Altering memories for emotional experiences

MARIJN KROES
The research in this thesis was carried out at the Donders Institute for Brain, Cognition and Behaviour of the Radboud University Nijmegen, the Netherlands, at the Wellcome Trust Centre for Neuroimaging of University College London, United Kingdom, and at the department of Clinical, Educational, and Health Psychology of the University College London, United Kingdom, and financially supported by research grants from the Netherlands Organisation for Scientific Research (NWO 918.66.613) the European Research Council (ERC-2010-AdG 268800) and DSM Nutritional Products Ltd to G. Fernandez, a Wellcome Trust Programme Grant to R. Dolan, and a Wellcome Trust Grant programme Grant to C. Brewin.

A digital version of this thesis and supplementary information (includes all tables) can be obtained by sending an email to m.c.w.kroes@gmail.com.

This work is licenced under CreativeCommons, it may be transmitted and reproduced for non-commercial purposes only, if no alterations are made and the author is attributed.
Altering memories for emotional experiences

Proefschrift

Ter verkrijging van de graad van doctor aan de Radboud Universiteit Nijmegen op gezag van de rector magnificus Prof. Mr. S.C.J.J. Kortmann, volgens besluit van het college van decanen in het openbaar te verdedigen op donderdag 5 september 2013 om 10:30 precies

Door

Marijn Christoffel Waander Kroes

Geboren op 5 maart 1983

Te Veenendaal
**Promotoren:**

Prof. Dr. G. Fernandez

Prof. Dr. B. A. Strange (Universidad Politecnica de Madrid)

**Copromotor:**

Dr. G. A. van Wingen

**Manuscript committee:**

Prof. Dr. R. Cools

Prof. Dr. H-C. Pape (Westfalische Wilhelms-Universitat)

Dr. S. Daselaar
# TABLE OF CONTENTS

1. Introduction 12

2. Structural brain abnormalities common to posttraumatic stress disorder and depression 22

3. Association between flashbacks and structural brain abnormalities in posttraumatic stress disorder 40

4. An fMRI investigation of posttraumatic flashbacks 56
   - Interlude 78

5. Food can lift mood by affecting mood-regulating neurocircuits 80
   - Interlude 98

6. Dynamic neural systems enable adaptive, flexible memories 100
   - Interlude 154

7. Emotion causes targeted forgetting of established memories 156

8. Evidence for reconsolidation of emotional episodic memory in humans 182

9. Adrenergic blockade during memory retrieval in humans evokes a sustained reduction of declarative emotional memory enhancement 198
10. Adrenergic blockade affects the neural network of extinction learning in humans, causes the loss of fear, and prevents the return of fear

11. The 5-HTT genetic polymorphism affects the neural network of fear learning

12. Discussion

- Dutch Summary
- References
- Donders Series
- Acknowledgements
- Publication List
- Curriculum Vitae
“The arduous events are outside time, either because the immediate past is as if disconnected from the future, or because the parts which form these events do not seem to be consecutive”

-Emma Zunz, from The Aleph- 

Jorge Luis Borges, 1949
A bright orange flash of light. KABOOM!!! ... And everything goes dark. I am blown out of my seat. Glass shattering, metal twisting and all I can hear is a deafening beep in my ears. Time slows down; the seconds seem to last forever, everything happens in an instant. I am dirty, covered in black stuff. I feel bruised; blood is starting to gush down from cuts in my face. My clothes are ripped and I can feel the sharp pain of splintered glass piercing my skin. An acidic smell starts to fill my nostrils. I can taste metallic smoke that smothers my breathing as it fills my lungs. The heat of a rolling fire burns the hair off my arms. Slowly sound returns, muffled but audible, I can hear people screaming. They’re screaming, their screams cut through the smoke and I realize they’re burning, they’re burning alive! The heat is coming closer, rolls over me, and envelops me. I realize the fire is closing in on me. This is the end of me. It can’t be long now. Dear Anna, I love you. I hope you know that. I am going to die, but I hope you know that. The flames allow me to see through the thick smoke. Next to me a heavily bleeding man crawls over the floor. A bit further, in the flicker of the flames, I see a woman’s leg, dismembered from her body. Why did we have that argument? It’s so stupid. No time to think about that now. She knows I love her. Too bad we won’t see each other again though. It’s the end for me. I look up; metal dark and twisted with jagged edges pointing up at the ceiling reveals a ten feet hole. A man climbs out through the hole, looks back, shouts at me: "Come on man, get up! We need to get out of here! I try to lift myself of the floor. I get up. An excruciating pain shoots down my leg. My vision turns blurry as white stars flash in front of my eyes. "Hurry up man! This way! We can get out this way!" The pain in my leg is unbearable, but I can do this. I have to do this. Slowly I stumble forwards towards the hole. The man grabs my hand and helps me through the hole. We get out. A light shimmers in the distance of the tunnel. It allows me to see through the smoke. The tracks are covered with bodies.
They don’t move, I know they’re dead. We walk towards the light in the distance. We see more people walking in front of us. "Stay of the middle of the track!" says a man in a fluorescent vest who has suddenly appeared. The man points: "The platform is that way, only a little further." I stumble past him as I feel the energy draining from my body. "Are you alright?" I hear. And everything fades.

Slowly I return. "Are you alright Ben?" The armchair I am sitting in is comfortable. I can feel my heartbeat slowing down, while I look around the office. This is the third time this has happened to me today. "Are you alright Ben?" Anna looks at me and I say yes, and highlight the sentence on the piece of paper the researcher across the table has put in front of me.
1 GENERAL INTRODUCTION.
1.1 Introduction

Ben’s story is fiction, but inspired by the traumatic events the patients suffering from post-traumatic stress disorder (PTSD), that took part in studies that are described in the first three chapters of this thesis, wrote down. The event described in the story reflects a flashback, moments in which PTSD patient relive the experience of their trauma as if it happens to them in real-time. This dissociative symptom is specific to PTSD, highly stressful and disturbing. Imagine what it must be like to experience multiple times per day as if the London underground train you are in gets blown up in a terrorist attack.

Emotional experiences are remembered well (1, 2), which is adaptive to survival but emotional memory enhancement can also be maladaptive and have pathological consequences (3), such as in the case of Ben. Affective disorders, including anxiety disorders such as PTSD, rank among the leading causes of disability worldwide, significantly contributing to the global disease burden (4). Therefore, optimizing understanding and treatment of affective disorders is of great importance.

Over the last decades, research on patients and studies on fear learning and memory in healthy humans and non-human animals has provided a wealth of insight into the neural mechanisms that support enhanced memory for emotional experiences and the dysfunctions that occur in this network in pathological states. A central role in the acquisition and expression of emotions is envisioned for the amygdala (5-7), a collection of nuclei deep in the brain (Figure 1.1 red). The amygdala is considered to be critical for the formation of associations between stimuli and emotional outcomes, and projects to brain regions such as the brainstem (Figure 1.1 light green) and hypothalamus that initiate the defensive behaviours (e.g. freezing, fleeing, fighting), autonomic changes (e.g. changes in heart rate and blood pressure), and endocrine changes (e.g. release of neuromodulators and stress hormones) that are part of an emotional response (5, 8). Additionally, emotions are accompanied by changes in action readiness coordinated by the striatum (Figure 1.1 blue).
Furthermore, emotional responses are often context-dependent and contextual information is thought to be provided by the hippocampus ((9), Figure 1.1 green). Control over fear and regulation of the amygdala is considered to be mediated by the medial prefrontal cortex (Figure 1.1 pink), and specifically the ventral (lower) part of this region (10-12). The cingulate cortex (Figure 1.1 purple) appears to be important for the integration of emotional responses with higher cognitive aspects of emotions, and regulation of emotions (13, 14), and the lateral prefrontal cortex (Figure 1.1 yellow) with attention regulation towards emotional stimuli (15). Finally, the insula (Figure 1.1 orange) has been implicated in pain perception (16), the anticipation of aversive stimuli (17), with the introceptive awareness of visceral responses (18, 19), and subjective feeling states (20). Thus, a set of brain regions interacts as a network supporting learning, expression, and regulation of emotions. Interestingly, it is this network that shows evidence of dysfunction in pathological states like PTSD (21-24). Furthermore, effective treatment by cognitive behavioural therapy and psycho-pharmaceutical treatment appears to restore functioning of this network (25, 26).

The effectiveness of pharmacological treatment in affective disorders provided the first indications for a critical role of neuromodulators such as serotonin, dopamine and noradrenaline in emotional responses and the enhancement for emotional experiences (27). As mentioned, emotions are accompanied by endocrine responses. Within the brain, the neuromodulator serotonin is released from the raphe nuclei, dopamine from the ventral tegmental area, and noradrenaline from the
locus coeruleus (12). Additionally, through the so-called hypothalamus-pituitary-adrenal pathway stress hormones such as cortisol are released peripherally and feedback on the brain (28). These neuromodulators and stress hormones play a critical role in the expression of emotions and modulate memory resulting in enhanced memory for emotional experiences (29-31). More specifically, research has shown that memories are initially labile and sensitive to disruption but stabilize over time, a process known as consolidation (30). Neuromodulators such as
noradrenaline are thought to affect both encoding and consolidation processes leading to enhanced memory for emotional experiences (30, 32, 33). The classical view of memory formation (Figure 1.2A) dictates that incoming information is first encoded, stabilizes over time during consolidation, and is subsequently retrieved (30). Further, the most prominent model of memory enhancement by emotions (Figure 1.2 B) considers that emotional or salient stimuli activate the amygdala and noradrenergic system that contribute to the enhanced processing of the sensory features of a stimulus at the time of encoding (34-36). Subsequently, the amygdala is thought to upregulate neural processing in the hippocampus (37, 38) through a noradrenergic mechanism (29, 33, 39, 40).

For a long time a central dogma in neuroscience was that once memories were consolidated they were relative insensitive to further change or disturbances (30). This was considered to contribute to the difficulty in treating affective disorders and the high rates of relapse (41). For example, during cognitive behavioural therapy a patient may learn to control his fears, but this is thought to be the result of the formation of a new ‘safety’ memory that competes with the old fear memory for behavioural expression (42, 43). However, the risk of the return of fear is probable, because the original fear memory is not altered and still exists. Recent research on reconsolidation questions the stability of memory following consolidation (44). Reconsolidation refers to the process whereby memories become instable upon reactivation and require restabilization to be maintained (45-50). Reconsolidation has attracted much interest as it principally allows the alteration or deletion of memory for emotional experiences opening up new avenues for the treatment of affective disorders.

Beyond the classical view of memory formation (Figure 1.2 A) and the dominant model for memory enhancement by emotions (Figure 1.2 B), little is known about the neural mechanisms that support memory retrieval, and more specifically about the neuromodulatory systems that may enhance retrieval of emotional memories. Further, the neural
mechanisms that determine the fate of memories following retrieval are highly underexplored. Following retrieval new memory traces may be formed that compete with old memory traces (e.g. extinction), the retrieved memory may be encoded again in combination with new information (secondary encoding), or may destabilize and require restabilization (reconsolidation). The neural mechanisms that support these memory functions and the effects of neuromodulators on these mechanisms are topic of investigation in this thesis. By extending the classical view of memory formation and the dominant model that support memory enhancement by emotions and increasing fundamental understanding of the neural mechanisms that support memory for emotional experiences we hope to gain novel insight into the possibilities to alter memory for emotional experiences. In addition, by studying the overlap between these neural systems and the brain abnormalities observed in patients with anxiety disorders insight may be obtained into the development of novel treatment methods.

In sum, this thesis addresses the question what happens in the brains of patients with anxiety disorders, specifically related to maladaptive memories for emotional experiences. Further, the possibilities to alter memory for emotional experiences are explored. A prominent focus here is on the possibility to target specific memory processes in combination with pharmacological medication. Combined, the studies in this thesis are aimed at increasing fundamental understanding of the neural mechanisms that support memory for emotional experiences by extending the classical view of memory formation and the dominant model for memory enhancement by emotions. Together this thesis aims to provide further understanding and open up novel avenues for optimized treatment for patients like Ben.

1.2 Thesis Outline

The first part of this dissertation describes a set of studies investigating structural and functional brain abnormalities in patients with PTSD or depression (Chapter 2). Both patient groups are found to exhibit changes
in brain structure in regions implicated in memory and emotional regulation. Further, a symptom specific to PTSD is the re-experiencing of traumatic events from memory during so-called flashback episodes. The prevalence and occurrence of flashbacks were found to be associated with structural (Chapter 3) and functional (Chapter 4) brain abnormalities in regions implicated in memory and emotion regulation. Having identified brain regions displaying abnormalities in affective disorders, the question rises whether it is possible to alter functioning in these regions as it relates to emotional processes.

Generally, affective disorders, including PTSD and depression, are treated with antidepressant medication and their actions have been implicated to affect the monoamine systems of serotonin and noradrenaline. In the second part of the thesis we describe a study showing that food optimized to increase the uptake of the serotonin precursor tryptophan into the brain, lifts mood by affecting mood-regulating neurocircuits via a serotonergic mechanism (Chapter 5). Thus, by targeting the serotonergic system it is possible to affect emotion related processes in brain regions that show abnormal responses in affective disorders, yet an open question remains whether these neural systems could be targeted to achieve permanent alteration of memory for emotional experiences.

Although antidepressant medication and psychotherapy can reduce symptom severity in affective disorders, both treatments leave the patients susceptible to relapse. The ability to reduce or erase memory for a traumatic experience is of considerable interest, as this may prevent relapse. This is especially true in cases where the aetiology of disease onset can be attributed to a specific emotional experience. Although the idea that well consolidated memories are stable and insensitive to changes has been a central dogma in neuroscience for decades, recent neuroscientific studies indicate that memories remain flexible even after consolidation. In the third section of this thesis the dynamic neural systems that enable adaptive and flexible memories are discussed (Chapter 6). The dynamic neural systems that enable adaptive and flexible memories principally open up novel avenues to permanently
alter memory for emotional experiences, yet an outstanding question is whether this can be achieved in practice.

In the final section of this thesis studies aimed at permanently altering or erasing memories for emotional experiences are presented. The first study provides evidence that episodic memory can be disrupted upon reactivation in humans, that this process is time-dependent and persists for at least six days, in line with a reconsolidation hypothesis (Chapter 7). The second study shows that reconsolidation of episodic memory for emotional experiences can be disrupted by electroconvulsive shock in humans, and thus provides causal evidence that interference with brain function after reactivation can erase memories by altering a reconsolidation process (Chapter 8). Erasing memories may prevent relapse, yet erasing or attenuating only the emotional aspect of a memory trace with treatment would be more optimal. The noradrenergic system has long been implicated in the enhancement of memory for emotional experiences, and more recently also in the reconsolidation of memories for emotional events. In a third study we show that administration of the beta-adrenergic receptor antagonist propranolol at memory retrieval abolishes an episodic memory enhancement for emotional items, and critically, that this effect persists in the absence of propranolol 24h later (Chapter 9). Although memories are susceptible to alteration upon reactivation, possibly through reconsolidation mechanisms, most cognitive behavioural therapies employed in treating affective disorder incorporate methods where patients are repeatedly required to retrieve memories of emotional experiences and actively learn to reduce their fear responses. These treatments rely on extinction learning mechanisms. Alternative hypotheses predicted that noradrenaline blockade during extinction learning would either result in a loss of fear or in the persistence of fear. Testing these hypotheses in a fourth study, we show that beta-adrenergic blockade abolishes differentially conditioned responses during extinction learning and subsequently prevents the return of fear 24 hours later. These effects are attributable to specific changes in the neural network involved in extinction learning (Chapter 10). This finding provides support to an
optimized treatment of affective disorders by specific combinations of psychotherapy and psychopharmacology. Optimization of treatments also requires understanding of vulnerability factors such as genetic variations on brain functioning as this holds the promise to tailor psychiatric treatment to individual needs. In Chapter 11, we show that the serotonin transported-linked polymorphic region (5-HTTLPR) that has been associated with anxiety and depressive disorders may increase vulnerability by determining functioning of the neural network involved in fear and safety learning. Together Chapters 7-11 provide evidence for the occurrence of reconsolidation in humans, reveal a role for noradrenaline in the retrieval of memory for emotional experiences and a deterministic effect of noradrenaline on the faith of such memories following retrieval, and influences of noradrenaline and serotonin on the neural networks that support extinction. These studies highlight the possibility to alter memory for emotional experiences.

This thesis will conclude with a discussion (Chapter 12) of the described studies and present and extension of the classical view of memory formation and an innovative neural model of memory enhancement by emotions via neuromodulatory systems as revealed by the work in this thesis, which significantly advances the field of cognitive neuroscience of memory and emotions. In addition, the overlap between the neural systems that support the alteration of memory for emotional experiences and the brain abnormalities observed in patients with affective disorders will be highlighted. Furthermore, the limitations of the presented studies will be discussed together with outstanding questions for future research.

Together this thesis provides insight into the neural mechanisms that underlie emotional memory and contribute to the aetiology and persistence of affective disorders, presents evidence for the ability to alter memory for emotional experiences, and hopes to contribute to the development of novel strategies to optimize treatment of affective disorders.
2 STRUCTURAL BRAIN ABNORMALITIES COMMON TO POSTTRAUMATIC STRESS DISORDER AND DEPRESSION.

2.1 Published as:


2.2 Abstract

Posttraumatic stress disorder (PTSD) and major depression (MDD) are associated with reductions in brain volume in markedly similar areas. To date no volumetric studies have directly contrasted these conditions. We investigated which grey matter reductions would be uniquely associated with each disorder. We also investigated more subtle independent effects: specifically, correlations between brain volume and self-report measures of psychopathology. We obtained structural magnetic resonance imaging scans from participants with PTSD (N=24), MDD (N=29) and healthy controls (N=29) exposed to trauma. Participants completed standardized self-report measures of anxiety and depression. We used voxel-based morphometry to identify associated volumetric changes. The clinical groups had regions of markedly smaller volume, particularly in prefrontal areas, but did not differ from each other. Greater self-reported anxiety was inversely related to volume in several areas, particularly the inferior temporal cortex, among patients with PTSD, but was associated with some volume increases in patients with MDD. Greater self-reported depression showed similar effects, being inversely related to brain volume in patients with PTSD but positively related to volume in the cuneus and precuneus of those with major depression. As such, we identified commonalities in areas of brain volume in patients with PTSD and those with MDD, suggesting that existing findings concerning volume reductions may not be specific to PTSD but rather related to features of the disorder that are shared with other conditions, such as MDD. More subtle differences between patients with PTSD and those with MDD were represented by distinct structural correlates of self-reported anxiety and depression.
2.3 Introduction

Structural brain abnormalities in individuals with posttraumatic stress disorder (PTSD) and major depression show a marked overlap. Thus, meta-analyses of structural MRI studies on individuals with both PTSD and unipolar depression have consistently identified reductions in hippocampal size relative to healthy controls (51-53). This overlap is hard to interpret since PTSD and depression have a number of symptoms in common and are frequently comorbid (54). For example, in one study reduced hippocampal size was only found in depressed individuals if they also reported early childhood trauma, but two thirds of this group had comorbid PTSD (55). For these reasons it is not known at present to what extent structural changes found in PTSD are accounted for by comorbid depression. The present study used voxel-based morphometry (VBM) to test directly for the first time to what extent there are commonalities in brain volume reductions and whether there is evidence for structural changes that are unique to one or other disorder.

In addition to reduced hippocampal size, PTSD has been associated with reductions in dorsal anterior cingulate cortex (56-59), and insular cortex (56, 60, 61). Differences in rostral ACC have also been reported (60, 62), although one study suggested apparent differences in the latter were more a reflection of shape than volume (61). In addition to a smaller ACC, a meta-analysis (51) found evidence for a reduction in the left amygdala of adult PTSD patients, and in the frontal/prefrontal cortex (but not the hippocampus) of paediatric PTSD samples. As such it has been suggested that PTSD is associated with abnormalities in multiple frontal-limbic structures. Studies of unipolar and psychotic depression have similarly found reduced volume in several regions of ACC (25, 63), and more specifically in subgenual ACC (64-69). Other areas of reduced grey matter concentration or reduced volume include the amygdala (67, 69-71), orbitofrontal cortex (63, 72, 73), thalamus (73, 74), striatum (74, 75), and several additional prefrontal areas (25).
Understanding these similar sets of results in both diagnoses is made difficult by the fact that there is typically a high level of comorbidity between PTSD and other anxiety disorders, and more particularly with depression. Any attempt to identify ‘pure’ cases of PTSD would therefore result in a small, highly unrepresentative sample (76). Another complication is that because of a substantial degree of overlap in the diagnostic criteria many of these comorbidities may be more apparent than real. Attempts to control for comorbidity in consequence risk removing effects that are actually those of interest. Our strategy, based on our previous work (77), was therefore to compare PTSD patients regardless of comorbid depression with a similar group who met criteria for unipolar depression in the absence of PTSD. Our intention was to have two groups typical of those seen in clinical practice, matched as closely as possible on all relevant variables, e.g. severity of depression, use of antidepressants, and early trauma, and differing only in the presence or absence of PTSD. A third, psychiatrically healthy, group provided a control for the effects of trauma exposure.

We hypothesized decreased grey matter volume in fronto-limbic structures and specifically the hippocampus and ACC in both clinical groups compared to the trauma control group. Previous research has also suggested the existence of more subtle structural effects relevant to psychopathology, independent of overall differences in volume, such as a correlation between self-reported anxiety and hippocampal volumes in depressed patients and healthy controls (78). Both self-reported anxiety and depression are likely to be elevated, relative to controls, in samples with PTSD and major depressive disorder. We therefore sought to investigate whether any group differences in brain structure were accounted for by current levels of anxiety and depression and whether, independent of overall difference in volume, correlations between brain structure and self-reported anxiety and depression differed in the depressed and PTSD samples. In order to improve the sensitivity of VBM we used the Dartel algorithm, an improved registration method applied to one of the pre-processing steps, which has recently been developed within SPM. A study with acutely depressed patients confirmed that VBM-
DARTEL was more sensitive than standard VBM in detecting hippocampal abnormalities (79).

2.4 Methods

2.4.1 Participants

T1-weighted magnetic resonance images of 82 subjects were collected as part of two functional magnetic resonance imaging studies. Participants were right-handed, under age 50, were not currently abusing substances, and had no history of head injury, neurological disorders, or other major medical conditions. The PTSD group consisted of 24 individuals (9 men) with an average age of 35.9 years (range 25-47 years). All met diagnostic criteria for current PTSD when assessed with the Structured Clinical Interview for DSM-IV (80). They had experienced a range of traumas, including exposure to traumatic death (n=11), robbery/assault (n=5), accidents (n=1), abuse in childhood (n=6), and other (n=1). Ten were currently taking antidepressant medication. The depressed group consisted of 29 individuals (8 men) with an average age of 33.4 years (range 24 to 47 years). All met diagnostic criteria for current major depressive disorder but not for PTSD when assessed with the SCID. They reported exposure to traumatic death (n=4), robbery/assault (n=5), and abuse in childhood (n=3). Twenty-two were currently taking antidepressant medication. The trauma control group consisted of 29 individuals (13 men) with an average age of 32.45 years (range 21-49 years). None met diagnostic criteria for any psychiatric disorder when assessed with the SCID and none were currently taking any psychotropic medication. They reported exposure to traumatic death (n=13), robbery/assault (n=7), accidents (n=7), and other (n=2). All participants gave written informed consent, and all procedures were approved by the National Hospital for Neurology and Neurosurgery & Institute of Neurology Joint Research Ethics Committee.
2.5 Measures

The Beck Anxiety Inventory (81) is a widely used self-report measure of the severity of physiological and cognitive components of anxiety. It contains 21 items that are scored on a 4-point scale (possible range 0-63). Participants were instructed to rate their mood over the past week.

The Beck Depression Inventory 2 (82) is a widely-used 21-item self-report measure of depression severity. It contains 21 items that are scored on a 4-point scale (possible range 0-63). Participants were instructed to rate their mood over the past week.

The Posttraumatic Diagnostic Scale (83) is a widely-used 17-item measure for use after a specific traumatic event. It assesses the frequency with which each symptom of PTSD has been experienced over the past month. Items are scored on a 4-point scale (possible range 0-51).

2.6 MRI acquisition

T1-weighted images were acquired on a Siemens 1.5 T Sonata whole body scanner (Siemens Medical Systems, Erlangen, Germany), using a whole-body coil for RF transmission and a head coil for signal reception. Whole-brain structural scans were acquired using a Modified Driven Equilibrium Fourier Transform (MDEFT) sequence (84) with optimized parameters as described in the literature (85). For each volunteer, 176 sagittal partitions were acquired with an image matrix of 256x224 (Read x Phase). Two-fold over-sampling was performed in read direction (head/foot direction) to prevent aliasing. The isotropic spatial resolution was 1 mm. Relevant imaging parameters were TR/TE/TI = 12.24 ms/3.56 ms/530 ms, BW = 106 Hz/Px, $\Phi = 23^\circ$. The total duration was 12 min. Spin tagging in the neck was performed to avoid flow artefacts in the vicinity of blood vessels. The flip angle of the tagging pulse was chosen to be $110^\circ$ to account for B1 losses in the neck. Special RF excitation pulses were used to compensate for B1 inhomogeneities of the transmit coil in superior/inferior (86) and anterior/posterior (87) directions. Images were reconstructed by
performing a standard 3D Fourier Transform, followed by modulus calculation. No data filtering was applied either in k-space or in the image domain.

2.7 MRI Processing

T1-weighted image registration was achieved using the diffeomorphic registration algorithm implemented in the DARTEL toolbox (88) for SPM5 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). Dartel has been found to optimize the sensitivity of such analyses (79, 89). First, the anatomical images were manually reoriented so that the mm coordinate of the AC matched the origin [0,0,0], and the orientation approximated MNI space. Next T1-weighted images were classified into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) using the segmentation routine implemented in SPM5 applying standard settings. The resulting parameter files were imported into the DARTEL procedure to produce rigidly aligned grey and white matter tissue classes resliced to 1.5x1.5x1.5 mm voxel size. After the affine transformation the rigidly aligned tissue class images were used to estimate the nonlinear deformations to best align all images. During this estimation stage, DARTEL iterates between building a template and registering tissue class images using the template. The resulting flow-fields, which specify the parameterized deformations, were used to warp white and grey matter images for each subject. The spatially normalized images were rescaled by the Jacobian determinants of the deformations, using 64 time points for solving the partial differential equations. 7-th degree B-spline interpolation was used when writing the normalized images. In order to obtain meaningful coordinates of volume alterations, the final DARTEL template was normalized to MNI space and the resulting deformations applied to the grey matter images of each subject using a Matlab script downloaded from the SPM email list (https://www.jiscmail.ac.uk/cgi-bin/wa.exe?A2=ind0807&L=SPM&P=R34150%20032749). Finally the images were smoothed using an 8x8x8 mm Gaussian smoothing kernel. The input features for the subsequent analysis were the smoothed, modulated and normalized GM images.
2.8 Statistical analyses

Grey matter group differences were modelled using a factorial design with the factor Group having 3 levels (PTSD, depressed, trauma control). We followed up a significant main effect of Group was followed up with comparisons between individual groups, inclusively masked with the results of the original Group effect (at uncorrected mask $p = 0.05$) to avoid type-I errors. Measurements were assumed to be independent between levels, but variance was assumed to be unequal. Error covariance components were estimated by restricted maximum likelihood (REML) by SPM5, and used for inference and to adjust the degrees of freedom. We used absolute threshold masking at 0.05 to specify the voxels within the image volume that were to be assessed. As non-stationarity is problematic in cluster-level inferences in VBM data (90, 91) and requires adjustment of cluster sizes according to local smoothness of data (90, 92), we adjusted our cluster-level results according to this method as implemented in the VBM5 toolbox developed by C. Gaser (http://dbm.neuro.uni-jena.de/vbm/download/). Additionally, we created hippocampus and ACC region-of-interest (ROI) masks using the PickAtlas toolbox (93) and applied as small volume correction. For each region two masks were created, one based on the WFU atlas and one on the AAL atlas (94).

To examine the explanatory power of the continuous measures of anxiety and depression, the above analyses were repeated including the BAI and BDI as covariates, centred to the overall mean. Finally, to investigate differential effects of BAI and BDI on PTSD and depression, we converted the BAI and BDI scores to z-scores for each group separately. We excluded the control group from this analysis as there were floor effects on the BAI and BDI and hence little explanatory variance for either measure. The z-scores (with scores for one group reversed, i.e. times -1) were then used to specify regressors testing the interaction of Group (PTSD, depressed) with anxiety and depression respectively, with effects thresholded at $p<0.001$. The effects of group differences in antidepressant use were additionally controlled for in the analyses. In order to restrict type-I
errors, significant interactions were saved as threshold images and used to inclusively mask explanatory analyses of the effects of the BAI and BDI within each group separately.

2.9 Results

2.9.1 Participants

PTSD group consisted of 24 individuals (9 men) with a mean age of 35.9 (range 25-47) years. They had experienced a range of traumas, including exposure to traumatic death (n = 11), robbery/assault (n = 5), accidents (n = 1), abuse in childhood (n = 6) and other (n = 1). Ten were currently taking antidepressant medication. The major depression group consisted of 29 individuals (8 men) with a mean age of 33.4 (range 24-47) years. They reported exposure to traumatic death (n = 4), robbery/assault (n = 5) and abuse in childhood (n = 3). Twenty-two were currently taking antidepressant medication. The control group consisted of 29 individuals (13 men) with a mean age of 32.45 (range 21-49) years. They reported exposure to traumatic death (n = 13), robbery/assault (n = 7), accidents (n = 7) and other (n = 2).

2.9.2 Demographic and symptom measures

The groups did not differ in sex composition, $X^2_{2} = 1.87$ (p = 0.39), or age, $F_{2,79} = 1.77$, (p = 0.18). There were no group differences in estimated lifetime units of alcohol consumed, $F_{2,77} = 2.92$ (p = 0.06), but the depressed group was significantly more likely than the PTSD group to be taking antidepressant medication, $X^2_{1} = 6.40$, P = 0.011. We conducted an initial group (PTSD, Depression) x medication (on medication, off medication) analysis test for effects of antidepressant treatment on grey matter volume. No cluster-level significant main effect or interaction involving medication was found. In comparisons of grey matter volume between the PTSD and major depression groups, use of medication was modelled as a covariate of no interest owing to the relative small groups (medication-free depression patients , n = 7). The PTSD group scored
2. Structural brain abnormalities common to posttraumatic stress disorder and depression

Scores in the PTSD group were slightly in excess of those expected according to the authors of the PDS (83). There were group differences on the BAI ($F_{2, 78} = 36.79, P < 0.001$, PTSD mean: 30.39, SD: 14.44; depressed mean: 18.76, SD: 9.45; trauma control mean: 6.14, SD: 5.91). Post-hoc least significant difference tests indicated that all three groups differed significantly from each other. There were also group differences on the BDI ($F_{2, 79} = 76.76, P < 0.001$, PTSD mean: 31.12, SD: 11.49; depressed mean: 30.00, SD: 8.34; trauma control mean: 5.21, SD: 6.55). Post-hoc least significant difference tests indicated that the controls scored significantly lower than the two clinical groups, who did not differ from each other. The patients with PTSD were at least mildly depressed, scoring in the range 14-61 on the BDI. Thus the two clinical groups were matched on relevant variables including severity of depression and differed primarily on the variable of interest, PTSD symptoms.

2.9.3 VBM group differences: Combined clinical groups show reduced volume compared to the control group

The one-way ANOVA showed a significant overall effect of Group. Tests did not indicate any significant differences in brain volume between the depressed and PTSD groups, and subsequent analysis therefore focused
on the contrast between the combined clinical groups and the trauma control group. As shown in Table S1.1 and Figure 2.1, the clinical groups show clusters of decreased brain volume compared with controls in the middle cingulate cortex; in the medial prefrontal gyrus extending from the medial portions of the superior frontal gyrus at the precentral gyrus to the orbitofrontal cortex and anterior cingulate cortex; and in an area within the right dorsolateral prefrontal cortex spanning the middle frontal gyrus and superior frontal gyrus. Several other areas are significant at cluster level but fail to survive FWE correction. Applying small volume corrections (SVC) reveals an additional cluster of volumetric reduction in the ACC.
2.9.4 BAI and BDI scores explain away volumetric reductions in combined clinical groups compared with the control group

To identify explanations for the observed volume reduction, we investigated morphological differences between the combined clinical groups and the control group controlling for individual scores on the BAI and BDI. These contrasts yielded no voxel-level significant results which survived FWE correction for the whole-brain volume. Additionally, no differences were found using SVC for the ACC and hippocampus. Inclusion of continuous psychopathology scores as a covariate thus explained away group differences.

2.9.5 Group x BAI interaction effect

To investigate differences between the groups in correlations between brain structure and self-reported anxiety, independent of overall differences in brain volume, we tested the Group (PTSD, depressed) x BAI interaction. This revealed a cluster including anterior areas of the left inferior temporal gyrus and middle temporal gyrus spanning the temporal pole; a left sided cluster including the fusiform gyrus, lingual gyrus and cerebellar vermis; a right sided cluster across the anterior inferior temporal gyrus and temporal pole; and a posterior cluster in the inferior temporal gyrus spreading to the middle temporal gyrus (Table S1.2 and
Supplemental Figure 2.1). Several other areas are significant at cluster level but fail to survive FWE correction.

Further analysis of this interaction by assessing volumetric changes within each group individually, inclusively masked by the overall interaction effect, reveals that for the PTSD group BAI scores are inversely associated with grey matter volumes in the fusiform gyrus, middle frontal gyrus, inferior temporal gyrus, transverse gyrus, cerebellum, middle temporal gyrus, putamen, and frontal operculum (Table S1.3 and Figure 2.2). In contrast, the depressed group shows a positive association between BAI scores and grey matter volumes in middle temporal gyrus and superior frontal gyrus (Table 2.4 and Figure 2.3).

2.9.6 Group x BDI interaction effect

To investigate similar differences between the groups in correlations between brain structure and self-reported depression, we tested the Group (PTSD, depressed) x BDI interaction on grey matter volume. This revealed a cluster including the right lingual gyrus, bilateral cuneus and right precuneus; and a cluster within the left fusiform gyrus (Table 2.5,
2. Structural brain abnormalities common to posttraumatic stress disorder and depression

Supplemental Figure 2.2). Several other areas are significant at cluster level but fail to survive FWE correction.

Following the masking approach used to examine the earlier interaction, BDI scores in the PTSD group were inversely associated with grey matter volumes in the superior frontal gyrus, fusiform gyrus, lingual gyrus, and cuneus (Table 2.6 and Figure 2.4). In contrast, BDI scores in the depressed group were positively associated with grey matter volumes in the cuneus and precuneus, fusiform gyrus, and middle frontal gyrus (Table 2.7 and Figure 2.5).

2.10 Discussion

To the best of our knowledge, this is the first study to have directly compared brain volumes in matched samples of PTSD and depressed patients, with similar levels of depression and differing only in the presence of PTSD. Consistent with the marked overlap in areas of brain volume reduction identified in previous studies of each disorder separately, we were not able to find evidence for unique structural changes associated with PTSD. Volume reductions were found in common
predominantly in frontal areas such as the dorsal ACC, consistent with prior studies of PTSD (56-59) and depression (63). Reductions in rostral ACC also mirrored previous findings in PTSD (60, 62) and depression (64, 65, 67-69). Other areas in which we detected reduced volume in common included medial orbitofrontal cortex, as previously implicated in PTSD (95) and depression (63, 72, 73). Reduced DLPFC volume has previously been found in young adults exposed to harsh physical punishment (96) and PTSD patients show abnormal recruitment of DLPFC activity during a working-memory updating task (97). The DLPFC is also implicated in the pathophysiology of depression (25).

Of note is the absence of evidence for hippocampal reduction in the clinical samples. Findings concerning hippocampal volume have been heterogeneous, with many studies of PTSD failing to report abnormalities in this area (98-102). There may be as yet undetected associations such that specific trauma histories, e.g. prolonged or childhood trauma, are more likely to be associated with changes in these areas. Hippocampal reduction may also be a preexisting vulnerability factor rather than a general consequence of PTSD (103).

As previously noted, (62) the convergent structural neuroimaging findings in PTSD and depression raise issues about whether models of the neurocircuitry of PTSD (104) reflect common vulnerabilities to different types of psychopathology. Similar models have also been put forward to account for mood alterations in depression (25, 66). Our findings could be taken to suggest that many of the neurobiological observations in PTSD do indeed relate to common mechanisms of emotional control rather than being specific to PTSD, or alternatively that they are better explained as being due to comorbid depression.

Despite the overall lack of differences in brain volume, PTSD and depression could be distinguished by more subtle effects involving the association between scores on the continuous measures of anxiety and depression and brain volume. This adds to a previous study that found differences in functional responses to script-driven imagery in
PTSD patients with and without comorbid major depression (105). In our data, BAI scores were inversely associated in PTSD with volumes in a number of areas, primarily involving the inferior temporal lobes but also including frontal areas, the frontal operculum, and other temporal areas. Neuroanatomical experiments in monkeys have shown that the orbital and medial prefrontal cortex is associated with an extended cortical circuit (“the orbital prefrontal network”), which includes visual association areas in the inferior temporal cortex and somatic-sensory association areas in the insula and frontal operculum (25). The functions of this network, which is likely to be implicated in PTSD, are thought to include sensory integration and affective coding.

Integration of visual stimuli is performed by the inferior temporal structures that form part of the ventral visual stream, allowing visual events to be placed with a general temporal and spatial context. Unlike the dorsal visual stream, which supports egocentric representations of visual experience that retain a close link with the original perceptual input, visual processing in the ventral stream and related structures allows allocentric representations of scenes to be manipulated and combined in novel ways (106, 107). It has recently been proposed (108, 109) that flashbacks in PTSD arise from egocentric trauma representations within the dorsal stream that have received inadequate contextualization within the ventral visual stream. The fact that flashbacks are a symptom associated with PTSD and not depression (110) may be related to the differential patterns of correlation identified in this study.

BDI scores were inversely associated in PTSD with volumes in similar regions, as well as in the superior frontal gyrus. This area has been implicated in the selection of action sets (111), and response inhibition (112). More strikingly, however, BDI scores were positively associated in the depressed group with volumes in the cuneus and precuneus, areas involved in visual processing and imagery. It is well-established that intrusive imagery involving past life stressors is extremely common in depression (110), and more frequent intrusions are regularly associated with higher depression scores on the BDI (113, 114).
Among the limitations of our study was that the sample consisted of individuals with PTSD arising from a variety of sources and was too small to permit the analysis of subgroups. Although this suggests the results are more likely to be generalizable, it leaves open the possibility that different results would be obtained in more homogeneous samples. A limitation of our analysis of the group effect of antidepressant use was the relatively small groups (medication-free patients with depression, n=7). Much of the available literature is based on samples of combat veterans or sexual abuse survivors, and this may account for the fact that we did not detect volume changes in some previously identified areas such as the hippocampus. Given the evidence that prior trauma, particularly in childhood, is a risk factor for PTSD (115, 116), and that cumulative childhood trauma predicts greater symptom complexity (117), future studies should collect separate samples of PTSD following single and multiple traumatic exposure, and separately evaluate the effects of childhood trauma.

In summary, we were unable to identify differences between regions of reduced brain volumes specifically associated with PTSD and depression. Overall, areas of volume reduction are very similar; suggesting some shared deficits in mechanisms for regulating emotion. Nevertheless, each disorder reveals a distinct pattern in the relation of brain volume to continuous measures of anxiety and depression, with anxiety inversely associated with inferior temporal volumes in PTSD and depression positively associated with volumes in the cuneus and precuneus of depressed patients. The findings underscore the fact that despite high levels of comorbidity PTSD and depression appear to be different kinds of disorders, and emphasize the need for more investigations that study both conditions simultaneously and attempt to isolate their unique pathophysiological mechanisms.

Competing interests

None declared for Mr. Kroes or Drs. Rugg and Whalley. Dr. Brewin declared having received grant and travel support from the Wellcome Trust.
Author contributions

Drs. Rugg, Whalley and Brewin designed the study. Dr. Whalley acquired data. MSc. Kroes and Dr. Brewin analyzed data, and wrote the article, which Drs. Rugg and Whalley reviewed. All authors approved publication of the article.
3 ASSOCIATION BETWEEN FLASHBACKS AND STRUCTURAL BRAIN ABNORMALITIES IN POSTTRAUMATIC STRESS DISORDER.

3.1 Published as:


3.2 Abstract

Posttraumatic stress disorder (PTSD) is reliably associated with reduced brain volume relative to healthy controls, in areas similar to those found in depression. We investigated whether in a PTSD sample brain volumes in these areas were related to reporting specific symptoms of PTSD or to overall symptom severity. Structural MRI scans were obtained from 28 participants diagnosed with PTSD according to DSM-IV-TR. Participants reported the extent of individual PTSD symptoms using the Posttraumatic Diagnostic Scale. Voxel-based morphometry applying the Dartel algorithm implemented within SPM5 was used to identify volumetric changes, related to PTSD total, symptom cluster, and individual symptom scores. Brain volume was unrelated to overall PTSD severity, but greater reexperiencing scores predicted reduced volumes in the middle temporal and inferior occipital cortices. Increased reports of flashbacks predicted reduced volume in the insula/parietal operculum and in the inferior temporal gyrus. The data illustrate the value of analyses at the symptom level within a patient population to supplement group comparisons of patients and healthy controls.
3.3 Introduction

The study of structural brain variation in posttraumatic stress disorder (PTSD) has almost exclusively been based on comparisons of overall volume levels with samples exposed to trauma but who have not developed PTSD. The abnormalities that have been found, mainly reduced volumes in frontal and limbic areas, are similar to those implicated in major depressive disorder, raising the question of whether they correspond to specific characteristics of PTSD or reflect common difficulties, for example in emotion regulation (25, 118). A complementary strategy is to investigate the relationship between brain volume and the individual symptom clusters and symptoms of PTSD. This addresses the question of whether certain brain areas are relevant to variations in clinical presentation within PTSD samples, including overall severity. Such relations may exist regardless of whether these areas are reduced in volume compared to controls. The results may help to explain the function within PTSD of areas that are already known to be reduced in volume. They may also identify new areas that could be considered as regions of interest in future investigations, or indicate how specific brain regions of normal volume contribute to symptom expression.

Meta-analyses of structural MRI studies on individuals with posttraumatic stress disorder (PTSD) have consistently identified reductions in brain volume (51, 52), most prominently in the hippocampus. In addition, reductions in dorsal anterior cingulate cortex (ACC) (56-59), rostral ACC (60, 119), and insular cortex (56, 60, 61) have been implicated in PTSD. A meta-analysis (51) found further evidence for a reduction in the left amygdala of adult PTSD patients, and in the frontal cortex (but not the hippocampus) of paediatric PTSD samples. The results of individual studies of brain volume in PTSD have shown wide variability, however, usually attributed to the type of trauma, prescan duration of the illness, or MRI methodology (120, 121). Studies of depressed patients have identified volume reductions in similar areas, including subgenual, pregenual, and posterior ACC, dorsal medial/lateral PFC, orbital and ventrolateral PFC, hippocampus, and amygdala (122).
Unlike other anxiety disorders and major depression, there are as yet no “core symptoms” that are essential for a diagnosis of PTSD in the Diagnostic and Statistical Manual of Mental Disorders (DSM-IV-TR), with individuals rather having to report one out of five symptoms from cluster B (reexperiencing), three out of seven symptoms from cluster C (avoidance and numbing), and two out of five symptoms from cluster D (hyperarousal). As a result individual clinical presentations may vary widely, creating diagnostic heterogeneity. Faced with this heterogeneity, as well as with the frequently noted high levels of comorbidity, PTSD researchers have suggested that some symptoms are more highly diagnostic of the condition than others. The two groups of symptoms so identified are the avoidance and numbing symptoms (123) and the reexperiencing symptoms of flashbacks and traumatic nightmares (110).

In the current study we investigated whether structural brain variation in a PTSD sample was related to overall symptom severity and to the intensity of symptoms in specific PTSD clusters such as avoidance/numbing and reexperiencing. We applied voxel-based morphometry to the structural MRI images using SPM5 and the Dartel toolbox, which has been found to optimize the sensitivity of such analyses (79, 89). We demonstrate no grey matter volume alterations related to overall PTSD severity, but do report an inverse relation between grey matter volume in the middle temporal and inferior occipital cortices and reexperiencing symptom cluster scores. Of the individual reexperiencing symptoms, reports of flashbacks were related to volume reductions, predominantly in the inferior temporal gyrus and parietal operculum/insula. Limited evidence was found for volumetric changes related to reports of nightmares and experiencing distress on trauma reminders.

