 MB Amsterdam, The Netherlands. Tel.: +31 20 4441728; fax: +31 20 4449636.
E-mail addresses: saskiawolfensberger@hotmail.com, spa.wolfensberger@vumc.nl (S.P.A. Wolfensberger).

0301-0511/$ – see front matter © 2008 Elsevier B.V. All rights reserved.
doi:10.1016/j.biopsycho.2008.01.002

compound risk score for anxiety and depression based on these ratings was shown to have strong predictive validity for clinical anxiety and clinical depression in this population (Middeldorp et al., 2004) as assessed by the Composite International Diagnostic Interview, a well-validated instrument to assess these disorders (Andrews and Peters, 1998). High-risk twins were included only, however, when they were not clinically depressed or anxious at the time of MRI recording, as assessed by a structured psychiatric interview.

About 60% of the risk for anxiety and depressive disorders can be attributed to environmental factors (Kendler and Prescott, 1999; Sullivan et al., 2000) and interaction of genetic and environmental factors has also been reported (Caspi et al., 2000; Eley et al., 2004; Grabe et al., 2005; Kaufman et al., 2006; Kendler et al., 1986). It is currently unclear whether environmental risk factors have the same pathogenic effects as genetic risk factors. It is quite possible that these two types of risk impact on emotional processing but through entirely different routes. To separate the effects of genetic and environmental risk factors on emotional biases, we here employed the concordant and discordant twin pair design described earlier (de Geus et al., 2007). To detect the effects of environmental risk factors, performance and brain activation were compared in MZ twin pairs that were discordant for the risk for anxiety and depression. In these pairs, one twin scored high on neuroticism, anxiety, and depression in longitudinal surveys, whereas the co-twin scored low on these measures. Because monozygotic twins are (nearly) always 100% identical at the DNA sequence level (Boomsma et al., 2002), this discordance in the risk for anxiety and depression must arise from differential exposure to unique environmental influences. Differences in performance and brain activation during encoding and retrieval of emotionally salient words between the high-risk twin and the low-risk co-twin, therefore, must reflect environmental effects on emotional biases. To detect the effects of genetic risk factors on emotional biases we compared the performance and brain activation during encoding and recognition in twin pairs concordant for low risk with that in twin pairs concordant for high risk, i.e., both MZ twins of the pair scored low or high on neuroticism, anxiety, and depression in longitudinal surveys. The contrast between these two groups mainly reflects a difference in genetic risk rather than in environmental risk because parents of the high-risk twins also scored very high on neuroticism, anxiety, and depression in the longitudinal surveys, whereas parents of the low-risk twins scored very low (de Geus et al., 2007).

Neuroimaging studies of verbal encoding and recognition to date have found that the main neural correlates of semantic processing and retrieval of words are the left inferior frontal gyrus in healthy volunteers (LIFG, Fletcher and Henson, 2001), reflecting semantic elaboration in working memory, and the medial temporal lobe (MTL), including the fusiform gyrus and hippocampus (Squire and Knowlton, 2000), the latter thought to reflect intermediate memory storage. Taken the hypothesized emotional bias, we expected twins at high risk for anxiety and depression to show higher levels of semantic elaboration at encoding, reflected by differential activity in LIFG and/or MTL, and better memory performance during recognition for negative words, again paired to differential activity in LIFG and/or MTL. For neutral stimuli we did not expect to see an effect of the risk for anxiety and depression on either performance or brain activation and for positive stimuli we expected effects in the opposite direction. The concordant/discordant twin design further enabled us to explore whether environmental and genetic risk factors for anxiety and depression influence performance and brain activation during verbal encoding and recognition in different ways.

2. Material and methods

2.1. Subjects

Selection of subjects has been described in detail by de Geus et al. (2007). Briefly, in a sample of 2455 same sex twin pairs registered in the Netherlands Twin Registry, a compound risk score for anxiety and depression was computed based on a genetic factor analysis of longitudinal survey data on trait anxiety, depression, neuroticism and somatic anxiety (for details see Boomsma et al., 2000). The surveys were collected in 1991, 1993, 1997 and 2000. An average risk score was computed for each twin across all available surveys (which could vary from one to four surveys).

Discordant MZ twin pairs were considered eligible for participation if both were right-handed, the high-risk subject had a risk score at least 0.5 standard deviation above the mean, and the score of high- and low-risk twins were at least 2 standard deviations apart. To estimate the relative risk for actual anxiety and depression disorder in twins ascertained by these criteria, we selected all subjects with comparable low- or high-risk scores from the larger sample of 1256 subjects that underwent a CIDI interview in 2000 (Middeldorp et al., 2004). To resemble the observed mean risk scores in our discordant twins (-0.62 vs. 0.97, S.D. = 0.74), we selected subjects with a risk score of at least 1.5 standard deviation above the mean (1.06) on the 1991, 1993, and 1997 surveys and subjects with a risk score of at least 0.5 standard deviations below the mean (-0.32). In this sample the relative risk to receive a lifetime depression diagnosis in the high-scoring subjects was 11.8 compared to low-scoring subjects and 40.1 for generalized anxiety disorder.

A total of 31 discordant MZ twin pairs met our criteria, of which we invited 17 pairs because they lived near Amsterdam and had filled out our most recent survey. Two pairs were excluded because one of the members had epilepsy or was pregnant. Four pairs refused to participate, mainly out of time constraints. One pair turned out to be dizygotic. This left a final 10 MZ pairs who were extremely discordant for the risk for anxiety and depression. Notably, the risk scores of the high-risk twin in these final 10 pairs did not differ from the risk scores in the high-risk twin from the original 31 selected pairs.

Concordant MZ twin pairs were considered eligible for participation if both were right-handed and both their risk scores were at least 0.8 standard deviation above (high risk) or below (low risk) the mean risk score. An even more stringent selection of extreme risk scores was possible here compared to the discordant MZ twin selection, because extreme scoring concordant MZ pairs are much less rare than extremely discordant pairs. This yielded 115 concordant high-risk and 137 concordant low-risk pairs, of which we invited 48 pairs that lived near Amsterdam and had filled out our most recent survey. Five pairs were excluded because one of the members had a medical illness or was pregnant. Twenty-one pairs refused to participate, mainly out of time constraints. This left a final 15 MZ pairs who were concordant for low risk and 7 MZ pairs who were concordant for high risk for anxiety and depression. Of the total group of 64 MZ, 28 were male and the mean age was 30 years (range 20–42 years).

