

**Title:** Is negative self-referent bias an endophenotype for depression? An fMRI study of emotional self-referent words in twins at high vs. low risk of depression

**Short title:** Self-referent memory in twins at high vs. low risk of depression

**Authors:** Miskowiak KW, Larsen JE, Harmer CJ, Siebner HR, Kessing LV, Macoveanu J\*, Vinberg M\*

*\*Contributed equally as last authors.*

**Affiliations:** JEL, JM, LVK, MV and KWM: Mental Health Services - Capital Region of Denmark, Psychiatric Centre Copenhagen, Copenhagen University Hospital, Rigshospitalet, Denmark, Emails: [Kamilla@miskowiak.dk](mailto:Kamilla@miskowiak.dk), [julia-evelyn@live.dk](mailto:julia-evelyn@live.dk), [Julian.macoveanu@regionh.dk](mailto:Julian.macoveanu@regionh.dk), [lars.vedel.kessing@regionh.dk](mailto:lars.vedel.kessing@regionh.dk), [maj.vinberg@regionh.dk](mailto:maj.vinberg@regionh.dk),

KWM: Department of Psychology, University of Copenhagen, Denmark

HRS: Danish Research Centre for Magnetic Resonance, Centre for Functional and Diagnostic Imaging and Research, Copenhagen University Hospital Hvidovre, Hvidovre, Denmark, Email: [hartwig.siebner@drcmr.dk](mailto:hartwig.siebner@drcmr.dk)

HRS: Department of Neurology, Copenhagen University Hospital Bispebjerg, Denmark

CH: Department of Psychiatry, University of Oxford, United Kingdom, email: [Catherine.harmer@psych.ox.ac.uk](mailto:Catherine.harmer@psych.ox.ac.uk)

**Corresponding author:** Kamilla Woznica Miskowiak, Psychiatric Centre Copenhagen, Copenhagen University Hospital, Rigshospitalet, Blegdamsvej 9, DK-2100 Copenhagen, Denmark. Email: [Kamilla@miskowiak.dk](mailto:Kamilla@miskowiak.dk)

**Word count:** Abstract: 232/ Article body (excl. abstract, acknowledgments and references): 3,750

**Figures/tables:** 0/2

**Key words:** Familial risk, depression, endophenotype, cognition, fMRI

## Abstract

**Background:** Negative cognitive bias and aberrant neural processing of self-referent emotional words seem to be trait-marks of depression. However, it is unclear whether these neurocognitive changes are present in unaffected first-degree relatives and constitute an illness endophenotype.

**Methods:** Fifty-three healthy, never-depressed monozygotic or dizygotic twins with a co-twin history of depression (high-risk group:  $n=26$ ) or no first-degree family history of depression (low-risk group:  $n=27$ ) underwent neurocognitive testing and functional magnetic imaging (fMRI) as part of a follow-up cohort study. Participants performed a self-referent emotional word categorisation task and free word recall task followed by a recognition task during fMRI. Participants also completed questionnaires assessing mood, personality traits and coping strategies.

**Results:** High-risk and low-risk twins (age, mean $\pm$ SD:  $40\pm 11$ ) were well-balanced for demographic variables, mood, coping and neuroticism. High-risk twins showed lower accuracy during self-referent categorization of emotional words independent of valence and more false recollections of negative words than low-risk twins during free recall. Functional MRI yielded no differences between high-risk and low-risk twins in retrieval-specific neural activity for positive or negative words or during the recognition of negative versus positive words within the hippocampus or prefrontal cortex.

**Conclusions:** The subtle display of negative recall bias is consistent with the hypothesis that self-referent negative memory bias is an endophenotype for depression. High-risk twins' lower categorisation accuracy adds to the evidence for valence-independent cognitive deficits in individuals at familial risk for depression.

## Introduction

Major depressive disorder (MDD) is a recurrent mental disorder affecting approximately one in five individuals in their lifetime (Kessler, 2003). While MDD is now the number one contributor to the global burden of disease (Global Burden of Disease, WHO), the aetiology and genetic origins of the disorder remain unclear. The identification of endophenotypes for MDD (i.e., the expressions of genes within behaviour) has therefore been proposed as critical for better insight into the illness aetiology, risk stratification, and development of new prevention strategies for MDD (Hasler et al., 2004; Peterson et al., 2014). An endophenotype is independent of illness state (i.e. trait-markers), and present in non-affected family members at a greater degree than in the normal population (i.e., associated with genetic predisposition for depression) (Gottesman and Gould, 2003).

Many studies have related negative information processing bias to depression including greater attention to and memory for negative versus positive emotional information (see reviews by Gotlib and Joorman, 2010; Miskowiak and Carvalho, 2014). This negative cognitive bias is associated with greater illness severity (Beeves and Carver, 2003; Johnson et al., 2007) and increased risk of depressive relapse (Bouhuys et al., 1999). In particular, emerging evidence points to negative bias in the encoding and recall of *self-referent* information as a potential depression endophenotype. **Greater encoding and recall of negative than positive self-referent adjectives has been observed not only in acutely depressed patients (Bradley and Mathews, 1983; Gotlib et al., 2004; Harmer et al., 2009) but also during periods of remission (Romero, Sanchez, and Vazquez, 2014).** At a neural level, patients display **enhanced activity in the rostral anterior cingulate and medial prefrontal cortex (mPFC) during encoding of negative self-referent words (Yoshimura et al., 2010), reduced hippocampal activation during encoding of positive words (van Tol et al., 2012) and** deficient activation of the ventrolateral **PFC (vlPFC)** and cuneus during retrieval of autobiographical memories (Foland-Ross et al., 2014). **However, it is not well elucidated whether negative cognitive**

**bias and aberrant neural processing of self-referent emotional words is a result of the depression, a scar effect, or a risk factor predisposing trait for depression.**