### 3.4 Methods and Materials

#### 3.4.1 Participants

Participants were a mixed sample of individuals meeting diagnostic criteria for current PTSD. The majority of participants were outpatients.
recruited from a specialist traumatic stress clinic, where diagnoses were independently confirmed, and the remainder from advertisements. Exclusion criteria were a history of head injury, neurological disorders, psychosis, or other major medical conditions, as well as current substance abuse. PTSD diagnoses were confirmed with the Structured Clinical Interview for DSM-IV (80), administered by a trained postdoctoral psychologist under the supervision of a clinical psychologist expert in trauma and PTSD. Demographic and clinical characteristics of the 28 right-handed individuals who met inclusion criteria are reported in Table 3.1. All patients gave written informed consent, and all procedures were approved by the National Hospital for Neurology and Neurosurgery & Institute of Neurology Joint Research Ethics Committee.

3.4.2 Measures

The Posttraumatic Diagnostic Scale (PDS) (83) is a widely-used self-report measure with high reliability and validity. Items measuring each of the 17 PTSD symptoms are rated on a 0-3 scale and can be summed to yield subscores for the reexperiencing, avoidance/numbing, and hyperarousal symptom clusters as well as a total score. In this sample the mean PDS score (see Table 3.1) was above the average of 33.59 scored by individuals diagnosed with DSM-IV PTSD in the original validation sample (83). PDS scores were higher in those currently on antidepressant medication, r(28) = .50, p < .01.

The Beck Depression Inventory 2 (82) is a widely-used 21-item self-report measure of depression severity. Items are scored on a 4-point scale (possible range 0-63). Participants were instructed to rate their mood over the past week. The average score in the sample (see Table 3.1) indicated a mood state associated with severe depression (81).

3.4.3 MRI acquisition

T1-weighted images were acquired on a Siemens 1.5 T Sonata whole body scanner (Siemens Medical Systems, Erlangen, Germany), using a whole-
body coil for RF transmission and a head coil for signal reception. Whole-brain structural scans were acquired using a Modified Driven Equilibrium Fourier Transform (MDEFT) sequence (84) with optimized parameters as described in the literature (85). For each volunteer, 176 sagittal partitions were acquired with an image matrix of 256x224 (Read x Phase). Two-fold over-sampling was performed in read direction (head/foot direction) to prevent aliasing. The isotropic spatial resolution was 1 mm. Relevant imaging parameters were TR/TE/TI = 12.24 ms/3.56 ms/530 ms, BW = 106 Hz/Px, $\alpha = 23^\circ$. The total duration was 12 min. Spin tagging in the neck was performed to avoid flow artefacts in the vicinity of blood vessels. The flip angle of the tagging pulse was chosen to be $110^\circ$ to account for $B_1$ losses in the neck. Special RF excitation pulses were used to compensate for $B_1$ inhomogeneities of the transmit coil in superior/inferior (86) and anterior/posterior (87) directions. Images were reconstructed by performing a standard 3D Fourier Transform, followed by modulus calculation. No data filtering was applied either in k-space or in the image domain.

3.4.4 MRI Processing

T1-weighted image registration was achieved using the diffeomorphic registration algorithm implemented in the Dartel toolbox (88) for SPM5 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). First, the anatomical images were manually reoriented so that the mm coordinate of the AC matched the origin [0,0,0], and the orientation approximated MNI space. Next T1-weighted images were classified into grey matter (GM), white matter (WM) and cerebrospinal fluid (CSF) using the segmentation routine implemented in SPM5 applying standard settings. The resulting parameter files were imported into the Dartel procedure to produce rigidly aligned grey and white matter tissue classes resliced to 1.5x1.5x1.5 mm voxel size. After the affine transformation the rigidly aligned tissue class images were used to estimate the nonlinear deformations to best align all images. During this estimation stage, Dartel iterates between building a template and registering tissue class images using the template. The resulting flow-
fields, which specify the parameterized deformations, were used to warp white and grey matter images for each subject. The spatially normalized images were rescaled by the Jacobian determinants of the deformations, using 64 time points for solving the partial differential equations. 7-th degree B-spline interpolation was used when writing the normalized images. In order to obtain meaningful coordinates of activation, the final Dartel template was normalized to MNI space and the resulting deformations applied to the grey matter images of each subject using a Matlab script downloaded from the SPM email list (https://www.jiscmail.ac.uk/cgi-bin/wa.exe?A2=ind0807&L=SPM&P=R34150%20032749). Finally the images were smoothing using an 8x8x8 mm Gaussian smoothing kernel. The input features for the subsequent analysis were the smoothed, modulated and normalized GM images.

3.4.5 Analysis

Regression was used to relate PTSD total, symptom cluster, and individual symptom scores to brain morphology. Where appropriate covariate vector scores were centred to the overall mean and no interactions were modelled. Multiple regression was used to analyze the unique effects of each symptom cluster controlling for the other two clusters. Given the relatively small variance in the symptom scores more detailed analyses at the symptom level used simple regression. We report regions that survive cluster-level corrections for multiple comparisons (family-wise error, FWE) across the whole brain at p<0.05, with individual voxels thresholded at p<0.001, unless otherwise specified. As non-stationarity is problematic in cluster-level inferences in VBM data (90, 91) requiring adjustment of cluster sizes according to local smoothness of data (90, 92), we have adjusted our cluster level results according to this method as implemented in the VBM5 toolbox developed by C. Gaser (http://dbm.neuro.uni-jena.de/vbm/download/). One subject was considered an outlier with an overall PDS score of 19, more than twice the standard deviation below the mean, but repetition of the analysis excluding this person did not affect the results.
3.5 Results

3.5.1 Overall Severity of PTSD

Analysis of the total PDS score did not reveal any significant grey matter volume decreases or increases that survived cluster level correction (see above).

3.5.2 PTSD Symptom Clusters

Grey matter volume alterations were uniquely predicted by reexperiencing scores, but not by any other symptom cluster scores. Reexperiencing scores predicted reduced grey matter volume in the left middle temporal gyrus and right inferior occipital gyrus (Table 3.2, Figure 3.1). For illustrative purposes the adjusted data at the peak voxel (Y-axis) plotted against the reexperiencing scores (X-axis) are depicted in Figure 3.1. The plot shows a strong linear effect indicating that the volume...
Figure 3.2 Reductions in brain volume with increasing flashback scores

Increased flashback scores predict decreased grey matter volume in the inferior temporal gyrus (A), parietal operculum (B), cerebellum (C), superior medial gyrus (D), and middle temporal gyrus (E). Results are displayed on the normalized Dartel grey matter template (image thresholded at p < .001, uncor.).

Right: The adjusted responses at the peak voxel for each region displayed (Y-axis) are plotted against the flashback scores (X-axis) for each individual subject.
reductions are associated with overall increases in reexperiencing scores. Although two possible outliers can be detected, exclusion of these cases increased the strength of the effect.

In additional analyses (Tables 3.3-3.6) we examined the effects of controlling for age, gender, education level, and Beck Depression Inventory scores. Reductions in the middle temporal area were still observed when controlling for BDI scores, but reduction in right occipital areas were no longer detected. In all additional analyses there was strong and consistent evidence for a relationship between reexperiencing and a reduction in the size of right inferior temporal cortex (BA20).

### 3.5.3 PTSD Reexperiencing Symptoms

Simple regressions conducted on the five reexperiencing symptoms separately revealed marked effects only for flashback scores (symptom B3). Increasing reports of flashbacks predicted reduced grey matter volume in the right inferior temporal gyrus and left parietal operculum (Table 3.7, Figure 3.2) as well as a cerebellar region. In addition reduced volumes within the medial extent of the superior frontal gyrus, the right middle temporal gyrus, the left inferior temporal gyrus, the right inferior occipital gyrus, and the left inferior frontal gyrus were cluster-level significant but did not survive whole-brain correction. Again for illustrative purposes plotting the adjusted data at the peak voxel against the flashback scores of any of the detected voxel-clusters reveals strong linear effects indicating that the volume reductions are associated with overall increases in flashback scores and are not driven by outliers. Flashback scores were unrelated to gender, age, or lifetime units of alcohol, largest r(28) = .17, p > .10.

In additional analyses (Tables 3.8-3.11) we examined the effects of controlling for age, gender, education level, and Beck Depression Inventory scores. Reductions in the right inferior temporal gyrus (BA20) and left parietal operculum, but not the cerebellum, were still consistently observed. In contrast, increasing reports of nightmares predicted volume
reductions in the right precentral gyrus and right inferior occipital gyrus that failed to reach significance after whole-brain correction (Table 3.12). Increasing reports of distress on trauma reminders predicted significant volume reductions in the right cerebellum. They also predicted volume reductions in the right superior orbital gyrus but these did not survive whole-brain correction (Table 3.13).

3.6 Discussion

Previous structural brain imaging studies of patients suffering either from PTSD or depression have identified reduced volume in similar frontal and limbic areas, including the hippocampus. Contrary to the assumption that the overall level of PTSD symptoms is linearly related to structural changes in the brain, we were unable to find any evidence that volume in these areas was related to overall symptom severity in a sample all diagnosed with PTSD. As previously noted, volumetric alterations related to symptom levels within a PTSD sample may or may not be present despite overall volume differences relative to healthy samples. PTSD is a highly heterogeneous disorder, encompassing intrusive cognition, deliberate avoidance, emotional numbing, and increased arousal. Our data indicated that smaller grey matter volumes were specifically related to reporting more symptoms from the reexperiencing cluster and, within that cluster, predominantly with flashback symptoms.

Symptom cluster analyses revealed two areas where volume was inversely related to increased reexperiencing scores, one being the left middle temporal gyrus (BA 21). This area has previously been linked to semantic processing (124) and to the retrieval of autobiographical memories (125), as well as being part of the 'default network' of brain areas that are active at rest and are thought to be involved in processing self-related information and introspective memory retrieval and has been found to be disregulated in PTSD (126-128). A recent functional MRI study of individuals with PTSD related to early life trauma found significantly reduced middle temporal activity in BA 21 compared to healthy controls (129). Stress-induced dissociative psychopathology has also been linked to
reduced middle temporal activity (130). The other area with significantly reduced volume was the right inferior occipital gyrus (BA 18). This area is part of the ventral visual stream, involved in the selection and classification of visual information. As explained in greater detail below, it has been suggested that in PTSD processing of traumatic scenes by the ventral stream is impaired (131).

Flashbacks were the symptom within the reexperiencing cluster to be most clearly related to reductions in brain volume. Flashbacks refer to intrusive memories in which the traumatic incident, including accompanying emotions and bodily reactions, are subjectively reexperienced in the present. These features, along with the fact that retrieval is involuntary, have generally been attributed to the underlying memory representations lacking adequate contextualization (104, 132-134). Compared to ordinary autobiographical memories flashbacks are typified by prominent sensory and somatosensory features such as pain (107, 135). They may vary in intensity from a transient sense of reliving the past while simultaneously being aware of the present to a full-blown reliving accompanied by a complete loss of connection with the outside world. This sense of reliving in the present distinguishes the intrusive sensory memories found in PTSD from those found in other disorders such as depression (136).

We found that flashbacks were inversely related to volume in an area including the insula and the parietal operculum. The insula has previously found to have a reduced volume in PTSD (56, 60, 61), and both areas are strongly implicated in somatosensory processing. A functional MRI investigation of PTSD patients using PET has reported that flashback intensity was positively related to bilateral activity in the insula, as well as being negatively related to activity in the bilateral superior frontal cortices, right medial temporal cortex, and right fusiform cortex (137).

We found flashback scores to be inversely related to the volume of lateral ventral regions of the right temporal cortex. The inferior temporal cortex is part of the ventral visual stream, and is thought to be involved
in coding abstract features of visual scenes. Along with other temporal lobe structures such as the hippocampus, activity in this region allows events to be placed with a general temporal and spatial context. Unlike the dorsal stream, which supports egocentric representations of visual experience that retain a close link with the original perceptual input, visual processing in the ventral stream and related structures allows allocentric representations of scenes to be manipulated and combined in novel ways (106, 107). Consistent with evidence that PTSD may be associated with a preexisting deficit in allocentric processing (103), it has recently been proposed that flashbacks in PTSD arise from egocentric trauma representations within the dorsal stream that have received inadequate contextualization within the ventral visual stream (131). An inverse association between more flashbacks and the volume of ventral temporal cortex is consistent with this theory.

Among the limitations of the present findings are the modest sample size and use of 1.5T MRI, both of which may have reduced the power to detect other significant effects. We also draw attention to the lower reliability and relatively small variance in the scores of the independent variable when analyses are conducted at the level of individual symptoms. This underscores the need for replication and could be overcome in future studies by including a more extensive assessment involving multiple items targeting the symptoms of most theoretical interest. Our sample also consisted of individuals with PTSD arising from a variety of sources and was too small to permit the analysis of subgroups. Although this suggests the results are more likely to be generalizable, it leaves open the possibility that, for example, effects may be stronger in civilian than military trauma. The data are also silent about whether there are volume differences in the identified areas compared to healthy controls, but they do identify regions of interest for future studies. A final limitation is uncertainty about the possible direction of any causal effects. Our data are consistent both with the experience of repeated reexperiencing or flashbacks causing reductions in brain volume and with preexisting volume reductions in these regions leading to increased symptoms,
although recent human imaging data suggests that prolonged stress can have a causal long-lasting effect on brain function (138).

In contrast to what might have been predicted from previous research, our data did not identify the hippocampus as a site in which brain volume varied with symptom intensity. There are various explanations of this. For example, hippocampal reduction has also been found in several other disorders, particularly depression, and may be a marker for more general psychopathology rather than PTSD symptoms specifically. Alternatively, the relation between hippocampal volume and symptom expression may be more variable across participants. Our data did nevertheless identify brain regions that may be involved in specific PTSD symptoms. Further, they suggest that variability in the findings of structural imaging studies in PTSD, and conceivably in other disorders as well, may be accounted for by in part by heterogeneity in the constituent symptoms.

In summary, volumetric studies at the symptom level have the ability to complement studies of participants defined by an overall syndrome. Because the analyses focus on relation to individual symptoms, they may identify areas where volumetric reductions are significant predictors of symptom expression without characterizing the disorder as a whole (and without necessarily demonstrating differences with healthy controls). We have shown that such analyses can identify new regions of interest that are consistent with a neurobiological model of flashbacks (103, 139).

**Competing interests**

The authors declare that they have no conflicts of interest concerning this article

**Acknowledgements**

The research was funded by a Wellcome Trust project grant number 074145.
Author contributions

Drs. Rugg, Whalley and Brewin designed the study. Dr. Whalley acquired the data, which MSc. Kroes and Dr. Brewin analyzed. MSc. Kroes and Dr. Brewin wrote the article, which Drs. Rugg and Whalley reviewed. All authors approved publication of the article.
3. ASSOCIATION BETWEEN FLASHBACKS AND STRUCTURAL BRAIN ABNORMALITIES IN POSTTRAUMATIC STRESS DISORDER
4 **AN FMRI INVESTIGATION OF POSTTRAUMATIC FLASHBACKS.**

4.1 *Published as:*


4.2 **Abstract**

Flashbacks are a defining feature of posttraumatic stress disorder (PTSD), but there have been few studies of their neural basis. We tested predictions from a dual representation model of PTSD that, compared with ordinary episodic memories of the same traumatic event, flashbacks would be associated with activity in dorsal visual stream and related areas rather than in the medial temporal lobe. Participants with PTSD, with depression but not PTSD, and healthy controls were scanned during a recognition task with personally relevant stimuli. The contrast of flashbacks versus ordinary episodic trauma memories in PTSD was associated with increased activation in sensory and motor areas including the insula, precentral gyrus, supplementary motor area, and mid-occipital cortex. The same contrast was associated with decreased activation in the midbrain, parahippocampal gyrus, and precuneus/posterior cingulate cortex. The results are discussed in terms of theories of PTSD and dual-process models of recognition.
4.3 Introduction

Post-traumatic stress disorder (PTSD) has frequently been characterised as a disorder of memory (140, 141). PTSD patients typically experience flashbacks, involuntary sensory images of the traumatic scene. These images are vivid, detailed, and lack temporal context, being re-experienced as though they were happening in the present (135). Flashbacks co-exist with ordinary episodic memories of the trauma that can be deliberately retrieved and communicated. The dual representation theory of PTSD (132, 133) makes an unusual prediction that, despite flashbacks being extremely vivid, they should involve less rather than more medial temporal lobe (MTL) activity. We carried out a preliminary test of this hypothesis by contrasting neural responses to stimuli eliciting either flashbacks or ordinary episodic memories of the same traumatic event in PTSD patients.

Previous research on the functional neuroanatomy of voluntary autobiographical memory has identified a core network of predominantly left-lateralized regions including the prefrontal cortex, MTL (in particular the hippocampus), and posterior cingulate (142). Emotional memories, in contrast, are associated with bilateral activation and engage additional areas such as the amygdala and insula (142, 143). Emotion is thought to enhance recollection, and increased recollective qualities such as level of detail, personal significance, and emotionality have been found to be accompanied by increased MTL and hippocampal activity (144-146).

Research on healthy participants therefore suggests that particularly detailed and emotional memories should be accompanied by high levels of activation in networks including hippocampal and parahippocampal regions. In contrast, the dual representation theory of PTSD distinguishes normal episodic memories, supported by flexible, contextualised representations, from flashbacks, supported by representations that are inflexible and lacking in context (131). Although there is no formal definition of a flashback, they are commonly thought to involve an intense sensory and emotional re-experiencing of the traumatic event
that exists on a continuum, ranging from complete loss of awareness of surroundings to a milder experience of reliving in the present (140). This continuum view is part of the definition proposed for the upcoming fifth revision of the American Psychiatric Association Diagnostic and Statistical Manual (DSM-V) (147). In the context of a basically intact episodic memory system, (visual) flashbacks are hypothesized (131) to reflect the dominance during the traumatic event of activity in the dorsal visual stream, extending from posterior visual to superior parietal regions, that processes sensation-near representations of experience designed to facilitate action. These dorsal visual stream representations are thought to be strongly associated with activity in motor cortex as well as with activity in the insula and amygdala. Dorsal stream activity is hypothesized to take precedence over activity in the ventral visual stream, including inferior and middle temporal regions, that ordinarily provides memories with their context.

Most previous neuroimaging studies of PTSD have employed the script-driven imagery paradigm, in which participants simultaneously recall and imagine a traumatic event, a process which typically elicits additional involuntary trauma memories (148). A case study was reported (149) in which a Vietnam veteran experienced a flashback during the perfusion phase of a SPECT scan. Relative to his own baseline, and to that of a PTSD sample who did not experience flashbacks, he exhibited decreases in blood flow to a wide range of cortical and subcortical regions. More recently subjective flashback intensity was correlated with rCBF in eleven PTSD patients (137). They observed flashback-related increases in left inferior frontal cortex and bilateral insula, and flashback-related decreases in right medial temporal, right fusiform, and bilateral superior frontal cortices. Increased flashback reports in PTSD patients have also been found to be correlated with reduced brain volume in the left insula/parietal operculum and in the right inferior temporal gyrus (150).

What are so far lacking are studies that directly compare flashbacks with ordinary episodic memories of the same traumatic event. In this study PTSD patients wrote a narrative account of their traumatic event, and
identified which sections were and were not accompanied by flashbacks (151, 152). We subsequently presented them with words and phrases from these flashback and episodic memory sections mixed with stimuli from another narrative. In order to determine which patterns of observed activity were unique to PTSD, we employed control groups of depressed patients and healthy controls who had experienced a traumatic event. We predicted that recognition of own versus other stimuli should show common effects across groups in previously identified areas, reflecting an intact episodic memory system. We also predicted that the flashback-episodic trauma memory contrast in PTSD patients would be distinguished from trauma memory processing in the two control groups by an absence of common effects, and by increased dorsal stream and decreased ventral stream activity.

4.4 Methods and Materials

4.4.1 Participants

Participants were 39 right-handed individuals without a history of head injury, neurological disorders, or other major medical conditions, assessed using the Structured Clinical Interview for DSM-IV (80). Ten patients met DSM-IV criteria for current PTSD (PTSD group). Fifteen participants had experienced traumatic events similar in magnitude to the PTSD group, but had not developed PTSD (trauma-exposed Control group). Fourteen patients meeting DSM-IV criteria for current major depression but not PTSD were also tested (Depressed group). Patients in the PTSD and control groups had experienced a range of traumas, including involvement in the July 7th 2005 London bombings and other terrorist attacks (PTSD= 2; control=7), road traffic accidents (PTSD= 1; control=3), interpersonal violence (PTSD= 5; control=3), the 2004 Asian tsunami (PTSD=2; control=1). Time since index trauma ranged from 2 months to 24 years, with a median of 2 years. Depressed patients had experienced a range of severe negative life events including some meeting PTSD Criterion A.
4. An fMRI Investigation of Posttraumatic Flashbacks

4.4.2 Task

All participants visited the lab approximately a week prior to the scan and wrote a narrative account of their traumatic or most distressing event, starting from just before they knew something was wrong until the point where the event had resolved. Participants were prompted to include description of what they could see, hear, touch, smell, taste, and feel at each stage of the incident, and also to describe their thoughts and emotions. For participants who reported multiple distressing experiences the event was chosen which bothered them most. At completion participants were invited to highlight any sections of the narrative during the writing of which they had experienced flashbacks. Flashbacks were defined for participants in the following way: “A type of memory that you experience as markedly different from those memories of an event that you can retrieve at will. The difference might be a marked sense of reliving of the traumatic experience(s). Some report complete reliving, whereas others report more momentary or partial reliving of perhaps just one aspect of the original experience. For some, flashback memories take them by surprise or swamp their mind. Finally, some report a sense of time-distortion and, for example, react to the flashback memory as though it was an event that was happening in the present.” Only PTSD patients identified flashbacks. Numbers of flashbacks were not counted separately.

Participants completed an autobiographical retrieval task, using stimuli culled from the written narratives, whilst brain activity was measured with fMRI. Participants were presented with items from their own narrative (Own), interspersed with items from another PTSD patient’s narrative thematically unrelated to their own (Other), and were asked to identify Own items. This type of cue-word task has previously been used in neuroimaging studies of autobiographical retrieval (153-155). For PTSD participants, Own items were additionally either associated with flashbacks they had of events (Flashback), or were simply normal episodic memories of the event (Episodic). Analyses were conducted to investigate
successful autobiographical retrieval (Own>Other) and activity specific to flashbacks (OwnFlashback>OwnEpisodic).

4.4.3 Stimuli

Words (task 1) and phrases (task 2) were included to increase the number of stimuli that could be used without directly repeating items. Words and phrases (typically 2-8 words long) were chosen to be descriptive of key components of each narrative. Words were matched to 'master lists' (see later) on variables of word frequency and number of letters per word, and phrases was matched on the variables of number of words and letters per sentence (examples of master list items are given in supplementary materials). Two ‘master lists’ were generated to serve as control stimuli. One participant with PTSD had survived the July 2005 London bombings, the other had survived the December 2004 Asian Tsunami. If a participant being tested had experienced one of these events then the alternative list was used. Certain words such as ‘blood’ or ‘helpless’ were common to many narratives, and words on the master list were substituted out on a case-by-case basis whenever overlap was identified.

4.4.4 Procedure

Stimuli were presented via a mirror mounted on the head coil of the scanner, in direct view of the supine participant, at a distance of approximately 50cm from the projection screen. Participants used an MR-compatible button-box in their right hand to respond and were instructed to respond as quickly and accurately as possible. For both tasks the presentation of an item was preceded by an asterisk (*) for 500ms, followed by the item for (1000ms in task 1, 1700ms in task 2), followed by a fixation cross for (1000ms in task 1, and 2000ms in task 2). These sequences of events gave SOAs for tasks 1 and 2 of 3500ms, and 4200ms respectively. 'Null events' consisted of a fixation cross presented for an entire SOA.
In task 1 72 Own words and 72 Other words from one of the master lists were presented in the centre of the screen in uppercase Arial font. 72 null-events were included whereby a fixation cross was presented for an entire SOA. Task 2 was identical but with a longer SOA to accommodate time for additional reading. Sixty Own phrases were presented along with 60 Other phrases in lowercase Arial font. After the scan PTSD participants were given lists of all the stimuli they had seen during the test and were required to identify whether or not each item had led to the experience of a flashback during the scan.

### 4.4.5 MRI acquisition

MRI data were acquired on a 1.5 T whole body scanner (Magnetom Sonata, Siemens Medical, Erlangen, Germany) operated with an 8-channel phased array receive coil and the standard body transmit coil. The manufacturer’s standard automatic 3D-shim procedure was performed at the beginning of each experiment. The participants were scanned with a single-shot gradient-echo EPI sequence sensitive to the blood-oxygen level dependent (BOLD) effect using the following imaging parameters: 30 oblique transverse slices, slice thickness = 2.5 mm, gap between slices = 1.25 mm, repetition time TR = 3 s, flip angle $\alpha = 90^\circ$, echo time TE = 50 ms, readout bandwidth BW = 2298 Hz/pixel, bandwidth in PE direction BWPE = 31.3 Hz/pixel, phase-encoding (PE) direction anterior-posterior, field of view FOV = 192x192 mm$^2$, matrix size 64x64, fat suppression. BOLD sensitivity losses in the orbitofrontal cortex and the amygdala due to susceptibility artifacts were minimised by applying a z-shim gradient moment of -0.8 mT/m*ms, a slice tilt of -300 and a positive PE gradient polarity (156). EPI magnitude images were reconstructed from the complex k-space raw data using a generalised reconstruction method based on the measured EPI k-space trajectory to minimise ghosting (157), combining the single coil images by sum of squares (158). EPI data acquisition was monitored on-line using a real-time reconstruction and quality assurance system (159). Data were acquired during two separate sessions, with the first five volumes of each session discarded to allow for T1 equilibration effects. A magnetic (B0) field map image was collected before the first session.
and was used to unwarp the echo-planar images (160, 161). Subjects were placed in a light head restraint within the scanner to limit head movement during acquisition. A 3D MDEFT T1-weighted structural image was also acquired following the functional acquisition for superimposing statistical maps over anatomy. Whole-brain structural scans were acquired using a Modified Driven Equilibrium Fourier Transform (MDEFT) sequence (Ugurbil et al, 1993) with optimized parameters as described in the literature (85). For each volunteer, 176 sagittal partitions were acquired with an image matrix of 256x224 (Read x Phase). Two-fold over-sampling was performed in read direction (head/foot direction) to prevent aliasing. The isotropic spatial resolution was 1 mm. Relevant imaging parameters were: TR/TE/TI = 12.24 ms/3.56 ms/530 ms.

4.4.6 MRI Processing

fMRI data were processed and analysed using the statistical software package SPM5 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). The first 5 EPI volumes were discarded to allow for T1 equilibration. The remaining functional images from each subject were realigned using rigid-body transformation to correct for head movements to the mean functional image using 7nd-degree B-spline interpolation, unwarped, and coregistered to the anatomical T1-weighted MR image using a normalized mutual information function. Next, structural images were segmented and normalized into a common stereotactastic space (MNI 152 T1-template). Subsequently, the normalization parameters were applied to the functional images and these were resampled to 3 *3 *3-mm3 isotropic voxels 7nd-degree B-spline interpolation. Finally, spatial smoothing was applied with an isotropic 3D Gaussian kernel of 8 mm full-width half-maximum.

4.4.7 Analysis

Statistical analysis was performed in two stages of a mixed effects model. In the first stage the neural activity was modeled by a stick function (impulse event) at stimulus onset (task 1 - words) or an epoch with a
duration of 1.5s (task 2 - phrases). The ensuing BOLD response was modeled with a canonical haemodynamic response function (HRF) (162). An AR(1) model was used to estimate and correct for non-sphericity of the error covariance (163). For each voxel, the image time-series was high pass filtered to 1/128 Hz.

The resulting general linear model (GLM) was used to obtain parameter estimates representing the activity elicited by the events of interest. Four event types were defined, consisting of correct responses to 'Own Flashback' items, correct responses to 'Own Episodic' item, correct rejections to 'Other Flashback' items, and correct rejections to 'Other Episodic' items. For PTSD patients the Flashback/Episodic distinction was made by combining pre- and post-scan reports of which items were associated with flashbacks. Thus for this group any items which led to reports of a flashback, either during the narrative or while being scanned, were assigned to the 'own flashback' category. For control and depressed participants items were randomly assigned to 'flashback' and 'episodic' categories in order to allow identical data analysis procedures. Participants’ performance was high enough such that it was not possible to model misses. Also included for each session were six covariates to capture residual movement-related artifacts (three rigid-body translation, and three for rotation).

Contrasts for effects of interest were specified at the first level, and entered in to a full-factorial model at the second level. Run 1 (words) and run 2 (phrases) were included as separate sessions within the second-level model, but to maximize power, and because we had no a-priori hypotheses requiring separate analyses, all results presented here are reported for the combination of words and sentences tasks. Based on an analysis of demographic data covariates of no interest 'length of time since trauma, and'BDI' were included in the model.

All effects were thresholded at p<0.001 with an extent of k>5 unless otherwise specified. Between-group effects were generated by interrogating the model discussed above. Common effects (effects
common to the PTSD, control and depressed groups) were obtained by an inclusive mask of activity in each of the three groups, each thresholded at $p<0.025$ to give a conjoint threshold, via Fisher’s procedure, of $p<0.001$ (conjunction null) (164).

4.5 Results

4.5.1 Demographics and behavioural data

Patient demographics and scores on clinical measures are given in Table 4.1. Groups did not differ according to age or the age at which they left full-time education, but post-hoc tests indicated that the depressed group had experienced their index event significantly longer ago than the controls. BDI and BAI scores for the PTSD and Depressed groups did not differ, but both were significantly higher than the controls. Participants in the PTSD group scored significantly higher than controls on the PDS. Nine of the depressed group, 3 of the PTSD group, and none of the control group were taking antidepressant medication.

Memory performance is given in Table 4.2. Pr provides an unbiased estimate of accuracy in the response to old and new items, where higher values correspond to greater accuracy. Br is an index of response bias, the tendency to respond "old" or "new" regardless of accuracy (165).

There was a main effect of task on Pr, ($F_{1,36} = 283.42, p < 0.001$), but no main effect of group or task by group interaction. Pr was significantly higher for phrases than for words ($t(38) = 16.77, p<0.001$, phrases mean 0.79, SD 0.09; words mean 0.55, SD 0.12). There was also a main effect of task on Br ($F_{1,36} = 17.02, p< 0.001$), but no main effect of group or task by group interaction. Br was significantly higher for words than for phrases, ($t(38) = 4.366, p<0.001$, words mean 0.47, SD 0.14; phrases mean 0.32, SD 0.22). Response bias for the words task was neutral but for the sentences task was $<0.5$, indicating a propensity for participants to respond "new". This pattern is driven by the very low rate of false alarms: participants were very accurate in correctly rejecting phrases from another
individual's narrative but were more cautious identifying phrases as being from their own narrative.

Patients in the PTSD group reported flashbacks to an average of 50.4 out of a total of 264 individual stimuli. Of these, an average of 31.6 were to words or phrases that had been reported as eliciting a flashback during the narrative task, and an average of 15.5 were to words or phrases that had not been reported as eliciting a flashback during the narrative task. The remaining 3.3 flashbacks were to words or phrases from the control (Other) narrative. As the focus of interest was the experience of a flashback, all these items were included in the analyses.

### 4.5.2 Common fMRI results of successful retrieval

Analyses were first directed at finding activity common to all three groups during the successful retrieval of autobiographical memories. Subsequent analyses aimed to delineate between-group differences. Unless otherwise specified all contrasts reported here include the variables'BDI and 'time since trauma as covariates of no interest. Effects common to all three groups for the Own>Other contrast (i.e hits > correct rejections) are given in Table 4.3 and Figure 4.1. There was significant activity in the left middle, superior medial, and superior frontal cortices. Bilateral posterior cingulate, right caudate, left insula, left retrosplenial cortex, and left middle temporal cortex also demonstrated activation. Minimal between-group effects were observed for the Own>Other contrast. Relative to the depressed and control groups, the PTSD group demonstrated increased activity in a 5-voxel region of the right mid-cingulate cortex ($x=9$, $y=9$, $z=45$), and decreased activity in an 8-voxel region of the right inferior frontal gyrus ($x=45$, $y=33$, $z=-3$).

### 4.5.3 Increased fMRI effects for flashback vs episodic memory retrieval in the PTSD group

We next examined activity associated with the OwnFlashback>OwnEpisodic contrast. There were no effects common to all three groups. Table 4.4 and
Figure 4.2 report activations associated with flashbacks versus episodic memories in the PTSD group only. Flashback-associated increases in activations were evident in left anterior cingulate (ACC) and middle occipital cortices, right supplementary motor area (SMA) and medial prefrontal cortex, and bilaterally in precentral areas, inferior parietal (supramarginal) cortices, and the insula.

4.5.4 fMRI effects flashback vs episodic memory retrieval across groups

We further investigated regions demonstrating flashback-associated increases or decreases in the PTSD group relative to the control and depressed groups. These are given in Table 4.5 and Figure 4.3. Parameter estimates for each group are given in Figure 4.4. There was a substantial overlap with the preceding analysis, PTSD patients showing significantly greater flashback-specific activity in regions of the right SMA, left middle
4. An fMRI Investigation of Posttraumatic Flashbacks

PTSD

OwnFlashback > OwnEpisodic

Figure 4.2: Flashback-specific activity in the PTSD group.
Results thresholded at p<0.001, k>5.

occipital gyrus and precentral cortex, and right insula. In addition PTSD
Flashback-related increases: PTSD > (Depressed + Control)

- R Insula
- L Middle Occipital Gyrus
- L Precentral Gyrus
- R SMA
- Midbrain
- R Precuneus / PCC

Flashback-related decreases: (Depressed + Control) > PTSD

- y=21, z=15
- y=-9, y=6
- x=3, y=-27
- L Parahippocampal Gyrus

Figure 4.3: Flashback-specific between-groups differences.
Results thresholded at p<0.001, k>5.
patients showed reduced flashback-specific activity in the midbrain, left fusiform / parahippocampal gyrus, and right precuneus / PCC.

4.6 Discussion

Behavioural results indicated that participants in all groups were able to identify their own autobiographical events with a high degree of accuracy. This contrasts with evidence for an overall deficit in PTSD in learning neutral material (166). Effects common to all three groups during the successful retrieval of autobiographical memories were observed in several prefrontal areas as well as the posterior cingulate, precuneus, and retrosplenial cortex, regions identified as primary areas of the autobiographical memory system (142, 167). As in previous studies of emotional memory (142, 143), insula activity was also common to all groups. Between-group differences for the Own>Other contrast were much more limited. Relative to the other groups PTSD patients showed increased activity in the right mid-cingulate cortex, and reduced activity...
in the right inferior frontal cortex. This limited pattern of differential activity is consistent with previous work (77) in suggesting a substantially intact neural system supporting episodic retrieval.

As predicted, given the absence of flashbacks reported by depressed patients and controls, there were no effects common to all three groups for the OwnFlashback>OwnEpisodic contrast. Relative to the depressed and control groups, PTSD participants exhibited flashback-specific increases in the right insula, left mid-occipital cortex, left precentral cortex, and right SMA. Consistent with the hypotheses, many of these interconnected areas are involved in translating dorsal visual stream activity into motor movements. For example, a connectivity analysis (168) linked body orientation in space to the SMA, precentral gyrus, and related areas.

Analyses within the PTSD group alone found additional evidence for greater activation of motor cortex and other closely connected regions involved in response selection such as dorsal ACC. This is of particular interest as dorsal ACC is thought to modulate fear expression in humans (169). It appears to be hyperresponsive in PTSD, and there is evidence for hyperresponsivity in some areas of dorsal ACC being a familial risk factor for PTSD (170). Other analyses in the PTSD group also revealed bilateral activation in the supramarginal gyrus. This inferior parietal area is hypothesized to be involved in the capture of attention bottom-up by behaviourally-relevant stimuli (171). The 'Attention to Memory model' specifically hypothesized that involuntary remembering would activate this area (171), but to our knowledge our findings are among the first to confirm their prediction.

The PTSD flashback-specific decreases were observed in the midbrain, the left parahippocampal gyrus, and the right precuneus. The precuneus is a complex area subserving a number of different functions. Whereas the anterior precuneus is mainly connected to sensorimotor areas, also including a projection to the insula, the central precuneus is viewed as a cognitive/association area (172). It has been argued that the precuneus
4. AN FMRI INVESTIGATION OF POSTTRAUMATIC FLASHBACKS

is part of an important network subserving awareness, and that during aversive sensations there may be efforts to terminate self-reflection, resulting in decreased processing in the precuneus (173). Flashbacks are certainly experienced as aversive by PTSD patients, as indicated by the fact that avoidance of involuntary trauma memories forms part of the diagnostic criteria. Alternatively, reduced precuneus activation may be connected with the involuntary nature of memory retrieval (174). The parahippocampal gyrus is a ventral visual stream region thought to be involved in allocentric scene information (106) and predicted to be less active during flashbacks (131).

The midbrain region contains a large number of nuclei critical to the neuromodulation of brain function. Additionally this region is part of a stress-related brain network that is modulated by noradrenaline (175), and noradrenaline levels are increased in PTSD (176). The alterations in midbrain processing for flashbacks in PTSD patients may reflect stress-related changes in attention and arousal.

The neural activity distinguishing the OwnFlashback>OwnEpisodic contrast in PTSD patients has some striking similarities with the neural correlates of recollection and familiarity (177). These authors identified precuneus activity that was common to recollection and familiarity, but found that it was lower for familiarity than for recollection. They also reported that hippocampal and parahippocampal activation was associated with recollection but not familiarity, whereas supplementary motor area activity was more strongly associated with familiarity than with recollection. This suggests there may be some connection between the experience of a flashback and a familiarity judgement.

Memory studies using healthy controls and stimuli of low to average personal significance have found that recollection of the learning event, accompanied by its contextual detail, is normally associated with greater MTL activity. In addition the posterior midline region, medial prefrontal cortex, and posterior parietal cortex contribute to a separate network related to decisions concerning retrieval success (178-180). Although
it is unclear whether activity in these regions reflects episodic long-term memory or simply retrieval confidence, it has been suggested that they could support judgments that a retrieval cue was familiar in the absence of actual recollection. For example, the mnemonic accumulator hypothesis (180) proposed that this region integrates and accumulates available mnemonic evidence relevant to decisions about whether a cue is old versus new.

It is natural to assume that very intensely experienced memories such as flashbacks would exemplify a particularly strong form of recollection, particularly since in the average laboratory study familiarity judgments are usually made in the absence of any detail. Recent evidence indicates, however, that in some instances familiarity judgements may also be very strong (181). Posttraumatic stress disorder is similarly characterized by richly detailed and persistent images that can potentially provide strong mnemonic evidence that a traumatic event has occurred. It is possible that when such images are retrieved they are treated like the cues presented in standard recognition studies, only eliciting a decision about their familiarity rather than any further elaboration and recollection. Consistent with the idea that elaboration of a very negative memory would be aversive, a cardinal symptom of PTSD is avoidance of thoughts and memories concerning the trauma, a tendency that is specifically counteracted in trauma-focused cognitive-behavior therapy (182). We propose therefore that successful identification of flashbacks is based on a strong feeling of familiarity. Full recollection is not required because the detailed sensory dorsal-stream representations enable the individual to identify the images rapidly and with great confidence without further memory search.

This preliminary study was subject to a number of significant limitations. The number of PTSD patients that could be included in this study is relatively small. Possibly due to this small sample size standard statistical analyses used for fMRI data yielded limited results. As we did wish to explore the data to its fullest extent, inclusive masking analyses were used. Although this method did yield results they are in need of
replication. Further, participants had experienced a heterogeneous mix of trauma exposure, and they included a mix of medicated and unmedicated patients. Although it has been argued that neuroimaging studies should include medicated and unmedicated patients in order to reflect the general PTSD population (76), not enough is known about potential interactions of drug and task. There were insufficient numbers in this study to examine drug effects statistically. There was also significant comorbidity in our patient sample, with high levels of depressive symptoms, although our use of a separate depressed group as well as the inclusion of depression as a covariate attempted to control for this.

Not all PTSD participants had flashbacks in the scanner. Limiting our analyses purely to items which were associated with in-scanner flashbacks would have been desirable but there were insufficient data. Our ‘flashback’ category was therefore an amalgam of items which led to flashbacks during the writing of the narrative and items which caused a flashback in the scanner (generally an Own but occasionally an Other item). As mentioned, confirmation of our findings must await replication with a larger sample that can offer more statistical power. Finally, we did not gather data about how participants were responding to flashbacks. There is a considerable literature examining attempts to suppress or reappraise emotional memories which could be applied to this question.

Our results support the idea that the episodic memory system in PTSD is likely intact, and provide preliminary evidence that flashbacks may be associated with increases in activity associated with dorsal visual stream and related areas coupled with decreases in ventral stream activity. Our data also support previous indications that PTSD patients can discriminate the occurrence of flashbacks during written narratives (151, 152). Consistent with the fact that the parts of narratives that involve flashbacks contain more motion words (152), the increased activations in numerous areas of motor cortex may suggest that flashbacks are a form of memory that facilitates action on the environment (such as fight or flight).
The data may also have a bearing on the controversy concerning dual-process models of recognition memory. Although comparisons of recollection and familiarity are often confounded with strong and weak memories respectively (183, 184), these two processes may nevertheless have separate neural substrates (177, 185, 186). Our findings suggest that even intense autobiographical memories, experienced with extreme clarity and vividness, may sometimes demonstrate a neural signature that more closely resembles familiarity than recollection.

Competing interests

The authors declare that they have no conflicts of interest concerning this article.

Acknowledgements

This work was supported by Wellcome Trust grant no. 074145. We would like to thank Natalie Jacoby, Lotte Meteyard, Trishna Patel, and Lorna Stewart for their assistance in conducting this study.

Author contributions

Dr. Rugg, Whalley and Brewin designed the study. Dr. Whalley acquired the data. Dr. Whalley and MSc. Kroes analyzed the data. Dr. Whalley, MSc. Kroes, and Dr. Brewin wrote the article. All authors approved publication of the article.
In Chapter 2, 3, and 4 we have provided results that indicate that PTSD and depression are accompanied by specific structural and functional alterations of the neural network that supports emotions and memory. One question that arises from these findings is whether it would be possible to change functioning of the neural network and possibly alleviate symptomatology. Pharmacological treatments may provide the possibility to affect neural functioning, but other options may also be available. In Chapter 5 we will explore the possibility to lift mood by affecting mood-regulating neurocircuits with a specific food product.
FOOD CAN LIFT MOOD BY AFFECTING MOOD-REGULATING NEUROCIRCUITS.

5.1 Under review as:


5.2 Abstract

It is commonly assumed that food can affect mood. One prevalent notion is that food containing tryptophan increases serotonin levels in the brain and alters neural processing in mood-regulating neurocircuits. However, tryptophan competes with other long-neutral-amino-acids (LNAA) for transport across the blood-brain-barrier, a limitation that can be mitigated by increasing the tryptophan/LNAA ratio. We therefore tested whether a drink with a favourable tryptophan/LNAA ratio improves mood and modulates specific brain processes as assessed by functional magnetic resonance imaging (fMRI). In a double-blind, placebo-controlled crossover study we show that one serving of this drink increases the tryptophan/LNAA ratio in blood plasma, lifts mood in young, healthy women and alters task-specific and resting-state processing in brain regions implicated in mood regulation. Thus, we provide initial, controlled evidence for a food effect on mood, and mood-related neural processing via a serotonergic mechanism.
5.3 Introduction

Primarily, food is required to meet basic nutritional requirements. However, in societies where these requirements appear safely met, attention has shifted towards healthy diets including biologically active components potentially reducing disease risks and optimizing physical as well as mental well-being (187). Stress-resilience and good, stable mood are essential to both health and well-being and there is a common assumption which holds that food can improve mood as readily revealed by entering "Food" and "Mood" into an internet search engine. However, this link is poorly investigated empirically in humans and no proven mechanism that can explain how food affects mood directly is known. Food could potentially affect mood in many ways. For example, food consumption may increase feelings of satiety and vigour through systems regulating energy homeostasis such as glucose, insulin, leptin, and ghrelin, but may also affect hedonic experiences through opioid and dopamine release (188). More directly, biologically active nutritional ingredients could affect neural processes in brain regions central to mood regulation like the prefrontal cortex, cingulate cortex, amygdala, hippocampus and striatum (3, 189). Mood disturbances and attenuations in this neural network have been associated with altered serotonin levels (190-192). Therefore, it has been proposed, based on pharmacological research, that food containing the serotonin precursor tryptophan (Trp) increases serotonin levels in the brain, and modulates processing in the neurocircuit regulating mood in a way that is beneficial for mood (192).