2.2. Procedure

Subjects visited the outpatient MR unit and experimental procedures were explained in detail. Twins from the same pair always came on the same day. Twins were randomly assigned to an MRI scan session or a psychometric session. After about 90 min twins switched between sessions. During the

psychometric session, cognitive abilities and current psychiatric diagnostic state were assessed. All subjects were interviewed using the Composite International Diagnostic Interview (Peters and Andrews, 1995; Wittchen, 1994), a clinical interview that assessed the occurrence of a current or recent depressive episode (6 months). The Montgomery Asberg Depression Rating Scale and the Beck Depression Inventory were used to assess depressive symptom characteristics and severity scores (Beck et al., 1961; Montgomery and Asberg, 1979). Furthermore, the state version of the State-Trait Anxiety Inventory (Spielberger et al., 1970) was administered pre- and post-scanning. Verbal comprehension (IQ) and working memory (forward and backward recall scores of digit scan) were assessed using subtests of the Wechsler Adult Intelligence Scale (Wechsler, 1997). Finally, social support was measured using the Duke-UNC questionnaire (Broadhead et al., 1988) and subjects were asked to recall the occurrence of 21 major life events. These included individual (e.g. maltreatment, disease, financial problems, job strain, relational problems) events as well as network-related events (e.g. disease or loss of close kin). Subjects were asked to locate these events in four temporal valences, i.e. whether they occurred the last 6 months, between 6 and 12 months, between 1 and 5 years or more than 5 years ago. For each event they indicated the impact on their lives on a 10-point visual analogue scale ranging from 'no impact' to 'extreme impact'.

During each MRI session, which lasted 45 min, subjects performed two tasks using emotionally relevant stimuli (faces and words). At first, a verbal memory task took place, followed by an emotional faces paradigm, results of which will be reported elsewhere. Between the encoding and recognition phase of the memory paradigm a structural MR scan was performed. At the end of both sessions subjects were debriefed and received sets of buccal swabs to collect mucosal cells for DNA extraction. DNA was used to confirm zygosity using 11 highly polymorphic markers. The ethical review board of the VU medical centre approved the study and all participants provided written informed consent.

2.3. Task paradigm

Prior to scanning, all participants practiced the event-related encoding paradigm outside the scanner on a personal computer. In the scanner, a device with response buttons was positioned near the right hand of the participant. During all task blocks, participants had to respond in a forced choice fashion by pushing the left button with the index finger or the middle button with the middle finger or the right button with the ring finger. Stimuli were presented in a self-paced fashion, although a time limit of 5 s was maintained in case of non-responses. The event-related retrieval paradigm followed 15 min after the encoding phase and was explained and practiced inside the scanner and presented by surprise.

Participants were requested to judge the valence of the word during the encoding phase. During recognition, they were requested to indicate whether or not each word had appeared previously in the encoding task. On each trial, response options were indicated at the bottom of the screen by an arrow pointing to the left ('⋘negative' in the encoding; '⋘seen' in the retrieval), and middle ('⋙neutral⋘', '⋙probably seen⋘') and right ('positive⋙'; 'not seen⋙'). Beforehand, it was explained that pushing the left button corresponded to the left arrow, etc. Scores were only registered when participants responded within the 5-s time limit. To increase experimental power and to minimize expectancy effects, a variable interstimulus interval (200–600 ms, offset to onset) was used.

Forty baseline trials were added to each phase (encoding and recognition). During a baseline trial, participants were presented with a cue to press either the left ('⋘left') or the middle ('⋘middle⋙') or the right button ('right⋙'). All types of baseline stimuli were presented equally often.

2.4. Material

The stimuli consisted of 240 Dutch words and 80 baseline stimuli. One third of the words had a negative connotation, one third of the words had a neutral connotation and one third of the words had a positive connotation. The valence of the stimulus material was validated in a perceptual clarification task (Ter Laak, 1992, unpublished Master's thesis), in which these words were recognized most consistently and rapidly as negative, neutral and positive words under conditions of minimal presentation conditions. A subset of 40 positive, 40 neutral and 40 negative words were randomly selected and ordered for the affective evaluation task (encoding). All three subsets were matched for word length, word type (verbs, adjectives and nouns) and frequency of usage. The use of abstract words was avoided. For the recognition phase, another subset of 40 positive, 40 neutral and 40 negative words were randomly selected in the same way and designated as 'new' words.

Before the experiment, the trials were randomly intermixed into blocks. For the encoding phase, 20 blocks consisting of 2 positive, 2 negative, 2 neutral words and 2 baselines were randomised. The resulting random sequence of 160 stimuli was used for all subjects to optimise comparison between subjects. For the recognition phase 20 blocks consisting of 2 positive, 2 neutral and 2 negative words from the encoding phase and 2 positive, 2 neutral and 2 negative new words and 2 baselines were randomised. Half of the words were 'new' words and half were old words presented during the encoding phase. As for encoding, the same random sequence of 280 stimuli was generated once and then used for all subjects. To prevent primacy and recency effects in the recognition task, three buffer words proceeded and followed the encoding stimuli. In addition, to let the participants get used to the unexpected recognition task, three buffer words preceded this phase. No feedback regarding the answers was provided during the experiment.

2.5. Performance measures

Performance during the encoding phase was measured by the error rates and mean reaction times (RT) for words that were later correctly recognized as 'seen' with certainty. In the recognition phase, sensitivity (Pr = proportion hits − proportion false alarms) and response bias Br (=false alarms/(1 − Pr)) were calculated for encoded negative, neutral and positive items, according to two-high-threshold theory (Snodgrass and Corwin, 1988).

2.6. Image acquisition

Magnetic resonance imaging was performed on a 1.5-T Sonata MR system (Siemens, Erlangen, Germany) with a standard RF receiver head coil. Stimuli were generated by a Pentium PC and projected on a screen at the end of the scanner table, which was seen through a mirror mounted above the subject's head. Two magnet-compatible four-key response boxes were used to record subject's performance and reaction times (RTs). To reduce motion artefacts, the subject's head was immobilized using foam pads.

For functional MRI, an echo planar imaging sequence (TR 3.04 s, TE 45 ms, matrix 64 × 64, field of view 192 mm × 192 mm, flip angle 90°) was used, creating transversal whole-brain acquisitions (35 slices, 3 × 3-mm in-plane resolution, 2.5-mm slice thickness, 0.5-mm interslice gap).

2.7. Data analysis

Differences in the questionnaire- and interview-based variables were examined by a mixed model ANOVA (MIXED SPSS) with group (discordant low risk, discordant high risk, concordant low risk, concordant high risk) as a fixed factor and family as a random factor to account for within-family dependence. With respect to performance, error rates and mean reaction times (RT) for correct responses during encoding and hits and false alarms as well as the two signal detection measures during recognition were tested by a similar MIXED ANOVA that additionally added word valence (negative, neutral and positive) as a second fixed factor. Primary planned contrasts in main or interaction effects involving groups were the comparison of the low-risk twin versus the high-risk co-twin in discordant pairs and the concordant low-risk twin pairs versus the concordant high-risk twins.