Studies of non-affected siblings and off-spring of patients with MDD, have reported negative bias in different aspects of emotional information processing. Specifically, these at-risk individuals display heightened attention to negative facial expressions (Joormann et al., 2007), greater attentional interference by sad distractor stimuli in an affective go/no-go task (Feder et al., 2011), attentional preference for depression-relevant adjectives (e.g. “helpless”, “demotivated”; Alloy et al., 1997) **and more negative interpretations of ambiguous words and stories** (Dearing and Gotlieb, 2009). **In contrast, a single study of encoding and recall of emotional adjectives in individuals at familial risk of depression found no negative bias at a behavioural level (Mannie et al, 2007).** Nevertheless, there is a lack of studies investigating *self-referent* information processing and recall of non-affected family members at **familial** risk. A few functional magnetic resonance imaging (fMRI) studies in non-affected first degree relatives have shown increased amygdala activation (Lisicka et al., 2013) and stronger negative functional coupling with fronto-parietal regions (Miskowiak et al., 2014) during emotional face processing. However, no fMRI study has yet assessed the neuronal underpinnings of memory for self-referent emotional words in non-affected first-degree relatives. **It is therefore unclear whether** aberrant neural response during encoding and retrieval of self-referent emotional information constitutes an illness endophenotype.

Neurocognitive and fMRI assessments of non-affected high-risk twins offer a strong methodology for an investigation of neurocognitive endophenotypes (Phillips, 2007). We have previously investigated neural and cognitive response to emotional faces in monozygotic (MZ) twins at high risk versus low risk for depression (Miskowiak et al., 2015). The present fMRI study aimed to investigate categorisation of and memory for emotional self-referent information in these healthy MZ and a group of healthy dizygotic (DZ) twins with or without a

co-twin history of depression (high-risk vs. low-risk groups) as part of a larger follow-up twin cohort study (Vinberg et al., 2013). We hypothesized that compared to low-risk twins, high-risk twins would display (i) lower accuracy in the categorization or and/or memory for positive vs. negative self-referent words and (ii) generally poorer memory (independent of valence), and that this would be accompanied by (iii) reduced hippocampal activation during recognition of positive relative to negative words and (iv) increased activity in a during retrieval of positive words or all emotional words (reflecting greater strain on the attention control network).

## Method and materials

### *Participant characteristics and recruitment*

After a baseline study conducted from May 2003 to September 2005 (Christensen et al., 2017), the participants were followed longitudinally at 6-month intervals for a period of seven years to obtain information on the potential onset of affective disorder. Of the 234 participants in the original baseline study, 218 participated in the follow-up assessment from 2010-2012 (Vinberg et al., 2013). The study was approved by the Danish Ministry of Health, Danish Scientific Ethics Committee and Data Protection Agency (see Vinberg et al., 2013).

In brief, all participants underwent a telephone interview at the 5-7 years follow-up. A SCAN interview was performed if, according to the telephone interview, the participants: 1) had any contact with a psychologist or psychiatrist, 2) had been on sickness leave because of psychological problems, 3) were prescribed any **psychoactive drugs**. **A diagnostic interview using the Schedule for Clinical Assessment in Neuropsychiatry (SCAN) (Wing et al., 1990)** was also conducted if 4) their answers to the questionnaires — the Beck Depression Inventory 21 Items (BDI-21; Beck et al., 1961) and the Mood Disorder Questionnaire (MDQ) — raised the suspicion of onset of psychiatric disorder or 5) they had been diagnosed with a first psychiatric

diagnosis in the Danish Psychiatric Central Research Register during follow-up (this was possible as the personal identification numbers of all participants were linked to the Danish Psychiatric Central Research Register). Exclusion criteria for this fMRI study were personal history of psychiatric or organic brain illness (for a complete description, see Vinberg et al., 2013).

### *Study design*

The research was conducted at the Danish Research Centre for Magnetic Resonance (DRCMR), Copenhagen University Hospital Hvidovre, Denmark, between 12:00 and 21:00 hours. The participants first completed the set of questionnaires (description below). After this, they were given an emotional categorization task modified from the Emotional Test Battery (ETB; P1Vital, Oxford) on a laptop computer followed by a free-recall test 20 minutes later. This was followed by an fMRI scan, during which participants performed an emotional word recognition task (details below), an emotional faces task as well as an affective picture task. Data from the latter two tasks are presented elsewhere. The experimenters were blind to participants' risk status throughout the data collection, management and analysis.

### *Emotional categorisation and memory tasks*

A list of 180 personality trait words was constructed from the Anderson's list of personality trait words (Anderson, 1968). The list consisted of an equal number of unambiguously positive and negative words (e.g. perceptive, talented, generous, selfish, hostile, and pompous) matched on length, written frequency (Francis and Kucera, 1982), and meaningfulness (Anderson, 1968). For the purpose of this task, the list was translated to Danish and back-translated with permission of the author of the task (CJH). In the categorisation task, 90 words (45 positive and 45 negative) were displayed one at a time for 500 ms. Participants were instructed to categorize

the words as likeable or unlikeable in self-referential terms as quickly and accurately as possible by pressing corresponding keys on the keyboard with their index fingers. Specifically, they were asked to imagine that they overheard someone talking about them and decide whether they would be happy or sad if these words were used about them. The total task time was six minutes. Participants' accuracy and response times to positive and negative words were recorded using Superlab software.

After 15 minutes participants performed a delayed free-recall task in which they were instructed to recall as many words as possible in any order from the categorisation task. The total number of correctly recalled positive and negative words and the number of memory intrusions (false recollections) were recorded.

In the emotional recognition task during fMRI, participants were shown the 90 previously presented positive and negative words ('old words') as well as 90 new matched positive and negative words ('new words') from the Anderson's list of personality trait words. Participants were instructed to indicate as quickly and accurately as possible whether or not they had seen the words or not by pressing corresponding keys on the response pad. Each trial consisted of a fixation cross shown for 500 ms immediately replaced by a personality trait word shown for 500 ms. Words were presented in random order, the inter-trial interval (ITI) varied between 4000 and 9000 ms, leading to a total task time of 12 minutes. Accuracy and response times for correctly recognised words and misclassifications were recorded with e-prime software version 1.2 (Psychology Software Tools Inc., USA).