In line with this hypothesis, reducing serotonin through tryptophan depletion is well-known to affect mood negatively (190-192). However, the results of such drug studies do not simply generalize to food. Food is a product or substance that can reasonably be expected to be ingested by humans and normally occurs in the existing food chain (193). Moreover, attempting to lift mood by ingesting food with high tryptophan levels is not as straightforward. Tryptophan competes with other long-neutral-amino-acids (LNAA) for a transport molecule that allows entry into the brain (194), thus limiting the possibility of food to increase brain
tryptophan levels. Although dietary effects on plasma Trp/LNAA ratio have been reported, these, and the resulting behavioural changes, have been modest (192). However, the ingestion of food with optimized Trp/LNAA ratio produced by a hydrolyzed egg protein can result in substantial increases in plasma Trp/LNAA ratio. Preliminary data indicate that such nutrition may improve mood and motor control, and reduce cortisol responses to stress (192), yet the neurobiological mechanisms underlying these effects are currently unknown.

We therefore tested whether a Test-drink with a Trp/LNAA ratio favourable to Trp uptake into the brain leads to mood improvements and changes in a set of neural processes implicated in mood regulation. Given the networks implicated in mood disturbances, of specific interest here are the ventromedial prefrontal cortex (vMPFC), cingulate cortex, striatum, and amygdala. Considering recent suggestions on the functional interaction of serotonin and dopamine (191), we predicted that increasing tonic serotonin levels by heightened availability augments neural responses in areas exhibiting cognitive control over negative affect, down-scales activity in the striatum related to response-priming during reward anticipation, and reduces emotion associated responses in the amygdala. Such a set of functional changes is assumed to result in mood improvements through inhibition of negative thoughts or, in other words, a general cognitive bias towards positive and away from negative emotions.

5.4 Methods and Materials

A brief description of the methods and materials is presented in this section. For a full and detailed description, please refer to the Supplemental Information.

5.4.1 Participants

Thirty-two healthy young women (age-range: 18-39 yr, mean: 22.387, SD: 3.955) were included in the study, and were tested during the second
week of their menstrual cycle. The study was approved by the institutional ethics committee (CMO Regio Arnhem-Nijmegen, The Netherlands) and all subjects provided written informed consent.

5.4.2 Food

Subjects consumed drinks (300ml) containing an equal amount of basis protein, but differential tryptophan and LNAA concentrations. The Control-drink contained 20g casein protein hydrolysate with 0.4g Trp and 10g Long-Neutral-Amino-Acids (LNAA; Valine, Isoleucine, Leucine, Tyrosine, and Phenylalanine). The Test-drink (lumiVidaTM: DSM Nutritional Product, Switzerland) contained a hydrolyzed enzymatic digest of egg white with 0.8g Trp and 4g LNAA. A sweetener (0.10g acesulfame) was added to both drinks to make the formulations more palatable.

5.4.3 Behavioural tasks

Participants were subjected to three behavioural tasks while BOLD fMRI data was acquired. The reward anticipation task (Figure S4.1) is a modified version of the monetary incentive delay task (195) in which cues are presented that signal trials that are either potentially rewarding or non-rewarding. Subjects respond to the appearance of a target by pressing a button, and if pressed fast enough receive a euro on reward trials, but not on non-reward trials. The threat-of-shock task (Figure S4.2) was based on a previous fear-potentiated startle paradigm (196). In the task, cues were presented that either signalled periods during which an electrical shock might be received or signalled periods of safety. A mild electrical shock was delivered only once in the first session. In the Emotion processing paradigm (Figure S4.3) subjects were presented with blocks of trials consisting of angry and fearful stimuli in Emotion blocks, or ellipses (that consisted of scrambles of the same face stimuli) in visuo-motor Control blocks (197). Subjects indicated which of the two images presented at the bottom of the screen matched a target at the top of the screen in term of orientation (Control trial) or emotional expression (Emotion trial). The Baseline checkerboard paradigm (Figure
5.4.4 Blood plasma Trp/LNAA ratio analyses

Blood samples were collected in duplicates of 5 ml vacutainer tubes containing sodium heparine and centrifuged at 1550 g for 5 min at 4°C. The supernatant lithium heparine plasma (750 µl) was mixed with 120 µl sulfasalicyl acid in duplo and stored at -80°C until analysis. Plasma amino acid analysis was conducted with high-pressure liquid chromatography (HPLC), making use of a 2- to 3- µm Bischof Spherisorb ODS II column. The plasma tryptophan to LNAA ratio was calculated by dividing the plasma tryptophan concentration (in µmol/L) by the sum of the other long neutral amino acids, i.e. valine, isoleucine, leucine, tyrosine, and phenylalaline, and averaging over the two samples.

5.4.5 Mood questionnaires

The short-form Profile of Mood States questionnaire (POMS-SF) measured mood along 5 dimensions: Fatigue, Tension, Depression, Anger, and Vigor (198). Total negative mood was calculated from these domains as Depression + Tension + Anger + Fatigue + (24-Vigor) (199). The State-Trait Anxiety Inventory (STAI) measured anxiety in a specific situation (state) or as a more general long lasting quality (trait) (200), where the latter was applied in this study. The Well-being Questionnaire measured current physiological distress. All reported T-tests were tested two-sided.

5.4.6 Functional magnetic resonance image acquisition

MR data was acquired on a 1.5 T MR scanner (Avanto, Siemens, Medical, Erlangen, Germany) equipped with an 8-channel transmit-receiver head coil. Functional images related to the tasks described and resting state were acquired with gradient echo-planar imaging (EPI) sensitive to the blood-oxygenation level dependent (BOLD) response. A structural
scan with a 3D magnetization-prepared rapid gradient echo (MPRAGE) sequence was acquired for normalization procedures.

5.4.7 Functional magnetic resonance data analyses

fMRI data were processed and analysed using the statistical software package SPM5 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). The first 5 EPI volumes were discarded to allow for T1 equilibration. The remaining functional images from each subject were realigned using rigid-body transformation to the mean functional image, and coregistered to the anatomical T1-weighted MR image. Next, images were slice-time corrected to the mean slice, subsequently spatially normalized into a common stereotactic space (MNI 152 T1-template) and resampled to 2 x 2 x 2-mm3 isotropic voxels, and finally smoothed with an isotropic 3D Gaussian kernel of 8 mm FWHM.

The data were modelled voxel-wise, using a general linear model. Condition-specific effects were modelled by trains of stick- or boxcar functions and convolved with the canonical hemodynamic response basis function of SPM5. Additionally, realignment parameters were included to model potential movement artefacts. The data were high-pass filtered (cut-off 128s) to remove low-frequency signal drifts, and a first-order autoregressive model was used to model the remaining serial correlations. Contrast images of from testing parameter estimates encoding condition-specific effects were created for each subject and session. The single-subject contrast images were entered into voxel-wise one-sample t-tests to assess main effects of task (See Supplemental Information for model descriptions), and paired-samples t-tests to assess task by drug interactions, implemented in a second-level random effects analysis. We report regions that survive cluster-level correction for multiple comparisons (family-wise error, FWE) across the whole brain at p<0.05 using an initial height threshold of p < 0.001. See Supplementary Information for model description of each task. Seed-region analysis was used to assess functional connectivity of the vMPFC with the rest of the brain under resting state conditions.
5.4.8 Procedures

5.4.8.1 Visit 1: Screening

During the first visit subjects signed an informed consent form upon which they filled out a demographics and medical history questionnaire, and underwent a physical screening. When subject met all inclusion criteria and none of the exclusion criteria they were randomly assigned to an order of drink administration in a double blind cross-over paradigm (Figure 5.1).

5.4.8.2 Visit 2 & 3: Experimental manipulation

On the day of the second and third visit subjects had a light breakfast free of dairy products or caffeine. Number of days since last menstruation was documented and vital signs were measured. Then, at time-point 1 (T1) POMS and STAI-S questionnaires were administered and a blood sample for Trp/LNAA determination was obtained (Figure 5.1). Subsequently, the
Test- or Control-drink was consumed. Thawing of samples took place in a fridge over night before the start of the study. Thawed samples were consumed within 4 hrs after thawing. Samples were consumed cold. After 90 minutes, at time-point 2 (T2), POMS and STAI questionnaires were again administered and another blood sample for Trp/LNAA determination obtained. Next, subjects entered the MR scanner and tasks were conducted in the following order: Emotional processing task, Resting-state scan, Reward anticipation task, structural T1-weighted scan, Fear processing task, Checkerboard task (Figure S5.1-S5.4). Tasks were presented using Presentation software (Neurobehavioral Systems, Inc., Albany, CA, USA). Upon exiting the scanner (approximately 150 minutes after drink ingestion), at time-point 3 (T3) subjects filled out POMS and STAI questionnaires, a blood sample for Trp/LNAA determination was obtained and a well-being questionnaire administered. Visit 2 followed the same procedure with the exception that the other Test-/Control-drink was consumed and no structural T1-weighted scan was obtained. Following completion of the study subjects were debriefed on the aims and details of the study.

5.5 Results

5.5.1 Test-drink consumption increases blood plasma Trp/LNAA ratio

The test-drink substantially increased plasma Trp/LNAA ratio 90 min (T2) after ingestion, and this remained raised until the end of the study procedures (T3; Figure 5.2A). Testing for the percentage change in Trp/LNAA ratio relative to baseline (T1) at time-point 2 and 3 (ΔT2, ΔT3), a Food (Test-drink, Control-drink) x Time (ΔT2, ΔT3) repeated measure ANOVA revealed a main effect of Food (F₁,₂₉ = 337.35, P < 0.001), a main effect of Time (F₁,₂₉ = 240.59, P < 0.001), and an interaction effect between Food x Time (F₁,₂₉ = 257.02, P < 0.001). Paired T-tests revealed a Trp/LNAA ratio increase following Test-drink consumption relative to Control-drink consumption for both ΔT2 (t(29) = 26.03, P < 0.001; Test-drink Mean: 119.72, s.e.m.: 4.67, Control-drink Mean: -21.15, s.e.m.: 1.60)
Food can lift mood by affecting mood-regulating neurocircuits

Mean plasma Trp/LNAA ratio

![Graph showing the ratio of tryptophan to other long-neutral-amino-acids (Trp/LNAA) in plasma following Test-drink consumption.](image)

**A** Mean plasma Trp/LNAA ratio

- **Test-drink**
- **Control-drink**

**B** Individual % change in plasma Trp/LNAA ratio

![Graph showing individual % change in plasma Trp/LNAA ratio](image)

Figure 5.2 The ratio of tryptophan to other long-neutral-amino-acids (Trp/LNAA) in plasma is increased following Test-drink consumption.

(A) Following Test-drink consumption the mean Trp/LNAA ratio in plasma is more than twofold increased. (B) Change in plasma Trp/LNAA ratio is highly consistent across subjects. For each subject changes in Trp/LNAA ratio relative to time point 1 (ΔT2, ΔT3) are depicted following Test-drink (dotted lines) and Control-drink (solid lines) consumption. Subjects are coded by the same colour lines. T1: prior to ingestion, T2: pre-scan, 90 minutes after ingestion, T3: post-scan, 150 minutes later. Error bars reflect s.e.m. in all figures.

and ΔT3 (t(29) = 9.43, P < 0.001; Test-drink Mean: 43.84, s.e.m.: 6.07, Control-drink Mean: -23.12, s.e.m.: 2.00), and a significant difference between ΔT2 and ΔT3 following Test-drink consumption (t(29) = 16.48, P < 0.001). On average, the consumption of the Test-drink caused a more than
2-fold increase in Trp/LNAA ratio plasma levels. Although some variance existed in the changes induced by the Test-drink, the increase was highly consistent across subjects (Figure 5.2B).

5.5.2 Test-drink consumption lifts mood

Having demonstrated a positive effect of Test-drink on plasma Trp/LNAA ratio, we next determined whether the Test-drink affected mood. In line with our hypothesis, mood improved following Test-drink consumption (Figure 5.3). Assessing the changes in Negative Mood, a Food (Test-drink, Control-drink) x Time (ΔT2, ΔT3) repeated measure ANOVA on a multi-dimensional Negative Mood score based on the short-form of the Profile of Mood States questionnaire (POMS-SF) revealed a main effect of Food ($F_{1,29} = 8.77, P = 0.006$), with no other main effects or interactions (Figure 5.3). Paired T-tests revealed reduced Negative Mood following Test-drink consumption.
consumption for ΔT2 compared to Control-drink consumption (t(29) = -2.13, P < 0.042; Test-drink ΔT2 Mean: -9.35, s.e.m.: 3.22, Control-drink ΔT2 Mean: -1.22, s.e.m.: 2.71), and reduced Negative mood following Test-drink consumption for ΔT3 compared to Control-drink consumption (t(29) = -2.69, P < 0.012; Test-drink ΔT3 Mean: -7.20, s.e.m.: 4.45, Control-drink ΔT3 Mean: 6.65, s.e.m.: 5.37). Comparable, albeit more modest, results were obtained for anxiety, in the sense that state anxiety scores appeared to be reduced after Test-drink consumption (see supplemental Information and Figure S5.5). Thus, food that increased the plasma Trp/LNAA ratio resulted in a clear improvement of mood.

If the drink would have affected the physical state of subjects, this could have contributed indirectly to possible differences in mood and neural processing. However, this interpretation is unlikely as Test-drink consumption did not affect ratings of well-being reflecting measures of physiological distress, such as nausea, dizziness, headache, etc. (Wilcoxon signed ranks test; Z = -1.507, P = 0.132).

### 5.5.3 Test-drink consumption affects neural processing of reward anticipation

In view of the positive effects of Test-drink on serum Trp/LNAA ratios and mood, we next examined Test-drink effects on neuronal responses in the network implicated in mood regulation. We first investigated the influence of drink consumption on reward related neural processing, using a reaction time task that is modulated by the anticipation of reward (195). Consistent with previous studies (195, 201), across all participants this task successfully engaged striatal regions, middle cingulate cortex, supplementary motor areas, and middle frontal gyrus (Figure 5.4A and Table 5.1). Critically, Test-drink consumption reduced reward anticipation-related activity within the dorsal caudate nucleus (Figure 5.4A and Table 1; all statistics are corrected for multiple comparisons across the whole brain (P<0.05)), without affecting speed or accuracy of responses (see supporting online material). The region affected by Test-drink consumption is within those regions showing a main effect of reward...
Figure 5.4 Food consumption affects neural processing in regions involved in mood regulation.

Test-drink consumption reduces responses during reward anticipation in the dorsal caudate nucleus (A), increases responses during threat-of-shock in the middle cingulate cortex (B), and increases connectivity between the vMPFC and lateral PFC under resting-state conditions (C). (A&B): Main effect of task (Grey) and Task x Food interaction (Yellow-Red). (C) Positive (Green) and negative (Blue) correlations with vMPFC activity, and the effect of the Test-drink on connectivity with the vMPFC (Yellow-Red). Bars indicate T-values of main effects, activation cluster significant at p < 0.05 FWE corrected or SVC, are displayed overlaid
5. Food can lift mood by affecting mood-regulating neurocircuits

5.5.4 Test-drink consumption affects neural processing of fear

Next we tested for the effect of Test-drink consumption on neural processing of fear, by using a threat-of-shock procedure. Consistent with previous studies, this task activated anterior insula, cingulate cortex, amygdala, medial frontal gyrus, striatal regions, and cerebellum, i.e. the ‘pain matrix’ (196, 202), (Figure 5.4B and Table 5.2). Crucially, greater activity in the middle cingulate gyrus during threat periods was detected following Test-drink consumption (Figure 5.4B and Table 5.2). Again, this region was within the network involved in fear processing (Figure 5.4B). Thus, Test-drink consumption influenced neural processing associated with fear.

5.5.5 Test-drink consumption affects neural resting-state connectivity

Finally, as connectivity patterns with the vMPFC under resting-state conditions have been implicated in mood regulation (26), we examined the influence of Test-drink consumption on resting-state connectivity with this region. Functional connectivity analysis revealed a network of regions displaying positive correlations with blood oxygenation level dependent (BOLD) signal fluctuations of the vMPFC known as the ‘default mode network’, which included the middle orbitofrontal cortex, posterior- and anterior- cingulate cortex, insula, inferior and medial parietal cortex, superior frontal cortex, cerebellum, amygdala, hippocampus, and striatum. Negative correlations were observed for the precentral gyrus, inferior parietal cortex, middle occipital cortex, inferior- and superior-temporal cortices (Figure 5.4C and Table 5.3). Critically, greater positive correlation with the BOLD signal fluctuation of the vMPFC following
Test-drink consumption was detected in the inferior-, and middle- frontal gyrus (Figure 5.4C and Table 5.3). Hence, Test-drink consumption affected resting-state connectivity of the vMPFC with the lateral PFC. Ingestion of the Test-drink thus influenced functional connectivity between brain regions associated with mood regulation.

5.5.6 Test-drink consumption did not affect neural responses on the emotion processing task and control task

Test-drink consumption was not found to affect neural responses on the emotion processing task and baseline checkerboard task (Figure S5.6a and Table 5.4; Figure S5.6b and Table 5.5, respectively). For result and discussion see Supplemental Information.

5.6 Discussion

Results of blood plasma, subjective mood questionnaires, and brain activity collectively suggest that food targeted at raising Trp/LNAA levels can improve mood and affect neural processing in neurocircuits implicated in mood regulation. We thus present initial empirical support for the common assumption that food can improve mood by affecting brain function via a serotonin mediated mechanism.

These effects occurred regardless of subjects ability to report the type of drink consumed, and cannot be explained by changes in physical well-being following drink consumption. The effects were obtained with a serving size that does not exceed normal amount of food consumption, yet increases in Trp/LNAA ratios were detected that are as large as observed following synthetic tryptophan administration (203). As this large change in Trp/LNAA ratio would increase the probability of tryptophan to be transported into the brain (194), these changes presumably resulted in increased central serotonin levels (192). Further, the effects were obtained in a healthy subject population; effects of food on mood may be even greater in populations experiencing reduced mood
(192). As such we infer that these results are relevant to normal human food consumption.

Food modulated brain activity on a specific set of tasks probing the neural circuitry of mood regulation. Following Test-drink consumption we found reduced activity in the dorsal caudate nucleus in a reward anticipation task. This region has been associated with higher order cognitive processes (204), and its connectivity pattern allows for an integration of initial emotional responses with cognitive information and motor output (201). In support of this suggestion, the dorsal striatum has been associated with the monitoring of value over long delays, a process improved by higher serotonin levels and reflected in plasma tryptophan levels (190). Thus, reduced dorsal caudate signals by increased serotonin levels may echo a more stable maintenance of goals over longer periods of time, which may go along with increased concentration and reduced distractibility (191).

Similarly, we demonstrate that Test-drink consumption resulted in an increase of activity in the middle cingulate cortex during fear processing. This region has been associated with the integration of emotional responses with cognitive information and motor output (205), and often been associated with the cognitive control over emotional responses (14). To our knowledge we provide the first evidence for the involvement of serotonin in the anticipation of aversive outcomes in a threat-of shock task. Our findings fit with reports of increased activity in the cingulate cortex following SSRI treatment of patients with mood disorders (3). Increased activation in the cingulate area reported here has been associated with lower neuroticism scores (196), found to correlate with a drop in negative affect due to reappraisal (14), decreased amygdala activity and physiological arousal responses (14). Given these reports, it is interesting to note that we found reduced amygdala activity with increases in middle cingulate activity under resting state conditions (see supplemental Information, Figure S5.4c, and Table 5.6). Hence, Test-drink consumption altered neural processing of fear in a network associated with the regulation of negative mood.
Finally, Test-drink consumption increased resting-state connectivity between the vMPFC and lateral PFC. The involvement of the vMPFC in control over emotional responses, fear and anxiety is well described (11), and connectivity patterns of the vMPFC under resting-state conditions have been implicated in mood regulation (26). The lateral PFC has been suggested to provide cognitive control over attention directed to emotional processing and in doing so regulates the vMPFC’s inhibitory control over the amygdala and emotional responses (15). Additionally, the lateral PFC has been associated with reappraisal of emotional stimuli, leading to reduced amygdala responses and decreased negative affect (14). Thus, Test-drink consumption lead to changes in a network involved in cognitive control over affect centred on the vMPFC.

Although the diverse behavioural tests elicited commonly expected behaviour amongst the subjects, the Test-drink did not affect behaviour. Additionally, the drink had no effects on a task probing general neural processing. Hence, food induced differences in behaviour or general neural functioning are unlikely to explain reported brain results.

These findings generate specific questions for future research: May other food substances also affect mood, possibly through other mechanisms than serotonin? May other parts of the neural network involved in mood regulation also be modulated by the food tested, and do these effects extend beyond cognitive control over mood? And what determines the individual differences in food induced mood improvements and brain activity? Regardless of these questions, the results of this study show that food increasing plasma Trp/LNAA ratios improves mood and alters associated brain function. In conclusion, we provide a mechanistic account for food in controlling mood by affecting brain function via a serotonergic mechanism. These results provide support to reducing disease risks, improving stress resilience, and optimizing physical and mental well-being through specialized diets.

**Competing interests**
This work was partially supported by a research grant from DSM Nutritional Products Ltd to G. Fernández. The tested drinks were provided by DSM Nutritional Products Ltd. J. Wittwer, J. Kloek, and H. Mohajeri are employed by DSM Nutritional Products LTD, but had no role in the acquisition or analyses of the data.

Acknowledgements

This work was supported by research grants from DSM Nutritional Products Ltd and the Netherlands Organisation for Scientific Research (NWO) to G. Fernández. All authors contributed to the design of the study. M. Kroes conducted the study. M. Kroes and G. van Wingen analyzed the data. M. Kroes, G. van Wingen, and G. Fernández wrote the manuscript. All authors agreed upon the manuscript. We would like to thank Roshan Cools, Hanneke den Ouden, and Bryan Strange for useful discussions of this work, and Sabine Kooijman for study coordination support.

Author contributions

All authors contributed to the design of the study. MSc Kroes acquired and analyzed the data. MSc Kroes and Dr. Fernández wrote the manuscript. All authors reviewed the manuscript.
Interlude B.

In Chapter 2, 3, and 4 we have provided results that indicate that PTSD and depression are accompanied by specific structural and functional alterations of the neural network that supports emotions and memory. In Chapter 5 we have shown that food targeted at raising Trp/LNAA levels can improve mood and affect neural processing in neurocircuits implicated in mood regulation. Although food or pharmacological treatment can improve mood and relieve symptoms of anxiety disorders it does not affect the root cause of such disorders. Considering that memory for emotional experiences contribute to the aetiology and persistence of memory one strategy would be to alter memory for such emotional experiences. In order to do so we need to understand under which situations memories are sensitive to alterations. In Chapter 6 we will explore the dynamic neural systems that enable flexible and adaptive memories. Insight into these dynamic systems will provide us with target mechanisms to alter memory for emotional experiences.
6  DYNAMIC NEURAL SYSTEMS ENABLE ADAPTIVE, FLEXIBLE MEMORIES.

6.1  Published as:


6.2  Abstract

Almost all studies on memory formation have implicitly put forward a rather static view on memory. However, memories are not stable but sensitive to changes over time. Here we argue that memory alterations arise from the inherent predictive function of memory. Within this framework, we draw an analogy between the lateral temporal - lateral prefrontal system that supports prediction based on simple stimulus-response associations and propose that a similar system centring on the hippocampus and medial prefrontal cortex (mPFC) exists for complex episodic memories. We consider the hippocampus to be elementary for regularity detection and the mPFC for regularity storage together with response options, which form the basis of abstract knowledge. As such, abstract knowledge can come to guide behaviour in novel situations that only share partial overlap with episodic experiences that have given rise to the formation of abstract knowledge. Furthermore, we suggest that systems consolidation and sleep contribute to the formation of abstract knowledge, and that abstract knowledge can function as pre-existing schemas to the encoding of novel memories. Finally, we discuss that reconsolidation supports the updating of memories to optimize prediction. We accentuate that memory formation requires dynamic interactions between brain regions, and that rapid formation of detailed memories depends on synaptic weight changes, whereas rather stable abstract knowledge is supported by cortico-cortical rewiring. Together, we attempt explaining that apparent memory alterations and distortions are adaptive.
6.3 Introduction

We are able to maintain memories over long periods of time. These memories may even originate in childhood, and often involve great detail. Thinking about such detailed childhood memories the first author of this review can remember a very specific occurrence involving a theme-park, knights, and his grandfather (see Figure 6.1A). This ability to retain information over extended amounts of time has fascinated scholars for centuries. The investigation of the biological substrates of memory (‘the search for the engram’) has lead to amazing discoveries on how information is encoded, stabilized, and becomes integrated in the architecture of the nervous system over time. Over the last 60 years, the search for the engram has zoomed in on the neurobiological basis of memory on an ever decreasing spatial scale. Starting from the scale of the whole brain (206), down to a focus on the hippocampus (207), and individual cells in the hippocampus (208), to the scale of synapses (209, 210), single proteins (211, 212), and methyl groups (213). Additionally, research has shown that memories are initially labile and sensitive to disruption but become stabilized over time, a process known as consolidation (30). This body of neuroscientific work implicitly suggests that stable engrams are stored in the brain enabling memories to persist for prolonged periods of time. Yet, it has long been recognized that memories are not accurate representations of past experiences and are sensitive to changes over time. In this review we will take memory alterations as an example, as we think it reveals important characteristics of the dynamic nature of memory, and the neural systems that support memory formation.

In trying to reconstruct the before mentioned episode from his life, the first author asked his brother to describe the event as he remembered it (Figure 6.1B). His brother’s description of the episode is largely similar in that he describes the same people, location, and objects. Yet, he remembers aspects that the first author has forgotten, such as them having practiced the show. Even more striking is the fact that there are various discrepancies. For example, the brother remembers himself as
6. Dynamic neural systems enable adaptive, flexible memories

As children, my brother and I had been to a medieval theme-park and had seen a knight's tournament which had included the ‘evil’ black-capped knight, who, importantly for the story here, had repeatedly spit on the ground from his horse. In the following weeks, we had re-enacted this tournament in our back yard. This is what I remember from ‘the incident’:

“We had been to the ‘Land van ooit’ (’the land of once’). My brother and I were displaying the knight-fight between the ‘red’ and ‘black’ knight to our grandfather, aunt, uncle, mom and dad. They were sitting in the backyard in white plastic garden chairs... The area in the backyard runs slightly uphill... I must have been the black knight and must therefore have been wearing the black cape. We were sword fighting... so we must have been carrying wooden swords... I don't actually remember the spitting incident itself, but I must have paraded by as if on a horse... and while passing the audience spat grandpa in the face. I’m not quite sure what happened next, but it probably caused some turmoil...”

As children, my brother and I had been to a medieval theme-park and had seen a knight’s tournament which had included the ‘evil’ black-capped knight, who, importantly for the story here, had repeatedly spit on the ground from his horse. In the following weeks, we had re-enacted this tournament in our back yard. This is what I remember from ‘the incident’:

“Marijn is the black knight, I am the blue knight, Aunt, uncle, grandpa, mom, and dad were sitting in folding chairs in front of a sitting area. The grass is very green. We had practiced the ‘show’. I was going to win (at least that was the deal), because the black knight was the bad guy. I don't actually remember the spitting incident itself, but I must have paraded by as if on a horse... and while passing the audience spat grandpa in the face. I’m not quite sure what happened next, but it probably caused some turmoil...”

“A: The first author’s memory for a specific episodic experience. B: The first author’s brother’s memory for this episodic experience. C: A photograph from the first author’s family album that relates to his memory for this episodic experience. D: Memory for this episodic experience of the first author’s father. Similarities (green), omissions (yellow), and discrepancies (red) detected between the first author’s memory and that of his brother and father.

In trying to resolve the differences between their memories, the first author went looking for a photo taken during that day he remembered being in a family photo album. Although he did find the photo he remembered, it was not taken on the day of ‘the incident’, and featured a neighbour as well (Figure 6.1C). Many of the details both his brother and the first author remember are present in this photo; the wooden swords, the back yard, a grown-up watching the show, etc. This example clearly shows that memories are not veridical representations of experience, and may become distorted over time, for example through integration of information from other sources and experiences.
Such memory inaccuracies are well known in psychological literature. Generally, although we are often quite confident about the correctness of our memories, these may be highly inaccurate at the same time or even false (214, 215). Hence, our memory is not a very reliable source for accurate reconstruction of past episodes. As an unreliable memory system may have serious consequences (216, 217) this raises the question what could possibly be the use of an inaccurate memory system? A hint at why memory is not an exact reproduction of experiences comes from a description of a man named Solomon Shereshevsky who had an incredible ability to accurately reproduce experiences from memory, yet experienced great difficulty detecting continuity in the world, abstracting knowledge across experiences, and predicting future events (218, 219). The ability of absolute memory seems to come at the cost of abstraction, which is necessary to form knowledge that is no longer specific to a given experience, but can be applied to novel situations. From this we can infer that the function of memory is not to reproduce accurately past experiences, but to abstract knowledge from experiences to predict the outcome of future events given the current situation. The need for abstraction from episodic experiences is twofold. First, an overly specific memory system may hinder generalization across experiences, as is the case for Shereshevsky. Second, a memory system that stores every experience as a veridical memory trace may run into a capacity problem (220). Yet, abstraction from episodic experiences and retention of specificity may not be mutually exclusive, for example details of a highly salient experience may be preferentially retained.

An intuitively appealing assumption thus is that memory serves the function of predicting the future (221). Taking this one step further, it has been suggested that an organism has to use sensory input to plan a future motor output, and this sensori-motor predictive function may be the sole reason for the existence of a nervous system (222). More specifically, prediction has been proposed to be an elementary function that occurs at each level of the neural computational hierarchy. Within the neural hierarchy ‘higher’ areas are thought to form predictions over incoming information from ‘lower’ areas. The difference between
predictions and incoming information (prediction error) is thought to be minimized through recurrent interactions within the system (223-227). Perception arises when the neural activity settles into a stable state in which the prediction error throughout the system is minimal. Similarly, actions are executed to match predictions, and again the difference between perceptual predictions and action outcomes are continuously tried to be minimized in a dynamic iterative cycle (224). Such a system requires a current state to be compared to a previous state, which has to be maintained online or stored in memory. The basic function of memory to serve prediction may thus be elementary, and complex episodic memory may have arisen from this function in phylogeny. Although this view is in need of further scientific support, we consider that viewing memory as having a predictive function can provide insight into the neural mechanism that support memory formation. Throughout this review we will highlight studies in which behaviour is guided by predictions that arise from memory for episodic experiences.

We adopt the premises that memory serves a predictive function, and that memory arises within a neural sensori-motor hierarchy. By taking autobiographical memory as an example we extend a previous proposal on predictive coding and episodic memory (228). This ‘predictive interactive multiple memory systems’ (PIMMS) model is a more in depth description of episodic memory as a predictive system, and shares with our view that multiple interactive memory systems exist in the brain. The PIMMS model focuses on differences between familiarity and recollective experiences, and on interactions between the hippocampus, entorhinal cortex, and the ventral visual system. Importantly, the formation of abstract knowledge in the PIMMS model is limited to predictive memory for item categories. We focus on more conceptual, abstract knowledge formation from episodic experiences and envision a critical role for the medial prefrontal cortex (mPFC) in this process. Although the mPFC has often been implicated in memory formation few theories on its role in memory exist. As such we provide a novel theory on the role of the mPFC in memory. The second section of this review focuses on the extraction of regularities across episodic experiences that are critical for a predictive
memory system, and we propose that the hippocampus is particularly suited to perform this function. In the third section we discuss how extracted regularities may be integrated with behavioural output. We will draw an analogy with simple stimulus-response learning that is supported by the lateral prefrontal cortex (lPFC) suggesting that the mPFC plays a critical role in linking regularities extracted across complex episodic experiences to behavioural output, and in doing so forms the basis of abstract knowledge in interaction with the hippocampus. In section four and five studies on systems consolidation and sleep are discussed, the interactions between the hippocampus and mPFC further explored, and we propose that systems consolidation and sleep aid the formation of abstract knowledge and the optimization of prediction from episodic memory. In section six we discuss relatively recent studies on schema memory and indicate that abstract knowledge comes to guide the integration of novel information again through interactions between the hippocampus and mPFC. In the final section we discuss studies on reconsolidation and suggest that memory alterations following memory reactivation functions to update predictions. Throughout this review we suggest that rapid memory formation and memory for details depend on synaptic weight changes, but that abstract knowledge formation is accompanied by changes in neural wiring. Additionally, we will highlight that memory formation requires recurrent and dynamic interactions between brain regions. Taken together this review provides insight into the neural mechanisms that might give rise to memory alterations, emphasizes the dynamic nature of memory, and proposes a neuroscientific and neuroanatomical model for an adaptive flexible episodic memory system.

Critically, a memory system serving prediction requires the integration of information across distinct experiences and the extraction of regularities. In the following section of this review we will propose that the hippocampus is particularly suited to extract regularities across episodic experiences.
6.4 Extraction of regularities

In the movie ‘Groundhog Day’ Bill Murray gets stuck in time, endlessly repeating the same day. As he is confronted with the exact same situation over and over again, he is able to adjust his behaviour based on previous experiences of the exact same situation. In reality veridical re-representation of the exact same situation never occurs, details of the experience are always different. In order to deal with this variability we need to base our predictions of the outcome of a situation on similarities between the current situation and previous experiences. Thus, a predictive memory system requires the rapid storage of experiences, the detection of regularities across experiences that are independent of the variance amongst experience, and the storage of extracted regularities forming the basis of abstract knowledge allowing predictions that can come to guide behaviour even when specific details of the current experience differ. Many studies have focussed on predictive coding for simple stimulus-response associations and have shown that interactions between the lateral temporal and lateral PFC critically support this function. Only recently have studies started to indicate that the ability of predictive coding for complex episodic memories arise through dynamic interactions between the hippocampus and neocortex, with a specific role for the mPFC in integrating abstract representations across modalities with behavioural output. In drawing an analogy between the lateral temporal-lateral PFC system that supports simple stimulus-response associations, and a hippocampus-mPFC centred network that supports complex episodic memory we aim to put forward a role for the mPFC in memory formation and extend predictive coding to apply to episodic and autobiographical memory. Considering this framework, it might be that the distortions observed for the memory of ‘the incident’ have arisen from the inherent sensori-motor predictive function of memory. Here we will first describe the neural systems that support the detection of regularities and the formation of abstract knowledge from episodic experience where we envision a critical role for the hippocampus.
Half a century of neuropsychological research has led to the consensus that multiple memory systems exist that are supported by distinct neural structures. Furthermore, memory functions are present in all processing hierarchies of the brain, and can operate both independently and in interaction (207, 228-233). A specific role for the hippocampus at the centre of an episodic memory system (memory for specific experiences) is envisioned, whereas the neocortex is implicated in the storage of memory for facts and general knowledge (semantic memory). The acquisition of factual knowledge may arise from the detection of regularities across episodic experiences. This implicates that the hippocampus is also important for the acquisition of factual knowledge. However, the role of the hippocampus in semantic memory is highly debated (234). Although amnesic patients with hippocampal damage are unable to acquire new episodic memory, several scholars have reported that these patients can acquire semantic knowledge during development (235), or over many learning sessions (235-239). Additionally, their memories are extremely rigid and not easily applied to novel settings (237, 240, 241). The role of the hippocampus in semantic memory remains a topic of debate. Nevertheless, it appears that the hippocampal episodic memory capabilities give rise to formation of new abstract knowledge, which in turn can become stored within the neocortex and retrieved relatively independent of the hippocampus (231, 238, 239). This raises the question with which specific characteristics the hippocampus is endowed that bestows upon it such a central role in episodic memory and the formation of abstract memory representations.

We consider several characteristics of the hippocampus to make it an ideal candidate structure to detect and extract regularities across episodic experiences and, hence, support the formation of abstract knowledge. These are the external connectivity of the hippocampus with the rest of the brain, the internal connectivity of the hippocampus, and a high degree of associative plasticity. Before describing these characteristics we wish to clarify several assumptions. First, a consequence of the predictive coding framework is that memory encoding and retrieval can no longer be conceived of as separate processes, but
Dynamic neural systems enable adaptive, flexible memories that occur through recurrent interactions (228, 242). This implicates that brain areas or neural functions (e.g. pattern completion and pattern separation) that have classically been associated with either memory retrieval or memory encoding now interact to support both functions. Second, we suggest that the hippocampus is able to detect regularities across episodic experiences and is able to orthogonalize these regularities from details of an experience. The hippocampus may then drive the neocortex to store these regularities more permanently. Critically, the formation of abstract knowledge then arises through interactions between these regions. We will return to this issue in the next sections, particularly when discussing studies on systems consolidation and sleep.

The first important characteristic of the hippocampus to the detection and extraction of regularities from episodic experiences is its extrinsic connectivity. Information entering the brain is transmitted through a hierarchy of unimodal and polymodal cortical areas, integrating and associating different features into increasingly more complex, yet unitary, representations (243, 244). Although memory functions exist at each stage of this processing hierarchy (244), such memory is specific to the modality of processing (245). The complexity of episodic memory requires the integration and association of information from different processing modalities and is considered to exist within a widespread distributed cortical processing network (245-248). However, within the neocortex only a small portion of all possible connections between neurons physically exist and physical connections between distant cortical regions is limited (249, 250). Therefore, cortical connectivity might be too sparse to allow associations between all elements of information linked to one episode and processed in distributed neocortical areas (251). In contrast, the hippocampus can be conceived of as a supramodal processing area as it receives convergent input from all polymodal areas and many subcortical regions, and projects back to most regions from which it receives input (243, 252). The connectivity pattern of the hippocampus thus allows integration of information from distributed neo- and subcortical regions that is critical for the detection of regularities across episodic experiences.
A second important characteristic of the hippocampus is its intrinsic connectivity that forms a random network allowing arbitrary associations between details processed in distributed cortical areas (252-254). Additionally, the intrinsic architecture gives rise to the information processing functions - pattern completion and pattern separation (248, 251, 255, 256) - enabling the creation of distinguishable memory representations, the completion of memory representations from partial input, and comparison of different memory representations to each other, detecting differences and commonalities between them (248, 251, 257, 258). As mentioned above, we propose that pattern separation and pattern completion occur in reiterative cycles whereby incoming information can be rapidly stored, compared to retrieved memories, regularities between these representations can be detected, and novel memory representations for regularities can be formed.

The extrinsic and intrinsic connectivity of the hippocampus enable the detection of regularities. Electrophysiological studies support such a role for the hippocampus. So called 'place cells' exhibit high firing rates at specific locations in the environment forming an allocentric spatial map of the environment (208), and have been proposed to be the neural basis for cognitive maps allowing the distinction between and storage of memories (259). However, as place fields of hippocampal place cells prove not to be as stable as initially thought, but remap relative to sensory cues, and show differential responses dependent on task demands, rewards, delays, and behavioural sequences, it has alternatively been suggested that place cells do not code for specific locations, but for relational representations or conjunctions between stimuli and behavioural options that will vary with the position of an animal in space, effectively forming a 'memory space' (260). The hippocampal system is then able to identify common elements across episodes, and these commonalties are thought to link episodes, composing a network of memories that allows one to bridge between experiences. The construction of multi-level networks of memories can be viewed as the emergence of a 'semantic space' forming abstract knowledge from many overlapping episodic representations (260-262). These functions can arise though attractor dynamics in the
neural processing hierarchy (263, 264). In neural networks, attractor states are conceived to be relatively stable activity patterns supported by previous plasticity changes towards which new similar activity patterns will converge. Additionally, sequences of attractor states can form temporal episodes (265). Hippocampal conjunction or regularity cells can form attractor states or 'nodes' that bring together information from distributed neocortical areas and across experiences (266). Additionally, the existence of attractor states in other brain regions would support dynamic interactions between them. In line with this suggestion, attractor dynamics in the hippocampus (267, 268), neocortex (269, 270), and other brain areas (271) have been reported. The existence of attractor dynamics highlight that episodic memory is supported by bidirectional dynamic interactions between brain regions. The idea that other interactions between brain regions also follow attractor dynamics is appealing. Given the scope of this review, we will, however, only discuss attractor dynamics when studies have explicitly investigated this issue. In conclusion, electrophysiological studies appear to support a role for hippocampal cells in the detection and storage of regularities that gives rise to abstract knowledge, possibly by forming attractor states to incoming information.

A third important characteristic of the hippocampus in the detection of regularities is its high degree of associative plasticity. Many episodic experiences only occur once, but their significance may only become apparent afterwards. Therefore, it is important to store episodic experiences as they occur, i.e. in real time. Such 'automatic recording of experience' is thought to be supported by the plasticity of the hippocampal NMDA dependent long-term potentiation system (272), enabling rapid weight changes through the strengthening of existing synapses (249). The rapid storage of novel experience enables the detection of regularities between a current environmental situation and existing memory representations.

However, an automatic recording system places high demands on the storage capacity of the hippocampus (251), and would rapidly run out of storage space for new experiences (272-274). Therefore, memory
selection procedures are envisioned. For instance, familiar information may not reach the hippocampus (242) or hippocampal memory traces may decay rapidly and only a proportion of traces are thus permanently selected for stabilization (272, 275). It would be beneficial if environmental regularities detected amongst hippocampal associations were eventually stored within a system less sensitive to interference and decay. Computational modelling studies have shown that a memory system storing environmental regularities should have a slow learning rate as to prevent over-writing of memories (276) and requires gradual, interleaved and repeated exposure to previously stored information and novel information (261, 262, 276). This has lead to the proposal that the hippocampus rapidly forms memory representations and subsequently replays these to the neocortex, enabling repeated exposure to, and integration of, previously stored and novel memory representations initiating a gradual development of connectivity between cortical modules (272, 276). As mentioned, cortical connectivity is sparse. Therefore, the formation of inter-cortical memory networks might require the creation of novel connections, not the strengthening of existing connections. Such re-wiring is more resistant to interference than weight changes (249) and thus provides a more robust memory representation. However, rewiring is resource costly and therefore only worthwhile if the induced changes represent regularities across experience that should, and do not, rapidly change over time. Through such a process, environmental regularities eventually become stored within inter-cortical connectivity, possibly independent of the hippocampus, and form the substrates of abstract knowledge. We will return to this matter in section four when discussing studies on systems consolidation. Given the sparseness of cortico-cortical connectivity it is unlikely that cortical memory representations are able to link all arbitrary details of a given experience, and might thus be fairly rigid and come at a loss of specific details. However, we consider that neocortical regularity representations interact with hippocampal processing, allowing the flexible application of regularity representations.

In summary, specific neural systems support the detection of regularities and the formation of abstract knowledge. Neuropsychological literature
distinguishes between hippocampus dependent episodic memory and semantic memory that is supported by the neocortex. We have argued that episodic memory gives rise to abstract knowledge (that is akin to semantic memory). Specific characteristics of the hippocampus are considered to support the detection of regularities. Hippocampal connectivity allows the integration, separation, and comparison of information from distributed brain regions and across experiences. In line with this idea, hippocampal cells appear to encode conjunctions and regularities, which can bring together associated information within and between episodic experiences. Further, the hippocampus allows the rapid encoding of information through weight changes supporting the detection of regularities between memories and new experiences. Finally, regularity representations are proposed to become stored in the neocortex through re-wiring, forming a stable substrate of knowledge yet losing detail and flexibility. We thus propose that the hippocampus detects regularities and drives the formation of abstract knowledge in the neocortex. Alternatively, the hippocampus may support the formation of multiple episodic memory traces within the neocortex upon which the neocortex performs abstraction across representations. This matter remains unresolved to date. However, we favour the former explanation and will highlight studies in support of this hypothesis throughout this review.