Imaging data were analysed using SPM2 (Statistical Parametric Mapping; Wellcome Department of Cognitive Neurology, London, UK). After the first two volumes were discarded to allow for magnetic saturation, time series were corrected for differences in slice acquisition times and realigned. Next, data were warped to MNI space as defined by the SPM EPI template, and spatially smoothed using an 8-mm FWHM Gaussian kernel. After spatial preprocessing, data were analysed using delta functions convolved with a canonical hemodynamic response function to model responses to each stimulus type. For each

subject, weighted contrasts were computed for 4 simple main effects: this was done across all word valences (correctly remembered words vs. baseline), and for each word valence separately (correctly remembered positive/negative/neutral words vs. baseline). The resulting contrast images were entered into second-level (random effects) analyses for between subject comparisons. A paired *t*-test was used for the comparison of the low-risk twin versus the high-risk co-twin in discordant pairs (Twin) and a one-way ANOVA for the comparison of concordant low- and high-risk pairs (Group). SPM2's non-sphericity option was used in the latter analysis to account for within-pair correlated observations. Main effects of all words across groups are reported at $p < 0.05$ FDR-corrected for multiple comparisons with an extent threshold of five voxels (Genovese et al., 2002). Group interactions across word valence and interactive effects of Group/Twin by Word valence (negative vs. baseline, neutral vs. baseline, positive vs. baseline; masked with the relevant main effect) are reported at $p < 0.001$ uncorrected, unless indicated otherwise, also with an extent threshold of five voxels.

3. Results

Demographic data and depression/anxiety ratings (BDI, MADRS, STAI-state) obtained at the time of the scanning session are listed in Table 1. Age and male/female distribution were not significantly different across the groups. Mixed ANOVA confirmed that the concordant low-risk pairs scored significantly lower on the MADRS, BDI, and state anxiety measures than the concordant high-risk pairs. Within the discordant pairs, the MADRS, BDI, and state anxiety measures all showed significant intra-pair differences in the expected direction. Only one subject in a concordant high-risk pair received a current diagnosis of depression using these instruments. Furthermore, one subject currently used anti-depressant medication (SSRI). This was the subject with the high anxious depression score from a discordant pair. These two subjects were excluded from all further analyses. Data from one concordant low-risk pair were lost due to technical problems during scanning.

3.1. Behavioral data

3.1.1. Encoding phase (affective classification)

Across all four groups (concordant low risk, concordant high risk, discordant low risk, and discordant high risk) reaction times for words later correctly recognized with certainty were fastest for negative words (mean reaction time negative words: 1106 ± 251 ms vs. neutral words: 1332 ± 345 ms and positive words: 1235 ± 331 ms; $F(2, 127.3) = 16.15$, $p < 0.001$). Error rate for the affective classification of successfully recognized words was also lower for negative words (negative words: 9% vs. neutral words: 27% and positive words: 21%; $F(2, 121.8) = 26.3$, $p < 0.001$). Importantly, no main group or group × valence interaction effects were found suggesting that affective classification does not differ as a function of the risk for anxiety and depression.

3.1.2. Recognition phase (retrieval)

Performance data were missing during retrieval for four subjects and these were excluded from further analyses of the performance data. Hit rates, false alarm rates (FA), sensitivity (Pr) scores and response bias (Br) for the remaining 56 subjects are shown in Table 2. Valence had a significant effect on the number of hits ($F(2, 125.4 = 3.39$, $p = 0.039$), false alarms ($F(2, 126.0) = 25.2$, $p < 0.001$), sensitivity ($F(2, 126.5) = 7.1$, $p = 0.001$), and Br ($F(2, 124.6) = 22.8$, $p < 0.001$). More hits as well as false alarms were made on positive and negative words than on neutral trials, and neutral trials had a smaller response bias. These effects were due to the fact that neutral trials were recognized with certainty much less often than positive or negative words. Sensitivity, however, was significantly lower for negative words than for either neutral or positive words.

A main effect of group was found on sensitivity only ($F(3, 37.8) = 6.2$, $p = 0.002$), which did not interact with valence. Post hoc comparisons revealed that, across all word valences, a significant difference between the high-risk twin and the low risk twin from discordant pairs was found ($p < 0.001$). A trend in a similar direction was seen when comparing the concordant high-risk pairs with the concordant low-risk pairs ($p = 0.079$). In all instances the high-risk subjects performed better. Across all word valences they had higher hit rates and lower false alarm rates than the low-risk subjects resulting in higher sensitivity scores.

Table 1
Characteristics of the twins at the time of MRI scanning

	Low risk concordant twins ($N = 30$)	Discordant twin pairs		High risk concordant twins ($N = 14$)
		Low risk twin ($N = 10$)	High risk twin ($N = 10$)	
Male/female, no.	14/16	4/6	4/6	6/8
Age, mean (year)	30.9	30.6	30.6	26.1
BDI depression, mean (S.D.)[a,b]	1.1 (1.5)	2.7 (2.5)	9.7 (10.6)	8.0 (6.0)
MADRS, mean (S.D.)[a,b]	0.43 (1.5)	2.1 (2.5)	5.0 (3.9)	5.1 (7.8)
STAI state anxiety before scan session, mean (S.D.)[a,b]	27.7 (5.0)	31.4 (5.3)	37.0 (7.3)	36.7 (9.8)
STAI state anxiety after scan session, mean (S.D.)[a,b]	25.1 (4.7)	28.3 (5.5)	37.1 (8.1)	33.8 (8.4)
Verbal IQ subscale, mean (S.D.)	13.1 (2.5)	14.5 (2.2)	14.6 (2.2)	12.9 (3.9)
Working memory IQ subscale, mean (S.D.)	7.6 (1.3)	8.9 (2.4)	8.3 (3.2)	8.6 (1.6)

[a] Significant difference between low risk and high risk concordants.
[b] Significant intra-pair difference in discordant twins.