### *fMRI data acquisition*

MRI data were collected at the DRCMR, with a 3 T Siemens Trio MR scanner using an eight-channel head array coil. Blood oxygen-level dependent (BOLD)-sensitive fMRI used a T2\*-weighted gradient echo spiral echo-planar imaging (EPI) sequence with an echo time (TE) of

30 ms and repetition time (TR) of 2.49 ms. In the MZ study, the flip angle was 20° (which minimizes physiological noise; Gonzalez-Castillo et al., 2011), while it was 90° in the parallel DZ study. The fMRI sequences included 279 brain volumes acquired in a single session, each consisting of 93 slices with a slice thickness of 3 mm and a field of view (FOV) of 192 by 192 mm using a 64 by 64 grid. Since in the present study we included data from the MZ and DZ twins from the two parallel studies, we included the zygosity as a factor in the statistical models which would also account for the different flip angles for the fMRI data across participants. High-resolution 3D structural T1-weighted spin echo images were obtained after the first session of BOLD fMRI (TI=800, TE=3.93, TR=1540 ms, flip angle 9°; 256 by 256 FOV; 192 slices). The present analyses focused on the effect of risk shared by MZ and DZ twins rather than the effect of zygosity, since the latter comparison would have been confounded by differences in fMRI data acquisition parameters between the MZ and DZ studies.

### *fMRI data analysis*

Functional MRI data processing was carried out with the FMRIB Software Library version (FSL version 6.00) ([www.fmrib.ox.ac.uk/fsl](http://www.fmrib.ox.ac.uk/fsl); Smith et al., 2004) to pre-process and analyse fMRI data. Pre-processing included image realignment, non-brain removal, motion correction, spatial normalization and spatial smoothing (Gaussian kernel, 5 mm full-width-half-maximum). The time series in each session were high pass-filtered (to a maximum of 0.008 Hz).

An event-related design was employed to explore the rapid BOLD responses triggered by recognition of personality trait words. Events of interest were defined according to whether or not the word was presented in the previous categorization task (old vs. new), whether the word was positive or negative, and whether the word was correctly identified as old (hit) or new (correct rejection). Four experimental conditions – ‘positive/hits’, positive/correct rejections,



negative/hits and negative/correct rejections – were defined as events occurring at the word presentations. We set up individual-level GLM models that included the four event regressors convolved with a canonical hemodynamic response function (Jezzard et al., 2001) and local autocorrelation corrections (Woolrich et al., 2001).

We investigated the hypothesis that familial risk of depression is associated with increased retrieval-specific activity for negative words and/or reduced retrieval-specific for positive words by contrasting neural activity to (i) positive/hits with positive/correct rejections and (ii) negative/hits with negative/correct rejections (i.e. risk group by task interactions). In addition, we contrasted (iii) negative/hits with positive/hits to investigate potential between-group differences in neural activity during valence-specific memory processes. At the group level, the data for the three contrasts were included in separate GLM models estimated with nonparametric permutation-based inference ( $n=5000$ ) using the ‘randomize’ algorithm in FSL. Bilateral hippocampi and the PFC were defined using maps included in the Harvard-Oxford Subcortical Structural Atlas and thresholded at 5%. The statistical inference at group level was restricted within the defined VOIs. In addition, whole-brain exploratory analysis was carried out to assess potential activity differences between high-risk and low-risk groups during retrieval of positive and negative words in other brain regions. Significant clusters were identified using the Threshold-Free Cluster Enhancement method at corrected  $p < 0.05$ . Foci of peak cluster activation were localized using a standard anatomical atlas (Talairach and Tournoux, 1988).

*Mood, subjective state, neuroticism and coping styles*

**Mood, neuroticism and coping styles were assessed in order to examine whether potential differences between high-risk and low-risk groups in self-referent encoding and memory occurred in the absence of group differences in these characteristics.** Mood and subjective

state were assessed with the BDI-21(Beck et al., 1961), the State and Trait Anxiety Inventory and visual analogue scales (VAS) of relevant **subjective** states (happiness, sadness, arousal, anxiety, dizziness, nausea). Neuroticism was assessed with the Eysenck Personality Questionnaire (EPQ; Beck et al., 1961; Eysenck and Eysenck, 1975), and coping styles measured with the Coping Inventory for Stressful Situation CISS (Endler and Parker, 1990).

### *Statistical analysis of behavioural and mood data*

Data from the emotional categorization task (accuracy and response times), free recall (total recall and memory intrusions) and recognition (accuracy, response bias and speed) were analysed with repeated measures analysis of covariance (ANOVA) with valence (positive, negative) as the within-participant factor and risk status (high-risk, low-risk) as between-participant factor. Signal detection theory was used for the recognition task to obtain a measure of memory accuracy corrected for the participants' response tendency according to the formula:  $d' = (\text{number of hits} + 0.5/\text{number of targets} + 1) - (\text{number of false alarms} + 0.5/\text{number of distractors} + 1)$ . Response bias – which reflects participants' tendency to categorize word as old rather than new in cases where of doubt – was calculated as:  $rb = (\text{number of false alarms} + 0.5/\text{number of distractors} + 1)/(1 - Pr)$ . Mood ratings, subjective state and coping styles were analysed with independent samples t-tests. All statistical analyses were performed with SPSS software v.23.0 (SPSS Inc., USA).