We have taken the point of view that memory serves a predictive function, and arises within a neural sensori-motor hierarchy, as we think this can provide insight into the neural mechanisms that supports formation of flexible memories. Within this framework an open question that remains is how extracted regularities are integrated with behavioural output to support sensori-motor prediction. As the hippocampus does not directly connect to the motor cortex, interactions with other areas are necessary to link extracted regularities from episodic experiences to motor outputs. In the next section we will draw an analogy between the lateral prefrontal cortex (lPFC) and medial prefrontal cortex (mPFC) to suggest that the mPFC is a critical region for the dynamic nature of memory. Further we will propose that the mPFC interacts with the hippocampus to form abstract knowledge and support sensori-motor prediction.
6.5 From regularities to response

In order to understand how to map from episodic memories to behavioural output we will draw an analogy with the neural systems that support rule learning for simple stimulus-response mapping. Rules are arbitrary associations between disparate but behaviourally related information (277), where a particular stimulus instructs a fixed response. Importantly, rule learning has been shown to occur through dynamic interactions in a neural hierarchy of brain regions that process stimulus information, and regions that process motor outputs. The lPFC shows a high degree of connectivity with the lateral temporal, entorhinal, and parahippocampal cortices (278), areas that process and support recognition memory for item and category information (279, 280), but notably not the hippocampus (281). Single cell recordings have revealed that neurons in the lPFC form integrated representations of cues, categories, rewards, task demands, and behavioural responses, and increase activity with increasing rule familiarity (282, 283). Additionally, a graded hierarchy exists with regard to the integration of stimulus, response, and rule representations. Neurons in the inferior and lateral temporal cortices show the strongest representations of stimuli, neurons in the lPFC integrate stimuli and motor output options, and neurons in the premotor cortex exhibit the strongest representation of rules and motor output options (282, 284, 285). Finally, a fast rule learning system has been observed in the basal ganglia, whereas the lPFC reflects a slower learning system (284, 286). Human imaging studies have revealed similar hierarchical interactions between the lPFC, lateral-, and inferior temporal cortices, and basal ganglia (287, 288), where increasingly higher order interactions between stimulus-response mapping require increasingly more anterior prefrontal areas (289, 290). Furthermore, the dependency of simple rule learning on cortical areas increases with experience and time through dynamic interactions between brain regions (291, 292).

The animal and human findings on rule learning mimic suggestions from computational modelling that have shown that within a neural hierarchy the higher levels are important for maintaining contextual information,
6. Dynamic neural systems enable adaptive, flexible memories

Communicating contextual information to lower processing levels, and in doing so guide neural processing at lower levels relative to behavioural response options (293-295). Thus, the PFC is considered to hold online behavioural options and dynamically form attractor networks towards which processing at lower levels settle. In line with this suggestion it has been shown that activity in the medial premotor cortex that reflects choice options, can be maintained by few neurons over a delay period, and these neurons can support the recovery of the other neurons within the same attractor network at the time when a behavioural response is required (270).

Together these findings indicate that the lPFC plays a cardinal role in simple rule learning. The lPFC interacts with regions that represent single item and category information, e.g. the lateral/inferior temporal cortex, and regions representing motor outputs, e.g. PMC, and appear to dynamically form attractor states within a graded hierarchy. Although simple stimulus-response rules can be applied to novel stimuli (296), such rules are generally rigid and do not efficiently generalize to novel situations when feature overlap is limited (297, 298). Thus, for predictions based on regularities detected amongst complex multi-feature configurations of episodic experiences a different system appears

Figure 6.2: Analogy between two memory systems supporting prediction

A: The lateral temporal-lateral prefrontal system supports simple rule learning. B: An analogous hippocampal-mPFC system hypothesized to support a predictive memory system for abstract rules based on complex episodic experiences. A graded hierarchy is proposed to exist for both systems through which stimuli, rules, and behavioural output are dynamically integrated.
necessary. As the neural system that supports simple rule learning complies with the neural hierarchy of sensori-motor prediction, we reason that a similar hierarchy may give rise to a system that supports prediction based on complex episodic experiences. We propose that such a system arises through interactions between the hippocampus and medial prefrontal cortex (mPFC) analogous to the lateral temporal-IPFC interactions described for simple rule learning (Figure 6.2). As discussed, predictions from episodic experiences require responses based on detected regularities that are invariant of specific stimuli. Considering that we do not wish to introduce new terminology to the field, we will refer to response prediction learning from complex episodic experiences as abstract rule learning.

Together with the hippocampus, the mPFC has often been implicated in episodic memory in both humans and animals (299-301). Importantly, in monkeys abstract rule learning that is independent of particular stimuli depends on the mPFC (302), where the principle sulcus was found to support working memory maintenance, the orbitofrontal cortex was implicated in the updating of rules based on rewards, and the anterior cingulate cortex played a role in representing highly integrated and context dependent abstract rule representations. In rodents, neurons in the infralimbic and prelimbic cortices, regions thought to be the homologue of the primate ventromedial prefrontal cortex and anterior cingulate cortex (303, 304), change responses when switching between behavioural strategies (305) that depend on either the hippocampus or basal ganglia (306), but not when a consistent strategy is followed and the behavioural output is different (305), indicating that infralimbic neurons initiate strategy switches whereas prelimbic neurons enforce execution of current strategies. Abstract rule representations of episodic experiences in the mPFC thus appear to interact with other brain regions, including the hippocampus and basal ganglia, to drive optimal. Two recent studies provide further inside into the dynamic interactions between the hippocampus and mPFC that give rise to prediction based on abstract rules detected across episodic experiences.
In a human fMRI study, subjects had to predict 'the weather' based on visual stimuli (297). Subjects could learn direct stimulus configurations to response mappings, which was associated with increased activation of the vMPFC, posterior cingulate cortex (PCC), parahippocampal cortex, ventral striatum, and amygdala. Subjects could also learn higher-order conceptual knowledge through the detection of regularities across experiences, which was reflected in activation of the hippocampus, vMPFC, and PCC, as well as greater functional coupling between the hippocampus and vMPFC (297). Interestingly, the parahippocampus was not found to reflect higher-order conceptual knowledge, which indicates that this area may be more concerned with specific associative stimulus pairings. Furthermore, when confronted with novel stimuli but similar underlying task rules, hippocampus, vMPFC, and PCC activity correlated with correct performance. Importantly, hippocampal activity during learning of the task rules with the original stimuli was correlated with performance when faced with these novel stimulus configurations (297). These findings clearly indicate that interactions between the hippocampus and mPFC contribute to the formation of abstract knowledge, and that this process is already initiated during learning. Interestingly, this suggests that the role of the hippocampus in abstract knowledge formation may not be a 'teacher' replaying individual memories to the neocortex so that the latter can extract regularities (276), but indicates that the hippocampus contributes to the formation of abstract knowledge by detecting regularities across episodic experiences (260). Further insight into the dynamic interactions between the hippocampus and mPFC during abstract rule learning comes from multi-unit recordings. Hippocampal and mPFC activity has been found to phase synchronize at theta frequencies, and this theta coupling has been associated with learning and memory performance (307). Neural oscillations are thought to coordinate interactions between information represented within the hippocampus and mPFC, possibly through the formation and dissipation of functional networks in the form of synchronized cell assemblies (246, 265, 308). When rats learn to shift flexibly between two rules in a spatial memory task (reflecting episodic memory, not simple stimulus-response learning), theta coherence between the hippocampus and mPFC
varies with learning and is accompanied by the formation of neural cell assemblies that are associated with memory (309). Specifically, theta frequency coherence between the hippocampus and mPFC increased at the choice point of a maze, i.e. the location at which the animal had to apply a learned abstract rule in order to predict in which direction to travel to obtain a reward, and was greatest once the animal had learned the correct rule. Additionally, this increase in theta coherence was accompanied by an increase in co-activation amongst mPFC cells that were modulated by hippocampal theta, which formed reliable subgroups of neurons, hence cell assemblies (309). The hippocampus appears to drive the completion of mPFC cell assemblies as during periods of high theta coherence, certain mPFC neurons would phase shift to the trough of theta, bringing together the neurons of mPFC assemblies in time. In turn mPFC assemblies may come to form attractor states to neural processing at other levels of the hierarchy. In line with this suggestion, only a subpopulation of mPFC cells is entrained by hippocampal-mPFC theta coupling. Entrained cells encode current relevant behaviours, whereas non-entrained cells become most active during errors and reward omissions, triggering a theta phase reset, which may signify a switch in the neural assemblies forming the current attractor state (310). The two studies discussed here (297, 309) critically indicate that episodic memory serves a predictive function that necessitates the detection of regularities across episodic experiences, and the formation and application of abstract knowledge. Furthermore, both studies indicate that it is the hippocampus that drives the mPFC and the formation of abstract knowledge.

In conclusion, we have drawn an analogy with the inferior temporal - lateral prefrontal neural system that supports simple rule learning to propose that prediction based on complex episodic experiences critically relies on interactions between the hippocampus and mPFC (Figure 6.2). We have proposed that during the formation of abstract rule representations the hippocampus detects regularities across episodic experiences and through neural coherence drives the mPFC to form cell assemblies representing integrations of detected regularities with
behavioural options and value associations. Furthermore, the flexible application of abstract rules to novel situations is considered to again require the hippocampus (297, 311, 312), a disability generally observed in hippocampal lesion patients (237, 240, 241). Finally, we conceive of a process whereby mPFC cell assemblies come to guide or bias memory processing akin to top-down retrieval or attention functions (309, 313, 314). Together, this dynamic system supports the formation of abstract knowledge that can nevertheless be applied flexibly. Hence the novelty of our thesis is that we extend the view of the predictive function of memory to episodic memory including autobiographical memory. Further, in assuming that hippocampal-mPFC interactions are part of a neural hierarchy that supports sensori-motor predictions, we put forward specific roles for both regions in the formation of abstract knowledge. This idea will be further explored in the next chapters where we will highlight the interactions between the hippocampus and mPFC.

Although the first authors’, his brother’s, and father’s (Figure 6.1D) recollection of the incident may differ in details, their memories and the photograph share many features. A memory system that functions to extract regularities and foster prediction, for example between the first author’s own episodic experience of ‘the incident’, his father’s telling of the story (a favourite party feature), and subsequent viewing of the photograph, may have contributed to alterations in the first author’s mnemonic representation of ‘the incident’. Novel memories may be formed based on regularities across episodic experiences that do not reflect veridical representation of experiences, and have led to some of the memory distortions observed. The distortions in the first author’s memory of the incident thus fit with the suggestion that the autobiographical memory system also serves a predictive function that extract regularities and forms abstract knowledge across episodic experiences. As such, memory alterations reflect the flexibility of an adaptive memory system. Although this proposal can explain why the different memories of the incident agree in general, it leaves the striking differences in details unexplained. One possibility that may explain these differences in details is that with the formation of abstract knowledge,
memory representations change and different details may be lost. In the next section we will focus on such changes in memory representations that may occur during so called systems consolidation.

6.6 Systems consolidation supports formation of abstract knowledge

The extraction of regularities across experiences could thus have contributed to the commonalities observed between the first authors, his brother’s, and father’s memory of 'the incident'. Yet, their memories markedly differ on specific details. It is well known that over time our ability to retrieve detailed aspects of an autobiographical experience diminishes, whereas abstract knowledge of experience is retained, and memory alterations become apparent (214, 215, 315). One process that may contribute to memory alterations over time is systems consolidation. Here we will discuss the neural basis of systems consolidation and emphasize an overlap with the neural systems that support the formation of abstract knowledge, highlighting interactions between the hippocampus and mPFC.

Systems consolidation refers to a process in which the hippocampus initially stores experiences by linking together neocortical representations, and subsequently reinstates experience dependent patterns of neural activity in neocortical circuits causing strengthening of cortico-cortical connections. This is thought to be an iterative process resulting in memories to become stored in a neocortical trace that can be retrieved independent of the hippocampus (248, 316, 317). An extension of systems consolidation theory holds that for certain memories the mPFC can take over the indexing function of the hippocampus (318). An alternative theory to standard systems consolidation proposes that the reactivation of memory representations automatically results in the novel encoding of the reactivated representations incorporating current contextual information resulting in multiple traces of a memory (319). The formation of such related memories facilitates the extraction of regularities across episodic experiences that is thought to become
stored neocortically, yet the details of memory representations are considered to remain hippocampally dependent (319, 320). The question whether episodic memories may ever become fully independent of the hippocampus has sparked intense debate, and a wealth of patient and imaging studies have been put forward in support of either side of the argument (for reviews see (316, 317, 319-322)). However, lesion studies in animals have yielded results that conflict with both theories (323). To us it appears that both recent and remote memory performance for complex episodic experiences at least benefits from hippocampal involvement that may be determined by task demands (323, 324). Note that the systems consolidation debate mirrors the discussion on neural systems supporting episodic and semantic memory from which we derived our hypothesis that the hippocampus rapidly stores episodic experiences, detects regularities and drives the formation of abstract knowledge in the neocortex. Furthermore, the system consolidation discussion also reflects the idea that abstract knowledge may become stored in the neocortex but that the flexible application of knowledge to specific details relies on the hippocampus, again in line with our proposal.

Both standard system consolidation theory and multiple trace theory assume that the differences observed between recent and remote memories are due to changes of the memory trace that have occurred over time. An alternative to the differences observed between recent and remote memory is that these are not caused by the passage of time per se, but are attributable to a memory system that functions to extract regularities and fosters sensori-motor prediction. According to this view, systems consolidation effects do not reflect differences in the memory trace retrieved as a result of a retrieval cue, but differences in abstract knowledge that interacts with the retrieval cue (325). Differences in abstract knowledge at the time of recent and remote memory retrieval would affect the formation of novel episodic memory traces, the detection of regularities and the development of abstract knowledge. This is conceived to occur through the reinstatement of experience dependent activations consistent with the hippocampal-mPFC interactions we have described for regularity to response mapping for episodic experiences.
In this section we will discuss studies on systems consolidation, further explore the interactions between the mPFC and hippocampus, and the role of the mPFC in memory. Where relevant we will highlight evidence in favour or against the three hypotheses on systems consolidation.

Studies on systems consolidation have revealed that remote memory formation requires interactions between the hippocampus and neocortex in the post-encoding consolidation period that drive plasticity changes in the neural network. Both lesion and inactivation studies have shown a double dissociation where recent memory depends on hippocampus but over time remote memory is supported by the mPFC (326-329). At the same time as memory comes to depend on the mPFC, neurons within the mPFC develop differential firing rates for specific memory representations, also in the absence of further training indicating that synaptic connections within the mPFC or between the mPFC and other brain areas change due to intrinsic activity, possibly through the reinstatement of experience dependent activation in the delay interval (326). According to the systems consolidation hypothesis, reinstatement, or replay, would lead to a strengthening of cortico-cortical connections rendering the hippocampal memory index no longer necessary. As such, systems level consolidation should develop through an interaction between the hippocampus and neocortical areas, and the mPFC in particular, requiring synaptic modifications between these connections (318). Although NMDA receptor blockade does not impair memory retrieval (330), genetic reversible inactivation of NMDA receptor in CA1 in the week after learning, but not at later times after learning impaired remote memory (331). In contrast, AMPA receptor activity within the hippocampus is necessary both in the first and second week after learning for remote memory formation (332), which suggests that for remote memory formation basal neural activity is necessary within the hippocampus, but neural plasticity is only necessary in the output region of the hippocampus (333). Furthermore, reversible NMDA receptor blockade in the mPFC during the two weeks following learning, but not during later weeks, impairs remote memory (326). Interestingly, the direct projections from CA1 to the mPFC display NMDA receptor-dependent
LTP (334), and the impairment of remote memory by mPFC NMDA receptor blockade in the two-week post learning period is accompanied by a deficiency in the ability to induce LTP in vivo within the CA1-mPFC pathway that is attributable to post-synaptic plasticity mechanisms, not presynaptic basal synaptic transmission alterations (329). Hence, in line with our proposal hippocampal activity appears to drive synaptic plasticity in the mPFC. In addition to the mPFC, the posterior parietal cortex, and retrosplenial cortex show an increase in immediate early gene (IEG) expression over time, as memory becomes to depend less on the hippocampus (327, 328). Within the posterior cingulate cortex, IEG expression shows a shift from deep cortical layers to superficial layers over time, consistent with a shift towards greater cortico-cortical connectivity (335). Furthermore, genetically modified mice that have relatively normal CaMKII levels in the hippocampus but not the neocortex (CaMKII+/− mice), exhibit normal recent but impaired remote memory, and this is accompanied by a decrease in IEG expression in the mPFC. In these animals LTP can normally be induced in the hippocampus, but although LTP can be induced in the neocortex, this rapidly decays back to baseline (336). These findings indicate that in the immediate post encoding consolidation period, interactions between the hippocampus and mPFC are necessary to instate changes in the mPFC which may come to support remote memory. Additionally, the shift of memory dependency from the hippocampus to the mPFC is not a simple information transfer, but is initiated in parallel and develops through concurrent hippocampal-mPFC interactions (328, 337). These findings provide support for theories that assume a transformation of memory with systems consolidation.

As we have described, the formation of abstract knowledge, given the sparseness of cortico-cortical connectivity, requires the integration of memory in cortical networks not just through weight changes of existing synapses but via rewiring through the formation of novel synapses (249). It has been proposed that connections between neocortical neurons are formed in a two-step procedure in which first a random growth of neocortical synapses occurs that can become potentiated in a second step if subsequently exited by the reinstatement of experience dependent
activity (338). Although the mPFC damage immediately after learning does not impair recent memory (326-328), temporary inactivation of the mPFC during encoding does impair remote memory (328). This suggests that during encoding synapses in the mPFC are 'tagged' for future enhancement through subsequent hippocampal-mPFC interactions (275, 328). Continuing this line of reasoning Lesburgueres and colleagues (2011) go on to show that preventing 'tag' setting through mPFC inactivation during the formation of one mnemonic association but not another, does not affect recent memory but impairs remote memory for the non-tagged experience, possibly by preventing histone modifications that serve as tag. Temporary inactivation of the mPFC during encoding also prevented subsequent morphological changes in the mPFC (328), and expression within the mPFC of a growth protein associated with novel synapse formation increases with the formation of remote memory (327). Additionally, the genetically modified CaMKII+/- mice exhibit reduced synaptogenesis in the mPFC (339). Similarly, transgenic mice in which PAK, a regulator of actin and therefore of synaptic remodelling, is inhibited in the postnatal forebrain show altered spine morphology in the neocortex but not the hippocampus and impaired remote but not recent memory (340). Interestingly, PAK activity depends on NMDA receptor activation and plasticity within cortical layers IV and II/III is retarded in PAK transgenic mice (340), suggestive of impaired cortico-cortical rewiring. Finally, disrupting spinogenesis in the mPFC impairs remote memory formation (341). Thus, the formation of remote memory is accompanied by the same type of cortical rewiring changes that support the formation and storage of abstract knowledge from episodic experiences. Although the studies discussed above fit with the thesis that systems consolidation supports the formation of abstract knowledge, it does not exclude standard systems consolidation theory or multiple trace theory, although it appears to favour theories that assume transformation of memory with systems consolidation.

As we have discussed, with the formation of abstract knowledge, through hippocampal-mPFC interactions, the mPFC can come to guide neural processing in other brain regions. With regard to systems consolidation
similar mechanisms appear to exist. First, the interaction between the hippocampus and mPFC are dynamic in that over time the hippocampus drives plasticity changes in the mPFC, which are at the same time accompanied by adjustments within the output of the hippocampus (331). The change in hippocampal output may be an activate inhibition initiated by the mPFC to prevent the encoding of already stored information, preventing redundancy in the system (342). Second, the formation abstract knowledge within the mPFC through hippocampal-prefrontal interactions can subsequently become important to the exchange of information that directs neocortical reorganization. In resonance with this idea, mPFC synchrony increases with learning and drives entorhinal to perirhinal communication supporting information exchange between the hippocampus and neocortex (343). The mPFC is thus considered to provide a binding function as it actively appears to optimize interactions between disparate brain regions to support the formation of flexible memories.

Neuroimaging studies have revealed that systems level reorganization also occurs for human episodic memory. Tracking memory for complex pictures over a three month period showed that retrieval of memory is associated with increased mPFC activity over time, whilst a decrease in hippocampal activity was observed (344). A similar increase in mPFC involvement in memory was observed when object-location associations were learned spaced over repeated episodes compared to massed learning (345), consistent with the suggestion that mPFC involvement increases over multiple reinstatements of experience dependent activity. Interestingly, this increase in mPFC involvement for stabilized memories is accompanied by a decrease in gamma activity at early visual areas, whilst stabilized memories are retrieved faster (346), suggesting that following systems level consolidation a reduction of detailed memory processing takes place. A study conducted over several days examining the effect of multiple retrievals of autobiographical episodic and semantic memories, revealed that although the hippocampus and mPFC show repetition suppression with increasing retrieval attempts, both regions exhibit greater activity for autobiographical than semantic memory at the first
A  mPFC

Hippocampus

Representation of details

B  mPFC

Hippocampus

Representation of details

C  mPFC

Hippocampus

Representation of details
retrieval attempt and after eight retrieval attempts (347). Additionally, the results appear to indicate that the differentiation in neural activation between autobiographical and semantic memory retrieval in the hippocampus does reduce over time, whereas this differentiation in the parahippocampus increases. Again these studies indicate that memories transform during systems consolidation, but also indicate that the
hippocampus remains involved when detailed episodic memories are retrieved.

In summary, the mechanisms underlying systems consolidation resemble those we have described for abstract knowledge formation from episodic experiences. The loss of details and retention of abstract knowledge over time is paralleled by a shift of memory dependency from the hippocampus to the neocortex, and specifically the mPFC. The hippocampus is thought to reinstate experience dependent neural activation in the neocortex in the post encoding consolidation period driving plasticity changes in the neocortex and between the hippocampus and neocortex. Interestingly, neocortical synapses that come to support memory over time are already tagged during learning for future strengthening. This suggests parallel memory formation in the hippocampus and neocortex, akin to the development of abstract knowledge during learning. Furthermore, the shift in memory dependency is considered to occur through a two-step procedure where in the neocortex random synaptic growth occurs that can become stabilized if this coincides with the reinstatement of neural activity supporting the stabilization of cortico-cortical connections. Over time regularity representation within the mPFC may take over the binding function of the hippocampus and come to guide neural processing in other regions. However, as the cortico-cortical connectivity remains sparse, this might come at a loss of the complexity of episodic memory. When details of episodic experiences are retrieved the hippocampus does appear to be involved, also in the retrieval of remote memories, possibly guided by abstract knowledge that is supported by a neocortical network centred on the mPFC (Figure 6.3).

In conclusion, studies on systems consolidation reveal that memory formation is a dynamic process. In line with our hypothesis systems level consolidation appears to support the formation of abstract knowledge through interactions between the hippocampus and mPFC proposed to be part of a sensori-motor predictive neural hierarchy. However, these studies do not provide definitive evidence that the formation of
abstract knowledge occurs with systems consolidation nor does it exclude standard systems or multiple trace theory and is therefore an important topic for future studies. Nevertheless, due to systems consolidation the memory of ‘the incident’ might thus have changed, whereby regularities between different experiences have become integrated and specific details lost. The loss of specific details and persistence of abstract knowledge indicates that a flexible memory system is adaptive, as specifics are lost but abstract knowledge that can be applied to multiple situations independent of particular details is retained. An open question that remains is when and how memory representations are transformed during a post-consolidation period. Sleep has been a long-suggested candidate for such memory transformation. In the next section studies on sleep will be discussed that provide further insight into the interactions between the hippocampus and mPFC that support the formation of abstract knowledge and emphasizes the dynamic nature of memory. Critically, these studies reveal that during sleep abstract knowledge is formed that supports a predictive function.

6.7 Sleep supports formation abstract knowledge

During sleep the hippocampus is considered to repeatedly reinstate experience dependent activations in the neocortex that drive changes in cortico-cortical connectivity and give rise to memory transformations. Importantly, sleep provides an optimal opportunity to reprocess memories, detect regularities, and map regularities to responses as during sleep memories are not subject to possible interference by incoming sensory information. We hypothesize that sleep contributes to the formation of abstract knowledge from episodic experiences and that studies on sleep and memory provide insight into the mechanisms of a dynamic memory system.

Supporting the idea that sleep plays a role in memory consolidation, sleep has been found to improve memory, more than a wake period of comparable length (for a review see (348). Recently, the re-exposure during sleep to smells or sounds associated with object-location learning
was found to enhance subsequent memory (349, 350). Although these latter studies provide causal evidence that specific memories can be reactivated and reprocessed during sleep and subsequently strengthened, they imply a rather passive function of sleep in memory consolidation, where sleep simply strengthens veridical memory representations. However, if sleep actively updates and reprocess memories serving the extraction of invariant regularities across experiences, sleep should result in qualitative changes in memory representations and the active consolidation of new memories (276, 348, 349, 351). Confirming this suggestion, subjects who had not discovered a hidden rule during a rule learning task were more likely to display knowledge of the hidden rule after a sleep period compared to subjects who had stayed awake (352). Thus, subjects demonstrated novel insight for information learned prior to sleep. Critically, this indicates that during sleep memories are reconfigured and regularities across episodic experiences can be detected to form new memories and the basis of abstract knowledge that supports a predictive function. However, the extraction of regularities during sleep may also lead to false memories. In a Deese-Roediger/McDermott paradigm subjects learn lists of words (e.g. door, glass, pane, shade, house, and curtain) which are semantically related to a critical word that is never shown (e.g. window). Subsequent testing on all words including the critical word leads to false memories (i.e. subjects indicate having studied ‘window’). Interestingly, sleep preferentially preserved memory for critical words (353). The mechanisms by which memories may be reprocessed and abstract knowledge is formed during sleep are starting to be uncovered. In the next paragraphs we will present studies that reveal interactions between the hippocampus and mPFC during sleep that we suggests gives rise to the formation of abstract knowledge.

Memory processing during sleep has been associated with specific electrophysiological mechanisms (253, 354). Importantly, neural oscillations are thought to provide a spatio-temporal structure for coordinated communication between neural cell assemblies within and across brain areas (354). During sleep interactions between cortical slow oscillations, thalamocortical spindles and hippocampal sharp wave ripple
events provide a mechanism to coordinate the output of reactivated hippocampal cell assemblies with ongoing neocortical activity and guide the hippocampal output to the neocortical circuits that gave rise to input to the hippocampus during the experience (354-356). In concordance, slow oscillations, spindles and sharp waves occurring during sleep have been found to correlate with memory performance, and interference with these mechanisms affects memory in animals and humans (357-364). As hippocampal sharp wave ripple events can become nested in the troughs of spindles (355), when the depolarized neocortical state is most sensitive to plasticity changes, this would be an optimal timing for memory interactions between the hippocampus and the neocortex. In line with this idea, sleep spindles have been detected in neocortical areas that were also involved in learning of a task (365) and their occurrence increases with the amount of difficulty of a task (360). Interactions between the hippocampus and neocortex during sleep may thus be critical to memory consolidation and the formation of abstract knowledge through reactivation of memories.

Electrophysiological studies have provided evidence for the reactivation of memories during sleep. During explorations, when learning of a route occurs, place cells fire in sequential order compressed in theta oscillations with a certain degree of overlap and as such allow for the representation of overlapping past, present, and future locations into single episodes (366-368). Seminal to the idea that memories are reprocessed during sleep has been the discovery that place cells within the hippocampus reactivate in the same spatio-temporal sequences during sleep as they had during a prior learning episode (369, 370) albeit at a faster time scale (371). Importantly, the forward nature of memory replay (370) implies a predictive memory function, instead of a mere recap of events.

In line with the suggestion that memory reactivation during sleep supports the formation of abstract knowledge in the neocortex, interactions between the hippocampus and visual cortex for replay of sequences have been found (356), as well as replay of cells in the
mPFC (372), striatum (373), several forebrain regions (374), and the locus coeruleus (357). Additionally, following learning of a unilateral motor task, sleep spindles also increase in the contralateral hemisphere (358), suggesting a generalization of experience dependent learning. Furthermore, replay of place cell sequences decreases over several post-learning sleep periods (370). These findings confirm to the idea that during sleep the hippocampus replays information acquired during the wake period to the neocortex, allowing for multiple interleaved exposures either suited to train a slow neocortical rule learning system that extracts regularities (276, 348, 375), or store regularities detected by the hippocampus. Hence, memory reactivations and interactions between the hippocampus and neocortex during sleep appear to be an ideal mechanism to the formation of abstract knowledge that might come to provide predictions and guide behaviour. Yet only recently has support for this suggestion been presented.

In section 3 we have described that during abstract rule learning from episodic experiences, neural cell assemblies are formed within the mPFC through interactions with the hippocampus that reflect abstract rules (309). Critically, such memories related to abstract rule learning have been found to reactivate during sleep through mPFC-hippocampal interactions (376). The reactivation of abstract rule related neural assemblies could be detected based on the similarity of the principle components between activity during learning and sleep. Interestingly, different components reactivate at different times during sleep, reflecting a possible interleaved learning schedule optimal to the formation of abstract knowledge (276). The replay of these cell assemblies occurs in concert with hippocampal sharp-wave ripple events, and the hippocampal activity appears to precede mPFC activation, indicating that the hippocampus drives the mPFC activity in line with the suggestion that following the detection of regularities, the hippocampus drives the formation of abstract knowledge in the neocortex. Importantly, the neural assemblies that activate during sharp-wave ripple events reflect assemblies that were active at the choice point of the maze, and the reactivation of these assemblies occurs more often when an animal
has learned the abstract rule (376). These results indicate that abstract
knowledge can be formed during sleep through hippocampal-mPFC
interactions, and may come to support sensori-motor prediction.

Human neuroimaging studies have also revealed memory related
reactivation during sleep of the hippocampus and neocortical regions
implicated in task learning (349, 377-380), and this reactivation activity
correlates with subsequent memory performance (379). Additionally,
functional hippocampal-mPFC connectivity increases following a night
of sleep and is related to memory performance (377, 378). Importantly,
subjects who gain insight into a hidden rule following a night of sleep
exhibit greater hippocampal-neocortical connectivity during learning
and an increase in mPFC activity following sleep (381). In contrast,
subjects who do not gain this abstract knowledge show no increase in
mPFC activity following subsequent sleep (381). Thus, the hippocampus
and mPFC might interact during sleep to detect regularities across
reactivations and form abstract knowledge. Interestingly, the formation
of false memories during sleep (353) has also been associated with
the hippocampal-mPFC network (381). Hence, interaction between the
hippocampus and the mPFC appear to give rise to abstract knowledge
that can come to support prediction, but this process can also give rise to
memory alterations and distortions.

With the formation of abstract knowledge, sleep may initiate the systems
consolidation shift of memory dependency from the hippocampus to the
neocortex. The decrease of memory related hippocampal activity over
time has been reported to correlate with the slow-wave sleep period after
learning, and is accompanied by an increase of mPFC activity over time
with regard to memory performance (344). During sleep, the mPFC may
support a binding function by bringing together distributed neocortical
memory processes. In a study where subjects had to learn face-location
associations, activity in the fusiform cortex (an area implicated in face
processing) was found to correlate with early visual and posterior
parietal areas (an area implicated in location processing) during sleep
stage 1. During sleep stage 2 the fusiform cortex switches towards
functional connectivity with the mPFC and middle temporal cortex, and this network overlapped with the neural networks that would come to support subsequent memory retrieval (382). Sleep appears to initiate changes in functional cortico-cortical connectivity that is associated with abstract knowledge formation and could take over a pointer function of the hippocampus (318). As such, the formation of a neural network supporting abstract knowledge appears to provide a structure that nests detailed memory representations.

The reduction of hippocampal involvement in memory has been proposed to be attributable to synaptic down-scaling during sleep (348, 383). Synapses that have been potentiated during a wake period are again weakened during sleep (384, 385). Synapses that are sensitive to regularities between experiences would have been strengthened more during learning and are more likely to survive synaptic downscaling than synapses potentiated by single experience that reflect details (375). Additionally, synaptic down-scaling affects weight but not wiring changes. Together this suggests that regularities are preferentially stored during sleep, but details of memory may be lost.

In summary, one function of sleep is considered to be the detection of regularities across experiences through the interleaved replay of episodic experiences thereby creating abstract knowledge that comes to guide future behaviour. Studies on memory and sleep critically reveal that the formation of memory is dynamic and that sleep can contribute to the optimization of prediction from episodic memory through the formation of abstract knowledge. In line with our thesis, the formation of abstract knowledge during sleep occurs through hippocampal-mPFC interactions and is associated with changes in cortico-cortical connectivity. Additionally, with the formation of abstract knowledge the mPFC brings together distributed cortical memory processes and takes over the binding function from the hippocampus. As such a hardwired neural network centred on the mPFC and reflecting abstract knowledge can provide structure to nested memory representations. Additionally, hippocampal involvement and details of a memory may be lost due to
synaptic down-scaling during sleep (Figure 6.3C). As such, the memory for 'the incident' might have become reconfigured during sleep, losing the specific details of the episode but maintaining or even creating memory for regularities between 'the incident' experience and other experiences.

Synaptic down-scaling is considered to free up capacity to allow new learning. In resonance with this idea, sleep facilitates new learning following sleep (382, 384-386), which is correlated with hippocampal connectivity during sleep (382). Hence, the integration of new memory within existing knowledge could also have contributed to the distortions of the first author’s own memory. In the next section we will discuss how abstract knowledge comes to guide the integration of new information into memory and more specifically that a hardwired neocortical network centred on the mPFC can provide elaborate information and as such affect hippocampal processing and new learning.

6.8 Abstract knowledge comes to guide integration new information

So far we have discussed how the regularities can be detected amongst episodic experiences, and how detected regularities form abstract knowledge over time, possibly during sleep, and come to guide behaviour. Additionally, the formation of abstract knowledge within the neocortex can develop to guide future memory encoding, possibly providing elaborate information to hippocampal processing. This allows the integration of current incoming sensory information with abstract knowledge, current goals, and task demands, as well as the formation of novel memory representations that are non-overlapping with existing representations (272, 299). The importance of existing knowledge in the formation of novel memories has long been recognized (272, 387, 388). Relevant to a learning situation, existing knowledge is considered to provide a prior state related to which incoming information is encoded. The prior state is thought to form dynamically and has been termed a 'schema' (387, 388), though Bartlett favoured the term 'organised setting'. Only recently have neuroscientists begun to investigate schemas in the
acquisition of novel information and have come to define a schema as a pre-existing network of interconnected neocortical representations (325), where we would like to add that schemas reflect abstract knowledge based on previously detected regularities and related to the current learning situation.

Initial suggestions that prior experience could alter novel learning came from studies showing that rats reared in a 'village', thereby providing extensive experience to a complex environment, could still learn routes within this environment following hippocampal lesions (389). More recent studies have shown that the speed with which memories become independent of the hippocampus and integrated within the neocortex depends on the presence of a corresponding schema (390). This implicates that the existence of prior information allows for a rapid integration of novel information and that the learning rate of the neocortex can be fast if information is congruent with a schema. These findings are suggestive that initial learning requires a slow neuronal growth process such as rewiring, but if wiring is in place and forming a schema with which novel incoming information is congruent, than novel information can be rapidly integrated through a fast learning process possibly reflecting weight changes. Furthermore, flexibility to adjust abstract knowledge requires new learning and is hippocampus dependent. As a consequence of these observations it has been suggested that systems consolidation might not so much be about time but about the need for specific changes, i.e. cortico-cortical rewiring or synaptic modification (391). In support of this idea learning of information in the presence of a schema is accompanied by up-regulation of immediate early genes (IEG) in the mPFC (337). Specifically, this up-regulation is greatest when new information is learned in the presence of a strong schema, but IEG up-regulation is greatest in the hippocampus when a completely new schema has to be learned, i.e. when the learning situation comprises most novelty. Furthermore, both suppression of activity (by blocking basic AMPA transmission with CNQX) and obstructing plasticity (with the NMDA receptor antagonist AP5) within the mPFC impairs learning of new information within the presence of a schema, yet only
suppression of activity not plasticity impairs the retrieval of these new memories (337). The existence of prior knowledge regarding a specific learning situation thus appears to boost the speed with which novel memories become hippocampus independent and integrated within the neocortex with a specific dependence on the mPFC. Yet to date it remains unclear whether there is no need for rewiring for new learning in the presence of a schema, and whether in this case memory formation would be insensitive to interference with spinogenesis. As before, these studies critically indicate that memory formation is dynamic, as it depends on the existence of prior knowledge. Additionally, these studies also highlight that the existence of abstract knowledge may aid the integration of new information to optimize predictions and guide behaviour. Similar to the discussion in the previous section of abstract rule learning, systems consolidation, and memory reactivations during sleep, schema memory critically involves interactions between the hippocampus and mPFC.

Human studies on schema consolidation have mirrored animal findings. An early positron emission tomography (PET) study has shown that the presentation of pictures relevant to interpreting a subsequent story aids understanding of an otherwise confusing text and is associated with increased activation of the mPFC and precuneus (392). More recent studies have investigated the dynamic interaction between the hippocampus and neocortex with regard to schema consolidation. Van Kesteren and colleagues (2010) effectively modulated prior knowledge, or schemas, by presenting subjects with a movie either in normal order (schema group) or in an order in which the scenes of the movie had been scrambled (no-schema group). When subjects viewed the last part of the movie in non-scrambled order on the next day, the schema group showed reduced functional connectivity between the hippocampus and mPFC. Interestingly, the subjects that exhibited better memory for schema related questions within the non-schema group showed also less hippocampal-mPFC interactions (393). These effects persisted in a post-learning encoding period, reflecting a change in schema-dependent consolidation. In a follow-up study, subjects learned associations between visual motifs, fabric textures, and words that were either incongruent
(acrylic-sock) or congruent (fleece-sweater) with common knowledge (394). When cued with the visual motifs, subjects were better at remembering congruent associates, which was accompanied by greater activity in the mPFC and task-specific somatosensory cortex during recall as well as stronger interactions between these regions. Thus, reduced hippocampal-mPFC connectivity during encoding is paralleled by increased mPFC-sensory cortex interactions at retrieval. This finding also indicates that the mPFC may take over a binding function from the hippocampus, as do other studies previously discussed in this review (343, 382). Novel memories appear to become rapidly integrated within existing neocortical networks through hippocampal-mPFC interactions if an overlap exists between novel incoming information and episodic regularities that we suggest to involve the mPFC. The mPFC may thus effectively form prior predictions over incoming information, and increase interactions amongst distributed neural regions that form a preexisting network. In line with this idea mnemonic congruency effects on perceptual processing have been detected at very early stages (100 ms) in the occipital cortex (395), have been associated with increased theta activity (396, 397), and greater theta phase locking between occipital regions, parahippocampus, retrosplenial cortex, and mPFC (398), where the mPFC is thought to exert prior expectations on visual processing (223). Similar to the formation of neural assemblies thought to form attractor states during abstract rules learning (309, 310).

Ultimately, the pre-existent network of interconnected neocortical representations based on regularities across episodic experiences should provide organizing principles guiding learning and behaviour. In section 3 we have described a study on the formation of abstract knowledge through hippocampal-mPFC interactions (297). In this study the authors also tested the application of acquired knowledge to novel situations. In a transfer test subjects were confronted with a choice paradigm with the same underlying rules as learned previously but with new stimulus-pairs. Thus, abstract knowledge based on regularities of prior experience served as a schema in a perceptually novel session effectively testing the application of detected regularities across episodic experience invariant
of the specific perceptual details. In this new learning session subjects showed superior performance that was correlated with the amount of prior abstract knowledge. The probability of correct responding in this transfer task was associated with increased activity in the hippocampus, mPFC, and precuneus, yet only the hippocampus showed a greater degree of activity with regard to correct performance on the transfer test than during initial learning (297). One can infer from this that regularities are detected by the hippocampus and become represented by the mPFC through dynamic hippocampal-mPFC interactions. In turn the regularity representations that dependent on the mPFC might provide prior expectation signals to the hippocampus which is critical for the flexible application of abstract task rules to novel situations. This finding reflects the idea that abstract knowledge depends on the neocortex and mPFC in particular, but that the retrieval of details and application of knowledge to new situations also requires the hippocampus.

The existence of schemas provides prior expectations that guide decision making and support the integration of novel memories into neocortical networks. However, schemas could also be responsible for the formation of false memories. Probing prior knowledge of people before presenting suggestive information increases the chances that suggestive information becomes incorporated into autobiographical memory creating false memories (214, 387). This is especially true for the integration of false details into a memory representation. It might thus be that the repeated exposures to retellings of 'the incident' and seeing the related picture have caused extraction of regularities across these experiences. Abstract knowledge of 'the incident' may have come to form a schema to guide novel memory formation during subsequent experiences related to 'the incident' and have led to integration of incorrect details into the memory for the event. Once more, the anecdote provides an example of a flexible memory system that supports prediction, but may result in inaccurate memories.

In conclusion, studies on schema consolidation indicate that prior knowledge guides new learning that optimizes subsequent behaviour.
Again schema consolidation studies emphasize that memory is a
dynamic process. Mirroring the discussion on abstracts knowledge,
systems consolidation, and sleep, schema consolidation occurs through
interactions between the hippocampus and mPFC, and implicates an
important role for the mPFC in abstract knowledge. We speculate that
the interactions between abstract knowledge and the formation of
new memories occur through dynamic interactions between the mPFC,
hippocampus and possibly other brain regions (Figure 6.3C). However, to
date little is known about the electrophysiology of schema consolidation
that can corroborate this claim. Furthermore, the rapid integration of
novel information into memory in the presence of a schema suggests
that this relies on weight changes of existing neocortical synapses, and
does not require neocortical rewiring. Again, future studies will have to
verify this hypothesis. Finally, the presence of abstract knowledge could
have caused a rapid integration of new experiences into the first author’s
memory of ‘the incident’, and as such, contributed to the observed
memory distortions. The alterations in the first author’s memory for the
incident thus indicates that autobiographical memory is flexible, and
that abstract knowledge may have optimized the integration of novel
information to form adaptive memory representations. A final possibility
we will discuss that could have contributed to these distortions, is the
possibility that certain aspects of the memory were erased or updated
when reactivated.

6.9 Memory reactivation and updating of predictions

Standard consolidation theory suggests that memories are initially labile
and sensitive to disruption but become stabilized over time and are
subsequently relatively insensitive to disruption (30). A dynamic memory
system would, however, benefit if memories remain sensitive to changes
if the outcome of a situation relative to a retrieved memory is different
from what was predicted, allowing for the optimization of prediction from
memory. In light with our hypothesis here, details of memories should
be more sensitive to updating than abstract knowledge, and updating of
episodic memories should again involve the hippocampus.
In the late 60’s and early 70’s of the last century, several studies showed that interfering with brain function immediately following retrieval of a memory caused subsequent amnesia of the memory (47, 48, 399). These findings regained interest when Nader and colleagues found that blocking protein synthesis in the amygdala following memory retrieval impaired memory a day later (50). Critically, this impairment was not observed if memory was not retrieved prior to protein synthesis blockade, nor was it observed immediately following the application of the protein synthesis inhibitor (50). These findings indicate that when a memory is recalled it is reactivated and becomes instable requiring a time dependent re-stabilization process that has been termed reconsolidation (48, 49, 400). Subsequent animal studies have indicated that complex hippocampal dependent memories also undergo a reconsolidation process when reactivated (401-403). In humans impairments of simple emotional associative memory upon reactivation have been reported (404-406) as well as disruptions of episodic memory following reactivation (407-412). However, several of the human studies do not meet the criteria to demonstrate a reconsolidation process, as they lack immediate memory tests or have given an amnesic agent prior to memory reactivation and are therefore unable to show a time-dependent process and may have interfered with the retrieval process or new learning (for a review see (413)). Nevertheless, reconsolidation does appear to occur for episodic memories, which could explain some of the alterations observed for the memory of ‘the incident’. Yet, the observed memory alterations are for specific details, but the first author has not lost memory for the episode completely. Our thesis predicts that his knowledge of the incident has come to depend on wiring changes whereas details of the memory have remained sensitive to synaptic weight changes and thus to alterations through reconsolidation.

In concordance with the weight versus wiring suggestion, several studies have reported that older memories and stronger memories are less sensitive to reconsolidation disruption, or can recover following disruption (414-417). In addition, reconsolidation is impaired by blocking molecular substrates that have been associated with synaptic
plasticity (for review see (400)), and is associated with changes in memory-related glutamate receptor expression (418, 419). Indeed, disrupting reconsolidation can result in a reduction of synaptic long term potentiation (420). Furthermore, upon reactivation memory destabilization depends on protein degradation from the synapse (421, 422). Thus, reconsolidation seems to be associated with changes in synaptic weight changes. In contrast, interfering with processes that are important to synapse formation and neural growth, such as BDNF or actin rearrangement, impairs new learning but not reconsolidation (421, 423). These findings mimic the results from studies on systems consolidation, sleep, and schema learning presented in the previous sections. To elaborate, studies within all these fields of memory research show that the formation of abstract knowledge is supported by neocortical rewiring, but that details of memories may be sustained by synaptic weight changes that are most sensitive to updating. Abstract knowledge that is supported by cortico-cortical connectivity should thus be less sensitive to disruption through reconsolidation. Nevertheless, disruption of reconsolidation for old memories have also been reported (401, 424), even for memories that are thought to have undergone systems consolidation at that time (401). Therefore, it remains an open question whether neocortical memories can be disrupted during reconsolidation, possibly via an impairment of critical nodes in the memory network, and which aspects of the memory network remain intact following disruptions. Amnesia following disruption of reconsolidation indicates that memory is dynamic but may not appear immediately useful.