Table 2
Mean proportions (S.D. in parentheses) of hit rates, false alarm rates (FA), sensitiviy (Pr) and response bias (Br) across all valences and groups and separately per group and word valence

	Concordant twins		Discordant twins		All groups
	Low risk	High risk	Low risk	High risk	
Hits					
Negative	0.72 (0.12)	0.74 (0.10)	0.72 (0.23)	0.73 (0.13)	0.72 (0.14)
Neutral	0.59 (0.19)	0.73 (0.15)	0.67 (0.16)	0.74 (0.15)	0.66 (0.17)
Positive	0.67 (0.17)	0.79 (0.09)	0.73 (0.23)	0.78 (0.10)	0.73 (0.16)
All words	0.66 (0.14)	0.75 (0.09)	0.71 (0.18)	0.75 (0.09)	0.71 (0.14)
FA					
Negative	0.23 (0.16)	0.20 (0.13)	0.22 (0.18)	0.16 (0.12)	0.21 (0.15)
Neutral	0.09 (0.07)	0.07 (0.06)	0.12 (0.16)	0.06 (0.04)	0.09 (0.09)
Positive	0.18 (0.14)	0.18 (0.11)	0.18 (0.22)	0.11 (0.07)	0.17 (0.15)
All words	0.17 (0.11)	0.15 (0.09)	0.17 (0.18)	0.11 (0.07)	0.15 (0.12)
Pr					
Negative	0.49 (0.17)	0.54 (0.16)	0.50 (0.17)	0.57 (0.14)	0.51 (0.16)
Neutral	0.50 (0.18)	0.66 (0.14)	0.55 (0.21)	0.68 (0.18)	0.58 (0.19)
Positive	0.49 (0.19)	0.62 (0.15)	0.56 (0.23)	0.67 (0.15)	0.56 (0.19)
All words	0.49 (0.17)	0.61 (0.13)	0.54 (0.18)	0.64 (0.14)	0.55 (0.16)
Br					
Negative	0.43 (0.21)	0.41 (0.18)	0.45 (0.32)	0.36 (0.20)	0.42 (0.22)
Neutral	0.20 (0.15)	0.23 (0.23)	0.22 (0.22)	0.18 (0.11)	0.21 (0.18)
Positive	0.35 (0.22)	0.45 (0.22)	0.39 (0.31)	0.31 (0.12)	0.37 (0.23)
All words	0.33 (0.16)	0.36 (0.17)	0.35 (0.27)	0.28 (0.10)	0.33 (0.18)

3.2. Imaging data

Results of random effects (RFX) analyses of the imaging data are summarized in Tables 3–5 and are shown in 3D visualizations and cross sections in Figs. 1 and 2.

3.2.1. Encoding phase

Fig. 1 shows the average BOLD effects across all 4 groups when comparing all correctly encoded words against baseline stimuli (glass brain).

Across groups, emotional classification of words compared to baseline was associated with robustly increased blood oxygenation dependent (BOLD) signal in left inferior frontal gyrus (LIFG). Furthermore, regions showing increased activity were bilateral hippocampus (HC), as well as dorsal anterior cingulate cortex (ACC), and visual processing areas, including bilateral occipital cortex and left fusiform gyrus, situated in the ventral route (cf. Table 3).

3.2.1.1. Modulation by the risk for anxiety and depression. Across word valences, we observed increased activity in LIFG (45, 42, 6; $Z = 3.34$) in the concordant high-risk group compared to the concordant low-risk group. Conversely, decreased left hippocampal activity was found in concordant high-risk group compared to concordant low-risk group ($Z = 3.15$). Across word valences, we did not find significant differences between high-risk discordant twins and their low-risk co-twins.

Interaction analyses by word valence (negative, positive and neutral; Table 4) showed increased activation during encoding of negative words in concordant high-risk twins in left occipital cortex ($Z = 3.20$). Conversely, decreased activation in left hippocampus was found. During encoding of neutral words, concordant high-risk twins showed increased activation in occipital cortex and left fusiform gyrus compared to concordant low-risk twins. At a slightly lower threshold ($p = 0.003$, $Z = 2.71$), more LIFG activation was also found in these high-risk twins. Interaction analyses of encoding of positive words showed increased activation in concordant high-risk twins in the LIFG and right amygdala. In discordant twins, only one

Table 3
Encoding: BOLD main effects for all correctly encoded words vs. baseline across groups at $p < 0.05$, FDR corrected for multiple comparisons

Region	x	y	z	Z value
Prefrontal				
Lateral inferior				
L	−48	27	3	6.82
L	−51	30	−6	5.94
L	−42	33	−3	4.86
Medial				
L	−18	54	30	3.31
Anterior cingulate				
L	−6	21	45	4.13
Medial temporal lobe				
R	21	−12	−21	5.04
L	−21	−24	−6	4.11
L	−18	−15	−21	3.66
L	−27	−12	−21	3.46
Occipital				
R	15	−99	12	6.21
L	−15	−90	9	6.02
R	6	−96	6	5.81

Table 4
Encoding: BOLD group interactions for word valences at $p < 0.001$ uncorrected

	Negative vs. baseline					Positive vs. baseline					Neutral vs. baseline				
	L/R	*x*	*y*	*z*	*Z* value	L/R	*x*	*y*	*z*	*Z* value	L/R	*x*	*y*	*z*	*Z* value
LC > HC															
Medial temporal	L	−21	−24	−18	3.09										
HC > LC															
Prefrontal															
Lateral inferior						L	−45	42	6	3.78					
						L	−42	36	−3	3.24					
Medial temporal											L	−15	−42	−3	3.61
Occipital	L	−15	−60	3	3.20*	L	−27	−96	18	3.55	R	9	−93	9	3.72
						L	−33	−72	−9	3.62					
											R	15	−63	−9	3.33
											R	39	−90	0	3.21
HD > LD															
Amygdala	R	27	−3	−18	3.22*										

HC = high concordant; LC = low concordant; HD = high discordant; LD = low discordant.

significant group by word valence interaction was found: for negative words more right amygdala was found in the high-risk twin compared to the low-risk co-twin ($Z = 3.22$).

3.2.2. Recognition phase

Fig. 2 depicts activation for all groups of the cortical network in response to correct recognition and correct rejection of encoded words across valences. Main regions for this comparison are listed in Table 5. Activity was left lateralized, like in the encoding task, although right hemisphere activation was more apparent in the recognition phase. We found the activation of the LIFG, dorsal ACC, and bilateral occipital cortex, areas that were also active during encoding. In addition, we found bilateral insular cortex activation.

3.2.2.1. Modulation by the risk for anxiety and depression. Across word valences, concordant high-risk twins showed increased LIFG compared to concordant low-risk twins ($Z = 3.41$), whereas in discordant twins we failed to observe significant effects of risk status. Interaction analyses with valence showed that compared to concordant low-risk twins, concordant high-risk twins showed increased activation in the LIFG only during negative words (Table 6). In addition, during retrieval of negative words increased LIFG and RIFG activation was found in discordant high-risk twins compared to their discordant co-twin albeit LIFG at a lower statistical level (Table 6).