## **Results**

### *Participant flow, demographics and mood*

Of the 234 participants who took part in the follow-up part of the twin cohort study (2010-

2012), 82 participated in fMRI and neurocognitive assessments. Data was missing for 29 participants due to technical problems with the scans (high-risk: N=2, low-risk: N=3), poor quality of the structural scans (high-risk: N=5, low-risk: N=3), and lost behavioural data (high-risk: N=10, low-risk: N=6). Full data sets were thus available and analysed for 53 twins (high-risk: N=26, low-risk: N=27) (see Table 1). Groups were comparable for age, gender and education levels ( $p$ -values $>0.23$ ) and showed no differences in mood and subjective state ( $p$ -values $>0.47$ ), neuroticism ( $p=0.22$ ) or coping styles ( $p$ -values $>0.52$ ) (see Table 1). Further, there was no significant association between familial risk status and zygosity ( $\chi^2=0.02$ ,  $p=0.89$ ).

### *Behavioural results*

#### Emotional categorisation

All participants showed faster response times during the categorisation of negative than positive words ( $F(1,51)=22.35$ ,  $p < 0.001$ ) but this did not differ between high-risk and low-risk groups ( $p \geq 0.26$ ). Categorisation accuracy was generally high (93% on average) for both positive and negative words across the entire cohort. Between-group comparisons revealed lower accuracy (independent of valence) in the high-risk than low-risk group ( $F(1,51)=4.70$ ,  $p=0.04$ ; see Table 2).

#### Free recall

Participants generally recalled more positive than negative self-referent words ( $F(1,50)=9.12$ ,  $p=0.004$ ; paired-samples  $t$ -test for positive vs. negative words:  $t(51)=3.05$ ,  $p=0.004$ ). There was a trend towards poorer recall in high-risk vs. low-risk twins, independently of valence ( $F(1,50)=3.64$ ,  $p=0.06$ ; see table 2).

In general, participants had more positive than negative memory intrusions (i.e., falsely recalled

words) ( $F(1,50)=6.09$ ,  $p=0.017$ ; posthoc paired samples t-test for positive vs. negative words:  $t=2.31$ ,  $df=51$ ,  $p=0.03$ ). High-risk twins had significantly more negative than positive memory intrusions compared to the low risk group ( $F(1,50)=8.06$ ,  $p=0.007$ ; posthoc t-tests  $\geq 0.07$ ). This significant interaction prevailed after adjustment for zygosity ( $F(1,49)=8.05$ ,  $p=0.007$ ). In addition, this analysis revealed a trend towards an effect of zygosity on memory intrusions ( $F(1,49)=3.47$ ,  $p=0.059$ ). Posthoc analyses within the MZ and DZ cohorts separately revealed an interaction between familial risk and memory intrusions within the MZ (but not the DZ) cohort, reflecting more negative memory intrusions in the MZ high-risk than low-risk twins ( $F(1,19)=10.63$ ,  $p=0.004$ ; posthoc t-tests for negative intrusions:  $t=-2.90$ ,  $df=19$ ,  $p=0.009$ ).

### Emotional recognition inside the scanner

Positive words were recognised with greater accuracy and speed than negative words across the entire cohort (Accuracy:  $F(1,50)=4.51$ ,  $p=0.04$ ; t-test for positive vs. negative:  $t=5.92$ ,  $df=51$ ,  $p<0.001$ ;  $d'$ :  $F(1,50)=22.82$ ,  $p<0.001$ ; t-test for positive vs. negative:  $t=-4.81$ ,  $df=51$ ,  $p<0.001$ ; RTs:  $F(1,50)=34.77$ ,  $p<0.001$ ; t-test for positive vs. negative:  $t=-8.38$ ,  $df=51$ ,  $p<0.001$ ). There was also a greater response bias for positive than negative words across the entire cohort (i.e., a greater tendency to categorise positive words as 'old' in cases of doubt) ( $F(1,51)=74.10$ ,  $p<0.001$ ; t-test for positive vs. negative:  $t=8.64$ ,  $df=51$ ,  $p<0.001$ ) (see Table 2). However, there were no differences between high-risk and low-risk groups in recognition accuracy, speed or response bias during recognition of emotional self-referent words ( $ps \geq 0.14$ ).

### *Functional imaging*

#### Volume of interest analysis

The left mPFC was significantly activated during recognition of all words independent of

valence, and all participants displayed greater bilateral medial and inferior PFC activation during recognition of old negative vs. old positive words (see cluster maxima in table 3).

There were no differences between the high-risk and low-risk groups in retrieval-specific activity in the hippocampi or PFC for positive or negative words or in response to old positive vs. old negative words.

### Explorative whole-brain analysis

There were no brain regions showing differences between high-risk and low-risk groups in retrieval-specific activity for positive or negative words, respectively, or in neural response to old negative vs. old positive words.

Across all participants, retrieval of positive and negative self-referent words activated a network including prefrontal cortex and bilateral hippocampi (see cluster maxima in Table 3). Retrieval-specific activity for negative words (negative/hits vs. negative/correct rejections) occurred in the superior temporal gyrus and lentiform nucleus, while retrieval-specific activity for positive words was seen in the left posterior cingulate gyrus and inferior parietal cortex. In addition, participants generally showed greater activation in right mPFC, hippocampus, ACC and parietal regions during the recognition of negative vs. positive words (see cluster maxima in Table 3).

## **Discussion**

This is the first fMRI study to investigate whether negative bias in categorisation and memory for self-referent information constitutes a potential neurocognitive endophenotype for depression in healthy, never-depressed twins at high or low familial risk of depression. Consistent with our hypothesis, high-risk twins showed a negative bias in the retrieval of

emotional self-referent words as reflected by more negative memory intrusions. In addition, high-risk twins showed poorer performance during self-referent categorisation of emotional words and a trend towards poorer recall of these words independent of word valence. In contrast with our hypothesis, no differences were observed between groups in the neural activity within the hippocampus or across other brain regions during retrieval of positive or negative words. The present effects occurred in the absence of differences between groups in mood, subjective state, neuroticism, or coping strategies.