Findings that memories could be strengthened upon reactivation implicated that the function of reconsolidation might be the updating of memory representations (425-428), thereby allowing for the optimization of behavioural prediction from memory. In support of this notion, reactivating memory in a setting that requires new learning renders a memory sensitive to disruption of reconsolidation, but not if the reactivation trial does not require new learning (403). A more recent study used a cellular approach to investigate the updating hypothesis by exploiting a dissociations between the dependency of new learning,
destabilization of memory, and re-stabilization on specific cellular mechanisms (421). Additionally, a specific learning paradigm was employed in which an animal’s memory of a context was updated with the incorporation of a shock association at reactivation. Interestingly, preventing the destabilization of synapses at reactivation, prevented the updating of the memory, but left the original memory for the context intact. In contrast, inhibiting the re-stabilization of synapses not only prevented the updating of the memory, but also impaired memory for the original context (421). Hence, under reconsolidation conditions memories can be updated to optimize the predictive function of memory.

Considering that amnesic agents used in animals to disrupt reconsolidation cannot be readily applied as experimental treatments in human subjects, most studies on reconsolidation in humans have used interference paradigms and have also indicated a role for reconsolidation in memory updating (413). Typically, in these studies subjects have to learn a set of items. In the following session memory is reactivated by reminding subjects of the learning set from the first session. After the reminder a new stimulus set is learned that is thought to interfere with the reactivated memory from session one. Using this procedure it has been shown that memory for the first set is impaired by learning of the new set when memory is reactivated 24 hours after learning, but not if memory is not reactivated (407, 408, 412) or if tested immediately after learning of the second set which reveals a time-dependent reconsolidation effect and excludes the possibility that the effect is attributable to secondary encoding (407, 408). Critically, subjects do not recall less items from the first set, but erroneously report items from the second set, thus indicating an integration of the two lists in memory (407). Additionally, instructing subjects after the reminder to actively update their memory causes subjects to remember items from the first set correctly in addition to the new information (412). In summary, both animal and human studies thus suggest that reconsolidation allows for the updating of memory for a known situation to a new outcome and, hence, supports memories predictive function. Throughout this review we have indicated that formation of flexible memories involves dynamic
interactions between brain regions. Recently studies have also revealed interactions between brain regions during reconsolidation.

Reconsolidation may also allow for interactions between brain regions to update memory. As we have discussed, the updating of a context memory with a fear association can occur through reconsolidation (421). Interestingly, fear conditioning critically depends on the amygdala, and the discussed updating impairment occurred through injections in the hippocampus (421), this indicates that interactions between the hippocampus and other brain regions occur during reconsolidation. Additionally, this finding concurs with our hypothesis that retrieval and updating of details of a memory involves the hippocampus. Furthermore, Kuhl and colleagues (2010) measured brain activity while subjects learned pairs of items (A-B). The presentation of AB pairs was preceded by cues indicating a high or low monetary reward for correct recall on a subsequent memory test. In a following study phase subjects memorized the same cues (A) but with new associate items (C). AC learning puts memory for AB pairs at risk of being forgotten. When assessing brain activity during AC learning that predicts AB memory on a final memory test following AC learning the investigators found hippocampal activation to be predictive for AB memory survival. Critically, greater activation in regions associated with reward processing at the time of AC learning predicted better memory for AB pairs on the final test. This indicates that hippocampal dependent memories can be reactivated together with brain regions that process their associated context, i.e. the reward (429) in line with animal findings (373). Thus, during memory reactivation interaction between brain regions also appear to occur and allow the updating of memories during reconsolidation (430).

The interactions between a stored memory representation and incoming information of the current situation, as well as the interactions between brain regions during reconsolidation may follow attractor dynamics. This implicates that reconsolidation only occurs when neural activation evoked by a reactivation session settles to an attractor state that represent the original memory. Two findings speak to this proposal.
First, reconsolidation only occurs if memory is reactivated in the original learning context (408, 431), thus the overlap between learning and re-exposure needs to be high enough to evoke memory reactivation. Second, under re-exposure regimes that do not evoke memory reactivation but new learning (e.g. extinction) reconsolidation of the original memory is not observed (402, 432). A recent modelling study exploited these findings to investigate attractor dynamics in reconsolidation (433). The authors created a neural network model that formed attractor states for specific memories following training. The network could be tested by presenting a partial cue and testing the overlap between the evoked state in the network and the specific memory attractor states. Importantly, synaptic weights in the network were adjusted following normal Hebbian plasticity rules, but importantly synapse destabilization was also modelled and caused by the difference between the retrieved attractor state and retrieval cue. Using this approach the authors were able to show that if a mismatch exists between the attractor state in which the network settles and the re-exposure cue, neuronal synapses that code for the mismatch will destabilize and require restabilization (433). Note that destabilization occurs as a consequence of mismatch detection; a process predicted to occur at each level of the neural hierarchy according to the sensori-motor predictive framework and could be seen as an extension of the PIMMS model (228). Interestingly, the authors used an auto-associative neural network and the results indicate that mismatch detection is critical to reconsolidation, as both have been suggested to be properties of the hippocampus, this structure may be critical to the reconsolidation process.

In summary, memories can become instable during reconsolidation and require restabilization upon reactivation. Studies on reconsolidation critically emphasize the dynamic nature of memory in that memories can be adjusted to optimize prediction from memory. We have presented studies that indicate that during reconsolidation synapses destabilize putting memories that depend on synaptic weight changes at risk of disruption. Memories that have been integrated through re-wiring changes in the cortico-cortical connectivity might be less
sensitive to reconsolidation disruption (Figure 6.3C). Furthermore, reconsolidation appears to serve the function of updating existing memory representations to changes in environmental contingencies, and this contributes to an optimized prediction from memory. Reconsolidation thus provide memory with the flexibility to adapt to new situations and could explain why details of the first author’s memory for 'the incident' have been lost or changed while his abstract knowledge of 'the incident' has remained relatively intact. Additionally, this could explain the differences observed between reconsolidation studies on simple associative memory and complex episodic memory (413). The former can be supported by plasticity of a limited number of local synapses and disruption of these could lead to complete memory loss, whereas remote episodic memory relies on distributed hardwired neocortical representations that would be less sensitive to full disruption. Future studies will have to reveal whether causal interference with neocortical brain function can disrupt distributed episodic memories, to which degree wiring changes are essential for restabilization of memory, and what degree of the cortical memory network remains intact following possible episodic memory disruption. Finally, we have discussed that memory reactivation can be accompanied by interactions between brain regions and that these may follow attractor dynamics. Specifically, mismatch between a retrieved memory attractor state and incoming information might open up the reconsolidation window, and the detection of this mismatch could be a principle function of the hippocampus in reconsolidation. To date the role of the mPFC in reconsolidation is unknown. According to our thesis abstract knowledge dependent on the mPFC would provide prior predictions over incoming information about a current situation, supporting mismatch detection by the hippocampus and initiation of reconsolidation. This suggestion is in need of empirical support. In conclusion, reconsolidation could have contributed to alteration of the ‘incident’ memory, and specifically have resulted in changes in details, highlighting that dynamic neural systems enable adaptive and flexible memories.
6.10 Discussion

In this review we have taken memory alterations as an example to reveal important characteristics of the dynamic nature of memory and the neural systems that support memory formation. We have tried to explain that memory distortions occur, because memories are not veridical representation of experience, but serve the function of predicting the future. Taking the point of view that memory serves a predictive function and arises in a neural hierarchy that supports sensorimotor prediction, we have tried to put forward a view on the dynamic neural systems that support the formation of flexible memories, and particularly propose specific functions for the hippocampus and mPFC in this process. Critically, prediction from episodic experiences requires integration of information from distinct experiences, the detections of regularities across experiences, and the formation of abstract knowledge that is invariant of specific features (Figure 6.3). In section two we have considered several characteristics of the hippocampus that support detection of regularities across experiences. Additionally, we discussed that the hippocampus allows rapid encoding of information through weight changes enabling the detection of regularities between memories and novel experiences. Following detection, regularities are thought to become stored in the neocortex through rewiring forming a relative stable substrate of abstract knowledge (Figure 6.3B). The formation of abstract knowledge implies a dynamic memory system as abstract knowledge no longer reflects a veridical memory representation. Abstract knowledge may come to support the predictive function of memory. However, behaviour based on predictions from episodic memory requires the integration between regularity representations and motor outputs. In section three we have drawn an analogy with the lateral temporal-lateral prefrontal system that supports simple rule learning to suggest that an analogous hippocampal-mPFC system supports a predictive memory system for abstract rules based on complex episodic experiences (Figure 6.2). Notably both these systems comply with the view of the brain as a sensory-motor predictive neural hierarchy. Specifically, our novel proposal is that the mPFC plays a critical role in the formation of
abstract knowledge by integrating regularities detected across episodic experiences with motor output to form abstract rules (Figure 6.3B). In the formation of abstract rule representations the hippocampus is considered to detect regularities and through dynamic interactions drive the mPFC to form cell assemblies representing integration of regularities and behavioural output options. In turn, the flexible application of abstract knowledge again requires interactions between the mPFC and hippocampus (Figure 6.3C). It is of interest to note that the neural network that we propose to support a predictive function for complex episodic experiences shows overlap with the neural networks implicated in prospective memory and imagination (434, 435). In section four and five we have emphasized an overlap between the neural systems that support the formation of abstract knowledge and those implicated in memory processing during systems consolidation and sleep, which provide further insight into the interactions between the hippocampus and mPFC and have lead us to suggest that systems consolidation and sleep contribute to the formation of abstract knowledge. During sleep the hippocampus is considered to reinstate experience dependent neural activations in the neocortex that enable the detection of regularities across experiences and initiate a shift in memory dependency from the hippocampus to the neocortex. This shift occurs with the formation of cortico-cortical connectivity through a two step procedure where random synaptic growth in the neocortex is stabilized if it coincides with the reinstatement of neural activity. Accompanying the formation of abstract knowledge, the mPFC may take over the memory binding function of the hippocampus, and come to guide neural processing in other brain regions (Figure 6.3B). Finally, during sleep synaptic down-scaling reduced the weights of synapses. This would results in a loss of weak synapses but the survival of strong synapses and rewired connections, and therefore a loss of memory details but retention of abstract knowledge (Figure 6.3C). In section six, we presented studies to suggest that the existence of abstract knowledge and a pre-existing network of interconnected neocortical representations can come to guide new learning. The presence of abstract knowledge is proposed to form a schema and cause a rapid integration of matching new experiences into the neocortex, possibly because this
rapid integration only requires weight changes (Figure 6.3C). Finally in section seven, we discussed that memory reactivation allows memories to be updated. This may be specific to details of memories that are only associated with synaptic weight changes (Figure 6.3C). Additionally, we have discussed that updating at reactivation might rely on interactions between the hippocampus and other brain areas that possibly follow attractor dynamics. Together these mechanisms contribute to the formation of abstract knowledge and support the extension of the predictive function of memory to episodic and autobiographical memory.

Our thesis makes specific predictions and raises specific questions to future studies. We propose that the hippocampus detects regularities and subsequently drives the neocortex and in particular the mPFC to form abstract knowledge. This suggests that detected regularities are immediately represented in the hippocampus but only at a later time in the mPFC. Both neuroimaging and electrophysiological studies employing decoding methods should thus be able to find regularity representations in the hippocampus before being able to detect these in the mPFC. Additionally, methods to study connectivity dynamics might reveal that the ability to detect regularity representations in the mPFC arises through hippocampal input. Similar study methods may be able to provide insight into the question whether the mPFC takes over a binding function from the hippocampus with systems consolidation, a consideration we support in this thesis. If so connectivity studies should be able to show that the mPFC affects effective connectivity between disparate brain regions, accompanied by an increase in abstract knowledge. An additional promising approach in our view is the study of network properties using graph theoretical approaches. A prediction based on our thesis is that the mPFC should exhibit greater characteristics of a hub with the development of abstract knowledge. Such an approach could also be beneficial when assessing which elements of the neural network dynamically change to support adaptive memory alterations, and which elements remain intact. With an increase in abstract knowledge we predict increases in effective connectivity between neocortical regions, and particularly with the mPFC, an increase in hubs in the network and
an overall decrease of path length. Additionally, we suggest that large parts of the network supporting episodic memory would remain intact following reconsolidation disruption, but details may be altered, or critical nodes or hubs might be lost resulting in 'catastrophic interference'. We have proposed that system consolidation contributes to the formation of abstract knowledge. As such we predict that over time memories become more abstract, more dependent on the neocortex and mPFC in particular. Additionally, over time memories lose details. However, if details of remote memories are recalled this generally involves the hippocampus. These suggestions are congruent with predictions from multiple trace theory, but conflict with standard systems consolidation theory. One possible consequence of our theory is that the application of abstract knowledge to novel situation increases as abstract knowledge loses its relation to specific details, to our knowledge this remains uninvestigated to date. We would however predict that abstract knowledge provides guidance to neural processing in these novel situations, reflected in effective connectivity from the mPFC to the hippocampus. Similarly, in schema learning studies the mPFC should exert effective control over the hippocampus when investigating the influence of abstract knowledge on the integration of novel information into memory. In our view the involvement of the hippocampus in memory is determined by the required information processing functions. If memories are retrieved that have become stored in a coherent neocortical network, these can be retrieved independent of the hippocampus. If, however, integration between memory representations that are not part of a single coherent neocortical network, or integration with novel incoming information is required, we predict that hippocampal involvement is critical. In our view neocortical memory representations are not 'semantic' in a classical way, as neocortical network can in principle also integrate details, but represent regularity structures. The involvement of the hippocampus in remote memory retrieval then does not reflect the ‘episodic-ness’ of a memory or vividness of a memory, but the interaction between abstract knowledge and mnemonic representations that are not part of the same neocortical representation. This view and these predictions conflict with predictions based on the multiple trace theory.
An important and unresolved issue is whether the hippocampus is necessary for the detection of regularities or is simply more optimized to do so. Studies on developmental amnesia suggest that such patients can acquire semantic knowledge (235), and thus that the hippocampus may not be essential for abstract knowledge formation. An interesting question is whether in these patients changes in neocortical function may allow the formation of abstract knowledge, for example through alterations in the degree of associative plasticity, and whether such changes persist in adulthood. Temporary inactivation of the hippocampus in animals by reversible genetic approaches or optogenetic methods may provide insight into this matter, as well as plasticity investigations in patients using transcranial magnetic stimulation.

As we have drawn an analogy between the lateral temporal-lateral PFC system and the hippocampal-mPFC system observations on the lPFC may also apply to the mPFC. One of these may be that within the mPFC a hierarchy exists where abstract rule representations show greater integration with motor option in more dorsal regions, and more anterior mPFC involvement is found with increasing complexity of abstract rules. The differentiation between the two systems that we have proposed is based on the type of information configurations processed. The lateral system can support mapping between unitary mnemonic representations and motor outputs, whereas mapping between complex, multimodal memory representations to motor outputs would require the hippocampus and mPFC. The involvement of the mPFC in memory tasks should thus follow the differentiation of tasks thought to be ‘hippocampally-dependent’. We would like to emphasize that we do not consider the distinction between the two systems to be a dichotomy, but envision extensive interactions between the two systems, and possibly even the existence of a gradient in the localization of information representations within the prefrontal cortex.

An issue we have repeatedly returned to throughout this review is a differentiation between synaptic weight changes and changes in neural rewiring. We have proposed that the weight changes support the rapid
storage of memories and details of memories. In contrast, we suggest that abstract knowledge formation relies more heavily on rewiring. This topic is scarcely investigated. However, several predictions can be made based on our proposal. First, abstract knowledge formation should suffer more from methods that impair rewiring, while leaving detailed memory formation relatively intact. Further, synaptic downscaling during sleep has been suggested to affect weight changes, and should thus reduce memory for details, but could be beneficial for abstract knowledge. Similarly, studies on wiring versus weight changes with regard to schema learning may answer the important question whether more rapid hippocampal independence of memory in the presence of a schema reflects a faster learning rate due to the presence of a schema, or simply 'less learning' as less neural rewiring has to be achieved to integrate a memory. Furthermore, with regard to reconsolidation, we suggest that memories that depend on weight changes are more likely to undergo reconsolidation. This means that abstract knowledge would be less sensitive to reconsolidation disruption than memory for details. Of importance here is that this entails that episodic memories that depend on a large distributed neural network are less likely to be fully altered during reconsolidation.

With regard to reconsolidation, very little is known about interactions between brain regions during reconsolidation, and to our knowledge no research has been conducted towards the involvement of the mPFC in reconsolidation. Reconsolidation may be initiated by mismatch detection and could be initiated by interaction between the mPFC and hippocampus. For reconsolidation of episodic memories abstract knowledge that depends on the mPFC may provide a prior probability over incoming information of a current episodic experience that is rapidly stored and subsequently compared by the hippocampus. As such interactions between the mPFC and the hippocampus may be critical for reconsolidation of episodic experience.

The proposal on hippocampus-mPFC interactions presented here arose from the intuitive assumption that episodic memory serves a
predictive function. However, few studies have provided evidence for this assumption, and we would welcome future attempts to investigate this matter.

In order to support prediction, we have suggested that memory is an adaptive dynamic system, and its flexibility may also be responsible for the distortions observed in the first author’s memory of ‘the incident’. In a final attempt to reconstruct this experience he also asked his aunt and uncle for a description of the event as both he, his brother, and father indicated them being present during that particular episode. This is their response:

"I've just consulted with your uncle, who we believe has a memory like an elephant, yet neither of us remember any of this."

Much to the first author’s surprise this either means that two people have completely lost their memories for the incident, or three people have a memory for an event that never happened. Whether true of false, at least the dynamic neural system that underlies adaptive, flexible episodic memory now allows the first author to predict one thing: Do not spit old people in the face.

Competing interests

The authors declare no conflict of interest.

Acknowledgements

This work was supported by a research grant from the Netherlands Organisation for Scientific Research (NWO) to G. Fernández.

Author contributions

MSc Kroes and Dr. Fernández wrote the manuscript.
Interlude C.

In Chapter 2, 3, and 4 we have provided results that indicate that PTSD and depression are accompanied by specific structural and functional alterations of the neural network that supports emotions and memory. In Chapter 5 we have shown that food targeted at raising Trp/LNAA levels can improve mood and affect neural processing in neurocircuits implicated in mood regulation. Although food or pharmacological treatment can improve mood and relieve symptoms of anxiety disorders it does not affect the root cause of such disorders. Considering that memory for emotional experiences contribute to the aetiology and persistence of memory one strategy would be to alter memory for such emotional experiences. In Chapter 6 we have explored the dynamic neural systems that enable flexible and adaptive memories. Insight into these dynamic systems provides us with target mechanisms to alter memory for emotional experiences. In Chapter 7-12 we will explore the possibility to alter memory for emotional experiences that may contribute to the aetiology or persistence of affective disorders.
7 EMOTION CAUSES TARGETED FORGETTING OF ESTABLISHED MEMORIES.

7.1 Published as:


7.2 Abstract

Reconsolidation postulates that reactivation of a memory trace renders it susceptible to disruption by treatments similar to those that impair initial memory consolidation. Despite evidence that implicit, or non-declarative, human memories can be disrupted at retrieval, a convincing demonstration of selective impairment in retrieval of target episodic memories following reactivation is lacking. In human subjects, we demonstrate that if reactivation of a verbal memory, through successful retrieval, is immediately followed by an emotionally aversive stimulus, a significant impairment is evident in its later recall. This effect is time-dependent and persists for at least six days. Thus, in line with a reconsolidation hypothesis, established human episodic memories can be selectively impaired following their retrieval.
7.3 Introduction

Following memory acquisition, there is a period during which memory traces undergo consolidation into long-term memory, a process involving synaptic protein synthesis and changes in gene expression (30). Until recently, a prevailing view was that once memory is consolidated, it is immune to manipulations that, prior to consolidation, impair its subsequent retrieval (30, 316). However, an early study (47) cast doubt on this account by demonstrating that electroconvulsive shock (ECS), known to impair consolidation (436), impairs expression of a previously consolidated memory if applied immediately after retrieval of that memory. This finding has been replicated and extended in animal models, leading to an hypothesis that retrieved memories must be reconsolidated in order to persist (49, 50, 437).

Convincing support for this phenomenon in humans is limited, primarily because amnesic agents such as protein synthesis inhibitors and ECS cannot be applied to healthy human subjects as experimental treatments (399). The existing studies supporting modulation of human memory following reactivation demonstrate impaired non-declarative forms of memory, namely fear conditioning (Schiller et al., 2010), fear potentiated startle (405, 406) or motor sequence learning (438). By contrast, reconsolidation impairment for specific declarative memories, in particular long-term memories for episodes that are accessible to conscious recollection (i.e. episodic memory), has not been demonstrated. Reconsolidation effects for human episodic memory are limited to decreased memory for a word list implied by a surrogate index of memory (439), or the integration of new list items into a previously learnt list following a reminder (407). Thus, these studies do not show selective impairment of a specific target episodic memory.

Recently, we described an emotion-induced retrograde amnesia (440), a manipulation that provides a potential technique for studying reconsolidation of specific human episodic memories. In brief, stimuli that precede an emotional event during encoding suffer a relative amnesia
Emotion causEs targEtEd forgEtting of EstablishEd mEmoriEs

in subsequent free or cued recall (440, 441). That is, whereas emotional (E) items show a well-described episodic memory enhancement (29), the preceding (E-1) items show a relative amnesia. Consequently, this manipulation (an emotional event) shares with interventions like ECS (436) and protein-synthesis inhibition (442) a fundamental ability to evoke retrograde disruption of memory.

Here, we modified the paradigm associated with emotion-induced retrograde memory disruption at encoding to demonstrate effects consistent with a reconsolidation interpretation. In two pilot studies, Experiments (Exp) 1 and 2, we established that the manipulation used in our reconsolidation experiments (Exp 3-5) could evoke emotion-induced retrograde disruption of verbal memory, as indexed by cued recall. In our previous studies, we employed aversive words as emotional stimuli. However, emotion-induced memory disruption is critically dependent on the amygdala (440) and the human amygdala shows strongest responses to affective facial expressions (443). On this basis, we elected to present fearful faces as the emotional stimuli in the current series of experiments. The critical finding from Exp 1 and 2 is that retrograde amnesia is evident 24 h after encoding, but not if recall is cued immediately after the study phase, indicating an effect on consolidation. Given the overlap in treatments that disrupt consolidation and reconsolidation (444), this paradigm provides a framework for selective impairment of a target episodic memory.

The proposal that retrieved memories become labile and must undergo reconsolidation, as expressed in the reconsolidation hypothesis, is highly controversial. Several studies have not replicated a post-retrieval memory impairment effect, while others have suggested alternative explanations for findings labelled as reconsolidation, particularly following observations of complete memory recovery within days (445-447). Animal experiments have led to specific criteria in order for memory disruption to be referred to as a reconsolidation impairment (44). Memory must be disrupted following reactivation, as indexed in a subsequent memory test (49). The impairment should not be attributable to retrieval failure.
or a reactivation-locked, temporary inability to access memory traces that dissipates over time (448). Impairment should, however, be time-dependent and not expressed in tests of immediate memory (50). Finally, the memory impairment must not be due to impaired novel encoding of any attribute of reactivated memories (449). We satisfy these criteria for disruption of reconsolidation within a series of human experiments (Exp 3-5) where we demonstrate emotion-induced disruption of memory following successful reactivation.

7.4 Methods and Materials

7.4.1 Participants

A total of 89 native English-speaking subjects completed Exp 1-5. All subjects gave informed consent and were free of neurological or psychiatric history. The study had full ethics approval. Twenty (10 males, 10 females, age-range: 19-37, mean: 25.6) and 14 (7 males, 7 females, age-range: 20-31, mean: 25.6) subjects completed Exp 1 and 2, respectively. Twenty, 15 and 20 subjects completed Exp 3, 4 and 5, respectively. We have previously shown that in subjects not familiar with the neutral nouns presented in these studies, the chance of correctly completing an item in the employed stimulus set is approximately 11% (411). On this basis, a performance criterion was set for our reconsolidation studies of recall hit rate minus false alarm rate greater than 10% for control nouns on Day 2. Thus, data from 16 subjects (8 males, 8 females, age-range: 21-35, mean: 27.2) were included in Exp 3, 11 subjects (5 males, 6 females, age-range: 22-30, mean: 25.6) in Exp 4 and 14 subjects (5 males, 1 left-handed, 9 females, age-range: 20-28, mean: 22.6) in Exp 5.

7.4.2 Stimuli

Verbal stimuli comprised a set of 400 nouns which were ‘stem-unique’ in that the first 3 letters were different for each of the nouns. Faces were selected from the Karolinska Directed Emotional Faces set (450), converted to gray scale and framed to exclude non-facial features.
An equal number of female and male faces were presented in each experiment.

7.4.3 Behavioural tasks and procedures

7.4.3.1 Exp 1

Encoding and cued recall sessions were conducted in the same environment, 24 h apart. On Day 1, subjects viewed 240 nouns, taken from the set of 400, and made a push-button response to indicate whether the noun described a living or nonliving entity. A face was presented after a random number of words (between 3 and 5; mean = 4) upon which
A  Day 1. Encoding

Day 2. Cued Recall

Day 3. Cued Recall

Week 2. Cued Recall

B

Day 1  Day 2  Day 3  Week 2
Exp 3: Encode  Test 1  Test 2  Test 3
Exp 4: Encode  Test 1  Test 2  Test 3
subjects were instructed to press a third button. Nouns and faces were each presented for 1 s, with a stimulus onset asynchrony (SOA) of 4 s. A total of 60 faces were presented, 30 neutral (N) and 30 fearful (E). Nouns were conditionalised according to their position relative to face presentation and labelled emotional (E) or neutral (N) ± 1, i.e. a noun that immediately preceded an emotional face is referred to as E-1. Nouns without positional assignment served as controls (C), with the constraint that these controls were more than 2 nouns before, and more than 1 noun after, a given face (Figure 7.1A). On Day 2, word stems were presented every 4 s (stimulus duration, 1 s) and subjects instructed to complete the stems to make a word from Day 1. In all experiments, nouns and word stems were presented in uppercase in random order.

7.4.3.2 Exp 2

This was identical to Exp 1, except that the encoding session was followed immediately by the cued recall task.

7.4.3.3 Exp 3

This experiment was conducted in the same environment, at the same time of day, across 4 separate days. On Day 1, subjects viewed 400 nouns at a rate of one every 4 s (stimulus duration, 1 s), and made a push-button response to indicate whether the noun described a living or nonliving entity. To promote retrieval success across the ensuing cued recall tests, the encoding task was repeated a total of 3 times. On Day 2, word stems were presented (stimulus duration 1 s, SOA 4 s) and subjects instructed to complete the stems out loud to make a word from Day 1, while avoiding guessing. A face was presented after a random number of word stems (between 3 and 5; mean = 4) upon which subjects were instructed to make a button press. A total of 80 faces were presented, 40 neutral (N) and
40 fearful (E). Day 3 and Week 2 (test day one week after Day 1) followed the same procedure as Day 2 except that no faces were presented (Figure 7.2A).

Word stems were conditionalised according to their position relative to face presentation on Day 2 and labelled emotional (E) or neutral (N) - 1, i.e. a stem that immediately preceded an emotional face is referred to as E-1. We also examined memory for E-2 and N-2 stems. Word stems without positional assignment served as controls (C), with the constraint that these controls were more than 2 word stems before, and more than 1 word stem after, a given face. Performance on Day 3 and Week 2 is expressed as the proportion of remembered Day 3 / Week 2 words that were remembered on Day 2.

7.4.3.4 Exp 4

This was identical to Exp 3 except that the experiment was conducted over 3 days, with Day 2 consisting of 2 cued recall sessions (Test 1 and Test 2) with Test 3 on Day 3 (Figure 7.2B).

7.4.3.5 Exp 5

This was identical to Exp 3 except that only 100 nouns, taken from the set of 400, were presented on Day 1. Furthermore, on Day 2 only 20 faces were presented, 10 neutral and 10 emotional, with each face presented above a neutral or emotional noun, respectively (Figure 7.6A). The neutral nouns presented with neutral faces were taken from the stimulus set of 400 nouns. The emotionally aversive nouns presented together with fearful faces were selected from a set used in our previous studies (33). The first 3 letters of these emotional nouns were different from the 100 nouns presented on Day 1. Given the lower number of critical items (E-1, N-1) in this experiment (only 10 emotional and neutral faces were presented on Day 2), we applied a further performance criteria for Exp 3. In view of the conditionality of Day 3 performance on Day 2, if only one E-1 or N-1 stem was recalled on Day 2 (i.e. 10% hit rate) the corresponding
conditional performance on Day 3 was excluded from our analyses. Thus 1 subject’s hit rate for E-1 stems, and 2 further subjects for N-1 stems, were excluded (E-1 effects remained significant if this E-1 hit rate was included).

7.4.4 Statistical analyses

In view of our strong a priori prediction that cued recall of E-1 nouns would differ relative to control and N-1 nouns, statistical analyses were constrained to planned comparisons. Thus, for all experiments we report separate paired t-tests comparing E-1 vs. control and N-1 cued recall. One-tailed significance is reported on the basis of the prediction that memory for E-1 nouns would be impaired. To test for persistence of reconsolidation effects over 1 week (Exp 3 and 5) we report separate word type x day (Day 3, Week 2) 2 x 2 ANOVAs for E-1 vs. control and N-1 nouns. For Exp 3-5 we did not predict any effect of face presentation on the first test of cued recall (Day 2). To demonstrate that this was indeed the case, we report, for Exp 3-5, emotion (E, N) x position (-2, -1, +1) 2 x 3 repeated measures ANOVAs for % Hits relative to controls on Day 2. Although our predictions pertain to E-1 cued recall, we include memory performance for E-2, N-2, E+1 and N+1 nouns in our plots of reconsolidation effects for descriptive purposes.

7.5 Results

7.5.1 Emotion-induced retrograde disruption of consolidation

In our previous studies of emotion-induced retrograde memory impairment at encoding (440), 14-word lists were presented with a stimulus onset asynchrony (SOA) of 3 s, the emotional (E) stimulus was an emotionally aversive noun and memory was assessed using free recall after a 30 s filled delay which followed each list. The nature of the reconsolidation task (Exp 3-5) required a longer SOA because presentation of the E stimulus occurred during cued recall, instead of verbal encoding. Thus, we employed an SOA of 4 s in the recall task to allow completion of cued recall before presentation of the E stimulus (a
fearful face). Because of the differences between this and our previous paradigm (440), we conducted a pilot study (Exp 1) to first establish that fearful faces, presented 4 s after the verbal stimulus, evoke emotion-induced retrograde amnesia of verbal encoding, as indexed by cued recall. Given that our main reconsolidation experiments were to be conducted on successive days, the interval between encoding and cued recall in this pilot study was 24 h.

On Day 1, healthy human subjects performed a semantic encoding task on visually presented ‘stem-unique’ nouns. The critical manipulation was that some of these nouns were followed by the visual presentation of a face. Faces could be either emotional (fearful) or neutral. Neutral (N) faces controlled for non-specific effects of face presentation during encoding. On Day 2, word stems, the first 3 letters of previously encoded words, were presented and subjects instructed to complete the stems out loud to make a word from Day 1 (Figure 7.1A). Figure 7.1B demonstrates a significant emotion-induced retrograde amnesia (i.e. E-1 effect) when cued recall is tested after a 24 h delay. Planned comparisons demonstrate a relative decrement of cued recall for E-1 cues (relative to control cues paired t-test t(19) = -3.181; p = 0.003 one-tailed; relative to N-1 cues paired t-test t(19) = -2.582; p = 0.009 one-tailed). Mean control noun recall was 23.3 % (s.e.m. 1.5).

Successful demonstrations of memory reconsolidation (50, 427, 446) use learning paradigms in which both consolidation and reconsolidation processes could be conclusively affected. To test whether emotion-induced amnesia induced at encoding with the experimental parameters used in Exp 1 reflects an effect consistent with impaired consolidation, we conducted a further pilot experiment. Consolidation impairment is time-dependent, and not expressed in tests of immediate memory retrieval (30). Thus, Exp 2 was identical to Exp 1 except that cued recall was tested immediately after encoding. Figure 7.1 C demonstrates that E-1 memory disruption is not present in immediate tests of cued recall. Planned comparisons confirm no significant effect of cued recall for E-1 cues relative to control cues (paired t-test t(13) = 0.545; p = 0.298 one-tailed)
or relative to N-1 nouns (paired t-test $t(13) = 0.000; p = 0.500$ one-tailed). Mean control noun recall was 28.5% (s.e.m. 2.1).

Thus, E-1 effects when tested by cued recall were only observed following a 24 h delay, and not on a test of immediate memory, an observation consistent with an emotion-induced retrograde disruption of consolidation. This finding is in contrast to our previously reported emotion-induced amnesia observed under free recall following a 30 s delay (440). This indicates that the emotional stimulus, in addition to disrupting consolidation, evokes an immediate interference with an additional process required for successful free recall. It is noteworthy in this context that in our original study (440), nouns were presented in semantically-related lists, suggesting that the immediate effects at free recall may reflect an emotion-induced disruption of a category-cued retrieval strategy for the E-1 noun.

7.5.2 Memory disruption following reactivation

Figure 7.2A illustrates the experimental protocol for Exp 3. On Day 1, healthy human subjects performed a semantic encoding task on visually presented verbal stimuli. On Day 2, subjects returned to perform a cued recall task. They were presented with stems of previously encoded nouns and instructed to complete the stems out loud. The critical manipulation was that some of these stems were followed by the visual presentation of either an emotional (fearful) or neutral face. Our working hypothesis was that emotional (E) faces would disrupt reconsolidation of the preceding retrieved word (E-1 words), indexed in a selective retrieval impairment on subsequent tests of cued recall. Neutral (N) faces controlled for non-specific effects of face presentation during the cued recall task. In view of the novel approach to studying reconsolidation in this paradigm, and unknown temporal profile of any potential effect, we also include cued recall performance for E-2 and N-2 word stems in the plots we present.

We did not expect an effect of face presentation on cued recall on Day 2, and Figure 7.3A demonstrates that this prediction was confirmed. A word
type (E, N) x position (-2, -1, +1) 2 x 3 repeated measures ANOVA (for % Hits relative to controls) revealed no significant main effect or interaction. The small decrement for E-1 relative to control is non significant (paired t-test $t(15) = -1.297; p = 0.214$ two-tailed), ensuring equal memory performance for the different word types immediately prior to the reconsolidation disrupting manipulation. To test whether emotional face presentation on Day 2 disrupted the memory trace of E-1 words, we repeated the cued recall task on Day 3. The memory task was identical to Day 2 except that faces were not presented.

Consistent with a reconsolidation hypothesis, a significant percentage of target nouns, i.e. those that preceded emotional faces (E-1 nouns), that were correctly recalled on Day 2 were no longer recalled on Day 3. Thus, on Day 3 there is a selective impairment of cued recall for E-1 cues.
relative to control noun cues (Figure 7.3B and C). Critically, this deficit is not observed for N-1 nouns, i.e. cues that on Day 2 were followed by a neutral face, indicating that it is the fearful emotional facial expression, and not the presentation of a face per se, that leads to impaired E-1 reconsolidation. Planned comparisons confirm the predicted relative decrement of cued recall for E-1 cues (relative to control cues paired t-test \( t(15) = -2.157; \ p = 0.024 \) one-tailed; relative to N-1 cues paired t-test \( t(15) = -1.880; \ p = 0.040 \) one-tailed). Thus, recall probability of an item previously correctly recalled can be specifically attenuated if, during memory re-activation, it is immediately followed by an emotional stimulus.

7.5.3 Veridical memory retrieval is required for emotion-induced subsequent memory impairment

To investigate the specificity of reconsolidation, we next tested whether a retrieval attempt alone immediately prior to an emotional stimulus is sufficient to produce impairment in subsequent memory testing. A small subset of words forgotten on first testing (Day 2) show spontaneous cued recall on second testing (Day 3). If a mere retrieval attempt on Day 2 for E-1 cues leads to impaired reconsolidation then it follows that recall of these words should be impaired on Day 3. Miss rates on Day 2 for E-1 and control cues were both 57.0% (SE 4.2 and 4.0, respectively). The proportion of these Day 2 misses that were correctly recalled on Day 3 was (mean, s.e.m.) 10.1% (1.9) and 8.5% (1.4) for E-1 and control nouns, respectively. A paired t-test confirmed no significant difference (\( t(15) = 0.864; \ p = 0.40 \) two-tailed). Thus, veridical memory reactivation, and not retrieval attempt, is required for emotion-induced retrograde impairment of reconsolidation.

7.5.4 Post-reactivation memory impairment persists over time.

Animal studies report inconsistent findings as to the durability of reconsolidation impairment with some reporting sustained reconsolidation impairment while others demonstrate spontaneous
In terms of the neurobiology of reconsolidation, reversibility favours a retrieval or performance interpretation, whereas lack of reversal would support a storage deficit (446, 449, 452). We therefore repeated the test of cued recall one week after the initial encoding session (Week 2). Consistent with this effect reflecting a persistent long-lasting deficit in reconsolidation, target words, whose cues were followed by an emotional face on Day 2, show impaired recall at Week 2 (Figure 7.4). A word type (E-1, Control) x day (Day 3, Week 2) ANOVA for % Hits demonstrates a significant effect of day $F_{1,14} = 7.848$, $p = 0.014$, an effect of word type at trend $F_{1,14} = 3.983$, $p = 0.066$, and, critically, no word type x day interaction. Similarly, a word type (E-1, N-1) x day (Day 3, Week 2) ANOVA for % Hits relative to control noun recall demonstrates an effect of word type at trend $F_{1,14} = 3.396$, $p = 0.087$, and no word type x day interaction. Thus, the effect we observe endures for at least one week (Exp 3).

Figure 7.4: Memory impairment for E-1 cues persists for at least one week (Exp 3). Cued recall (% Hits) on Day 3 and Week 2 for those word stems correctly recalled on Day 2. E-1 (red •), control nouns (black □).
Figure 7.5: Reconsolidation is time-dependent (Exp 3 and 4).
Cued recall (%) for Exp 4 (dashed lines) on Test 2 (Day 2) and Test 3 (Day 3) for those word stems correctly recalled on Test 1 (Day 2). Critically, there is no emotion-induced impairment of reconsolidation if tested immediately after induction. Performance from Exp 3 (solid lines) is plotted for comparison. E-1 (red •) and control (black □). The day during which Tests 1-3 were performed in Exp 3 and 4 is indicated below.

least 6 days and does not represent transient memory impairment for E-1 stimuli, arguing against a retrieval or performance deficit.

7.5.5 Memory impairment following reactivation is time-dependent

As stated above, for post-reactivation memory disruption to be judged as reconsolidation, the effect must not be due to impaired novel encoding of reactivated memories (449). Thus, it might be argued that the observed impairment simply reflects recalled E-1 words not benefiting from a second 'encoding' when retrieved on Day 2. Indeed, a counter argument to the reconsolidation hypothesis states that the reactivation of memory causes a second distinct memory trace to be formed (321, 449). To control
for this possibility we conducted a further experiment (Exp 4), identical in design to our previous experiment, except that on Day 2 subjects performed a second cued recall session immediately after the first, i.e. the initial cued recall session (Test 1) on Day 2, during which emotional and neutral faces were presented, was immediately followed by the cued recall session (Test 2) that, in Exp 3, occurred on Day3 (Figure 7.2B). Cued recall on Test 1 was, as in Exp 3, not affected by face presentation. A word type (E, N) x position (-2, -1, +1) 2 x 3 repeated measures ANOVA for % Hits relative to controls in Test 1 revealed no significant main effect or interaction.

Encoding is defined as the rapid acquisition of a memory trace, a process occurring prior to completion of consolidation (30, 316). Thus, if the impaired recall of E-1 stems observed on Day 3 in Exp 3 is simply attributable to lack of secondary encoding, the same effect should be present in the second test (Test 2) of cued recall in Exp 4. Figure 7.5 demonstrates that this was not the case. There is no decrement for E-1 stems on Test 2 if this comes immediately after Test 1 (E-1 relative to control noun stems paired t-test t(10) = -0.167; p = 0.436 one-tailed; relative to N-1 recall paired t-test t(10) = -0.333; p = 0.373 one-tailed). Thus, impaired recall of E-1 words is present 24 h after the emotional manipulation, supporting a claim that this reflects an emotion-induced impairment of reconsolidation and not a non-specific effect consequent upon impaired re-encoding.

7.5.6 Replication of initial findings

To demonstrate the robustness of E-1 reconsolidation impairment, we conducted a further experiment (Exp 5), identical to Exp 3 except that we now present fearful faces paired with an emotionally aversive noun (Figure 7.6A). Reliable emotion-induced amnesic effects are evoked by aversive nouns (440) and picture-noun pairings (453). We also decreased the number of words encoded on Day 1 to increase cued recall performance on Day 2, thereby increasing our sensitivity to memory performance on subsequent testing.
7. Emotion causes targeted forgetting of established memories

As in Exp 3, there was no significant effect of facial emotional expression or position on Day 2 performance (Figure 7.6B). An emotion (E, N) x position (-2, -1, +1) 2 x 3 repeated measures ANOVA (for % Hits relative to controls) on Day 2 revealed no significant main effect or interaction. As shown in Figure 7.6C, the pattern of emotion-induced reconsolidation impairment on Day 3 is identical to that observed in Exp 1 except that we now observe a deficit approximately two-times that seen in Exp 1. Planned comparisons confirm the predicted relative decrement of cued recall for E-1 cues (relative to control cues paired t-test t(12) = -2.147; p = 0.026 one-tailed).
Figure 7.6D provides further evidence that the reconsolidation impairment we observe persists over time. A word type (E-1, Control) x day (Day 3, Week 2) ANOVA for % Hits demonstrates a significant effect of word type \( F_{1,12} = 9.470, p = 0.010 \), and no significant main effect of day or word type x day interaction. A further word type x day ANOVA comparing E-1 and N-1 % Hits also demonstrates a significant effect of word type \( F_{1,10} = 5.795, p = 0.037 \). [N-1 % Hits (s.e.m.) for those word stems correctly recalled on Day 2 were 93.5 (4.3) for Day 3 and 83.1 (6.7) for Week 2]. Thus, the effect is not attributable to the presentation of a face-word pair per se as neutral pairings did not affect memory on Day 3 or Week 2.

7.5.7 Effect size

Whereas the effect size [Cohen's \( d = (\text{MeanControl} - \text{MeanE-1}) / \text{pooled standard deviation} \)] for E-1 memory impairment relative to control stems on Day 3 in Exp 3 is equal to 0.52, indicating a moderate effect, the corresponding effect size for Exp 5 is 0.87, demonstrating that this is a large effect. Corresponding effect sizes relative to N-1 cues are 0.61 and 0.82 for Exp 3 and 5, respectively.

7.5.8 False alarms

The false alarm rates for Exp 3-5 are given in Tables 1 - 3. Employing the Hit - False alarm rate as the dependent variable in our statistical analyses yielded similar results as described for Hit rate alone. For Exp 3, a word type (E-1, Control) x day (Day 3, Week 2) ANOVA for % Hits minus False alarms demonstrates a significant effect of day \( F_{1,14} = 10.065, p = 0.007 \), an effect of word type at trend \( F_{1,14} = 3.655, p = 0.077 \), and no word type x day interaction. The same test in Exp 5 reveals a significant effect of word type \( F_{1,13} = 6.800, p = 0.023 \).

7.5.9 Antrograde effects of emotional stimuli

In our original report of impaired E-1 noun free recall (440), emotional stimuli were aversive words presented in a neutral word list with a SOA
of 3 s. In this study (440), we observed no difference in recall of E+1 nouns relative to neutral controls. A further study employing the same experimental paradigm (441) also reported no E+1 memory impairment. However, a similar study, employing a SOA of 5 s, reported significantly impaired recall of stimuli presented immediately after aversive picture-noun pairings (453). The current experiments testing for the effects of emotional stimuli at encoding (Exp 1 and 2) employed a SOA of 4 s, i.e. intermediate between the SOAs used in previous studies. It is therefore interesting that recall performance for E+1 cues is numerically less than that for control nouns at both a delayed test (Exp 1, Figure 7.1B) and immediate test (Exp 2, Figure 7.1C) of cued recall. Although this apparent cued recall decrement does not reach significance in either Exp 1 (E+1 relative to control cues paired t-test t(19) = -1.328; p = 0.200 two-tailed) or Exp 2 (E+1 relative to control cues paired t-test t(13) = -1.120; p = 0.283 two-tailed), this observation suggests that SOAs longer than 3 s are more likely to lead to E+1 memory impairment if the emotional stimulus is presented at encoding. By contrast, Figures 6.3C and 6.6C demonstrate that if the emotional stimulus is presented immediately before successful retrieval of E+1 nouns, there is essentially no decrement of E+1 noun recall on subsequent testing (i.e., no reconsolidation impairment for E+1 nouns). This may, therefore, indicate greater resistance of established memories to anterograde disruption of reconsolidation evoked by an emotional stimulus, relative to anterograde disruption evoked at encoding.