Table 5
Retrieval: BOLD main effects for all correctly recognized/rejected words vs. baseline across groups at $p < 0.05$, FDR corrected for multiple comparisons

Region	*x*	*y*	*z*	*Z* value
Prefrontal				
Lateral inferior				
L	−45	27	21	4.05
L	−57	30	24	3.41
Anterior cingulate				
R	6	27	42	3.32
Insula				
L	−30	21	−3	4.59
R	30	24	−3	5.14
Occipital				
L	−39	−78	−12	5.91
L	−24	−93	12	5.63

4. Discussion

The goal of the present study was to investigate possible emotional biases in subjects who were at risk for anxiety and depression through either genetic or environmental factors, but not yet clinically affected. To this end, we employed an incidental encoding task followed by an explicit retrieval task for emotional versus neutral verbal information in monozygotic twins, that were either strongly concordant or very discordant for scores on neuroticism, anxiety, and depression in longitudinal surveys. We expected that the high-risk twins would be characterized by differential encoding and retrieval of emotional words, particularly negative words, both on a behavioral and a neural level.

In contrast to our expectation, behavioral data acquired during the encoding phase did not show differences between high-risk and low-risk twins. During recognition, we did observe improved memory performance in high-risk subjects compared to low-risk subjects, but this was true across all types of words, and not specific for negative words. Therefore, our behavioral data do not support the hypothesis that a negative response bias may be present in subjects at risk for depression/anxiety. This was true for subjects at risk through environmental factors (discordant pairs) as well through genetic factors (concordant pairs).

Our findings are at odds with previous studies showing increased recall of recently learned negatively valence words (Rinck and Becker, 2005) and decreased recognition of happy facial expressions (Gotlib et al., 2004; Surguladze et al., 2004) in depressed patients. These negative biases were not simply

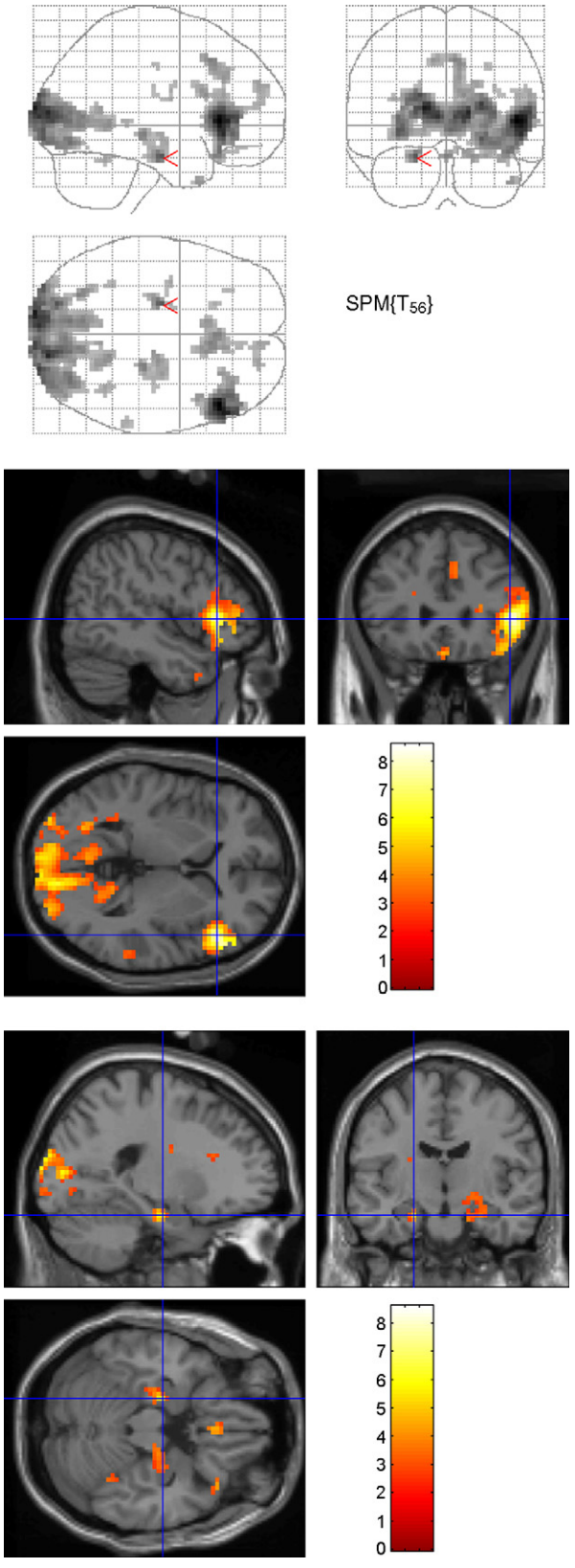

Fig. 1. Encoding. BOLD main effects across all 4 groups and across all words. Top: glass brain; middle and lower panel: 3D visualization.

due to depressive state at the time of scanning, as they tended to persist in patients recovered from clinical depression (Leppanen, 2006). Also, a recent study (Joormann et al., 2007) showed a clear negative bias in a dot probe task in young people at risk for depression through having a depressed mother, but not yet depressed. This latter study, however, employed a negative mood induction paradigm. Negative perceptual and memory biases following negative mood induction have also been found in healthy volunteers without a family history (Bouhuys et al., 1995).

Results of other studies do converge with ours. A recent study in non-depressed people at familiar risk for depression failed to reveal evidence for a negative emotional bias in either a facial expression or an emotion categorization task (Mannie et al., 2007). With regard to memory performance, Pine et al. (2004) did not find altered recall of emotional expressions in unaffected offspring of depressed parents. Furthermore, in a recent study in elderly non-depressed first-degree relatives of patients with major depression compared to controls, no evidence was found for a negative bias in facial expression recognition accuracy or word recall (Le Masurier et al., 2007). In this study, the depressed relative group responded faster during recognition of facial expressions of fear, but not for anger, disgust and sad faces. In addition, relatives responded slower when they had to recognize positive personality characteristics in a verbal categorization task, but there were no differences for negative words between groups (Le Masurier et al., 2007).

Taking together findings from the above studies with the present findings, we are led to conclude that subjects at risk for depression, either through genetic or environmental determinants, do not exhibit a negative emotional bias at the behavioral level, unless emotional material is presented immediately after negative mood induction. Only in a setting of lowered mood vulnerable individuals might respond with larger negative emotional biases. This would potentially place them at greater risk of experiencing more severe or prolonged dysphoric reactions in response to negative stimuli (Mannie et al., 2007). The above conclusion, however, is based only on the behavioral results and needs to be interpreted against the background of parallel responses at the neural level.