The demonstration of lower accuracy during self-referent categorization of emotional words and strong trend towards poorer recall in high-risk than low-risk twins is consistent with the previous demonstration of deficits in emotional categorisation and memory for self-referent words in patients' first-degree relatives (Mannie et al., 2007; Mannie et al., 2009). This poorer emotional cognitive function in high-risk twins is likely to confer enhanced risk of depression since lower cognitive reserve is associated with higher incidence of MDD (Downey et al., 2008) and we observed an association between general cognitive deficits and illness onset in our follow-up study of a large high-risk twin cohort (Vinberg, Miskowiak, and Kessing, 2013). Further, although mean age of participants was around 40 years of age, the high-risk twins still had increased risk of onset of affective disorder compared with low-risk twins (Vinberg et al., 2013).

The subtle display of negative bias in self-referent memory (as reflected by more false recollections of negative words) is consistent with the hypothesis that negative self-referent memory bias is a putative endophenotype. However, if we had corrected for multiple comparisons this difference between high-risk and low-risk twins would have rendered non-significant. The negative memory bias in high risk individuals should therefore be regarded as preliminary in nature. Indeed, negative memory bias seems to be a relatively robust state marker and - in situations of stress or negative affect – also a trait marker of depression but is

inconsistently observed in unaffected individuals at genetic risk (for review, see Miskowiak and Carvalho, 2014). Further, the absence of differential neural activity between high-risk and low-risk groups during the retrieval of positive and negative words contrasted with the hypothesised lower retrieval-related hippocampal and prefrontal activity for positive than negative words in high-risk twins. Indeed, the lack of any neural activity differences between groups in the exploratory whole-brain analyses was unexpected since fMRI BOLD activity may constitute a more sensitive measure of abnormal brain function than behavioural measures (e.g., Siegle et al., 2007). **The absence of differences in retrieval-related neural response between groups is likely to be a consequence of the only subtle display of negative memory bias in our high-risk twins, as indexed by more negative memory intrusions rather than greater retrieval negative vs. positive (target) words. Further, the inter-participant variability in the behavioural measures of self-referent memory was relatively large. This may be even higher for neuroimaging measurements, therefore requiring larger groups to capture significant differences.**

Strength of the research was the rigorous clinical evaluation during baseline and follow-up made by a psychiatrist (MV) using the SCAN research based interview. A further strength is our access to Danish registers which circumvented confounding factors such as acute or remitted depression in our sample, enabling inclusion of only healthy twins at high vs. low risk of depression. A limitation was the relatively high age of the included twins at the time of testing (mean $\pm$ SD: 40 $\pm$ 11 years). Specifically, this group of high-risk twins may also display compensatory mechanisms in addition to risk markers since they are still healthy despite their familial risk and agreed to participate in this time demanding research at both baseline and follow-up. Nevertheless, we have previously shown that there was significantly higher incidence of onset of psychiatric disorders in the high-risk vs. low-risk twins (21% vs. 6% respectively) during the seven years follow-up (n=234, mean age at baseline 39), It thus seemed

that the high-risk twins continued to be at high risk of developing depression during middle age and later life (Vinberg et al., 2013).

In conclusion, this fMRI study of unaffected twins at risk vs. low of depression suggests that valence-independent deficits and subtle negative bias in the processing for self-referent emotional information may constitute neurocognitive endophenotypes for depression.

### **Acknowledgements**

The study was funded by the Lundbeck Foundation. The funder had no role in the conception of the study, data collection, data analysis, interpretation or dissemination of the findings.

The Lundbeck Foundation and the Weimann Foundation are acknowledged for providing half of KWM's salary at the Psychiatric Center Copenhagen for her to do full-time research (Lundbeck Foundation 2012-2015; Weimann Foundation: 2015-2018). HRS was supported by a Grant of Excellence on the control of actions "ContAct" from the Lundbeck Foundation (R59 A5399) and the Novo Nordisk Foundation Interdisciplinary Synergy Program 2014 (NNF14OC0011413). The Simon Spies Foundation is acknowledged for donation of the Siemens Trio Scanner. The sponsors had no role in the design and conduct of the study; collection, management, analysis, and interpretation of the data; or preparation, review, and approval of the manuscript.

### **Declaration of Conflicting Interests**

KWM reports having received consultancy fees from Lundbeck and Allergan in the past three years. CJH has received consultancy fees from P1vital Ltd, Lundbeck, Servier and Eli-Lilly, and is a company director of Oxford Psychologists Ltd. CJH has also received grant income from GlaxoSmithKline, UCB Pharma, Janssen Inc, Lundbeck, Servier and Astra Zeneca. HRS has received honoraria as senior and reviewing editor for Neuroimage (Elsevier Publishers,



Amsterdam, Netherlands) and book editor from Springer Publishing, Stuttgart, Germany. He has received grant support from Niogen Idec, Denmark A/S. LVK reports having been a consultant for Lundbeck, AstraZeneca and Sunovion within the last 3 years. MV discloses consultancy fees from Lundbeck and Astra Zeneca within the last three years. JL and JM report no biomedical financial interests or potential conflicts of interest.