7.6 Discussion

We demonstrate selective impairment in reconsolidation of target episodic memories following successful reactivation. In a series of experiments, we systematically address the criteria generated by animal models for memory disruption to reflect reconsolidation impairment. We show that impaired reconsolidation occurs if an emotionally aversive stimulus is presented immediately after successful reactivation of a memory. Successful memory reactivation is critical, in that retrieval attempt is insufficient to impair subsequent cued recall. This retrograde effect is long-lasting, enduring at least one week, arguing against a
retrieval or performance deficit. Critically, the reconsolidation impairment we report is time-dependent in that it is observed after a 24 h delay, but not in immediate tests of cued recall. Our data therefore provide evidence that is consistent with reconsolidation of episodic memory, and its disruption, in humans.

Reconsolidation is postulated to be an adaptive update mechanism by which new information is incorporated into old memories (48, 403, 407, 422, 444, 447, 454); but see (455). Indeed, reactivation of an established episodic memory has been shown to modify its content (407, 456). We suggest that the reconsolidation effects reported here reflect a process whereby disrupting the integration of retrieved nouns by the presentation of an emotional face corrupts a pre-existent memory trace (403, 447). This is not simply disrupted re-encoding of recalled items because the reported reconsolidation impairment is time-dependent, being detectable only 24 h after it is induced, and not immediately following induction.

The exact mechanism by which a fearful face corrupts a pre-existing verbal memory remains to be determined. Our previous studies on emotional memory encoding demonstrate that impaired memory for E-1 items is coupled to enhanced memory for the emotional stimuli (440). These effects are critically dependent on both amygdala and central adrenergic activation, suggesting that while memory is enhanced if adrenergic activation occurs at the time of encoding, it may be impaired if occurring 3-6 s after the initial encoding event (440). If an analogous mechanism mediates the reconsolidation impairment observed in the current studies, this would predict enhanced memory for the fearful faces themselves. We tested this hypothesis in a further experiment (Supplementary Exp 1, see supplemental information) which followed a similar experimental protocol to Exp 5, with the addition of a surprise recognition memory test on faces presented on Day 2. Consistent with our hypothesis, reconsolidation impairment of E-1 word stems was associated with enhanced recollection of fearful, but not neutral, faces (see supplemental information).
We therefore suggest that the reconsolidation impairment we observe is mediated by an amygdala-dependent adrenergic release, evoked by fearful face presentation, which boosts face memory while corrupting a pre-existent memory trace reactivated sometime in the preceding 4 s. Validation of this proposed mechanism will require a demonstration that pharmacological blockade of the adrenergic system blocks reconsolidation impairment and enhanced memory for fearful faces. It is unlikely that reconsolidation impairment reflects physiological arousal evoked by fearful faces. Emotion-induced memory enhancements do not require peripheral adrenergic engagement (457). Furthermore, previous data demonstrate that although fearful faces elicit robust amygdala responses (443, 458), they do not elicit reliable increases in physiological measures of arousal relative to neutral faces (459). We suggest instead that the impaired reconsolidation effects we observe are likely to be mediated by rapid central adrenergic release within the amygdala.

An interesting additional observation from Exp 4 (Figure 7.5) is that normal retrieval of E-1 cues during Test 2 immediately after our reconsolidation disrupting manipulation (Test 1), prevents the emergence of impaired reconsolidation 24 h later (Test 3). There are two potential explanations for this finding. Firstly, correct retrieval at Test 2 may re-stabilise the memory trace before emotion-induced retrograde disruption of reconsolidation is complete. Alternatively, performance for E-1 cues on Test 3 might be supported by secondary encoding, the process whereby during every retrieval episode, retrieved mnemonic information, together with incoming sensory information, is stored as a novel memory trace (319). This latter explanation suggests that, although reconsolidation of the original verbal memory is disrupted, normal levels of retrieval at Test 3 reflect cued recall of the novel memory trace formed at Test 2.

A recent study reports a manipulation which completely impairs human fear conditioned memory following reactivation, while leaving declarative memory for the acquired contingency between the conditioned and unconditioned stimulus intact (405). This suggests that declarative memories are more resistant to reconsolidation disruption relative to
non-declarative forms. Whereas conditioned fear could reflect a discrete neural representation localised to the amygdala, declarative memories are thought to comprise a distributed cortical representation (316). This distributed representation may explain why our manipulation partially impairs reconsolidation, as opposed to the complete impairment following reactivation observed in human fear conditioning studies (405, 406), even though reactivation of distributed memory traces has been reported (429). However, despite the relative resistance of declarative memory to reconsolidation impairment, our manipulation achieves approximately 20% reduced retrieval probability of a target episodic memory by the second week of testing (Figure 7.6D). A similar effect magnitude was observed in a recent report suggesting reconsolidation of recalled autobiographical memory (410). In this study, new learning of Bartlett’s "War of the Ghosts" narrative after the reactivation of a personal, autobiographical neutral experience impaired subsequent memory for that experience. Interestingly, the "War of the Ghosts" narrative has been shown to be mildly emotionally arousing (460), which may, in view of the effects we report, have contributed to the reconsolidation impairment reported by (410).

A further factor contributing to the smaller magnitude of reconsolidation impairment described here, relative to that observed in tests of non-declarative memory (405, 406), is inter-subject variability. Averaging over subjects we observe significant reconsolidation impairment in both Exp 3 and 5. However, of the 16 subjects included in our analyses of Exp 3, 10 showed reconsolidation impairment (mean 12.9% decrease relative to control noun recall on Day 3), 1 subject showed no differential E-1 memory and 5 showed a slight E-1 memory enhancement. Similarly, of the 14 subjects included in Exp 5, 9 showed reconsolidation impairment (mean 27.8% decrease relative to control noun recall on Day 3) whereas 2 subjects showed no differential E-1 memory and 3 showed E-1 memory enhancement. This variability mirrors the discrepancies between studies reporting an impairing retrograde effect of emotion on memory encoding (440, 441, 453) and those demonstrating a retrograde enhancement (461, 462). With respect to E-1 effects at encoding, it has been shown that
the magnitude of memory impairment depends on subjects’ genotype (463), level of anxiety trait (441), and gender, with females showing greater impairment than males (440). Whether these factors also influence emotion-induced retrograde reconsolidation impairment is yet to be determined. Both males and females were included in the current experiments, (and both showed reconsolidation impairment) although sample sizes were insufficient to make inferences on gender effects. As described above, the time interval between neutral and emotional stimuli, i.e. SOA, may influence the magnitude of reconsolidation impairment, which could in turn interact with other variables such as genotype, temperament and gender. Furthermore, future studies which measure the degree of arousal elicited by specific emotional stimuli, and how this varies between subjects, will also contribute to our understanding of inter-subject variability in E-1 reconsolidation impairment. It is unlikely in the current study that individuals’ general memory ability contributed to inter-individual differences in E-1 effects. The correlations between control noun cued recall on first testing (Day 2) and E-1 memory relative to control nouns on Day 3 were not significant in Exp 3 (Pearson’s r = 0.02; p = 0.930) or Exp 5 (Pearson’s r = -0.02; p = 0.948).

Reconsolidation is proposed to provide a principled basis for therapeutic interventions in conditions such as post-traumatic stress disorder (PTSD), where the aim is to selectively decrease the strength of aversive memories for traumatic events (44, 404, 464, 465). Although we address reconsolidation of neutral as opposed to aversive memory, the critical observation here is that reconsolidation of specific human episodic memories can be selectively impaired. Moreover, this can be achieved via a non-invasive psychological manipulation. Future studies might usefully extend these observations to the context of human emotional episodic memory, with testing intervals longer than 1 week necessary to determine the clinical potential of these effects. In conclusion, we demonstrate a targeted, selective reduction in established episodic memory, a finding that provides a potential framework for treatments aimed at reducing unwanted human memories.
Competing interests

The authors declare no conflict of interest.

Acknowledgements

This research was supported by a Wellcome Trust Programme Grant to R.D. We thank M Rugg for assistance with stimuli.

Author contributions

Dr. Strange and MSc Kroes designed of the study. MSc Kroes, Dr. Strange, and BSc Fan acquired the data. Dr. Strange and MSc Kroes analyzed the data. Dr. Strange, MSc Kroes and Dr. Dolan wrote the manuscript.
7. EMOTION CAUSES TARGETED FORGETTING OF ESTABLISHED MEMORIES
8 EVIDENCE FOR RECONSOLIDATION OF EMOTIONAL EPISODIC MEMORY IN HUMANS.

8.1 Under review as:


8.2 Abstract

Despite accumulating evidence for a reconsolidation process in animals, support in humans, especially for episodic memory is limited. Here we show in a within subjects-design that a single application of electroconvulsive therapy following memory reactivation in patients with unipolar depression disrupts reactivated, but not non-reactivated, memories for an emotional episode in a time-dependent manner. Our results satisfy critical criteria to provide evidence for reconsolidation of emotional episodic memories in humans.
8.3 Introduction

The reconsolidation hypothesis postulates that upon reactivation, memory traces return to a labile state and are once again susceptible to disruption by treatments similar to those that impair initial memory consolidation. The criteria to provide compelling evidence for the reconsolidation phenomena generated from non-human animal studies comprise 1) consolidated memory must be reactivated by a reminder cue, 2) the manipulation aimed at altering reconsolidation must be provided post-reactivation not prior, and 3) reconsolidation is a time-dependent process and therefore, memory should not be affected immediately but after a time window allowing reconsolidation to take place - usually tested after 24 h (50, 446, 449). Non-human animal studies on reconsolidation have generally employed invasive techniques to disrupt memory traces, consequently only lesser invasive methods have been translated to humans. Due to the use of slow-acting pharmacological compounds, or psychological manipulations such as interference learning, human reconsolidation studies often do not meet criteria 2 and 3 leaving them susceptible to alternative explanations, such as secondary encoding or source confusion (409, 411, 413, 466). The use of electroconvulsive therapy (ECT) to study reconsolidation in humans would circumvent these concerns. ECT has historically provided considerable insight into the nature of memory consolidation (467) and pioneering rodent studies have indicated that reactivated memory traces may be disrupted by electroconvulsive shock (47).

We therefore tested for an ECT-evoked disruption of episodic emotional memory reconsolidation in 42 patients with unipolar depression undergoing ECT. Patients were randomly assigned to one of three groups (Figure 8.1., Supplementary Table 8.1). One week after learning two emotionally aversive slide-show stories (Supplementary Figure 8.1.), memory for one of the two stories was reactivated. Immediately following memory reactivation, patients in groups A and B were anaesthesitized and received ECT, whereas patients in group C did not. We hypothesized that ECT would disrupt memory of the reactivated relative to the non-
reactivated story in group A when tested one day after reactivation and ECT, but not in group B when tested ~90min after reactivation and ECT.

### 8.4 Methods and Materials

#### 8.4.1 Participants

Over 3.5 years we investigated forty-two patients diagnosed with unipolar depression classified according to the DSM-IV-TR (468), who were undergoing ECT treatment and either at the end of an acute treatment cycle or receiving maintenance ECT (age-range: 34-84 yr, mean: 57.16 yr, s.e.m.: 3.96 yr). The study was approved by the institutional ethics committee (CMO Regio Arnhem-Nijmegen) and all patients provided written informed consent. Patients were pseudo-randomly assigned to one of three groups, equated for demographic, cognitive and clinical parameters (Figure 8.1, Supplementary Table 8.1). Patients were either medication free or under stable treatment throughout this study, had normal- or corrected-to-normal vision and hearing. Patients were free
of current or past relevant somatic or neurological disorders, and were free of a diagnosis of bipolar depression, schizophrenia or substance dependence disorders, and had a global assessment functioning score (GAF) >40. Patients were recruited from the Radboud University Nijmegen Medical Centre, and the Rijnstate Hospital in Arnhem, the Netherlands.

8.4.2 Stories

During an initial learning session, patients were shown two slide-shows depicting emotionally aversive stories. The to be memorized material consisted of slide shows accompanied by an auditory narrative forming an episode. To maximize memorability and ecological validity for clinical
settings, we selected high-arousing stories with negative valence. One story was identical to the one of Cahill and colleagues (469), adapted to the Dutch language (457, 469) and consisted of 11 slides. Of note, although a neutral version exists also, we only used the negative arousing version of the "Cahill Story". A second slide-show story was developed identical in structure, in valance and arousal ratings of the critical emotional slide, and matched in details (presence of human beings, objects/buildings, background, positions, etc.). Additionally, the narrative is identical in structure, grammar, arousing/neutral words, similar in word and syllable count, and voice emphasis (Figure 8.2). To reduce interference caused by learning two stories shortly after each other, and to allow specific reactivation of one story, we ensured that the two stories were clearly distinct. The new story consists of modern digital photographs, and the narrative is spoken by a female voice (the "Cahill Story" uses images from a more remote time and a male voice). Whereas the main characters in the "Cahill Story" are a mother and her son, the new story involves two twenty-something sisters. Presentation duration of the slide shows is 5 min each.

8.4.3 Memory reactivation

Following previous human reconsolidation studies, subjects received a brief memory reactivation to initiate reconsolidation (405, 406, 470-472). One week after the initial study session, memory for one of the two stories was reactivated by presenting the first slide of this story. To reactivate memory and initiate a reconsolidation process for one of the two stories, patients were presented with the first slide of one of the two stories. Parts of this slide, however, were masked by black-and-white checkerboard patterns. Patients were asked three questions on what was visible behind the mask. Upon answering a question, the related part of the mask was removed, until the entire slide was visible after three questions were answered. Answers were provided by free recall. If the patient was unable to answer freely, a two-alternative forced choice question was posed. Reactivation score was calculated as follows: number
of questions answered correctly by free recall * 2 + number of correctly answered questions by multiple choice.

8.4.4 Electroconvulsive stimulation

Immediately following memory reactivation, patients in groups A and B were anaesthetized and received ECT, whereas patients in group C did not. Electroconvulsive stimulation was administered right unilateral or bifrontotemporal with a brief pulse (0.25-0.5 ms), constant current (0.9 Ampere) apparatus with a maximum stimulus output of 1008 milliCoulombs, i.e. 200% (Thymatron System IV, Somatics Inc., IL, USA). The stimulus dosage was set at 2.5 times the initial seizure threshold (i.e., the stimulus dosage that elicited a motor seizure of at least 20 seconds established during the first ECT session) in bifrontotemporal and at 6 times the initial seizure threshold in right unilateral ECT. Anaesthesia was achieved with intravenous administration of etomidate (0.2-0.3 mg/kg) followed by succinylcholine (1.0 mg/kg) and positive pressure oxygen (100%) was provided until the return of spontaneous respiration.

8.4.5 Digit Symbol substitution Test (DSST)

To test for possible group differences in general cognitive functioning at the time of story memorization and memory testing, a digit symbol substitution test was administered. The DSST consists of nine digit-symbol pairs and a list of 115 symbols (473). The task is to write down the associated digits underneath the 115 symbols as fast and with as few errors as possible. The DSST was administered prior to the initial study session and prior to memory testing.

8.4.6 Multiple choice memory test

Following the order of the 11 presented slides, 3-5 multiple-choice questions per slide with 4 answer options were posed (2). To reduce the burden on patients, the number of questions was reduced from the original 78 to 40 questions per story. An example of one of the questions
8.4.7 Procedures

Upon signing informed consent patients were pseudo-randomly assigned to one of three groups (A,B,C; Figure 8.1). The randomization of subjects was done in such a way that the order of study of the stories, the to be reactivated story, and the order of story testing was balanced across groups.

8.4.7.1 Encoding

Patients completed the DSST. Next, the two slide shows were presented on a computer screen. Patients were explicitly instructed that their memory for the stories would subsequently be tested, but were blind to specifics of the study in terms or reconsolidation or memory reactivation. Order of presentation of the two stories was pseudo-randomized so that ordering was similar across groups.

8.4.7.2 Reactivation and ECT

One week after encoding, memory for one of the two stories was reactivated. For groups A and B this occurred in the operating room once all ECT preparations had been made. For group C reactivation took place within the same hospital office as where memorization had taken place. Immediately after memory reactivation patients in group A and B were anesthetized and received ECT. The delay between reactivation and start of ECT procedures did not differ between groups A and B (t(24) = 0.75, P = 0.461, group A mean: 4.46 min, s.e.m.: 0.46; group B mean: 4.00 min,
s.e.m.: 0.41). The reactivated story was pseudo-randomized so that the reactivated story was balanced across groups.

8.4.7.3 Memory Test Session

For group A and C memory was tested one day after memory reactivation and ECT, for group B this was done immediately after reactivation and recovery from ECT (mean: 103.77 min, s.e.m.: 7.84; range: 29-133 min). First patients completed the DSST, next the memory test was conducted. Multiple-choice questions were asked verbally by the experimenter and tests were completed for one story then for the next. The order of tested stories was pseudo-randomized so that this was balanced across groups.

8.5 Results

8.5.1 Subjects

Thirteen patients per group completed the study. Three patients dropped out because ECT treatment was stopped (N=2) or because the patient relapsed in psychosis (N=1). There were no group differences in terms of gender, age, number of prior administered ECT sessions, duration of the depressive episode, modified cumulative illness rating scores (CIRS (474)), global assessment of functioning scores (GAF (468)), or depression ratings (MADRS (475)) (Supplementary Table 8.1). Thus, the observed between-group differences in reactivated memories are not due to differences in cognitive or clinical status.

8.5.2 ECT procedure

ECT induced a satisfactory generalized seizure in all patients. Groups A and B did not differ in terms of motor seizure duration, electroencephalogram (EEG) seizure duration, applied current, electrode placement (right unilateral, bifrontotemporal), amount of administered etomidate or succinylcholine, and initial seizure threshold as was established at the first ECT session (Supplementary Table 8.2). Hence, the
observed between-group differences in reactivated memories are not due to differences in ECT parameters or anesthetic dose.

8.5.3 Memory reactivation does not differ between groups

All groups showed evidence of memory reactivation, i.e., memory performance at reactivation was above chance level as indexed by the memory reactivation score (one-sample t-test across all groups, \(t(36) = 7.53, P < 0.001\)), and groups did not differ in memory reactivation scores (Kruskal-Wallis for group \((A,B,C)\), \(H(2) = 1.77, P = 0.412\); group A mean: 3.08, s.e.m.: 0.35; group B: mean: 2.62, s.e.m.: 0.96; group C mean: 3.23; s.e.m.: 1.24). Therefore, the observed between-group differences in reactivated memories are not due to differences in strength of memory reactivation, and adequate memory reactivation principally allows the initiation of a reconsolidation process.
8.5.4

<table>
<thead>
<tr>
<th>Test:</th>
<th>Group</th>
<th>Reactivation</th>
<th>Reactivation x Group</th>
</tr>
</thead>
<tbody>
<tr>
<td>1Group = A, B, and C</td>
<td>$F_{2,36} = 6.50$</td>
<td>$F_{1,36} = 0.01$</td>
<td>$F_{2,36} = 11.18$</td>
</tr>
<tr>
<td></td>
<td>$P = 0.004$</td>
<td>$P = 0.926$</td>
<td>$P &lt; 0.001$</td>
</tr>
<tr>
<td>2Group = A and B only</td>
<td>$F_{1,24} = 1.28$</td>
<td>$F_{1,24} = 4.24$</td>
<td>$F_{1,24} = 4.29$</td>
</tr>
<tr>
<td></td>
<td>$P = 0.270$</td>
<td>$P = 0.050$</td>
<td>$P = 0.049$</td>
</tr>
</tbody>
</table>

Table 8.1: Statistics

Statistical test results. 1Group x Reactivation (reactivated story, non-reactivated story) repeated measures ANOVA; 2One-way ANOVA; 3Post-hoc pair-wise comparisons; 4Paired t-tests; 5One-sample t-test.

No difference in general cognitive functioning between groups as assessed by the DSST

A group (A, B, C) x time point (study, test) repeated measures ANOVA on DSST scores revealed no main effect of time point ($F_{1,31} = 0.03$, $P = 0.864$), group ($F_{1,31} = 0.52$, $P = 0.600$) or group x time point interaction ($F_{1,31} = 1.39$, $P = 0.265$). Hence, group differences in memory performance are unlikely to be due to group differences in general cognitive functioning.
8.5.5 ECT disrupts reactivated memory

As predicted, we observed ECT-evoked disruption of memory for the reactivated story when tested one day after reactivation in group A (Figure 8.3 red), but critically not on immediate memory testing in group B (Figure 8.3 blue). Specifically, memory for slides pertaining to the reactivated story in group A is not different from chance level, and impaired relative to that for the non-reactivated story (Table 8.1 in main text). By contrast, in group B, memory performance is not different between the two stories, and at a level equivalent to the non-reactivated story in group A. Taken together, the results from groups A and B accord with a view that reactivated emotional episodic memories are impaired by a single ECT session in a time-dependent fashion. Interestingly, our control group C (Figure 8.3 green), which followed the same protocol as group A except for not receiving anesthetics and ECT, showed better memory for the reactivated vs. non-reactivated story. Thus, reactivation is beneficial for memory in the absence of subsequent ECT.

To further examine the interaction effect of group x reactivation, present in the ANOVA comparing all three groups, we specifically tested whether this effect is driven by memory for reactivated or non-reactivated stories. Critically, there are no significant between-group differences in memory for non-reactivated stories, but an effect for reactivated stories (Table 8.1). Thus, memory performance for the non-reactivated story was not significantly modulated by ECT or the time interval between ECT and test. That group C showed better memory for reactivated material in the absence of ECT supports similar findings in animals (476).

8.5.6 No emotion by reactivation interactions

Emotional narratives accompanied both slide-stories learnt by patients. Each story can be separated into three phases of which the middle is considered most emotional and results in enhanced memory when compared to the same images accompanied by a neutral narrative (2). Testing for a reactivation (reactivated story, non-reactivated story) x
phase (1,2,3) effect within group A revealed a main effect of reactivation ($F_{1,12} = 8.75, P = 0.012$), but no main effect of phase ($F_{2,24} = 0.48, P = 0.0624$), or reactivation x phase interaction ($F_{2,24} = 2.40, P = 0.112$). Thus we observe no interaction between relative emotionality of the studies material and the disturbance of reactivated memory.

### 8.5.7 Memory performance is associated with illness and ECT parameters

Next we assessed whether screening scores or elements of ECT treatment were related to memory performance. Cumulative illness rating scale scores (CIRS) correlated with memory performance over all groups (Pearson $r = -0.32, N = 38, P = 0.047$), thus the lower the comorbid physical problems or illnesses the better memory performance. Limiting the analyses to group A and group B, we tested whether ECT treatment parameters were related to memory performance. An independent t-test showed that memory impairment was greater for bifrontotemporal electrode placement compared to right unilateral stimulation ($t(24) = 2.28, P = 0.032$, right unilateral mean: 0.36, s.e.m: 0.02; bifrontotemporal mean: 0.29, s.e.m.: 0.02), in line with previous reports (467, 477, 478).

Given that bilateral stimulation leads to more memory impairment, we specifically tested for a modulation of reconsolidation by electrode placement in group A. No effect was observed ($P > 0.05$), but this null finding may reflect the size of our sample (unilateral N=4, bilateral N=9). That bifrontotemporal electrode placement is associated with greater memory impairments than right unilateral electrode placement replicates previous findings (467, 477, 478). The reconsolidation impairment observed in group A was still evident when controlling for electrode placement. The relation between electrode placement and memory performance suggests that the observed memory impairments are a result of the electrical stimulation and/or the convulsion itself and not other elements of the ECT treatment such as the anaesthesia.
8.6 Discussion

In summary, we show in unipolar-depressed patients that a single ECT application following memory reactivation disrupts reactivated, but not non-reactivated, memory for an emotional episode. Additionally, we show that this effect is time-dependent, expressed only after a 24h period. In the absence of ECT, reactivation of memory actually benefits memory performance 24h later. The effects we report are therefore not due to an unspecific effect of time or reactivation. Our data provide evidence for a disruption of reactivated emotional episodic memories by invasive interference with normal neural activity. The novel findings are: 1) Our results satisfy critical criteria for demonstrating reconsolidation in humans, that have been difficult to meet by other methods, 2) the memory impairment/enhancement following reactivation cannot be due to new learning and thus be confounded by possible interference at the time of retrieval, 3) shows the loss of emotional episodic memory following reactivation extending findings on simple associative fear memories (413, 466). Episodic memory can contribute to the persistence of emotional responses (479), therefore the ability to disturb reconsolidation of emotional episodic memory is of great clinical importance. By circumventing the difficulties previously facing human reconsolidation studies, our results bridge a cross-species gap between human and non-human research, and provide critical support to previous human studies on reconsolidation (413, 466).

Our findings are in line with early studies on non-human animals and psychiatric patients that show a disruption of reactivated memory or clinical symptoms when followed by electroconvulsive stimulation, respectively (47, 399). Our data show that an extensive and ecologically valid emotional episodic memory task is sufficiently sensitive to detect reconsolidation impairment. By contrast, an early human study that showed no effects of ECT on reactivated memory also showed limited effects on the consolidation of item recognition and paired associative memory (480), which could reflect that these tests lacked the sensitivity to reveal reconsolidation effects.
Reconsolidation has received much interest due to its potential for permanently altering or erasing memories that contribute to the persistence of psychiatric disorders (481, 482). A translational significance of the present findings is the potential therapeutic advantage of evoking depressive ideations prior to ECT, in order to disrupt reconsolidation and therewith induce amnesia for these symptoms (399, 480). However, despite observing a relationship between memory performance and electrode placement but not with anesthetic dose, we cannot distinguish whether the applied electrical current or the anesthesia cause the impairment of reactivated memory. Regardless, our data emphasizes the dynamic nature of memory by demonstrating that the alteration of episodic memory for emotional experiences during reconsolidation is possible.

Competing interests

The authors declare no conflict of interest.

Acknowledgements

This work was supported by a research grant from the Netherlands Organisation for Scientific Research (NWO) to G. Fernández. We would like to thank O. Buno Heslinga for help with subject recruitment, S. Vosters for help in creating and piloting stimulus material, and A. van Stegeren for advice on the creation of the stimulus material.

Author contributions

MSc Kroes, Dr. Strange, Dr. Tendolkar, Dr. Van Wingen and Dr. Fernandez designed the study. MSc Kroes acquired and analyzed the data. MSc Kroes wrote the manuscript. Dr. Strange, Dr. Tendolkar, Dr. Van Wingen, Dr. Van Waarde and Dr. Fernandez reviewed and approved the manuscript.
8. EVIDENCE FOR RECONSOLIDATION OF EMOTIONAL EPISODIC MEMORY IN HUMANS
β-Adrenergic blockade during memory retrieval in humans evokes a sustained reduction of declarative emotional memory enhancement.

9.1 Published as:


9.2 Abstract

Memory enhancement for emotional events is dependent on amygdala activation and noradrenergic modulation during learning. A potential role for noradrenaline (NE) during retrieval of emotional memory is less well understood. Here we report that administration of the β-adrenergic receptor antagonist propranolol at retrieval abolishes a declarative memory enhancement for emotional items. Critically, this effect persists at a subsequent 24 hours memory test, in the absence of propranolol. Thus, these findings extend our current understanding of the role of NE in emotional memory to encompass effects at retrieval, and provide face validity to clinical interventions using β-adrenergic antagonists in conjunction with reactivation of unwanted memories in anxiety-related disorders.
9.3 Introduction

Enhanced episodic memory for emotional events (1, 483) is dependent on the amygdala (5). During initial acquisition or consolidation of an emotional memory, the amygdala is thought to upregulate neuronal processing in the hippocampus (37, 38) through a $\beta$-adrenergic mechanism (29, 33, 39). $\beta$-adrenergic blockade with the $\beta_1\beta_2$-antagonist propranolol selectively impairs long-term human episodic memory for emotionally arousing material without affecting memory for neutral material when administered prior to learning (440, 469). This widely accepted model of noradrenergic influences on human emotional memory, by focusing on encoding and consolidation (29), implicitly suggests that the role of noradrenaline does not extend to retrieval of emotional memories. Providing evidence for a role of NE in retrieval would therefore add considerably to a broader understanding of the nature of emotional influences on memory.

Although adaptive to survival, emotional memory enhancement may also have pathological consequences, as in anxiety-related disorders such as post-traumatic stress disorder (PTSD) (3). Once consolidated, memories are thought to be relatively insensitive to disruption or further modulation (30), rendering pharmacological treatments long after the traumatic event therapeutically ineffective. On the other hand, if neuromodulatory effects on emotional memory following consolidation were evident then this would provide a potential therapeutic approach for attenuating the retrieval of unwanted, established, emotional memories. As such, the application of NE antagonists might, for example, reduce the occurrence of intrusive memories such as flashbacks in PTSD. Identifying a role for NE during emotional memory retrieval would therefore be of considerable clinical significance.

Currently, the role of NE at retrieval is unclear. Human brain imaging during emotional memory retrieval demonstrates activation of the locus coeruleus (LC) (484), the origin of noradrenergic forebrain projections (485, 486). Other studies demonstrate amygdala engagement during
β-adrenergic blockade during memory retrieval in humans evokes a sustained reduction of declarative emotional memory enhancement. However, human psychopharmacological data fail to demonstrate blockade of enhanced emotional retrieval by propranolol given at recall (490, 491). This contrasts with observations in rodents that propranolol, given prior to retrieval of contextual conditioned fear, attenuates the fear response (492). One possible explanation for this discrepancy is that the memory paradigm employed in previous human studies (490, 491) did not evoke a robust emotional memory enhancement. Thus, the absence of consistent and robust emotional memory enhancements in these studies indicates a relatively marginal contribution of the adrenergic system in these tasks.

Figure 9.1: Experimental design

a. Depiction of the encoding task (left panel) and cued recall task (right panel). During encoding stem-unique nouns are presented with occasional emotional nouns (E) and perceptually distinct nouns (P). During the cued recall task three letter stems of the nouns encoded on the first day are presented. The stems of the perceptual nouns are not presented in a distinct font. b. Schematic of the study design over 3 days. Note that propranolol is only administered on Day 2. BP = blood pressure measurement.
such that an effect of propranolol could not be detected. Therefore, we conducted a double-blind, placebo controlled experiment testing the effect of propranolol on cued-recall of emotionally aversive verbal stimuli, using a stimulus set known to cause a marked emotional memory enhancement effect that is also known to be dependent on the amygdala and NE (33, 440).

Briefly, on Day 1 subjects encoded neutral control nouns (C) with occasional nouns being emotionally aversive (E nouns) or perceptually distinct (P nouns) (440) (Figure 9.1a). On Day 2, subjects were administered either propranolol (40 mg) or placebo and 90 minutes later engaged in a cued recall test (Figure 9.1b). To rule out the possibility that any modulation of emotional memory retrieval by propranolol is restricted to testing in the presence of drug we again tested subjects 24 hours later, in the absence of placebo or drug, using the cued recall test. We demonstrate enhanced memory for E relative to C nouns in the placebo group on both Day 2 and Day 3. Administration of propranolol at retrieval abolishes emotional memory enhancement on Day 2, an effect which is present at least 24 hours later on Day 3.

9.4 Methods and Materials

9.4.1 Participants

Twenty-four right handed native English speakers with normal/corrected to normal vision, without cardiorespiratory, neurological or psychiatric history participated in the study. The placebo and propranolol groups each comprised 12 subjects, 7 and 5 females respectively, with a mean age of 24.4 yrs (range: 19-34). One male subject in the propranolol group only completed Day 1 and Day 2 of the experiment. All subjects gave informed consent and the study had full ethics approval. The experiment was conducted on 3 consecutive days, with each session at the same time of day and in the same testing room for a given subject.
9.4.2 Procedures

9.4.2.1 Day 1

Blood pressure (BP) was measured and an electrocardiogram acquired for each participant on arrival. During encoding, subjects viewed 360 nouns presented visually in uppercase at a rate of one noun every 3 s (stimulus duration 1 s). The nouns consisted of 300 neutral nouns all presented in the same font, 30 nouns were emotionally aversive oddballs (E), presented in the same font as the neutral nouns, and 30 nouns were perceptual oddballs (P) presented in different fonts. E nouns were selected from the set employed in our previous studies (38, 440). Subjects engaged in a deep encoding task (493), making a push-button response to indicate whether the noun described a living or nonliving entity. Although we have previously shown emotional memory enhancement to be more pronounced following shallow encoding (440), a deep encoding task was chosen to ensure adequate memory performance on Day 2 and Day 3. Word presentation was randomized and oddballs were randomly presented after 4-6 (mean = 5) neutral words. Words were ‘stem-unique’ in that the first 3 letters were different for each of the 360 words. 120 words served as controls (C), with the constraint that these controls were more than 1 word before or after a given oddball.

9.4.2.2 Day 2

Blood pressure (BP) was again measured and subjects were then administered either an oral dose of 40mg propranolol or placebo (Vitamin C) in a double-blind experimental design. After 80 minutes subjects filled out the State form of the State-Trait Anxiety Inventory (STAIS) (200) and the Beck Depression Inventory-II (BDI) (82). Subjects then performed a surprise memory task on the words encoded on Day 1. In view of the kinetics of propranolol’s peak plasma concentration (1-2 h), the memory task started 90 min after drug administration. Randomly selected word stems (first 3 letters) for 240 of the 360 nouns encoded on Day 1 were presented visually in uppercase in random order. Of these,
β-adrenergic blockade during memory retrieval in humans evokes a sustained reduction of declarative emotional memory enhancement

200 were neutral, 20 were emotional oddballs, and 20 were perceptual oddballs. All nouns were presented in the same font, including nouns that were perceptually distinct on Day 1, i.e. P nouns. Subjects were instructed to complete the stems to make a word encoded on Day 1. Presentation of stimuli was continuous and stems were presented every 4 s with a stimulus duration of 1 s. We employed a cued recall test of retrieval in order to increase recall performance from a large study list.

9.4.2.3 Day 3

This followed the same procedure as Day 2 except that no drug was administrated on Day 3. Furthermore, 360 stems were presented; the 240 presented on Day 2 plus stems pertaining to the remaining 100 neutral words, 10 emotional oddballs, and 10 perceptual oddballs from Day 1 that had not been presented on Day 2.

9.5 Results

Propranolol did not affect cued recall of control nouns (Figure 9.2a). A Group (‘placebo’, ‘propranolol’) x Day (Day 2, Day 3) 2x2 repeated measures ANOVA on C noun recall revealed no main effects of Group (F_{1,10} = 0.706, P = 0.420) or Day (F_{1,10} = 2.124, P = 0.176) and no Group x Day interaction (F_{1,10} = 0.381, P = 0.551). No memory enhancement was
detected for perceptual oddballs (see supplemental Information) and as such remaining analyses were restricted to emotional vs neutral nouns. Importantly, recall performance for all words was significantly above chance as indicated in a separate experimental baseline cue-completion group (see Supplemental Information).

In our critical comparison, we tested for the effect of drug manipulation on E noun memory relative to subject-specific C noun recall (Figure 9.2a). A Group (‘placebo’, ‘propranolol’) x Day (Day 2, Day 3) x Noun type (emotional, control) 2x2x2 repeated measures ANOVA revealed a significant main effect of Group (F\textsubscript{1,10} = 9.491, P = 0.012) and Noun type (F\textsubscript{1,10} = 8.461, P = 0.016), and a significant Group x Noun type interaction (F\textsubscript{1,10} = 13.174, P = 0.005), with no other significant main effects or interactions.

Planned comparisons confirmed that on Day 2 the Placebo group showed enhanced recall for emotional nouns relative to control nouns (t(11) = 5.365, P < 0.001 (two-tailed)), while no emotional memory enhancement was evident in the Propranolol group (t(11) = 1.044, P = 0.319 (two-tailed)) (Figure 9.2b). On Day 3 the Placebo group continued to show enhanced recall for emotional relative to control nouns (t(11) = 2.702, P = 0.021 (two-tailed)). Critically, subjects who had received propranolol on the previous day persist in showing an absent emotional memory enhancement on Day 3 (t(11) = -0.202, P = 0.844 (two-tailed)) (Figure 9.2b). Directly comparing E noun recall between the Placebo and Propranolol group revealed reduced memory scores for the Propranolol group on Day 2 (t(22) = 2.900, P = 0.008 (two-tailed)) and on Day 3 (t(21) = 2.701, P = 0.013 (two-tailed)). In both groups, the E nouns remembered on Day 2 tended to be the same as those remembered on Day 3 (see Supplemental results, Figure S8.1).

Cued recall performance for perceptual nouns, as well as for nouns cued only on Day 3, is given in Supplemental Information (Figure S8.2). For the items that had not been presented on Day 2, drug manipulation did not affect mean recall of control nouns on Day 3 (t(21) = -1.177, P = 0.255 (two-tailed)). Critically, no recall differences between the drug groups were found for the new emotional nouns (t(21) = 0.153, P = 0.880 (two-tailed)).
On Day 2 there was no baseline systolic BP differences between groups (t(21) = -0.223, P = 0.826 (two-tailed)). In line with previous findings (491), we observe a drop in systolic BP in the Propranolol group (t(20) = 1.842, P = 0.04 (one tailed)) after 90 minutes, but no drop in diastolic BP. Note that at both time points BP measurements for a different Placebo subject was lost due to equipment failure.

9.6 Discussion

Our data provide a clear demonstration of a critical role for NE in the retrieval of emotional memory. We show that β-adrenergic blockade by propranolol at retrieval abolishes the declarative memory enhancement for emotional items. Secondly, we show that 24 hours after β-adrenergic blockade at retrieval, a sustained reduction of emotional item recall is still observed in a second retrieval test in the absence of propranolol. Our findings provide support for the possibility of lasting attenuation of emotional memory in therapeutic contexts by pharmacological interventions at retrieval.

Our finding of a role for NE in retrieval of emotional memory speaks directly to the discrepancy between neuroimaging data, which implicate the NE system in emotional memory retrieval (484), and human psychopharmacological studies which report no effect of propranolol administration at retrieval on the enhancing effect of emotion on memory (490, 491). The latter negative findings might be explained by an absence of robust memory enhancement for emotional items in these tasks, possibly due the presentation of equal numbers of neutral and emotional items and/or the presentation of emotional items of both positive and negative valence (491). The use of equal numbers of emotional and neutral items in these previous studies may have resulted in rapid amygdala habituation (494), attenuating the memory enhancement process. In the oddball task employed here the amygdala response to emotional items has been shown not to habituate over repeated exposure (495). Further, valence can enhance memory for low arousing items via prefrontal-based organisational strategies during retrieval that are
independent of the amygdala and noradrenergic modulation (496, 497). Indeed we note a suggestion (491) that a lack of an effect of propranolol might be attributable to a non-dependence of their task on the amygdala and/or NE.

Animal data indicate that a role for NE in emotional memory retrieval is time-specific. Emotional memory retrieval is affected by noradrenergic manipulation at 2 hours, but not 1 hour, after learning and this effect is absent if the study-test interval is extended to 4 days (492). The modulation of emotional memory by propranolol in humans may, therefore, also be time-dependent. Thus, a second explanation for the discrepancy between our findings and a previous study (491) might reflect the difference in study-test interval. This previous study, which found no effect of propranolol administration on emotional memory retrieval, tested retrieval 1 week after initial learning (491) compared to the delay-interval employed here of 24 hours. It will be of interest to test whether the effect of propranolol observed in the current study is still observed if memory is tested more than 4 days after learning.

A critical question pertains to the neurobiological basis for the dependency of emotional episodic memory retrieval on the β-adrenergic system. The current data do not directly address this but potential explanations include the requirement for a cue-evoked phasic rise in NE (34) such that there is congruence between NE signals at encoding and retrieval (498). Alternatively, the emotional memory trace may be dependent on tonic noradrenergic input (492, 499) such that engagement of β-adrenergic receptors is essential for retrieval. It is conceivable that tonic NE input is no longer required after completion of systems level consolidation (500), which concords with a time-limited role of NE in emotional memory retrieval derived from animal data (492).

In terms of a neuroanatomical basis for the effects we observe, neuroimaging studies demonstrate amygdala involvement in emotional memory retrieval. Amygdala activity has been reported during retrieval of neutral stimuli previously encoded in an emotional context (487, 489, 501,
The amygdala may directly support emotional memory retrieval or could achieve this through hippocampal modulation. Increased hippocampal-amygdala coupling during retrieval of information encoded in an emotional context has been demonstrated in human (488) and rodent (503) studies. We suggest, therefore, that the effects we observe result from a propranolol-evoked disruption of amygdala responses, or hippocampal-amygdala coupling.

Animal studies have provided evidence for a more general role of NE in retrieval (492, 504-506). Specifically, increasing NE activity through pharmacological manipulation or by stimulation of the LC can restore forgotten spatial memory in rats (504, 505), via engagement of β-adrenergic receptors (506). Thus, these studies extend the influence of NE at retrieval to include non-emotional memories. In contrast to this, the effect of propranolol in our task did not extend to cued recall of neutral words or items that were perceptually distinct at encoding (see Supplemental material), which may reflect the essential non-verbal nature of animal tasks or effective central propranolol concentrations in human versus animal studies.

Our second main finding is that β-adrenergic blockade at retrieval leads to a sustained reduction of emotional memory retrieval 24 hours later. The neurobiological basis for this effect could reflect a number of processes including a disruption of secondary encoding (449), a disruption of reconsolidation (50, 507) or an impairment of lingering consolidation (446). Secondary encoding is the process whereby during every retrieval episode, retrieved mnemonic information, together with incoming sensory information, is stored as a novel memory trace (319). This results in the accumulation of multiple, related, mnemonic traces of a single episode within a slightly altered neuronal and experiential context, which together aid the retrieval of an episode (319). Thus, blocking enhanced recall of emotional stimuli on Day 2 with propranolol may have prevented formation of a novel emotional memory trace, expressed as absent emotional memory enhancement in this group on Day 3.
Alternatively, impaired reconsolidation may account for, at least in part, the memory effects observed on Day 3. Reconsolidation refers to the idea that, upon reactivation, memories return to a labile state and are once again sensitive to disruption (50). It is also known that reconsolidation of emotional memories is disrupted by blocking NE within the amygdala (507). The long-term memory impairment we observe on Day 3 could be attributed to propranolol having blocked reconsolidation of the episodic enhancement process. As such, the mechanisms mediating the enhancement of memory for emotional items might not have been expressed. Typically, in reconsolidation experiments, memories need to be reactivated to return to a labile state. However, in the case of amygdala processing of fear information, there is evidence that fear memories can undergo reconsolidation even if not expressed (508). Thus, propranolol might have impaired the reconsolidation of emotional items not remembered on Day 2, reducing their chance of memory retrieval on Day 3.

The third possible explanation for our effects on Day 3 is propranolol-induced impaired lingering consolidation. The lingering consolidation hypothesis posits that iterations of reactivation and stabilization of memory traces are an essential feature of consolidation, which, in contradistinction to standard models of consolidation, is not time-limited (446). The memory effects observed on Day 3 could, therefore, be attributed to a disruption of ongoing consolidation. However, an argument against this explanation is the absence of any difference between propranolol and placebo groups on memory for emotional items on Day 3 that were not cued on Day 2. Thus, the neurobiological mechanisms underlying our second main finding, that emotional memory impairment evoked by propranolol on Day 2 persists to Day 3, are likely to be a failure of secondary encoding, impaired reconsolidation or a combination of these processes. Thus, blocking enhanced recall of emotional stimuli on Day 2 with propranolol may have prevented formation of a novel emotional memory trace, expressed as absent emotional memory enhancement in this group on Day 3.
The efficacy of pharmacological prevention of the acquisition of memories for traumatic experiences is limited by practicality (509, 510). The possibility of pharmacologically modulating established emotional memory at retrieval we demonstrate here supports a potential therapeutic approach of attenuating unwanted established emotional memories. Our data support recent clinical findings that propranolol administration to PTSD patients, following memory reactivation, reduces physiological responses to a subsequent script-driven trauma imagery session (404). We extend these findings of post-reactivation modulation of implicit measures of emotional memory by propranolol (404, 405) by demonstrating specific attenuation of retrieval of emotional episodic memory, without an impact on neutral memory. Thus, we provide a theoretical justification for the potential application of propranolol in reducing retrieval of traumatic memories that justify studies of clinical interventions with β-adrenergic antagonists during reactivation of traumatic memory in anxiety-related disorders.