For BOLD activation we also hypothesized differential results for low- and high-risk subjects, which we expected to be most prominent during negative emotional words. Overall, imaging results were in line with previous studies using this paradigm. During encoding, emotional classification of words compared to baseline was associated with a robustly increased BOLD signal in the dorsal ACC and the LIFG, the latter reflecting semantic elaboration of words. Furthermore, regions showing increased activity were the bilateral occipital cortex, bilateral hippocampal region, and the left posterior medial temporal cortex, probably reflecting the visual analysis of written word forms (orthographic processing). During explicit retrieval in the recognition phase we observed similar brain regions as in the encoding phase, although activity tended to be more bilateral compared to the relative left lateralization during encoding, and MTL activity was absent. These areas

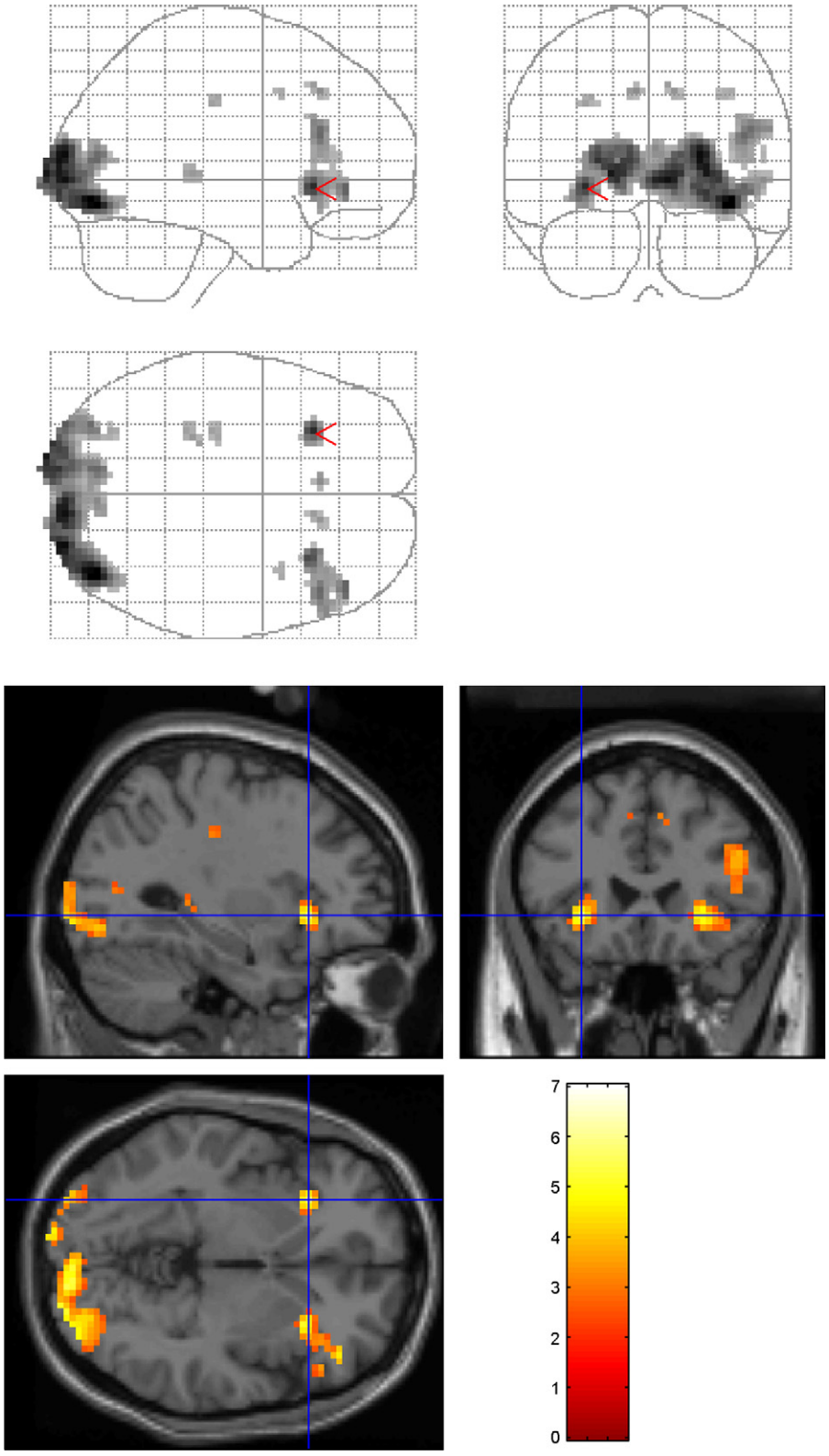

Fig. 2. Retrieval. BOLD main effect for all words across groups. Top: glass brain and lower panel: 3D visualization.

correspond closely to those reported in other fMRI studies of lexical decision tasks indicating the involvement of neural circuits responsible for semantic processing (Jobard et al., 2003) and verbal declarative memory (Buckner and Koutstaal, 1998). Although hippocampal activity has been reported previously in fMRI studies using a similar paradigm (Daselaar et al., 2004; de Ruiter et al., 2007), other investigators have failed to detect significant MTL activity. Therefore, it has been suggested that the hippocampus is mainly involved in associative memory rather than single-item memory (as in

Table 6
Retrieval: BOLD group interactions for word valences at $p < 0.001$ uncorrected

	Negative vs. baseline					Positive vs. baseline					Neutral vs. baseline				
	L/R	*x*	*y*	*z*	*Z* value	L/R	*x*	*y*	*z*	*Z* value	L/R	*x*	*y*	*z*	*Z* value
HC > LC															
Prefrontal															
Lateral inferior	L	−42	42	0	3.05*										
Insula											L	−36	33	0	3.90
HD > LD															
Prefrontal															
Lateral inferior	L	−54	33	6	2.65[#]										
	R	48	36	18	3.83										
Hippocampus											L	−36	−21	−21	3.37

HC = high concordant; LC = low concordant; HD = high discordant; LD = low discordant.
[#] $p = 0.004$.

our task), and that hippocampal activity observed during single-item memory tasks is due to e.g. novelty effects (Squire et al., 2004, for an overview). The issue is still undecided, as a recent study demonstrated that hippocampal damage may impair both associative and single-item memory (Gold et al., 2006), and that familiarity-based recognition for complex visual stimuli is supported by the hippocampus, as demonstrated by single-cell recordings (Rutishauser et al., 2006). Our data generally support the hypothesis that the hippocampus is involved during successful encoding of single items (words).