## References

- Alloy, L.B., Abramson, L.Y., Walshaw, P.D., and Neeren, A.M. (2006). Cognitive vulnerability to unipolar and bipolar mood disorders. *J. Soc. Clin. Psychol.* 25(7), 726-754.
- Alloy, L.B., Abramson, L.Y., Murray, LA, Whitehouse, WG, Hogan, ME (1997). Self-referent information-processing in individuals at high and low cognitive risk for depression. *Cogn. Emot.* 11, 539-568.
- Beevers, C.G., Carver, C.S. (2003). Attentional Bias and Mood Persistence as Prospective Predictors of Dysphoria. *Cognit Ther Res.* 27, 619–637.
- Bouhuys, A.L., Geerts, E., Gordijn, M.C. (1999) Depressed patients' perceptions of facial emotions in depressed and remitted states are associated with relapse: a longitudinal study. *J. Nerv. Ment. Dis.* 187, 595-602.
- Bradley, B., Mathews, A. (1983). Negative self-schemata in clinical depression. *Br. J. Clin. Psychol.* 22, 173-181.**
- Brittlebank, A.D, Scott, J, Williams, JM, Ferrier, IN (1993). Autobiographical memory in depression: state or trait marker? *Br. J. Psychiatry*, 162, 118-121.
- Browning, M., Holmes, E.A., Harmer, C.J. (2010). The modification of attentional bias to emotional information: A review of the techniques, mechanisms, and relevance to emotional disorders. *Cogn. Affect. Behav. Neurosci.* 10, 8-20.
- Cabeza, R., St Jacques, P. (2007). Functional neuroimaging of autobiographical memory. *Trends Cogn Sci.*, 11, 219-227.
- Dearing, K.F., Gotlib, I.H. (2009). Interpretation of ambiguous information in girls at risk for depression *J Abnorm Child Psychol.*, 37(1): 79-91.

Derry, P.A., Kuiper, N.A. (1981). Schematic processing and self-reference in clinical depression. *J Abnorm Psychol.* 90, 286-297.

Di Simplicio, M., Norbury, R., Harmer, C.J. (2012). Short-term antidepressant administration reduces negative self-referential processing in the medial prefrontal cortex in subjects at risk for depression. *Mol. Psychiatry.* 17, 503-510

Downey, L.A., Johnston, P.J., Hansen, K., Schembri, R., Stough, C., Tuckwell, V., Schweitzer, I. (2008). The relationship between emotional intelligence and depression in a clinical sample. *Eur. Psychiatry* 22, 93-98.

Endler, N.S., Parker, J.D.A. (1990). Multidimensional Assessment of Coping: A Critical Evaluation. *J. Pers. Soc. Psychol.* 58, 844-854.

Eysenck, H.J., Eysenck, S.B.G., 1975. *Manual of the Eysenck Personality Questionnaire.* Hodder and Stoughton, London.

Foland-Ross, L.C., Cooney, R.E., Joormann, J., Henry, M.L., Gotlib, I.H. (2014). Recalling happy memories in remitted depression: A neuroimaging investigation of the repair of sad mood. *Cogn. Affect. Behav. Neurosci.* 14, 818-826.

Fossati, P., Hevenor, S.J., Lepage, M., Graham, S.J., Grady, C., Keightley, M.L., Craik, F., Mayberg, H. (2004). Distributed self in episodic memory: neural correlates of successful retrieval of self-encoded positive and negative personality traits. *Neuroimage.* 22, 1596-1604.

Fritzsche, A., Dahme, B., Gotlib, I.H., Joormann, J., Magnussen, H., Watz, H., von Leupoldt A. (2010). Specificity of cognitive biases in patients with current depression and remitted depression and in patients with asthma. *Psychol. Med.* 40, 815-826.

**Gotlib, I.H., Krasnoperova, E., Yue, D.N., Joormann, J. (2004). Attentional biases for negative interpersonal stimuli in clinical depression. *J. Abnorm. Psychol.* 113, 121-135.**

Gotlib, I.H., Joormann, J. (2010). Cognition and depression: current status and future directions. *Annu. Rev. Clin. Psychol.* 6, 285-312.

Gottesman, I.I., Gould, T.D. (2003). The endophenotype concept in psychiatry: etymology and strategic intentions *Am. J. Psychiatry.* 160, 636-645.

Gotlib, I.H. and Cane, D.B. (1987). Construct accessibility and clinical depression: a longitudinal investigation. *J. Abnorm. Psychol.* 96(3): 199-204.

Hamlat, E.J., Connolly, S.L., Hamilton, J.L., Stange, J.P., Abramson, L.Y., Alloy, L.B. (2015). Rumination and overgeneral autobiographical memory in adolescents: An integration of cognitive vulnerabilities to depression. *J Youth Adolesc.* 44: 806-818.

**Harmer, C.J., Goodwin, G.M., Cowen, P.J. (2009). Why do antidepressants take so long to work? A cognitive neuropsychological model of antidepressant drug action. *Br J Psychiatry*, 195, 102-108.**

Hasler, G., Drevets, W.C., Manji, H.K., Charney, D.S.(2004). Discovering endophenotypes for major depression *Neuropsychopharmacology.* 29, 1765-1781.

Hasler, G., Northoff, G. (2011). Discovering imaging endophenotypes for major depression. *Mol. Psychiatry.* 16, 604-619.

Jezzard, P., Matthews, P.M., Smith, S.M. 2001. *Functional MRI an Introduction to Methods* Oxford Univ. Press, Oxford.

Johnson, S.L., Joormann, J., Gotlib, I.H. (2007). Does processing of emotional stimuli predict symptomatic improvement and diagnostic recovery from major depression? *Emotion.* 7, 201-206.

Joormann, J., Siemer, M., Gotlib, I.H. (2007). Mood regulation in depression: Differential effects of distraction and recall of happy memories on sad mood. *J. Abnorm. Psychol.* 116, 484-490.

Kuiper, N.A., MacDonald, M.R. (1982). Self and other perception in mild depressives. *Soc. Cogn.* 1, 223–239.

MacDonald, M.R., Kuiper, N.A. (1985). Efficiency and automaticity of self-schema processing in clinical depressives. *Mot. Emot.* 9, 171–184.

Mackinger, H.F., Pachinger, M.M., Leibetseder, M.M., Fartacek, R.R. (2000). Autobiographical memories in women remitted from major depression. *J. Abnorm. Psychol.* 109, 353–357.

Mannie, Z.N., Bristow, G.C., Harmer, C.J., Cowen, P.J. (2007). Impaired emotional categorisation in young people at increased familial risk of depression. *Neuropsychologia*. 45, 2975–2980.