**Competing interests**

The authors declare no conflict of interest.

**Author contributions**

MSc Kroes and Dr. Strange acquired and analyzed data. MSc Kroes, Dr. Strange and Dr. Dolan designed the study and wrote the manuscript.
10.1 In preparation as:


10.2 Abstract

The combined use of psychotherapy and noradrenergic antagonists has been advocated as beta-blockers can disrupt reactivated memory resulting in a loss of fear. However, animal studies have shown that beta-blockers can also impair consolidation of extinction memory, leading to increased fear. We therefore tested these opposing hypotheses in a double-blind placebo controlled functional magnetic resonance study and show that a single dose of propranolol prior to extinction learning eliminates contextual fear conditioned responses in healthy humans, causes a subsequent loss of fear and prevents the return of fear one day later in the absence of drug. Moreover, the elimination of differential conditioned responses is associated with absence of fear conditioned responses in the dorsal medial prefrontal cortex, whereas the loss of fear one day later is directly related to increased differential responses in the hippocampus. Our results extend our current understanding of the role of noradrenaline in contextual fear generalization, and provide face validity to treatments combining psychotherapy and beta-blockers.
10.3 Introduction

Mental illness such as anxiety disorders rank amongst the leading causes of disability worldwide, significantly contributing to the global disease burden (WHO, 2011). Therefore, optimizing treatment of these disorders is of great importance. Over the last decades antidepressants medication has been used to treat affective disorders and their actions have been implicated to affect the monoamine systems including noradrenaline (27). Yet, recent clinical studies have indicated that the combination of psychopharmacological treatments and psychotherapy is more effective than either treatment alone (511). However, little is known about the optimal combinations of pharmacological and psychological treatments and the mechanism that underlie optimization of treatment strategies.

Generally fear and anxiety disorders are treated by either pharmacological treatments or cognitive behavioural therapy including exposure therapy. During exposure therapy a fear response is extinguished or suppressed by repeatedly exposing a patient to a fear-evoking stimulus in a safe setting. The suppression of a fear response has been successfully modelled using Pavlovian conditioning. Here a neutral conditioned stimulus (CS, e.g. a light) comes to elicit a conditioned fear response (CR) after repeated pairing with an aversive unconditioned stimulus (US). The conditioned fear response can subsequently be reduced or suppressed through repeated exposure to the CS in the absence of the US in a procedure known as extinction learning. Importantly, extinction learning does not erase the CR memory, but reflects new learning, an association between the CS and safety. Thus, the conditioned memory trace and the extinction memory trace coexist and compete, such that the extinction memory inhibits expression of the CR (10, 512, 513). As such, although exposure therapy can lead to the reduction of fear responses in patients suffering from fear or anxiety disorders, patients are subject to spontaneous recovery of the fear response and to reinstatement of the fear response at subsequent re-exposure to the US. Therefore, a need exists for methods that more optimally reduce the expression of fear, the generalization of fear from
threatening situations to save contexts in which fear is maladaptive, and that prevent the return of fear following treatment.

Recent studies have shown that memories can be reactivated upon which they enter an instable state and require restabilization. The finding that this reconsolidation process can be disrupted causing an erasure of fear memories has sparked great interest as it may prevent the return of fear. Importantly, reconsolidation of fear memories can be disrupted by the application of the β1β2-noradrenergic antagonist propranolol in animals (507). Similarly, human studies have revealed that reactivation of fear memories in the presence of a beta-blocker can impair subsequent retrieval of these memories (405, 411). Interestingly, propranolol has also been found to impair retrieval of emotional memory (411, 514, 515). Such findings provide a theoretical justification for applying beta-blockers in combination with psychotherapy in order to reduce symptoms in anxiety disorders (404, 411, 464, 509). However, the methods by which a memory is reactivated and becomes susceptible to the disruption of reconsolidation, or is subject to extinction are very similar (432, 476). For example a brief single exposure to a CS is considered to evoke a reconsolidation process, yet a more prolonged exposure or repeated exposures to the CS evoke an extinction process and, hence, new learning. Propranolol in the presence of re-exposure to a CS could thus block reconsolidation resulting in the loss of fear, but could principally also impair extinction learning resulting in an increase of fear (402, 432, 516-518). Interestingly, several animal studies indicate that beta-blockers may interfere with the consolidation of extinction memory (519, 520), whilst others have found no effects (521) or find beta-blockers to reduce fear expression during extinction learning (514, 522). Currently, the effects of beta-blockers on extinction learning in humans remain uninvestigated. To provide face validity to the use of beta-blockers in combination with psychotherapy to treat patients with anxiety disorders it is critical to elucidate the possible influence of beta-blockers on extinction learning. From literature opposing hypothesis arise where the combination of beta-blockers with extinction learning either results in a subsequent increase or decrease of fear.
Investigating the effects of beta-blockers on the neural circuits that underlie fear and safety learning can provide insight in the mechanisms that contribute to the optimization of combined pharmacological and psychotherapeutical treatments. The neural circuit that supports fear and safety learning in pavlovian conditioning and extinction paradigms is well defined. The amygdala is considered to mediate (conditioned) responses to adversity and the storage of associative fear memory (10, 12, 512, 513), in close coordination with the dorsal medial prefrontal cortex (dMPFC) and midbrain regions. Following extinction learning the ventral medial prefrontal cortex (vMPFC) is thought to inhibit the fear response (10, 12, 512, 513). Additionally, extinction is context specific and the hippocampus is thought to provide this contextual information (523). Together, the idea is that during extinction learning a novel memory is formed through hippocampus-amygdala-vMPFC interactions, which, following consolidation, leads to a hippocampus-dependent vMPFC inhibition of the amygdala preventing the CR (513). Exactly this mechanism appears impaired in anxiety disorders (21). Further, this implies that optimal pharmacotherapies should support the flexible adjustment of specific conditioned fear responses during extinction learning by affecting the critical nodes in the neural network. Interestingly, recent work shows that the neural network involved in regulating extinction is affected by noradrenergic medication (175).

Considering the opposing hypothesis on the influence of beta-blockers on extinction, we tested whether a single administration of the beta-blocker propranolol affects extinction and the neural processes implicated in safety learning. Briefly, we conducted an experiment over three consecutive days (Figure 10.1a). On day 1, subjects were conditioned to a light in a specific context by pairing it with an electrical shock. Next, on Day 2 subjects received a single dose of propranolol and underwent an extinction paradigm. To test the long term effect of propranolol, subjects were tested on spontaneous recovery of the fear response and on the resistance to reinstatement of the fear response on Day 3, in the absence of propranolol. To investigate the influence of noradrenaline on the neural circuit underlying extinction learning brain activity was assessed.
10. β-ADRENERGIC BLOCKADE AFFECTS THE NEURAL NETWORK OF EXTINCTION LEARNING, CAUSES THE LOSS OF FEAR AND PREVENTS THE RETURN OF FEAR IN HUMANS

Figure 10.1: Experimental design and task
A) In a three day between-subjects design participants were conditioned on Day 1. On Day 2 subjects heart-rate (HR), blood pressure (Bp, pink boxes), and mood were measured at the start (T1), and subsequently received propranolol (40mg) or placebo. After 60 minutes (T2) heart-rate, blood pressure and mood were again assessed, and next subjects entered the MRI scanner and underwent an extinction task. Upon exiting the scanner (T3) heart-rate and mood was again measured. On Day 3 heart-rate and mood measures were acquired (T4), upon which subjects entered the scanner for a recall task testing spontaneous recovery, followed by a reinstatement procedure, and a reextinction task testing for the reinstatement of fear. At the end of Day 3 (T5) heart-rate, blood-pressure and mood were assessed and subjects received a contingency questionnaire. Skin conductance responses (SCR) was acquired on all three days, BOLD fMRI data was collected on Day 2 and Day 3. B) Subjects were exposed to a context conditioning task on Day 1. Following a fixation cross, a room was presented. Within the room a light ‘turned on’ either emitting yellow or blue light serving as conditioned stimuli. One of the lights was followed by a transcutaneous electrical shock 33% of the time, the other colour light was never followed by a shock. On Day 2 and Day 3 the task was identical to Day 1 except that a different room, but the same lights were presented, and neither of the colour lights was followed by a shock. C) Propranolol caused a drop in systolic blood pressure at T2 and T3 on Day 2 but no longer has an effect on Day 3.

with BOLD fMRI on Day 2 and Day 3. Considering the findings that beta-blockers can impair the consolidation of extinction learning in animals (519, 520), that propranolol blockade can abolish the retrieval of memory
for emotional memories (411, 522), and that reactivating emotional memory in the presence of propranolol can result in elimination of emotional memory 24h later (411, 507), we perceived of several possible outcomes of our study. First, propranolol would have no effect on extinction learning on Day 2, but would impair consolidation and result in increased spontaneous recovery and reinstatement of fear. Second, propranolol might reduce retrieval of the conditioned fear response on Day 2 and due to new learning lead to an attenuation of the return of fear on Day 3. Third, irrespective of a possible effect of propranolol on the retrieval of the conditioned fear response on Day 2, the reactivation of memory in the presence of propranolol may lead to a subsequent loss of fear on Day 3 as a result of disrupted reconsolidation. The second and third hypothesis both result in the loss of fear on Day 3 but may yield differences the possible return of fear, and would thus have different clinical implications.

10.4 Methods and Materials

10.4.1 Participants

Fifty-four healthy young participants with normal- or corrected-to-normal vision, normal uncorrected hearing, were included in the study (age-range: 19-26 yr, mean: 21.72, s.e.m.: 0.32). Subjects were pseudo-randomly assigned to one of the drug groups so that for each of four consecutive subjects two would receive propranolol and two placebo. Subjects and investigators were blind to drug conditions. Five subjects were excluded as they displayed no conditioned responses, three subjects could not complete the study due to scanner problems. The placebo group comprised 24 subjects (11 males, 13 females) and the propranolol group 22 subjects (8 males, 14 females), and one subject (female) in the propranolol group did not complete the reinstatement and re-extinction task.

All subjects gave informed consent. Subject had to meet all inclusion criteria and none of the exclusion criteria. These included that subjects
were free of neurological, cardiovascular, endocrine or psychiatric history, and MRI- or propranolol contraindications. Subjects were predominantly right-handed, female subjects used oral hormonal contraceptives, and were tested in the second week of their menstrual cycle. Additionally, subjects had no history with psychotropic medication, or hard drug use. Moreover, subjects had not used over-the-counter medication or cannabis in the 72 hours, or alcohol in the 24 hours prior to study, with the exception of oral contraceptives and paracetamol, and were not recipients of investigational products as part of research studies in the three months prior to the initial dose in this study. Furthermore, subjects did not donate blood in the two months prior to initial study dose. Finally, subjects did not consume more than 3 units of alcohol daily. The study was approved by the institutional ethics committee (CMO Regio Arnhem-Nijmegen, The Netherlands).

10.4.2 Medication

Subjects were administered either an oral dose of 40mg Propranolol HCl or placebo (microcrystalline cellulose) in capsules. Propranolol is a non-selective \( \beta \)-adrenergic receptor blocking agent that crosses the blood-brain barrier. It is a competitive antagonist which competes with \( \beta \)-adrenergic receptor stimulating agents (like adrenaline) for both \( \beta_1 \)- and \( \beta_2 \)-adrenergic receptor sites.

10.4.3 Tasks

10.4.3.1 Conditioning

As extinction learning in clinical situations occurs in a different context than that in which fear was acquired we employed a context conditioning task. A differential conditioning paradigm based on previous contextual conditioning paradigms was used with delay conditioning and partial reinforcement (524-526). Partial reinforcement was used to study the gradual development of fear learning, and slow down subsequent extinction learning (527, 528). Subjects were told that they would see on
a computer screen a picture of a room in which a light would turn on and emit a yellow or blue light while they could receive shocks. The level of the shocks was set before the experiment to a subjective intensity that was maximally uncomfortable without being painful. The instructions were to pay attention to the computer screen and subjects were told that a relationship existed between the stimuli and the shocks.

The context stimuli were images of two rooms with an identical lamp in them (Figure 10.1b). The CSs were images of the same rooms and lamp, but now the lamp emitted either a yellow or blue light. The US was a mild electric shock to the fingers (200msec asymmetrical biphasic pulse with a pulsewidth of 250 μs at 150 Hz.) that co-terminated with the CS+. On a given trial the context was presented for 11, 12, or 13 sec followed by a CS presentation for 4 sec, with an 11, 12, or 13 sec inter-trial interval during which a fixation cross was presented. During conditioning, one light colour was paired with the US on one-third of the trials (CS+), and the other colour was never paired with the US (CS-). The order of the different trial types was pseudo-randomized so that half of each trial type occurred during the first half of the task (Early Phase), whilst the other half of each trial type occurred during the second half of the task (Late Phase), the assignment of light colour into CS+ and CS- was counterbalanced across subjects. During, conditioning there were 12 presentations of each CS that did not co-terminate with the US, intermixed with an additional 6 CS+ trials that co-terminated with the US. The conditioning task was preceded by a habituation phase consisting of 6 trials, during which each CS type was presented three times. Subjects were instructed that no shocks would occur during the habituation phase. The habituation phase and conditioning task were separated by a visual instruction on a computer screen that the actual task was going to start.

10.4.3.2 Extinction

The extinction task was conducted one day after the conditioning task and followed the same procedure as the conditioning task except that the instructions were to pay attention to the computer screen and
subjects were told that the task would continue as before. Further, during the extinction task an image of a different room, but with the same lamp served as context (Figure 10.1b). The CSs were the same as during conditioning, both were presented 12 times but neither co-terminated with the US. Again, CS+ and CS- presentations were pseudo-randomized according to an Early- and Late Phase, and the assignment of the room picture that served as conditioning or extinction context was counterbalanced across subjects. A habituation phase preceded the extinction task.

10.4.3.3 Recall

The recall task was conducted one day after the extinction task and followed the same procedure as the extinction task (Figure 10.1b).

10.4.3.4 Reinstatement

The re-instatement procedure was conducted after the recall task and consisted of the continuous presentation of the extinction context during which four USs were administered (30, 100, 115, and 200 sec after context onset). Subjects received the same instructions as during the extinction task, but the shock level was not set again prior to reinstatement, and the task was not preceded by a habituation phase.

10.4.3.5 Reextinction

The re-extinction task was conducted after the reinstatement procedure and followed the same procedure as the extinction task, except that the shock level was not set again, and it was not preceded by a habituation phase (Figure 10.1b).

10.4.4 Skin conductance acquisition

Skin conductance was assessed with Ag/AgCL electrodes attached to the subject’s distal phalanges of the index and middle finger of the non-
dominant hand with standard NaCl electrolyte gel. The skin conductance signal was amplified using MR compatible BrainAmp MR and BrainAmp ExG MR (BrainProducts GmbH) within the MR environment, through an optical cable recorded outside the MR environment using BrainVision Recorder software. Data was continuously recorded at 5000 samples per second.

10.4.5 Skin conductance analyses

Skin conductance data was assessed using an in-house analysis programme written in Matlab (the MathWorks) and using FieldTrip (529). Data was low-pass filtered at 5Hz and resampled to 100Hz. The level of skin conductance responses was determined for each trial as the peak-to-peak amplitude difference in skin conductance of the largest deflection in the latency window from 0-8 s after stimulus onset. The raw skin conductance responses were square root transformed to normalize the distributions. Analyses are restricted to non-reinforced trials only.

10.4.6 fMRI sequences

MR data was acquired on a 3T MR scanner (TrioTim syngo, Siemens, Medical, Erlangen, Germany) equipped with 32-channel transmit-receiver head coil. The manufacturer's automatic 3D-shimming procedure was performed at the beginning of each experiment. Subjects were placed in a light head restraint within the scanner to limit head movements during acquisition.

10.4.6.1 Functional image acquisition

Functional images were acquired with multi-echo gradient echo-planar imaging (EPI) sensitive to the blood-oxygenation level dependent (BOLD) response using the following parameters: 38 oblique transverse slices, slice thickness = 2.5mm, interslice gap = 17%, repetition time (TR) = 2.32s, flip angle $\alpha = 90^\circ$, echo times (TE) = 9, 19.3, 30, and 40 ms (FOV) = 211 x 211 mm2, matrix size 64 x 64, fat suppression. A magnetic (B0) field map
was collected before each task and was used to unwarp the echo-planar images (160, 161, 530).

10.4.6.2 Structural image acquisition

A 3D magnetization-prepared rapid gradient echo (MPRAGE) image was acquired for normalization procedures using the following parameters: TR = 2300 ms, TE = 3.03 ms, 192 contiguous 1 mm slices, FOV = 256 x 256 mm2, matrix = 256x256.

10.4.7 fMRI analyses

10.4.7.1 Preprocessing

fMRI data were processed and analysed using the statistical software package SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). Fieldmaps were created for each subject using the Fielmap Toolbox (531). The first 6 EPI volumes were discarded to allow for T1 equilibration. The remaining functional images of the first echo series from each subject were realigned and unwarped using rigid-body transformation to correct for head movements to the mean functional image using 7th-degree B-spline interpolation. The subsequent realignment and unwarping parameters were applied to the other echo series. Next, the first 36 of the remaining volumes were used to calculate the signal to noise ratio of each echo for each voxel, and a voxel-wise weighting function was used to calculate weighted volumes for the remaining volumes. Hereafter, a three step coregistration procedure was used where first each subject’s anatomical T1-weighted MR image, and second each subjects’ mean-, and weighted functional images were coregistered to the MNI 152 T1-template, and third each subject’s anatomical T1-weighted MR image was coregistered to the mean functional image using a normalized mutual information function. Subsequently, the anatomical T1-weighted MR image was segmented. The weighted functional images were spatially normalized into a common stereotactic space (MNI 152 T1-template) by applying the normalization
parameters from the segmentation procedure, and resampled to 2 x 2 x 2-mm³ isotropic voxels using trilinear interpolation. Finally, spatial smoothing was applied with an isotropic 3D Gaussian kernel of 8 mm full-width half-maximum.

10.4.7.2 General statistics

The data were modelled voxel-wise, using a general linear model. Trial-specific effects were modelled by trains of stick- or boxcar functions and convolved with the canonical hemodynamic response basis function of SPM8. Additionally, realignment parameters were included to model potential movement artefacts. The data were high-pass filtered (cut-off 128s) to remove low-frequency signal drifts, and a first-order autoregressive model was used to model the remaining serial correlations (163). Contrast images of testing parameter estimates coding condition-specific effects were created for each subject. The single-subject contrast images were entered into a voxel-wise full factorial model to assess main effects of task and task by drug interactions, implemented in a second-level random effects analysis. We report regions that survive cluster-level correction for multiple-comparisons (family-wise error, FWE) across the whole brain at p<0.05 using an initial height threshold of p < 0.001, unless otherwise indicated.

10.4.7.3 Task specific model definitions

For the extinction and recall task regressors were created for CS+, CS-, and context presentations for the habituation, early, and late phase. In addition, a regressor modelling the occurrence of instructions was added resulting in 10 regressors. Respective durations of each trial were included in the model. The model for the reextinction task was the same with the exception that there was no habituation phase and this was thus not modelled resulting in 7 regressors. Contrast images for each regressor were created against an implicit baseline and taken forward to random effects second level analyses.
10.4.8 Contingency questionnaire

The contingency questionnaire measures memory for the contingencies between the CSs and US as experienced on each day of the experiment. Subjects are presented with the CSs and first on a 10-point likert scale how pleasant they experienced seeing a given CS, this is done for each day of the experiment in order. Next, they indicate how many USs they experienced with the presentation of a given CS, and indicate the percentage of times the given CS co-terminated with the US, this is done for each day of the experiment in order.

10.4.9 Procedures

10.4.9.1 Screening

Subjects signed an informed consent form upon which they filled out a demographics and medical history questionnaire. Next, a 12-lead electrocardiogram was obtained, heart-rate, blood pressure, body height and body weight was measured, and the body-mass index was calculated. Finally, the list of Threatening Experiences Questionnaire and STAI-T was administered. When subject met all inclusion criteria and none of the exclusion criteria they were randomly assigned to a subject number. Drug randomization in groups of four coupled with subjects numbers by an independent pharmacy ensured double blind drug assignment. Subjects returned at a later time for the experimental sessions which were conducted on three consecutive days (Figure 10.1a).

10.4.9.2 Day 1

Upon arrival subject washed their hands and their fingers were cleaned with rubbing alcohol. Ag/AgCl electrodes filled with standard NaCl electrolyte gel were attached to the non-dominant index and middle finger with Velcro straps. Subjects entered a dummy scanner (i.e. a non-functional remake of a MRI scanner), and underwent a staircase procedure to adjust the strength of electrical shocks in order to
determine a level where the shock felt maximally uncomfortable, but not painful. SCR responses were checked with a Valsava maneuver. Upon receiving instructions, the conditioning task was started, and SCR was recorded (Figure 10.1a). If subjects did not show conditioned responses they were replaced.

10.4.9.3 Day 2

On the second day, subjects’ heart-rate and blood pressure were measured and mood questionnaires (STAI, PANAS) were filled out (Figure 10.1a). Next propranolol or placebo was administered (time-point 1). For safety purposes heart-rate and blood pressure were assessed after 30 minutes. After 60 minutes (time-point 2) heart-rate and blood pressure were assessed and mood questionnaires were administered (STAI, PANAS). Subject’s hands were cleaned as described for session 1. Next, subjects entered the MR scanner, Ag/AgCl electrodes were attached and subjects underwent a staircase procedure to set the shock level as described for Day 1. SCR responses were checked with a Valsava maneuver. After acquiring a fieldmap, the extinction task was started, fMRI data acquired and SCR and HR recorded. Next, a structural T1-weighted scans was obtained. Subjects exited the MR scanner, heart-rate and blood pressure was measured and mood questionnaires (STAI, PANAS) administered (time-point 3).

10.4.9.4 Day 3

On the third day (Figure 10.1a), subjects’ heart-rate and blood pressure were measured and mood questionnaires (STAI, PANAS) were filled out (time-point 4). Subject’s hands were cleaned as described for Day 1. Next, subjects entered the MR scanner, Ag/AgCl electrodes were attached and subjects underwent a staircase procedure to set the shock level as described for Day 1. SCR responses were checked with a Valsava maneuver. The recall, reinstatement, and reextinction tasks were conducted in order and a fieldmap was acquired prior to each task. SCR was recorded during each task. Upon exiting the MR scanner, heart-
rate and blood pressure was measured and mood questionnaires (STAI, PANAS) were administered (time-point 5). Finally, subjects filled out the contingency questionnaire, and the vTCI. Following completion of the study subjects were debriefed on the aims and details of the study.

10.5 Results

Here the results of the critical analyses are provided. Full and detailed analyses can be found in the Supplementary Information.

10.5.1 Participants

The placebo and propranolol group did not differ in terms of age, STAI-T scores, HR, and BP measures as obtained during screening (See supplemental information and Table 10.1).

10.5.2 Propranolol administration affects blood pressure on Day 2

At the start of Day 2, there was no baseline systolic BP difference between groups (t(44) = -0.048; p = 0.962, two-tailed). In line with previous findings (411, 491), we observe a drop in systolic BP in the propranolol group after 60 min (t(21) = 3.571; p = 0.002; Mean time-point 1: 122, s.e.m: 2.975, Mean time-point 2: 113.45, s.e.m: 2.628), and after 105 minutes (t(21) = 2.670; p = 0.014, Mean time-point 1: 122, s.e.m: 2.975; Mean time-point 3: 116.32, s.e.m: 2.270), and group differences at trend after 60 minutes (t(44) = -1.874; p = 0.068, Mean Propranolol: 113.45, s.e.m: 2.628; Mean Placebo: 119.33, s.e.m: 1.794) and 105 minutes (t(44) = -1.711; p = 0.094, Mean Propranolol: 116.32, s.e.m: 2.270; Mean Placebo: 121.29, s.e.m: 1.850), but no differences in diastolic BP (Figure 10.1c top panel). Although both groups show a decrease in HR over time, no group differences were observed. Furthermore, no group differences in BP or HR measures are observed on Day 3 24h after propranolol administration (Figure 10.1c bottom panel).
10.5.3 SCR results

We tested for the effect of propranolol administration on skin conductance responses for each task using repeated measures ANOVAs. A full overview of ANOVA results can be found in Supplementary Table 10.2. Here only the critical main effects of Group or interaction effects with Group are provided.

10.5.3.1 Both drug groups show development of conditioned responses during the conditioning task

Critically we find an interaction effect of CStype x Phase ($F_{1, 44} = 5.503, P < 0.024$) revealing that both drug groups learn differential responses over time (Figure 10.2). Following up on this finding, we found that differential conditioned responses were learned over time as paired T-tests on the difference scores between average responses to the CS+ and CS- for the Early and Late phase revealed greater differential responding in the Late compared to the Early phase ($t(45) = -2.356, P = 0.023$). Follow up T-test revealed that subjects had greater responses to the CS+ than the CS- during the Early phase ($t(45) = 3.760, P < 0.001$), and Late phase ($t(45) = 6.134, P < 0.001$). Hence, greater skin conductance responses to the CS+ compared to the CS- were observed, this difference increases over time, but critically, no differences between drug groups is detected.

10.5.3.2 Propranolol eliminates differential conditioned responses during the extinction task

For the extinction task (Figure 10.2) we find an interaction effect of CStype x Phase x Group ($F_{1, 44} = 7.779, P = 0.008$), but no main effect of Group ($F_{1, 44} = 0.471, P = 0.496$). Next, we find that the groups differed in differential conditioned responses in the Early Phase ($F_{1, 44} = 10.120, P = 0.003$) but not the Late Phase ($F_{1, 44} = 0.006, P = 0.938$) as revealed by one-way ANOVAs on the difference scores between average responses to the CS+ and CS-. Follow up paired T-tests revealed greater responses to the CS+ compared to the CS- during the Early phase in the Placebo
10. β-ADRENERGIC BLOCKADE AFFECTS THE NEURAL NETWORK OF EXTINCTION LEARNING, CAUSES THE LOSS OF FEAR AND PREVENTS THE RETURN OF FEAR IN HUMANS

β-adrenergic blockade affects the neural network of extinction learning, causes the loss of fear and prevents the return of fear in humans.

Group (t(23) = 5.127, P < 0.001), but not in the Propranolol Group (t(21) = 1.054, P = 0.308), and no differences during the Late Phase in both the Placebo Group (t(23) = 1.147, P = 0.263), and the Propranolol Group (t(21) = 1.341, P = 0.194). Interestingly, the effects were not caused by absolute group differences in responses to CS+ or CS- trials in the Early nor Late phase as revealed by independent samples T-tests; Early Phase CS+ (t(44) = -1.138, P = 0.261), Early Phase CS- (t(44) = 0.493, P = 0.624), Late Phase CS+ (t(44) = -0.813, P = 0.420), Late Phase CS- (t(44) = -1.065, P = 0.293). Thus, whereas the Placebo group showed greater responses to the CS+ compared to the CS- during the Early Phase, this differential effect is eliminated in the Propranolol group. Further, the Propranolol group displayed reduced responses during the habituation trials, and a trend towards attenuated responses to context presentations (See Supplemental Information).

Figure 10.2: Results skin conductance

On Day 1 subject in both group initially do not differentiate between CS+ and CS- trials in the early phase (mean over first half of the trials), show greater responses to the CS+ than CS- during the late phase (average over the last half of the trials), indicating successful conditioning. Critically no group differences exist on Day 1. On Day 2 the placebo group shows a differential conditioned response during the early phase of extinction, but this differential response disappears over the course of extinction learning in the late phase. In contrast, the propranolol group shows no differential conditioned responses during the early phase of extinction, and also not during the late phase. On Day 3 the placebo group shows evidence for spontaneous recovery of fear during the early phase of the recall task (average over the first third of the trials), but this effect disappears over time. In contrast, the placebo group shows no evidence for spontaneous recovery of fear. Furthermore, following a reinstatement procedure, the placebo group shows a reinstatement of fear during the re-extinction task, but the propranolol group shows no return of fear. Placebo group (solid bars), propranolol group (open bars), CS+ (red), CS- (blue), error bars reflect s.e.m.
10.5.3.3 Propranolol causes a loss of fear during the recall task

As the recall paradigm is principally a second extinction session and follows the actual extinction task, extinction learning can be expected to occur rapidly. In order to maximize sensitivity to spontaneous recovery effects, we therefore recalculated for each CS type (CS+, CS-) the Early Phase as the average skin conductance response over the trials 1-4, a Middle Phase as the average skin conductance response over the trials 5-8, and the Late Phase as the average response over trial 9-12. Critically, for the recall task (Figure 10.2) we find a main effect of Group ($F_{1, 44} = 4.968, P = 0.031$) and an interaction effect of Phase x Group ($F_{1.676, 73.739} = 7.565, P = 0.002$). Follow up one-way ANOVAs on the average responses to the CS+ and CS- reveal that the Placebo group showed greater general responses compared to the Propranolol group in the Early Phase ($F_{1, 44} = 7.852, P = 0.008$) and Middle Phase at trend ($F_{1, 44} = 2.966, P = 0.090$) but not the Late Phase ($F_{1, 44} = 0.490, P = 0.487$). Further, the Placebo group exhibited greater differential conditioned responses than the Propranolol group in the Early Phase ($F_{1, 44} = 4.116, P = 0.049$) but not in the Middle Phase ($F_{1, 44} = 0.026, P = 0.872$) and not the Late Phase ($F_{1, 44} = 0.236, P = 0.629$) as revealed by one-way ANOVAs on the difference scores between average responses to the CS+ and CS-. Follow up, paired T-tests revealed that only the placebo group showed greater responses to the CS+ compared to the CS- during the Early phase ($t(23) = 5.127, P < 0.001$), but not the Propranolol Group ($t(21) = 1.515, P = 0.145$), and both the Placebo Group ($t(23) = 1.181, P = 0.250$) and Propranolol Group ($t(21) = 1.573, P = 0.131$) show no differences during the middle phase and the Placebo ($t(23) = 0.797, P = 0.434$) and Propranolol group ($t(21) = 0.490, P = 0.630$) showed no differences during the late phase. In conclusion, propranolol application prior to extinction learning on Day 2 caused a loss of fear responses 24h later in the absence of drug. More specifically, this effect was most pronounced during the early phase of extinction recall. Further, the Propranolol group showed elimination of differential responses during the early phase compared to the Placebo Group. Thus, whereas the placebo group shows spontaneous recovery the propranolol group does not.
10.5.3.4 Propranolol prevents the return of fear during the re-extinction task

Following previous work (406) we calculated a reinstatement score as the difference between the first re-extinction trial and the last trial of the recall task for the CS+ and CS- for each subject. For the reinstatement task (Figure 10.2) we find a main effect of Group ($F_{1, 43} = 5.438, P = 0.024$), and an interaction effect of CStype x Group ($F_{1, 43} = 5.560, P = 0.023$). Follow up one-way ANOVAs on the average responses to the CS+ and CS- reveal that the placebo group showed greater general responses compared to the Propranolol group ($F_{1, 44} = 5.438, P = 0.024$). Critically, the Placebo group exhibited a differential reinstatement effect at trend ($t(23) = 2.005, P = 0.057$), but not the Propranolol Group ($t(20) = -1.371, P = 0.186$). Independent Samples T-tests revealed that this effect is specific to reinstatement group differences to the CS+ ($t(37.021) = 2.991, P = 0.004$), not the CS- ($t(43) = 0.232, P = 0.818$). Critically, the reinstatement effect was only detected for the CS+ in the Placebo Group ($t(23) = 3.376, P = 0.003$), not the Propranolol Group ($t(20) = -0.396, P = 0.696$). In conclusion, propranolol application prior to extinction learning prevents the return of fear 24h later.

10.5.4 The Propranolol group underestimates the number of shock received on Day 1 at the end of the experiment

At the end of the experiment on Day 3 subjects estimated the number of shocks they had received following the presentation of each CS on each day (Figure 10.3A). A Day (Day1, Day2, Day3) x CStype (CS+, CS-) x Group (Propranolol, Placebo) 3x2x2 repeated measure ANOVA with Group as between subjects variable revealed a main effect of Day ($F_{2, 86} = 112.951, P < 0.001$), a main effect of CStype ($F_{1, 43} = 97.888, P < 0.001$), a main effect of Group at trend ($F_{1, 43} = 2.904, P = 0.096$), an interaction effect of Day x Group ($F_{2, 86} = 3.102, P = 0.050$), an interaction effect of CStype x Group ($F_{1, 43} = 7.303, P < 0.010$), an interaction effect of Day x CStype ($F_{2, 86} = 83.084, P < 0.001$), and an interaction effect of Day x CStype x Group ($F_{2, 86} = 3.484, P < 0.035$). Next we calculated the difference in estimated
Number of USs that were received following the CS+ and CS- for each day per subject. Testing for the effect of Group, independent samples T-tests revealed a group difference on Day 1 \( (t(44) = -2.340, P = 0.024) \), but no differences on Day 2 or Day 3. Following up on these findings, independent T-tests revealed that the Propranolol group estimated a lower number of USs received following CS+ presentation on Day 1 \( (t(44) = -2.580, P = 0.013) \), but no group difference for the CS- \( (t(44) = 0.407, P = 0.686) \). Conclusively, when asked after completion of the study on Day 3, subjects who were administered propranolol on Day 2 underestimated the number of shocks they had received following CS+ presentation on Day 1.

10.5.5 The propranolol group accepts higher intensity settings of the shocker and this affects persists 24h later

On each day the setting of the shocker level was documented. Although we find no Group differences in shock levels on Day 1 between-subjects variance is high. Therefore we corrected for baseline differences by calculating for each subject the relative change in shock level on Day 2.
and Day 3 as a percentage of the setting on Day 1. A Day (Day 2, Day 3) x Group (Propranolol, Placebo) 2x2 repeated measure ANOVA with Group as between subjects variable revealed a main effect of Group ($F_{1, 42} = 4.969, P = 0.031$), with no other main effects or interactions (Figure 10.3B). One-way ANOVAs revealed a greater change in the setting of the shock level from Day 1 to Day 2 for the Propranolol group than the Placebo group ($F_{1, 43} = 4.852, P = 0.033$), and for the change from Day 1 to Day 3 at trend ($F_{1, 42} = 3.411, P = 0.072$). Limiting the analyses to shock level settings within groups, paired T-tests reveal an increase in shock level setting in the Propranolol Group from Day 1 to Day 2 ($t(20) = -2.660, P = 0.015$), and from Day 1 to Day 3 ($t(20) = -2.210, P = 0.039$), but not in the Placebo Group from Day 1 to Day 2 ($t(23) = 0.756, P = 0.458$), and not from Day 1 to Day 3 ($t(22) = 0.569, P = 0.575$). Thus, propranolol application on Day 2 causes subjects to accept higher shock level settings, and this effect persists 24h later in the absence of drug.

10.5.6 fMRI results

We first tested which brain regions were involved in fear and safety memory for each task (Tables 10.3, 10.5, 10.7). Subsequently we tested how these regions were influenced by propranolol administration using orthogonal contrasts. Specifically, following previous region of interest analyses in the field (10, 524, 525), we extracted data from the regional clusters revealed by comparing CS+ trials to the CS- trials for each task. Here only analyses on the extracted data for regions that showed an effect of drug administration on differential conditioned responses are reported. Additional analyses for regions that do not show effects of propranolol are provided in Supplementary Information. Furthermore, we run for each task a Phase (Early, Late) x CStype (CS+, CS-) x Group (Propranolol, Placebo) 2x2x2 repeated measure ANOVA on the regions of interest. A full overview of ANOVA results can be found in Table 10.9. Here only the critical main effects of Group or interaction effects with Group are provided.
The dorsal mPFC and midbrain are affected by propranolol administration during extinction learning

Comparing CS+ trials to the CS- trials reveals differential BOLD signal in regions including the insula, dorsal mPFC, ventral mPFC, and midbrain (Table 10.3, Figure 10.4).

Testing for the effect of propranolol administration on dorsal mPFC responses (Figure 10.5) we find an interaction effect of CStype x Group ($F_{1,44} = 4.894, P = 0.032$). Critically, within the Placebo group paired T-tests revealed greater responses to the CS+ than CS- ($t(23) = 3.062, P = 0.006$), but no such effect was observed within the Propranolol Group ($t(21) = 0.223, P = 0.826$). Thus, the Placebo group exhibits greater dorsal mPFC responses to the CS+ than CS- whereas no such effect is detected in the Propranolol group. Further, within the Placebo group the differential
**10. β-ADRENERGIC BLOCKADE AFFECTS THE NEURAL NETWORK OF EXTINCTION LEARNING, CAUSES THE LOSS OF FEAR AND PREVENTS THE RETURN OF FEAR IN HUMANS**

**Figure 10.5: Results drug effects on fMRI**

Propranolol administration affects the neural network of extinction learning. Analyses on the extracted data from regions revealed by the main effects of task shows that during extinction learning propranolol eliminates differential conditioned responses in the dorsal mPFC whilst increasing differential responses in the midbrain. During the recall task the propranolol group shows differential conditioned responses in the hippocampus, an effect not observed in the placebo group. During the re-extinction task the propranolol group shows reduced differential conditioned responses in the dorsal mPFC. Placebo group (solid bars), propranolol group (open bars), CS+ (red), CS- (blue), early = average over the first half of the trials, late = average over second half of the trials, error bars reflect s.e.m.

Dorsal mPFC responses disappear over the course of extinction learning (See Supplementary Information).

Further, testing for the effect of propranolol on midbrain responses (Figure 10.5), revealed an interaction effect of CStype x Group ($F_{1, 44} = 4.509, P = 0.039$). Critically, within the Placebo group paired T-tests revealed no differential responses between the CS+ and CS- ($t_{23} = 1.045, P = 0.307$), whereas the Propranolol group showed greater responses to the CS+ than CS- ($t_{21} = 3.045, P = 0.006$). Thus, the Propranolol group exhibited greater midbrain responses to CS+ compared to CS- trials than the Placebo group.
The hippocampus is affected by propranolol administration during recall of extinction learning

Comparing CS+ trials to the CS- trials reveals differential BOLD signal in regions including the right insula, midbrain, hippocampus, and amygdala (Table 10.5, Figure 10.4).

Testing for the effect of propranolol administration on left hippocampus responses (Figure 10.5) we find an interaction effect of CStype x Group at trend ($F_{1,44} = 3.820, P = 0.053$), with no other main effects or interactions. Critically, the Placebo group showed no differences between the CS+ and CS- in the Early Phase ($t(23) = 0.097, P = 0.924$), and at trend in the Late Phase ($t(23) = -2.039, P = 0.053$), but we found clear differences in the Propranolol group in the Early Phase ($t(21) = -3.085, P = 0.006$) and the Late Phase ($t(21) = -2.615, P = 0.016$). In conclusion, the Propranolol group showed greater differential hippocampal responses during the early
β-adrenergic blockade affects the neural network of extinction learning, causes the loss of fear and prevents the return of fear in humans phase of recall, evidence for this differential response only arises in the late phase for the Placebo group.

10.5.6.3 The dorsal mPFC shows reduced differential responses during the re-extinction task.

Comparing CS+ trials to the CS- trials reveals differential BOLD signal in regions including the dorsal mPFC, and insula (Table 10.7, Figure 10.4).

Testing for the effect of propranolol administration on dorsal mPFC responses (Figure 10.5) revealed a main effect of Group ($F_{1, 43} = 4.371, P = 0.042$). When averaging over all conditions an independent samples T-test showed that the Placebo group displayed greater responses than the Propranolol group ($t(43) = 2.091, P = 0.042$). In conclusion, during re-extinction the Placebo group displayed greater dorsal mPFC responses.

10.5.7 fMRI BOLD responses in areas affected by propranolol correlate with SCR responses and US estimation

We find effects of propranolol on extinction learning, recall, and the recovery of fear, and additionally find similar effects of propranolol on hippocampus @ recall

Figure 10.7: Correlations BOLD fMRI and US estimation

The differential BOLD fMRI response in the hippocampus during recall correlates with the differential number of shock subjects estimate to have received on Day 1 following completion of the study. Thus, the greater the differential hippocampus response the lower the estimated number of received shocks.
brain function. As such we attempted to determine whether changes in brain function could be associated with changes in behaviour. Therefore, we correlated the difference in SCR scores to CS+ and CS- trials with the difference in BOLD fMRI responses to CS+ and CS- trials for those regions of interest that showed an effect of drug administration during the early task phases that showed the largest differences between CS+ and CS- trials (Figure 10.6). We find positive correlations between SCR responses and neural responses in the dorsal mPFC during the early phase of the extinction task ($r_s = 0.346, P = 0.020$), and during the early phase of the recall task at trend ($r_s = 0.276, P = 0.066$). Thus, greater differential dorsal mPFC responses are associated with greater differential SCR response. Furthermore, we find a negative correlation between hippocampus responses during the recall task and the subsequent reinstatement effect in SCR responses during the re-extinction task ($r_s = -0.337, P = 0.025$). Hence, greater differential hippocampus responses are related to smaller differential SCR responses. In summary, brain areas that are affected by propranolol administration show a relationship to the differential conditioned responses as expressed in SCR measures. These correlations were not significantly different between groups (using Fisher’s z-transformation). Considering these findings, we tested whether BOLD fMRI responses might also correlate with the US estimation for Day 1. We find that greater differences in US estimation between the CS+ and CS- correlate negatively with the hippocampus response during recall ($r_s = -0.292, P = 0.049$, Figure 10.7), and again this correlation does not differ between groups. Thus, greater differential hippocampus responses are also associated with a smaller number of US estimation.

10.5.8 Ventral striatal prediction error responses are not affected by propranolol administration and show no relation with the loss of fear

To further investigate the three different hypotheses, and distinguish between the neural mechanisms that may underlie our behavioural results we set out to investigate neural reinforcement learning signals as the three hypothesis yield different predictions. One explanation
for the loss of fear and prevention of the return of fear on Day 3 in the Propranolol group is a disruption of reconsolidation. Reconsolidation only occurs when memory is reactivated in a situation that requires new learning to update memory (403). This has lead to the suggestion that a reconsolidation is only initiated when there is a mismatch between the expectation of the retrieved memory and the outcome of the actual situation, i.e. a prediction error (532, 533). If the loss of fear and the absence of the return of fear that we report here is attributable to a disruption of reconsolidation, then a relationship between a neural prediction error response and the amount of loss of fear would provide supportive evidence for this explanation. An variant of the new learning hypotheses is that propranolol enhances extinction learning by affecting a reinforcement learning signal, that may subsequently also effect the consolidation process. If propranolol enhances extinction by affecting reinforcement learning signals we should be able detect a group difference in prediction error related neural responses. Therefore, in a complementary analysis we investigated the neural basis of errors in aversive prediction during the extinction task (see Supplemental Information). From a Rescorla-Wagner model (534) we generated a prediction error regressor for each subject, and, after taking covariates of no interest into account, group analyses revealed a cluster within the ventral striatum (Figure S10.4). An independent samples T-test on the extracted data from this cluster does (Figure S10.4), however, not reveal an effect on prediction error related BOLD responses (t(43) = 1.022, P = 0.313; Placebo Mean: 2.485, s.e.m., 0.672, Propranolol Mean: 1.432, s.e.m., 0.790). Investigating the relation between the ventral striatal prediction error response and the differential SCR response during extinction learning reveals that greater mean differential SCR responses are correlated with greater ventral striatal PE responses (r_s = 0.396, P = 0.007, Figure S9.4). Critically however, ventral striatal PE responses are not correlated with spontaneous recovery (r_s = 0.039, P = 0.801), nor the reinstatement of SCR responses on Day 3 (r_s = 0.136, P = 0.379). Additionally, we find no group differences in these correlation scores (using Fisher’s Z-transformation). Thus, although the neural prediction error response correlates with the reduction in SCR responses during
extinction learning we find no group differences in ventral striatal PE responses and no relation to the loss of fear or the return of fear on Day 3 and, hence, no support for enhanced extinction or a reconsolidation explanation.

10.6 Discussion

We show in a double-blind placebo controlled experiment that a single dose of the β1-β2-adrenergic receptor antagonist propranolol administered prior to extinction learning eliminates differential conditioned fear responses during extinction learning, causes a subsequent loss of fear and prevents the return of fear one day later. The elimination of conditioned responses during extinction learning was associated with a loss of responses to a fearful stimulus (CS+) in the dorsal mPFC and increased conditioned responses in the midbrain. The loss of fear one day later was accompanied by increased conditioned responses in the hippocampus during recall, whereas the absence of the return of fear was reflected in reduced conditioned responses in the dorsal mPFC. These findings support the second hypothesis that propranolol reduced retrieval of conditioned fear response during the extinction task, and due to new learning led to an absence of fear during the recall and re-extinction tasks. The combined used of noradrenergic medication and psychotherapy may thus be a beneficial strategy to treat anxiety disorders.