Our main interest was the modulation of activity in these areas by genetic and environmental risk for anxiety and depression. In contrast to analyses at the behavioral level we found clear evidence that brain activation during encoding and retrieval of emotional stimuli is indeed influenced by these risk factors. During encoding, we found increased activity in the LIFG in concordant high-risk twins compared to the concordant low-risk twins across all word valences. High- and low-risk individuals from discordant twin pairs did not show this difference. During recognition we again found increased activity in the LIFG in high-risk twins, this time in both discordant and concordant pairs. A parsimonious explanation for the increased LIFG activation in high-risk subjects is that it reflects their better performance during the recognition task. Previous studies in healthy volunteers have indicated that greater activation is found in left inferior frontal and fusiform regions for remembered than forgotten words (Baker et al., 2001). Importantly, however, the effect of risk on LIFG activation, but not on performance, was most evident in negatively valenced words. Therefore, our findings could indicate that high-risk subjects engage in increased elaboration of negative words compared to the low-risk subjects. Further evidence for differential effects of negative words on the low- and high-risk subjects was found in the amygdala response during encoding. In keeping with studies showing larger reactivity to negative emotional faces (Canli et al., 2005; Fu et al., 2004; Siegle et al., 2001; Surguladze et al., 2005) discordant high-risk twins showed larger right amygdala activation to negative words than their low-risk co-twin.

In conclusion, in the present fMRI study we failed to find evidence at the behavioral level for a negative response bias in subjects at high risk for anxiety and depression. At the neural level, increased LIFG activation by negative words was found in high-risk subjects, most prominently during retrieval. These fMRI results applied to subjects at high risk through either genetic or environmental risk factors. They suggest that fMRI activation of the LIFG in a verbal emotional memory task may be a useful vulnerability marker for anxiety and depression.

Acknowledgements

This study was supported by the Netherlands Organization for Scientific Research (NWO) grants 900-562-137, 904-61-090, 985-10-002, 904-61-193, 480-04-004 and 575-25-006, the Centre for Neurogenomics and Cognitive Research (CNCR) and the Centre for Medical Systems Biology (CMSB), a centre of excellence approved by the Netherlands Genomics Initiative/NWO. We acknowledge the valuable contribution of Marcel Jansen and Kim Baas to data collection in the twins.

Conflict of interest

All authors had no conflicts of interest, including specific financial interests and relationships and affiliations relevant to the subject of their manuscript.

References

Andrews, G., Peters, L., 1998. The psychometric properties of the Composite International Diagnostic Interview. Social Psychiatry and Psychiatric Epidemiology 33, 80–88.

Baker, J.T., Sanders, A.L., Maccotta, L., Buckner, R.L., 2001. Neural correlates of verbal memory encoding during semantic and structural processing tasks. Neuroreport 12, 1251–1256.

Beck, A.T., 1967. Depression: Causes and Treatment. University of Pennsylvania Press, Philadelphia.

Beck, A.T., Erbaugh, J., Ward, C.H., Mock, J., Mendelsohn, M., 1961. An inventory for measuring depression. Archives of General Psychiatry 4, 561–571.

Beevers, C.G., Wells, T.T., Miller, I.W., 2007. Predicting response to depression treatment: the role of negative cognition. Journal of Consulting and Clinical Psychology 75, 422–431.

Bhagwagar, Z., Cowen, P.J., Goodwin, G.M., Harmer, C.J., 2004. Normalization of enhanced fear recognition by acute SSRI treatment in subjects with a previous history of depression. American Journal of Psychiatry 161, 166–168.

Boomsma, D., Busjahn, A., Peltonen, L., 2002. Classical twin studies and beyond. Nature Reviews Genetics 3, 872–882.

Boomsma, D.I., Beem, A.L., van den, B.M., Dolan, C.V., Koopmans, J.R., Vink, J.M., et al., 2000. Netherlands twin family study of anxious depression (NETSAD). Twin Research 3, 323–334.

Bouhuys, A.L., Bloem, G.M., Groothuis, T.G.G., 1995. Induction of depressed and elated mood by music influences the perception of facial emotional expressions in healthy-subjects. Journal of Affective Disorders 33, 215–226.

Bradley, B.P., Mogg, K., Williams, R., 1995. Implicit and explicit memory for emotion-congruent information in clinical depression and anxiety. Behaviour Research and Therapy 33, 755–770.

Broadhead, W.E., Gehlbach, S.H., Degruy, F.V., Kaplan, B.H., 1988. The Duke-Unc Functional Social Support Questionnaire—measurement of social support in family medicine patients. Medical Care 26, 709–723.

Buckner, R.L., Koutstaal, W., 1998. Functional neuroimaging studies of encoding, priming, and explicit memory retrieval. Proceedings of the National Academy of Sciences of the United States of America 95, 891–898.

Canli, T., Cooney, R.E., Goldin, P., Shah, M., Sivers, H., Thomason, M.E., et al., 2005. Amygdala reactivity to emotional faces predicts improvement in major depression. Neuroreport 16, 1267–1270.

Caspi, A., Taylor, A., Moffitt, T.E., Plomin, R., 2000. Neighborhood deprivation affects children's mental health: environmental risks identified in a genetic design. Psychological Science 11, 338–342.

Daselaar, S.M., Veltman, D.J., Witter, M.P., 2004. Common pathway in the medial temporal lobe for storage and recovery of words as revealed by event-related functional MRI. Hippocampus 14, 163–169.

de Geus, E.J., Ent, D.V., Wolfensberger, S.P., Heutink, P., Hoogendijk, W.J., Boomsma, D.I., et al., 2007. Intrapair differences in hippocampal volume in monozygotic twins discordant for the risk for anxiety and depression. Biological Psychiatry 61, 1062–1071.

de Ruiter, M.B., Veltman, D.J., Phaf, R.H., van Dyck, R., 2007. Negative words enhance recognition in nonclinical high dissociators: an fMRI study. Neuroimage 37, 323–334.

Eley, T.C., Sugden, K., Corsico, A., Gregory, A.M., Sham, P., McGuffin, P., et al., 2004. Gene-environment interaction analysis of serotonin system markers with adolescent depression. Molecular Psychiatry 9, 908–915.

Fletcher, P.C., Henson, R.N.A., 2001. Frontal lobes and human memory—insights from functional neuroimaging. Brain 124, 849–881.

Fu, C.H.Y., Williams, S.C.R., Cleare, A.J., Brammer, M.J., Walsh, N.D., Kim, J., et al., 2004. Attenuation of the neural response to sad faces in major depression by antidepressant treatment—a prospective, event-related functional magnetic resonance imaging study. Archives of General Psychiatry 61, 877–889.

Genovese, C.R., Lazar, N.A., Nichols, T., 2002. Thresholding of statistical maps in functional neuroimaging using the false discovery rate. Neuroimage 15, 870–878.

Gold, J.J., Smith, C.N., Bayley, P.J., Shrager, Y., Brewer, J.B., Stark, C.E., et al., 2006. Item memory, source memory, and the medial temporal lobe: concordant findings from fMRI and memory-impaired patients. Proceedings of the National Academy of Sciences of the United States of America 103, 9351–9356.