Mannie, Z.N., Barnes, J., Bristow, G.C., Harmer, C.J., Cowen, P.J. (2009). Memory impairment in young women at increased risk of depression: influence of cortisol and 5-HTT genotype. *Psychol. Med.* 39, 757–762.

Miskowiak, K.W., Carvalho, A.F. (2014). 'Hot' cognition in major depressive disorder: a systematic review. *CNS Neurol. Disord. Drug. Targets*. 13, 1787–1803.

Miskowiak, K.W., Glerup, L., Vestbo, C., Harmer, C.J., Reinecke, A., Macoveanu, J., Vinberg, M. (2015). Different neural and cognitive response to emotional faces in healthy monozygotic twins at risk of depression. *Psychol. Med.* 45, 1447–1458.

Nolen-Hoeksema, S., Wisco, B.E., Lyubomirsky, S. (2008). Rethinking rumination. *Perspect. Psychol. Sci.* 3(5): 400–424.

Nolen-Hoeksema, S. (2000). The role of rumination in depressive disorders and mixed anxiety/depressive symptoms. *J Abnorm. Psychol.* 109, 504–511.

Nadel, L., Samsonovich, A., Ryan, L., Moscovitch, M. (2000). Multiple trace theory of human memory: computational, neuroimaging, and neuropsychological results. *Hippocampus*. 10, 352-368.

Oostrom, I., Franke, B., Vasquez, A.A., Rinck, M., Tendolkar, I., Verhagen, M., Janzing, J.G. (2013). Never-depressed females with a family history of depression demonstrate affective bias. *Psychiatry Res*. 205, 54-58.

Phillips, M.L. (2007). The emerging role of neuroimaging in psychiatry: characterizing treatment-relevant endophenotypes. *Am. J. Psychiatry*. 164, 697-699.

Phillips, M.L., Drevets, W.C., Rauch, S.L., Lane, R. (2003). Neurobiology of emotion perception I: The neural basis of normal emotion perception. *Biol. Psychiatry*. 54, 504-514.

Peterson, B.S., Wang, Z., Horga, G., Warner, V., Rutherford, B., Klahr, K.W., Hao, X. (2014). Discriminating risk and resilience endophenotypes from lifetime illness effects in familial major depressive disorder. *JAMA psychiatry*. 71, 136-148.

Price, J.L., Drevets, W.C. (2010). Neurocircuitry of mood disorders. *Neuropsychopharmacology*. 35, 192-216.

Romero, N., Sanchez, A., Vazquez, C. (2014). Memory biases in remitted depression: The role of negative cognitions at explicit and automatic processing levels. *J. Behav. Ther. Exp. Psychiatry*. 45, 128-135.

Siegle, G.J., Thompson, W., Carter, C.S., Steinhauer, S.R., Thase, M.E. (2007). Increased amygdala and decreased dorsolateral prefrontal BOLD responses in unipolar depression: related and independent features. *Biol. Psychiatry*. 61, 198-209.

Smith, K.A., Fairburn, C.G., Cowen, J. (1997). Relapse of depression after rapid depletion of tryptophan. *Lancet*, 349, 915-919.

Talairach, J., Tournoux, P. (1988). *Co-Planar Stereotaxic Atlas of the Human Brain*. Thieme

New York.

**van Tol, MJ, Demenescu, LR, van der Wee, NJ, KorteKaas, R, Marjan, M.A.N., Boer, J.A., Renken, R.J., van Buchem, M.A., Zitman, F.G., Aleman, A., Veltman, D.J. (2012). Functional magnetic resonance imaging correlates of emotional word encoding and recognition in depression and anxiety disorders. *Biol. Psychiatry* 71, 593-602.**

Vinberg, M., Miskowiak, K., Kessing, L.V. (2013). Risk markers for affective disorder, a seven-years follow up study of a twin cohort at low and high risk for affective disorder. *J. Psychiatr. Res.* 47, 565-571.

**Wing, J.K., Babor, T., Brugha, T., Burke, J., Cooper, J.E., Giel, R., Jablenski, A., Regier, D., Sartorius, N. (1990). SCAN. Schedules for Clinical Assessment in Neuropsychiatry. *Arch. Gen. Psychiatry* 47, 589–593.**

World Health Organization. Depression and other common mental disorders: Global health estimates. 2017. Available at: [http://www.who.int/mental\\_health/management/depression/prevalence\\_global\\_health\\_estimates/en/](http://www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/) (assessed 22 June 2017)

Woolrich, M.W., Behrens, T.E., Beckmann, C.F., Jenkinson, M., Smith, S.M. (2004). Multilevel linear modelling for fMRI group analysis using Bayesian inference. *NeuroImage*. 21, 1732–1747.

Yoshimura, S., Okamoto, Y., Onoda, K., Matsunaga, M., Ueda, K., Suzuki, SI. (2010). Rostral anterior cingulate cortex activity mediates the relationship between the depressive symptoms and the medial prefrontal cortex activity. *J. Affect. Disord.* 122, 76-85.

**Table 1.** Demographic information, mood and anxiety ratings, neuroticism and coping styles on the test day for high-risk and low-risk groups (n=53).

	Low-risk (n=27)	High-risk (n=26)	<i>p</i> value
Age, years, mean (SD)	38 (10)	42 (12)	0.23
Gender, no. Female (%)	17(63%)	14 (54%)	0.58
Neuroticism, mean (SD)	3.7 (3.2)	4.9 (4)	0.22
CISS, Coping style*			
Task-oriented, mean (SD)	33 (13)	31 (5)	0.47
Emotion-oriented, mean (SD)	40 (11)	42 (7)	0.63
Avoidance-oriented, mean (SD)	43 (6)	42 (6)	0.58
BDI-21, mean (SD)	2 (2)	2 (2)	0.86
STAI-state, mean (SD)**	29 (7)	28 (6)	0.80
STAI-trait, mean (SD)**	29 (8)	29 (7)	1.0
Vas of subjective state***			
Happiness, mean (SD)	31 (39)	30 (30)	0.84
Sadness, mean (SD)	6 (10)	5 (12)	0.78
Alertness, mean (SD)	31 (33)	30 (31)	0.94
Anxiety, mean (SD)	5 (12)	4 (12)	0.83
Dizziness, mean (SD)	3 (6)	4 (8)	0.52
Nausea, mean (SD)	3 (9)	2 (7)	0.73

SD, standard deviation; BDI-21, Beck Depression Inventory 21 items; CISS, Coping Inventory for Stressful Situations; STAI, State and Trait Anxiety Inventory; VAS, Visual Analogue Scale.

\*Including complete data for 48 participants (High-risk: N=24, low-risk: N=24)

\*\*Including complete data for 50 participants (High-risk: N=24, low-risk: N=26)

\*\*\*Including complete data for 51 participants (High-risk: N=25, low-risk: N=26)



**Table 2.** Emotional categorisation, free recall, recognition in high-risk and low-risk twins.

Task and performance measure	High-risk (N=26)	Low-risk (N=27)	Analysis	
			Main effect of Task [Task by Group]	
	Mean (SD)	Mean (SD)	<i>P</i>	<i>df</i>
<i>Emotional categorization</i>				
Accuracy (percent correct responses)				
Positive words, mean (SD)	92.1 (8.1)	96.1 (3.9)	0.08	1,51
Negative words, mean (SD)	90.0 (6.7)	93.3 (7.9)	<b>[0.04]<sup>a</sup></b>	
Response times (ms)				
Positive words, mean (SD)	891 (236)	832 (180)	<0.001	1,51
Negative words, mean (SD)	994 (236)	921 (248)	[0.74]	
<i>Emotional recall</i> (total number)				
Positive recalled, mean (SD)	3.3 (2.6)	4.2 (2.1)	0.004	1,50
Negative recalled, mean (SD)	2.3 (1.8)	3.3 (2.1)	<b>[0.06]<sup>a</sup></b>	
Positive invented, mean (SD)	2.2 (1.9)	3.2 (2.1)	0.02	1,50
Negative invented, mean (SD)	2.4 (1.9)	1.5 (1.4)	<b>[0.007]</b>	
<i>Emotional recognition</i>				
Accuracy (percent correct responses)				
Old positive words, mean (SD)	68.6 (12.8)	74.7 (15.9)	<0.001	1,51
Old negative words, mean (SD)	59.5 (15.1)	64.2 (15.5)	[0.74]	
New positive words, mean (SD)	50.6 (18.3)	51.8 (17.2)	<0.001	1,51
New negative words, mean (SD)	69.1 (19.1)	70.0 (21.0)	[0.96]	
Discrimination accuracy ( <i>d'</i> )				
Positive words, mean (SD)	0.19 (0.17)	0.26 (0.17)	<0.001	1,51
Negative words, mean (SD)	0.28 (0.19)	0.33 (0.21)	[0.67]	
Bias				
Positive words, mean (SD)	0.10 (0.16)	0.15 (0.15)	<0.001	1,51
Negative words, mean (SD)	-0.08 (0.18)	-0.06 (0.20)	[0.37]	
Response times (ms)				
Old positive words, mean (SD)	1136 (238)	1090 (214)	<0.001	1,51
Old negative words, mean (SD)	1246 (250)	1214 (264)	[0.61]	
New positive words, mean (SD)	1273 (260)	1251 (266)	0.04	1,51
New negative words, mean (SD)	1231 (247)	1218 (259)	[0.82]	

<sup>a</sup>Main effect of group. SD, standard deviation. \*Including complete data for 52 participants (High-risk: N=26, low-risk: N=26)

**Table 3.** Peak cluster activation in regions activated across all participants (n=53) during retrieval of positive and negative self-referent words in whole-brain analysis (main effects of task).

Regional response	Cluster size	Cluster	MNI Coordinates		
	(no. voxels)	P value	x	y	z
<b>All correct hits &gt; Baseline across the entire cohort</b>					
Left medial prefrontal gyrus (BA 10)*	3,128	<0.001	-36	46	-4
Right cerebellum	37,774	<0.001	28	-48	-40
Left precentral gyrus (BA 6)	269	0.031	-36	-10	66
<b>Positive/hits vs. positive/correct rejections across the entire cohort</b>					
Left Posterior Cingulate gyrus (BA 31)	439	0.009	-2	-64	20
Posterior Cingulate gyrus (BA 23)	77	0.035	0	-44	28
Left Inferior Parietal Gyrus (BA 40)	32	0.038	-42	-70	40
<b>Negative/hits vs. negative/correct rejections across the entire cohort</b>					
Left superior temporal gyrus (BA 22)	41,609	<0.001	-58	-60	16
Left lentiform nucleus	31	0.046	-12	4	0
<b>Old negative/hits vs. old positive/hits across the entire cohort</b>					
Left medial prefrontal gyrus (BA 6)*	355	0.015	-38	4	52
Left medial prefrontal gyrus (BA 10)*	121	0.027	-34	48	12
Right inferior frontal gyrus (BA 47)*	37	0.031	34	28	-8
Right middle frontal gyrus (BA 9)*	1,475	0.001	34	14	34
Left anterior cingulate gyrus (BA 24)	89	0.036	-4	30	28
Right anterior cingulate gyrus (BA 32)	24	0.043	14	32	18
Left parahippocampal gyrus (BA 30)	20	0.042	-12	-42	-8
Left inferior frontal cortex (BA 44)	1,132	0.002	-34	12	26
Right inferior parietal gyrus (BA 40)	21	0.045	62	-50	26
Left inferior parietal gyrus (BA 40)	1,842	0.007	-38	-36	26
Right precuneus (BA 7)	14,644	0.002	6	-72	48
BA: Brodmann area; MNI coordinates (x, y, z) refer to the point of peak activation within each cluster.					
* VOI analysis within the hippocampus and PFC					