Gotlib, I.H., Kasch, K.L., Traill, S., Joormann, J., Arnow, B.A., Johnson, S.L., 2004. Coherence and specificity of information-processing biases in depression and social phobia. Journal of Abnormal Psychology 113, 386–398.

Grabe, H.J., Lange, M., Wolff, B., Volzke, H., Lucht, M., Freyberger, H.J., et al., 2005. Mental and physical distress is modulated by a polymorphism in the 5-HT transporter gene interacting with social stressors and chronic disease burden. Molecular Psychiatry 10, 220–224.

Gur, R.C., Erwin, R.J., Gur, R.E., Zwil, A.S., Heimberg, C., Kraemer, H.C., 1992. Facial emotion discrimination. 2. Behavioral findings in depression. Psychiatry Research 42, 241–251.

Hayward, G., Goodwin, G.M., Cowen, P.J., Harmer, C.J., 2005. Low-dose tryptophan depletion in recovered depressed patients induces changes in cognitive processing without depressive symptoms. Biological Psychiatry 57, 517–524.

Jobard, G., Crivello, F., Tzourio-Mazoyer, N., 2003. Evaluation of the dual route theory of reading: a metanalysis of 35 neuroimaging studies. Neuroimage 20, 693–712.

Joormann, J., Talbot, L., Gotlib, I.H., 2007. Biased processing of emotional information in girls at risk for depression. Journal of Abnormal Psychology 116, 135–143.

Kaufman, J., Yang, B.Z., Douglas-Palumberi, H., Grasso, D., Lipschitz, D., Houshyar, S., et al., 2006. Brain-derived neurotrophic factor-5-HTTLPR gene interactions and environmental modifiers of depression in children. Biological Psychiatry 59, 673–680.

Kendler, K.S., Heath, A., Martin, N.G., Eaves, L.J., 1986. Symptoms of anxiety and depression in a volunteer twin population—the etiologic role of genetic and environmental-factors. Archives of General Psychiatry 43, 213–221.

Kendler, K.S., Prescott, C.A., 1999. A population-based twin study of lifetime major depression in men and women. Archives of General Psychiatry 56, 39–44.

Le Masurier, M., Cowen, P.J., Harmer, C.J., 2007. Emotional bias and waking salivary cortisol in relatives of patients with major depression. Psychological Medicine 37, 403–410.

Leppanen, J.M., 2006. Emotional information processing in mood disorders: a review of behavioral and neuroimaging findings. Current Opinion in Psychiatry 19, 34–39.

Mannie, Z.N., Harmer, C.J., Cowen, P.J., 2007. Increased waking salivary cortisol levels in young people at familial risk of depression. American Journal of Psychiatry 164, 617–621.

Middeldorp, C.M., Cath, D.C., Beem, A.L., Boomsma, D.I., 2004. Genetic epidemiology of depression in a selected population of Dutch twins and their siblings. Behavior Genetics 34, 652–653.

Montgomery, S.A., Asberg, M., 1979. New depression scale designed to be sensitive to change. British Journal of Psychiatry 134, 382–389.

Murphy, B.C., Shepard, S.A., Eisenberg, N., Fabes, R.A., Guthrie, I.K., 1999. Contemporaneous and longitudinal relations of dispositional sympathy to emotionality, regulation, and social functioning. Journal of Early Adolescence 19, 66–97.

Peters, L., Andrews, G., 1995. Procedural validity of the computerized version of the Composite International Diagnostic Interview (Cidi-Auto) in the anxiety disorders. Psychological Medicine 25, 1269–1280.

Phillips, M.L., Drevets, W.C., Rauch, S.L., Lane, R., 2003. Neurobiology of emotion perception. II. Implications for major psychiatric disorders. Biological Psychiatry 54, 515–528.

Pine, D.S., Lissek, S., Klein, R.G., Mannuzza, S., Moulton, J.L., Guardino, M., et al., 2004. Face-memory and emotion: associations with major depression in children and adolescents. Journal of Child Psychology and Psychiatry 45, 1199–1208.

Rinck, M., Becker, E.S., 2005. A comparison of attentional biases and memory biases in women with social phobia and major depression. Journal of Abnormal Psychology 114, 62–74.

Rutishauser, U., Mamelak, A.N., Schuman, E.M., 2006. Single-trial learning of novel stimuli by individual neurons of the human hippocampus-amygdala complex. Neuron 49, 805–813.

Siegle, G.J., Granholm, E., Ingram, R.E., Matt, G.E., 2001. Pupillary and reaction time measures of sustained processing of negative information in depression. Biological Psychiatry 49, 624–636.

Snodgrass, J.G., Corwin, J., 1988. Pragmatics of measuring recognition memory—applications to dementia and amnesia. Journal of Experimental Psychology-General 117, 34–50.

Spielberger, C.D., Gorsuch, R.L., Lushene, R.E., 1970. STAI Manual for the State-Trait Anxiety Inventory. Consulting Psychologists Press, Palo Alto, CA.

Squire, L.R., Knowlton, B.J., 2000. The medial temporal lobe, the hippocampus and the memory systems of the brain. In: Gazzaniga, M.S. (Ed.), The New Cognitive Neurosciences. The MIT Press, Cambridge, MA, pp. 765–779.

Squire, L.R., Stark, C.E.L., Clark, R.E., 2004. The medial temporal lobe. Annual Review of Neuroscience 27, 279–306.

Sullivan, P.F., Neale, M.C., Kendler, K.S., 2000. Genetic epidemiology of major depression: review and meta-analysis. American Journal of Psychiatry 157, 1552–1562.

Surguladze, S.A., Young, A.W., Senior, C., Brebion, G., Travis, M.J., Phillips, M.L., 2004. Recognition accuracy and response bias to happy and sad facial expressions in patients with major depression. Neuropsychology 18, 212–218.

Surguladze, S., Brammer, M.J., Keedwell, P., Giampietro, V., Young, A.W., Travis, M.J., et al., 2005. A differential pattern of neural response toward sad versus happy facial expressions in major depressive disorder. Biological Psychiatry 57, 201–209.

Teasdale, J.D., Russell, M.L., 1983. Differential-effects of induced mood on the recall of positive, negative and neutral words. British Journal of Clinical Psychology 22, 163–171.

Wechsler, D., 1997. WAIS-III Wechsler Adult Intelligence Scale. Psychological Corporation, San Antonio, Texas.

Wittchen, H.U., 1994. Reliability and validity studies of the who Composite International Diagnostic Interview (Cidi)—a critical-review. Journal of Psychiatric Research 28, 57–84.