The dorsal mPFC is considered to be the human homologue of the rodent prelimbic cortex (303, 304), and suggested to be a memory storage site that drives the sustained expression of conditioned fear responses (535-538). In line with our findings, firing rates of neurons in prelimbic cortex increase during conditioning, decrease during extinction (537, 538), and are reduced after systemic propranolol injections, resulting in reduced expression of conditioned fear (522). Similarly, we found that propranolol eliminated enhanced responses to the CS+ in the dorsal mPFC. We also observed greater differential conditioned responses for the propranolol group in a midbrain region that largely overlaps with the periaqueductal gray (PAG). The PAG has been implicated as an output structure for
specific reactions to fearful stimuli, such as freezing (5, 539). Activation of the PAG causes expression of fear responses (540), yet in contrast we found increased midbrain responses in the absence of fear responses following propranolol application. This finding is indicative of a potential compensatory mechanism for ineffective noradrenergic signalling. Furthermore, we observed that propranolol induced differential hippocampus responses during the recall task. Moreover, we already detected greater hippocampus responses to context presentations during the extinction task that persisted during the recall and re-extinction tasks. Previous studies suggested a critical role for the hippocampus in context representations (208) and context conditioning (524-526, 541). Critically, where we observed that the reduction of dorsal mPFC responses were related to reduced fear responses, greater hippocampus responses as detected in the propranolol group were related to both reduced measures of implicit and explicit fear. The latter findings suggest that the hippocampus provides a safety signal resulting in reduced fear.

Our results extend our current understanding of the role of noradrenaline in fear and safety learning and provide support for the second hypothesis that blocking noradrenaline reduces fear memory retrieval and prevents the return of fear due to new learning. Cue-evoked phasic rises in noradrenaline (34) appear to regulate the retrieval of fear memory via the dorsal mPFC (522, 535-538, 542), possibly via dynamic interactions with midbrain structures including the locus coeruleus. This brain region is the central source of noradrenergic projections (485) that has previously been implicated in extinction learning, retention, and retrieval (411, 492, 519, 521, 543-545), and perhaps also enhances the amygdala contribution to retrieval (487, 489, 502, 546). Considering that in our paradigm extinction, recall, and re-extinction occurred in a different context as in which conditioning took place, we hypothesize that the elimination of differential conditioned responses during extinction prevented the formation of an association between the new context and the retrieved fear memory or even caused the formation of a novel memory association between the new context and a safe situation (449). Here phasic noradrenaline release associated with the
expression of fear would function as an intrinsic US (547) resulting in the association of conditioned fear with a novel context effectively supporting fear generalization. Blocking the differential noradrenergic response prevents contextual fear generalization, and leads to an association of a new context with relative safety, as expressed in the introduction of a differential conditioned response in the hippocampus. This critically indicates that phasic noradrenergic signals are directly or indirectly involved in the modulation of hippocampal function in contextual fear and safety learning.

Alternatively noradrenergic blockade may result in a general reduction in fear expression by affecting output structures responsible for the behavioural expression of fear (548-552), or tonic noradrenergic input through engagement of β-adrenergic receptors may be essential for retrieval of fear memory (492, 499). Critically, during the extinction task we found an interaction effect between drug administration and conditioned stimuli but no main effect suggesting not a general decrease in fear responses or memory retrieval but an inability to differentiate between stimuli at the time of cue-evoked retrieval. Subsequently, we observed a loss of fear and no return of fear during the recall and re-extinction tasks, not merely an elimination of differential conditioned responses. Even though we detected a main effect of reduced responses during habituation trials and context presentations during the extinction task, no such general reduction of fear expression was observed for the recall and re-extinction tasks (see Supplemental Information). Hence, a general reduction of fear expression could only partially explain the results.

Furthermore, alternative explanations for our finding that β-adrenergic blockade during extinction learning results in a subsequent loss of fear and prevents the return of fear one day later are a disruption of reconsolidation (50, 507), impaired lingering consolidation (446), or enhanced extinction (446). Distinguishing between these alternative hypotheses is clinically relevant as a disruption of lingering consolidation or reconsolidation suggest a loss of the original fear memory, and
thus a reduced risk for the return of fear. Our findings argue against a reconsolidation or lingering consolidation interpretation as these would predict the loss of memory representations. Although the propranolol group displayed a loss of differential responses in the dorsal mPFC during the extinction task, only a decrease in conditioned responses was detected during the re-extinction task. Moreover, many brain regions still reflected conditioned responses despite a behavioural absence of fear. Furthermore, in contrast to the loss of memory representation we actually observe the introduction of a novel hippocampus memory representation in the propranolol group. The results of our prediction error analyses yield further arguments against the alternative explanations of reconsolidation and enhanced extinction learning. First, the loss of fear and the absence of the return of fear could have been the result of enhanced extinction due to increased reinforcement learning signals by propranolol administration. Second, reconsolidation has been postulated to be initiated in case of a mismatch between expectancy and outcome, i.e. a prediction error, at the time of memory reactivation (403, 532, 533). In contrast to these predictions we found ventral striatal prediction error related responses to be associated with extinction learning, but these were not affected by propranolol and had no relation with the loss of fear or the return of fear. Finally, the enhanced extinction hypotheses would predict greater ventral mPFC responses as this region is critical for the inhibition of fear; however, we observed no such effect. In conclusion, we found no supportive evidence for a disruption of reconsolidation, impaired lingering consolidation, or an enhanced extinction explanation. Interestingly, this conclusion also casts doubt on studies investigating reconsolidation by applying beta-blockers prior to memory reactivation and that lack tests of immediate memory.

Our findings speak to an important discrepancy in the literature on the influence of beta-blockers on extinction. Our results are in line with studies that report a role for noradrenaline in the retrieval of fear memory (411, 514, 515) and expression of fear during extinction learning (520, 522), but contrast with studies that find no effects on fear expression during extinction learning (519, 521, 553), and those studies
that report impaired consolidation of extinction due to noradrenergic blockade that results in increased fear one day later (519, 520, 553). The detection of an effect of noradrenergic blockade on fear memory retrieval was only possible because we employed a differential conditioning paradigm, whereas most animal studies have not used such a design (519-522). Also, we used systemic propranolol application and found effects on the dorsal mPFC and hippocampus in line with animal studies applying systemic administration or local infusions within these regions (522, 542), whereas several animal studies have performed local infusions in other regions (519, 520), possibly explaining discrepancies. Further, in our study extinction and memory retention was tested in a novel context and we have argued that the loss of fear and the absence of the return of fear is the result of new contextual safety learning. In contrast, animal studies that have reported impaired consolidation of extinction due to noradrenergic blockade have tested extinction memory in the same context as in which conditioning took place (519, 520, 553).

Finally, a recent human study reports effects of beta-blockers on explicit expectancy ratings of fear but not on an implicit measure of fear expression (553), supporting the idea that multiple memory systems (230, 233, 466) may be effected differently by noradrenergic blockade (476, 515). Our findings support these ideas, as we found that although subjects explicitly know which stimulus was threatening, and found this stimulus less pleasant, they no longer showed implicit fear responses following noradrenergic blockade. Moreover, while we detected a loss of fear memory representations in the dorsal mPFC, several other brain areas retained these representations. A critical question for future research pertains to the contribution of these regions to fear memory, and the possibility for the expression of fear to recover or the maintenance of a cognitive representation of fear that may be less sensitive to disturbances (466).

To conclude, our results provide supportive evidence for a method to reduce fear expression, prevent generalization of fear to contexts in which fear is maladaptive, and that prevents the return of fear following
treatment. Our results extend our current understanding of the role of noradrenaline in fear memory retrieval and the generalization of fear to novel contexts. Taken together we provide face validity to studies investigating the combined used of noradrenergic medication and psychotherapy aiming to improve treatment and reduce the disease burden of affective disorders.

Competing interests

The authors declare no conflict of interest.

Acknowledgements

We would like to thank S Kooijman for additional subject recruitment, screening and clinical coordination, and N Müller for contributions to reinforcement learning analyses.

Author contributions

MCW Kroes, GA van Wingen and G Fernández designed the study. MCW Kroes, KD Tona, and S Vogel acquired data. MCW Kroes analyzed data. MCW kroes and H den Ouden programmed the reinforcement learning algorithm. MCW Kroes and G Fernández wrote the manuscript. All authors agreed on the final manuscript.
11 THE SEROTONIN TRANSPORTER-LINKED PROMOTER REGION POLYMORPHISM MODULATES FEAR CONDITIONED RESPONSES IN THE HUMAN BRAIN.

11.1 In preparation as:


11.2 Abstract

Understanding vulnerability factors such as genetic variations on brain functioning holds the promise to tailor psychiatric treatment to individual needs. The serotonin transported-linked polymorphic region (5-HTTLPR) has been associated with anxiety and mood disorders and one hypothesis is that genetic variance in 5-HTTLPR contributes to susceptibility for affective disorders by modulating emotional learning. We therefore tested whether genetic variance in 5-HTTLPR determines emotional learning and modulates specific brain processes as assessed by functional magnetic resonance imaging. We show that 5-HTTLPR has no effect on Pavlovian fear conditioning and extinction as assessed with skin conductance responses, but does modulate functioning of the neural network that supports fear and safety learning. Furthermore, we show that critical nodes in the neural network of fear learning that are modulated by the variance in 5-HTTLPR are related to sympathetic fear responses, whereas other brain regions show a genetic variance-dependent differential association with coping strategies. Our findings suggest that 5-HTTLPR mediates emotional learning, thus contributing to a mechanistic understanding of the neural processes that determine individual susceptibility to affective disorders, and the development of optimized treatments for specific individuals.
11.3 Introduction

To tailor psychiatric treatment to individual needs recent research has focussed on understanding vulnerability factors for anxiety and mood disorders (554, 555). Particular attention has been directed towards the effects of genetic variations on functioning of brain regions involved in fear and mood (555-557). Understanding the effect of genetic variations in serotonin neurotransmission has been an important focus as studies have revealed associations between short (S)-allele carriers of the serotonin transporter-linked polymorphic region (5-HTTLPR) and anxiety-related traits as well as neuroticism (558), posttraumatic stress disorder (559, 560), and depression (561, 562).

The functional deletion/insertion polymorphism within 5-HTTLPR produces short (S') and long (L') repetitive sequences, which influences 5-HTT protein expression and function (558, 563). The functional consequence of the S'-allele has been associated with relatively less 5-HTT mRNA transcription and 5-HTT binding, reduced 5-HT platelet reuptake (563-565), and reduced 5-HT1a binding in the brain (566). Hence, genetic variance in 5-HTTLPR influences serotonergic signalling, but an open critical question pertains to the functional neural mechanism by which genetic variance in 5-HTTLPR affects susceptibility to the development of anxiety and depression disorders.

Emotional learning contributes to the development of anxiety disorders (3), and an early study showed an association between genetic variance in 5-HTTLPR and emotional memory in humans (463). An effective model for fear and anxiety has been Pavlovian fear conditioning, which allows the investigation of fear and safety learning (5). During Pavlovian conditioning a neutral conditioned stimulus (CS, e.g. a light) comes to elicit a conditioned fear response (CR) after repeated pairing with an aversive unconditioned stimulus (US). The conditioned fear response can subsequently be reduced or suppressed through repeated exposure to the CS in the absence of the US in a procedure known as extinction learning. Importantly, extinction learning does not erase the CR memory but
reflects new learning, namely the association between the CS and safety (567). Thus, the conditioned memory trace and the extinction memory trace coexist and compete, such that the extinction memory inhibits expression of the CR (10, 512, 513). Studies have reported increased sensitivity to conditioning and impaired extinction learning in S'-carriers of the 5-HTTLPR in both humans and animals (568-571), although other have not found effects on behavior (572, 573).

Investigating the influence of genetic variance in 5-HTTLPR on the neural network that supports fear and safety learning may provide mechanistic insight into individual pathophysiology that would allow personalized treatment. The neural circuit that supports fear and safety learning in Pavlovian conditioning and extinction paradigms is well defined (5, 518). The amygdala is considered to mediate (conditioned) responses to adversity and the storage of associative fear memory (10, 12, 512, 513), in close coordination with the dorsal medial prefrontal cortex (dMPFC), and midbrain regions. Following extinction learning the ventral medial prefrontal cortex (vMPFC) is thought to inhibit the fear response (10, 12, 512, 513). Additionally, extinction is context specific and the hippocampus is thought to provide this contextual information (523). Together, the idea is that during extinction learning a novel memory is formed through hippocampus-amygdala-MPFC interactions, which, following consolidation, leads to a hippocampus-dependent vMPFC inhibition of the amygdala preventing the CR (513). Exactly this mechanism appears impaired in anxiety disorders (21), and this network has been found to be affected by effective pharmacological treatment targeting serotonergic transmission in depression (3, 25). Beyond these regions, sensory cortical areas have been implicated in conditioned fear learning and have been found to be modulated by serotonergic treatment in animals (574, 575). Furthermore, recent studies provided initial evidence for an association between genetic variance in 5-HTTLPR and amygdala responses (576), and amygdala-vMPFC coupling (577, 578) during exposure to emotional events. More recent work in animals reported a change in amygdala-vMPFC communication via theta-synchronization as a consequence of 5-HTTLPR genotype during Pavlovian conditioning (570). Additionally, recent human
work has revealed enhanced conditioned responses in S'-carriers in the insula, dMPFC, and occipital cortex (572, 573), and decreased amygdala responses during conditioning and vMPFC responses during extinction learning (573), albeit in relative small sample sizes (579, 580).

Here we set out to test the influence of 5-HTTLPR genotype on the neural network of fear and safety learning. Moreover, we were particularly interested in the relation between elements of the neural network that may be modulated by 5-HTTLPR and their relation to fear responses and fear regulation. Briefly, 106 subjects were differentially conditioned to coloured squares by pairing with an electrical shock in a partial reinforcement regime. The conditioning task was immediately followed by an extinction task during which the same coloured squares were presented but no shocks were administered. To investigate the influence of 5-HTTLPR genotype on the neural network of fear and safety learning BOLD fMRI was acquired throughout the experiment as well as skin conductance responses.

11.4 Methods and Materials

11.4.1 Participants

One hundred and six predominantly right-handed healthy young male participants with normal- or corrected-to-normal vision, and normal uncorrected hearing were included in the study (age-range: 18-31 yr, mean: 21.87, s.e.m.: 2.63). Six subjects could not be genotyped, and one subject was excluded due to SCR data loss. The S'-allele carriers comprised 69 subjects and the L' L'-allele carriers 30 subjects. All subjects gave informed consent, were free of cardiovascular, neurological, endocrine, and psychiatric history. Subjects did not consume more than three alcoholic beverages daily, did not use psychotropic medication or recreational drugs weekly or more, and were not habitual smokers. Subjects did not consume alcohol in the 24h prior to testing and did not use psychotropic medication or recreational drugs in the 72 hours prior
11. The serotonin transporter-linked promoter region polymorphism modulates fear conditioned responses in the human brain

to testing. The study was approved by the institutional ethics committee (CMO Regio Arnhem-Nijmegen, The Netherlands).

11.4.2 Genotype

Genetic analyses were carried out at the Department of Human Genetics of the Radboud University Nijmegen Medical Centre, in a laboratory that has a quality certification according to CCKL criteria. High molecular weight DNA was isolated from saliva using Oragene containers (DNA Genotek, Ottawa, Ontario, Canada) according to the protocol supplied by the manufacturer. The 5-HTTLPR genotype was the genotype at the single-nucleotide polymorphism (rs25531) in this region, with alleles designated as LG, LA, and S (581). Because it is thought that the LG is comparable to the S-allele with regard to gene transcription and function, we reclassified the alleles on the basis of the resulting expression levels of the SLC6A4/5-HTT gene into L’L’, S’L’, and S’S’ genotypes (582). Testing for Hardy-Weinberg equilibrium did not show deviation from the expected distribution (P < 0.05).

11.4.2.1 STAI-T

The State-State Anxiety Inventory (STAI-T) measures anxiety as a general long lasting quality (trait) (200). Subjects indicate on a 4-point likert scale for 20 statements to which degree these apply to them.

11.4.2.2 CISS

The Coping Inventory for Stressful Situations (CISS) measures coping strategies in adult samples (583). Subjects indicate on a 5-point likert scale for 48 statements to which degree these apply to them. The CISS assesses three basic coping dimensions: Task-oriented coping (coping by altering the situation), emotion-oriented coping (coping by regulating emotional distress), and avoidance-oriented coping (coping by distraction or seeking other peoples company).
Figure 11.1: Conditioning-extinction task

Subjects were cue-conditioned to coloured squares. Following a fixation cross, either a blue or yellow square was presented serving as conditioned stimuli. During the conditioning part of the experiment one of the coloured squares was followed by a transcutaneous electrical shock 33% of the time, the other coloured square was never followed by a shock. During the extinction part of the experiment neither coloured squares were followed by a shock.
11.4.3 Conditioning-extinction task

A differential delay conditioning paradigm (Figure 11.1) with partial reinforcement was conducted (524-526). Partial reinforcement was used to study the gradual development of fear learning, and slow down subsequent extinction learning (527, 528). Subjects were told that they would see on a computer screen a yellow or blue square while they could receive shocks. The level of the shocks was set before the experiment to a subjective intensity that was maximally uncomfortable without being painful to the subject. The instructions were to pay attention to the computer screen and subjects were told that a relationship existed between the stimuli and the shocks. The US was an electric shock to the fingers (200msec asymmetrical biphasic pulse with a pulse width of 250 μs at 150 Hz.) that co-terminated with the CS+. On a given trial a CS was presented for 4 sec, with an 11, 12, or 13 sec inter-trial interval during which a fixation cross was presented. During conditioning, one square colour was paired with the US on one-third of the trials (CS+), and the other colour was never paired with the US (CS-). Thus, during conditioning each CS was presented 18 times, out of which 6 CS+ trials were reinforced. During extinction each CS was presented 12 times and never paired with a shock. Only non-reinforced trials were used for analyses, thus leaving 12 trials of each stimulus type per task. The order of the different trial types was pseudo-randomized so that half of each trial type occurred during the first half of the task (Early Phase), whilst the other half of each trial type occurred during the second half of the task (Late Phase). The assignment of light colour into CS+ and CS- was counterbalanced across subjects.

The conditioning task was preceded by a habituation phase consisting of 6 trials, during which each CS type was presented three times. Subjects were instructed that no shocks would occur during the habituation phase. The habituation phase and conditioning task were separated by a visual instruction on a computer screen indicating that the actual task was going to start. In between the conditioning and extinction task, subject saw an instruction screen informing them on a short (20sec) break after which the task would continue in the same manner.
11.4.4  Skin conductance acquisition

Skin conductance was assessed with Ag/AgCl electrodes attached to the subject’s distal phalanges of the index and middle finger of the non-dominant hand with standard NaCl electrolyte gel. The skin conductance signal was amplified using MR compatible BrainAmp MR and BrainAmp ExG MR (BrainProducts GmbH) within the MR environment, through an optical cable recorded outside the MR environment using BrainVision Recorder software. Data was continuously recorded at 5000 samples per second.

11.4.5  Skin conductance analyses

Skin conductance data was assessed using an in-house analysis programme written in Matlab (the MathWorks) and using FieldTrip (529). Data was low-pass filtered at 5Hz and resampled to 100Hz. The level of skin conductance responses was determined for each trial as the peak-to-peak amplitude difference in skin conductance of the largest deflection in the latency window from 0-8 s after stimulus onset. The raw skin conductance responses were square root transformed to normalize the distributions.

11.4.6  fMRI

MR data was acquired on a 1.5 T MR scanner (Avanto, Siemens, Medical, Erlangen, Germany) equipped with an 32-channel transmit-receiver head coil. The manufacturer’s automatic 3D-shimming procedure was performed at the beginning of each experiment. Subjects were placed in a light head restraint within the scanner to limit head movements during acquisition.

11.4.6.1  Functional image acquisition

Functional images were acquired with gradient echo-planar imaging (EPI) sensitive to the blood-oxygenation level dependent (BOLD) response using
the following parameters: 32 oblique transverse slices, slice thickness = 3.5mm, interslice gap = 10%, repetition time (TR) = 2.34s, flip angle $\alpha = 90^\circ$, echo times (TE) = 35 ms, (FOV) = 212 x 212 mm2, matrix size 64 x 64, fat suppression.

11.4.6.2 Structural image acquisition

A 3D magnetization-prepared rapid gradient echo (MPRAGE) image was acquired for normalization procedures using the following parameters: 176 slices, slice thickness = 1.0 mm, no slice gap, TR = 2730 ms, TE = 2.95 ms, FOV = 256 x 256 mm2, matrix = 256x256.

11.4.6.3 Functional imaging data analyses

11.4.6.4 Preprocessing

fMRI data were processed and analysed using the statistical software package SPM8 (Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm). The first 5 EPI volumes were discarded to allow for T1 equilibration. The remaining functional images from each subject were realigned using a three step coregistration procedure where first each subject’s anatomical T1-weighted MR image, and second each subjects’ mean-, and weighted functional images were coregistered to the MNI 152 T1-template, and third each subject’s anatomical T1-weighted MR image was coregistered to the mean functional image using a normalized mutual information function. Subsequently, the anatomical T1-weighted MR image was segmented. The weighted functional images were spatially normalized into a common stereotactic space (MNI 152 T1-template) by applying the normalization parameters from the segmentation procedure, and resampled to 2x2x2-mm3 isotropic voxels using trilinear interpolation. Finally, spatial smoothing was applied with an isotropic 3D Gaussian kernel of 8 mm full-width half-maximum.
11.4.6.5 General statistics

The data were modelled voxel-wise, using a general linear model. Trial-specific effects were modelled by trains of stick- or boxcar functions and convolved with the canonical hemodynamic response basis function of SPM8. Additionally, realignment parameters were included to model potential movement artefacts. The data were high-pass filtered (cut-off 128s) to remove low-frequency signal drifts, and a first-order autoregressive model was used to model the remaining serial correlations (163). Contrast images of testing parameter estimates coding condition-specific effects were created for each subject. The single-subject contrast images were entered into a voxel-wise full factorial model to assess main effects of tasks and tasks by genotype interactions, implemented in a second-level random effects analysis. We report regions that survive cluster-level correction for multiple-comparisons (family-wise error, FWE) across the whole brain at p<0.05 using an initial height threshold of p < 0.001, unless otherwise specified.

11.4.6.6 Task specific model definitions

For the conditioning-extinction task regressors were created for the CS+ and CS- for the habituation, and early and late phase per task. Reinforces CS+ trials were separately modelled for the early and late phases of each task. In addition, a regressor modelling the occurrence of shocks and a regressor modelling the presentation of instructions were added. Respective durations of each trial were included in the model. Contrast images for each regressor were created against an implicit baseline and taken forward to random effects second level analyses.

11.4.7 Procedures

Upon arrival subject washed their hands and their fingers were cleaned with rubbing alcohol. Ag/AgCl electrodes filled with standard NaCl electrolyte gel were attached to the non-dominant index and middle finger with Velcro straps. Subjects underwent a staircase procedure to
adjust the strength of electrical shocks in order to determine a level where the shock felt maximally uncomfortable, but not painful. SCR responses were checked with a Valsava maneuver. Upon receiving instructions, the conditioning-extinction task was started during which fMRI and SCR data was recorded. Following completion of the study subjects were debriefed on the aims and details of the study.

11.5 Results

Here the results of the critical analyses are provided. Full and detailed analyses can be found in the Supplementary Information.

11.5.1 Participants

The S’-carriers group and L’L-carriers group did not differ in terms of age or STAI-T scores (Table 11.1).

11.5.2 SCR results show that both genotypes develop conditioned responses during the conditioning task and lose fear with extinction training.

To test for possible baseline differences between genotypes we tested a CStype (CS+, CS-) x Trial (1,2,3) with Genotype (S’-carriers, L’L’-carriers) as a between subjects variable 2x3x2 repeated measure ANOVA on the habituation trials which revealed a main effect of Trial ($F_{2, 194} = 12.77, P < 0.001$). Further, A Trial (1,2,3,4,5,6,) x genotype (S’-carriers, L’L’-carriers) 6x2 repeated measure ANOVA on the responses to the US revealed a main effect of Trial ($F_{1, 97} = 59.08, P < 0.001$), with no other main effect or interaction. Thus, there were no baseline differences and no differences in response to shock application between genotypes.

For each CS type (CS+, CS-) skin conductance responses were averaged over the first half of the trials (Early Phase) or the last half of the trials (Late Phase) for each task (Conditioning, Extinction). A Task (Conditioning, Extinction) x Phase (Early, Late) x CStype (CS+, CS-) with Genotype (S’-
carriers, L’L’-carriers) as a between subjects variable 2x2x2x2 repeated measure ANOVA revealed a main effect of Task (F₁, 97 = 31.93, P < 0.001), Phase (F₁, 44 = 43.78, P < 0.001), CStype (F₁, 97 = 61.83, P < 0.001), and an interaction effect of CStype x Phase (F₁, 97 = 58.72, P < 0.001), Task x Phase (F₁, 97 = 6.82, P = 0.010), Task x CStype (F₁, 97 = 8.13, P = 0.005) and Task x Phase x CStype (F₁, 97 = 3.91, P = 0.051) with no other main effects or interactions (Figure 11.2). Thus, skin conductance responses to the CS+ and CS- differ and change over the course of conditioning and extinction learning.

To gain further insight into interactions revealed by the ANOVA test, we calculated the difference scores in responses to the CS+ and CS- per phase for each task per subject. Paired T-tests revealed greater response differences in the late phase of conditioning compared to the early phase (t(98) = -2.28, P = 0.025). Further, we found a decrease in differential conditioned responses from the late phase of conditioning to the late phase of extinction (t(98) = 4.72, P < 0.001), although we found no further change in differential conditioned responses during the extinction task (t(98) = -0.63, P = 0.522). Finally, during all phases of the experiment

Figure 11.2: Results skin conductance revealed no effect 5-HTTLPR genotype
Both genotypes developed differential conditioned fear responses during conditioning and showed a subsequent reduction in conditioned fear responses with extinction training. No differences between genotype were observed. Left: responses to CS+ and CS- trials are plotted per phase per task for both genotypes. Right: Difference scores between responses to CS+ and CS- trials per phase and task for each genotype are plotted. S’-carriers (solid bars), L’L’-carriers (open bars), CS+ (red), CS- (blue), Early = average over first half of trials of task per CS type, Late = average over second half of trials of task per CS type, error bars reflect s.e.m.
11. The serotonin transporter-linked promoter region polymorphism modulates fear conditioned responses in the human brain

**Figure 11.3: Results task on fMRI**

fMRI revealed the neural network of fear and safety learning. Activation (CS+ vs CS-, red) of the dMPFC, bilateral insula, and midbrain was detected, and a deactivation (CS- vs CS+, blue) of the vMPFC, and visual cortex during both conditioning and extinction. Bars indicate T-values of main effects, activation cluster are displayed overlaid on selective slices of a template brain, and thresholded at p<0.001. Display view follows neurological convention, i.e. right hemisphere is depicted on the right.

Subjects showed greater responses to the CS+ than CS- as revealed by paired T-tests (Early conditioning: (t(98) = 6.22, P < 0.001; Late conditioning: (t(98) = 8.49, P < 0.001; Early extinction: (t(98) = 4.80, P < 0.001; Late extinction: (t(98) = 5.32, P < 0.001). In conclusion, although subjects developed differential conditioned responses over the course of conditioning, and showed a reduction of fear responses with extinction training, we found no differences in skin conductance responses between genotypes.

11.5.3 Genetic variance in 5-HTTLPR is associated with changes in differential conditioned responses in the dorsal medial prefrontal cortex, bilateral insula and ventral visual cortex

We first tested which brain regions were involved in fear and safety memory. Subsequently, we tested how these regions were influenced by genetic variance in 5-HTTLPR using orthogonal contrasts. Specifically, following previous region of interest analyses in the field (10, 524, 525), we extracted data from the regional clusters revealed by comparing CS+ trials to the CS- trials for each task. Note, data extraction was limited.
to those a-prior hypothesized regions that also showed a main effect of CStype (CS+, CS-), which were the dMPFC, bilateral insula, visual cortex, and vMPFC (Table 11.2, Figure 11.3). For each region of interest we tested a Task (Conditioning, Extinction) x Phase (Early, Late) x CStype (CS+, CS-) x Genotype (S'-carriers, L'L'-carriers) 2x2x2x2 repeated measure ANOVA. Only the highest order interaction with Genotype or main effect
of Genotype are provided here, all other significant main effects and interactions can be found in Table 11.3.

The effect of genetic variance in 5-HTTLPR on the dorsal medial prefrontal cortex and bilateral insula followed a similar pattern. Following up the Task x CStype x Genotype interaction in the dMPFC (F\(_{1, 96} = 15.54, P < 0.001\)), right insula (F\(_{1, 96} = 15.54, P < 0.001\)), and left insula (F\(_{1, 96} = 16.10, P < 0.001\)), we calculated for each task the difference in responses to CS+ and CS- trials over both phases for each subject and subsequent one-way ANOVAs revealed greater differential responses in the S'-carriers during conditioning in the dMPFC (F\(_{1, 97} = 12.96, P = 0.001\)), right insula (F\(_{1, 97} = 19.60, P < 0.001\)), and left insula (F\(_{1, 97} = 15.74, P < 0.001\)), but not during extinction in the dMPFC (F\(_{1, 97} = 2.49, P = 0.118\)), right insula (F\(_{1, 97} = 0.49, P = 0.486\)), and left insula (F\(_{1, 97} = 1.47, P = 0.228\)). Following up on these findings, we calculated for each task the average in responses to CS+ and CS- trials over both phases for each subject and subsequent independent samples T-tests revealed greater responses to the CS- during conditioning in the L'L'-carriers in the dMPFC (t(97) = -2.68, P = 0.009; Mean S'-carriers: 0.13, s.e.m.: 0.03; L'L'-carriers mean: 0.28, s.e.m.: 0.06), right insula (t(97) = -4.09, P < 0.001; Mean S'-carriers: 0.12, s.e.m.: 0.03; L'L'-carriers mean: 0.34, s.e.m.: 0.04), and left insula (t(97) = -3.59, P = 0.001; Mean S'-carriers: 0.07, s.e.m.: 0.02; L'L'-carriers mean: 0.18, s.e.m.: 0.03). Thus, S'-carriers showed greater differential conditioned fear responses in the dMPFC and bilateral insula, but this appeared due to L'L'-carriers showing greater responses during conditioning to the CS- (Figure 11.4).

Following up the Task x CStype x Genotype interaction in the visual cortex (F\(_{1, 97} = 5.03, P = 0.027\)), we calculated for each task the difference in responses to CS+ and CS- trials over both phases for each subject and subsequent one-way ANOVAs revealed no difference in differential responses during conditioning (F\(_{1, 97} = 1.58, P = 0.21\)), but greater differential responses in the S'-carriers during extinction at trend (F\(_{1, 97} = 3.02, P = 0.085\)). Following up on this finding, we calculated for each task the average in responses to CS+ and CS- trials over both phases for each subject and subsequent independent samples T-tests revealed greater
responses to the CS+ during extinction in the L’L’-carriers (t(97) = -3.79, P < 0.001; Mean S’-carriers: 0.08, s.e.m.: 0.03; L’L’-carriers mean: 0.28, s.e.m.: 0.05), and increased responses to the CS- at trend (t(97) = -1.76, P < 0.001; Mean S’-carriers: 0.26, s.e.m.: 0.03; L’L’-carriers mean: 0.36, s.e.m.: 0.05). Finally, following up the main effect of Genotype (F1, 97 = 8.45, P = 0.005) we calculated the average response across conditions and a independent samples t-test revealed increased responses in the L’L’-carriers (t(97) = -2.91, P = 0.005; Mean S’-carriers: 0.24, s.e.m.: 0.02; L’L’-carriers mean: 0.36, s.e.m.: 0.04). In conclusion, S’-carriers showed greater differential conditioned responses in the visual cortex during extinction, but this effect appears due to greater responses to the CS+ during extinction in L’L’-carriers. Additionally, the L’L’-carriers displayed
greater responses across the experiment (Figure 11.4). For the vMPFC we found no interaction with genotype.

11.5.4 Bold fMRI - SCR correlation results

We found clear differences in neural functioning of the dMPFC, bilateral insula, and the visual cortex as determined by 5-HTTLPR genotype. The dMPFC and sensory cortex have been implicated as storage sites for fear memory (518, 535-538, 584), where the former has also been associated with fear expression (535-538). In contrast the insula has been suggested to play a role in the interoceptive awareness of visceral responses that lead to subjective feeling states. Despite an absence of genotype on behaviour we were interested in examining the relationship between neural functioning and behaviour and possible differences between genotypes. Therefore, we determined whether changes in brain function could be associated with changes in behaviour, and might have differential relations between genotypes. We correlated the difference in SCR scores to CS+ and CS- trials with the difference in BOLD fMRI
responses to CS+ and CS- trials from the regions of interest that showed effects of genotype per phase (Figure 11.5).

For the dMPFC we found a correlation between SCR responses and neural responses during the early phase of conditioning ($r_s = 0.30, P = 0.002$), and during the late phase of conditioning ($r_s = 0.30, P = 0.003$), but no correlations during extinction. Next, for the right insula we report a correlation between SCR responses and neural responses during the early phase of conditioning ($r_s = 0.33, P = 0.001$), and during the late phase of conditioning ($r_s = 0.39, P < 0.001$), during the early phase of extinction at trend ($r_s = 0.17, P = 0.099$), and during the late phase of extinction ($r_s = 0.26, P = 0.010$). Additionally, for the left insula we found a correlation between SCR responses and neural responses during the early phase of conditioning ($r_s = 0.26, P = 0.009$), and during the late phase of conditioning ($r_s = 0.21, P = 0.039$), but only during the late phase of extinction ($r_s = 0.26, P = 0.010$). Interestingly, for the visual cortex we observed no correlations between SCR responses and neural responses. Thus, neural responses in the dMPFC and bilateral insula that were determined by 5-HTTLPR genotype showed a positive relationship with SCR, whereas neural responses in the visual cortex did not. Moreover, we observed no differences in the BOLD fMRI and SCR relationship between genotypes (using Fisher’s z transformation).

11.5.5 Relationship between task-directed coping and neural responses in the visual cortex is determined by 5-HTTLPR genotype

The visual cortex showed reduced activity in response to CS+ trials compared to CS- trials. The differential responses in the visual cortex did not, however, show a relationship with fear reflected in SCR. One hypothesis could be that subjects diverted attention away from the CS+ resulting in decreased activation in the visual cortex. Therefore, we correlated the differential visual cortex activity with the Coping Inventory for Stressful Situation scores and found that higher task-directed coping scores were correlated with less differences in visual cortical activations.
during the Late Phase of conditioning ($r_s = -0.29$, $P = 0.003$). Interestingly, the S'-carriers showed a greater negative correlation ($r_s = -0.42$, $P < 0.001$) than the L’L'-carriers ($r_s = -0.04$, $P = 0.852$) as revealed by Fisher’s z transformation at trend ($Z = -1.8$, $P = 0.075$). Interestingly, we found no group differences in terms of CISS scores, but task-directed coping scores did correlate with differential SCR responses during the late phase of conditioning at trend ($r_s = -0.17$, $P = 0.078$). Thus, higher task-directed coping scores were associated with reduced differential neural responses in the visual cortex and this correlation was greater in ‘S-carriers. Furthermore, higher task-directed coping scores were associated with smaller differential SCR.

### 11.6 Discussion

In this study we show that a genetic variation in the serotonin transported-linked polymorphic region (5-HTTLPR) that has been associated with increased vulnerability to anxiety and depressive disorders (558-562) modulates functioning of the neural network that supports fear and safety learning.

Although we found that both S'-carriers and L’L'-carriers acquired conditioned fear and extinguish fear as measured by skin conductance responses, we found no differences between genotypes. This finding is in line with previous reports (569), but conflicts with others (572, 573). The skin conductance response is a highly variable measure of sympathetic fear responses, yet fear is a multidimensional emotion of which not all dimensions may be reflected in SCR analyses. Therefore, we further analysed possible genotype effects on neural activity despite an absence of genotype effects on behaviour. We found clear differences in neural functioning of the dMPFC, bilateral insula, thalamus, supramarginal gyrus, and the visual cortex as determined by 5-HTTLPR genotype. For the dMPFC and bilateral insula we observed greater differential responses during conditioning which appeared attributable to greater responses to the CS- during conditioning in the L’L'-carriers, in line with previous findings (568, 572, 573). The dMPFC is considered to be the
human homologue of the rodent prelimbic cortex (303, 304), which has been suggested to be a memory storage site that drives the sustained expression of conditioned fear responses (535-538). As differential conditioning was largely driven by a reduction in responses to the CS-over time, it appears that S'-carriers more rapidly acquired differential conditioned fear responses in an area critical for the expression and retention of fear. The insula has been implicated in the anticipation of aversive stimuli (17), with the introceptive awareness of visceral responses (18, 19), and subjective feeling states (20). Greater differential conditioned responses in the S'-carriers may thus imply a greater degree of subjective experienced fear. In line with a possible role for the dMPFC in the expression of fear and the insula in the interoceptive awareness of visceral responses that lead to subjective feeling states we found correlations between neural responses in these regions and skin conductance responses.

A different pattern was observed for the visual cortex where greater differential responses during extinction in the S'-carriers were observed which appeared due to greater responses to the CS+ in the L’L’-carriers during extinction. Furthermore, the L’-L’-carries showed a general increase in visual cortical responses across the tasks. Sensory cortices have been implicated as long-term storage sites for fear memory (518, 585), where sparse coding of memory representation has been observed (575) that could explain the relative deactivation in BOLD signal. Moreover, heightening serotonin level by pharmacological treatment has been found to increase plasticity in sensory cortices during extinction learning (574). Again, this suggests that genetic variance in 5-HTTLPR modulated the formation of long-term fear memories. Interestingly enough we found no relation between neural visual cortical responses and skin conductance responses, suggesting no direct involvement of the visual cortex in fear expression, at least not at short time intervals. However, we did find a relation between task-directed coping strategies and visual cortical responses, suggesting that differences in coping strategies and 5-HTTLP genetic variance could contribute to individual differences in fear and safety learning.
Throughout several regions of the neural network that supports fear and safety learning we found faster acquisition of conditioned fear responses in the S’-carriers that only arises in L’-L’ carriers later on. This finding is in line with the suggestion that 5-HTTLPR genotype determines the rate of adaptive behaviour where S’-carriers are thought to adapt more rapidly to a changing environment and hence exhibit faster learning (586). An impaired adaptation to a safe environment during extinction learning has been observed in posttraumatic stress disorder (21), and in 5HTT knock-out mice (587). This suggests that there may be interindividual differences determined by 5-HTTLPR variation in the efficacy of psychotherapy treatments that are based on extinction learning, such as exposure therapy. Interestingly, recent work on a different polymorphism in the serotonin transporter showed that a specific variant of this polymorphism is associated with impaired extinction retention, but critically that this impairment could be reversed by pharmacological treatment aimed at increasing serotonergic transmission (588). Again indicating that variants in genotypes related to serotonergic signalling may determine efficacy of specific treatments.

Our results indicate that 5-HTTLPR modulates emotional learning providing a possible mechanism for 5-HTTLPR contribution to susceptibility to affective disorders. However, a limitation of our study is that we only included men, whereas a recent study showed that 5-HTTLPR genotype and sex may interact (589). Furthermore, several studies have indicated an interaction between 5-HTTLPR and (early) life stress (570, 572, 573, 589), an issue not investigated here. Nevertheless, our findings contribute to the development of a mechanistic understanding of the neural processes that underlie individual susceptibility to anxiety and mood disorders, and the development of optimized treatments for specific individuals.

**Competing interests**

*The authors declare no conflict of interest.*
Author contributions

MCW Kroes, GA Wingen and G Fernández designed the study. MCW Kroes analyzed SCR data, MCW Kroes and F Klumpers analyzed fMRI data, B Franke performed genotyping and provided genetic expertise. MCW Kroes and G Fernández wrote the manuscript. All authors agreed on the final manuscript.
12 GENERAL DISCUSSION.
12.1 General conclusion

The intrusive memories in the form of flashbacks that patients like Ben experience are highly disturbing and disabling. Despite intensive research optimal treatments for affective disorders are lacking. This thesis has tried to contribute to a greater fundamental understanding of the neural mechanisms that underlie emotional memory and contribute to the aetiology and persistence of affective disorders, PTSD in particular, and to investigate novel methods to alter memories for emotional experiences that hopefully open up novel avenues for the development of optimized treatment.

To end this thesis a summary of the findings of the reported studies will be provided. Next, an extension of the classical view of memory formation and the dominant model of emotional memory will be provided. Hereafter, limitations and peculiarities of the work in this thesis will be discussed, open questions for future research will be presented, and finally the story of this thesis will come to an end.

12.2 Summary of findings

The aim of this thesis was to provide insight into the neural mechanisms that underlie emotional memory and contribute to the aetiology and persistence of affective disorders, presents evidence for the ability to alter memory for emotional experiences and to contribute to the development of novel strategies to optimize treatment of affective disorders. Here a summary of the findings of the studies in presented in this thesis is provided that will come to form the bases of the extension of the classical view of memory formation, and the dominant model of emotional memory.

The first chapters (Chapter 2-4) of this dissertation describe structural and functional brain abnormalities in PTSD and depression patients in regions implicated in memory and emotion regulation, specifically related to re-experiencing traumatic events from memory during so-
called flashback episodes. Next, we went on to show that food targeted to increase central serotonergic signalling lifts mood and affects mood-regulating neurocircuits (Chapter 5). Although the alleviation of psychiatric symptoms is beneficial, it does leave patients susceptible to relapse, as it does not alter emotional memories that contribute to the aetiology and persistence of affective disorder. It is therefore of considerable interest to permanently alter or even erasing memory for traumatic experiences, but this requires an understanding of the mechanisms that allow alteration of memory. A description of the dynamics of the neural system that enables adaptive and flexible memories (Chapter 6) provides insights into the mechanisms that contribute to flexible memory, including systems consolidation, sleep, schema consolidation, and reconsolidation, and provide hints at opportunities to alter memory. We first focussed our studies on the reconsolidation hypothesis, which postulates that upon reactivation, memory traces return to a labile state and are once again susceptible to disruption and require restabilization to be maintained. Evidence shows that (emotional) episodic memories can indeed be permanently altered upon reactivation in humans in line with a reconsolidation mechanism (Chapter 7 & 8). Furthermore treatment of affective disorders may be optimized by the combination of psychotherapy and pharmacology. The use of noradrenergic antagonists was found to abolish episodic memory enhancement for emotional experiences, and this memory decrement was found to persist one day later in the absence of drug (Chapter 9). However, single reactivation of emotional memory is hard to control in clinical settings and most psychotherapy methods incorporate methods in which patients are repeatedly required to retrieve memories of emotional experiences and actively learn to reduce fear responses and rely on extinction learning mechanisms. Noradrenergic antagonists may however impair consolidation of extinction. Yet, administration of a single dose of a noradrenergic antagonist prior to extinction eliminated differential fear responses during extinction learning, caused the loss of fear and prevented the return of fear one day later in the absence of drug (Chapter 10). This phenomenon was mirrored by specific alterations in the neural network of fear and safety learning. Finally optimization of treatment
may be obtained by increased understanding of vulnerability factors such as genetic variations on brain functioning as this holds the promise to tailor psychiatric treatment to individual needs. Genetic variation in the serotonin transported-linked polymorphic region (5-HTTLPR) that has been associated with anxiety and depressive disorders, was found to determine functioning of the neural network of fear and safety learning, and contributes to an understanding of the neural mechanisms underlying increase vulnerability to anxiety disorders (Chapter 11).

In summary, the results of the studies presented in this thesis show extended brain abnormalities in patients with affective disorders in the neural network that supports emotional memory. A theoretical overview of the dynamic neural systems that support adaptive and flexible memories highlights possible mechanisms that would allow the alteration of memory for such emotional experiences. Subsequently, evidence for the occurrence of reconsolidation of episodic memories in humans is provided. In addition, the results of the subsequent studies highlight that interventions with the noradrenergic system can cause a sustained loss of memory for emotional experiences and differentially modulates the dynamic neural network that support emotional memory. Finally, our results reveal a critical role for serotonin in the neural network of emotional memory. These results considerably advance the field, and in the next paragraph a discussion is provided that highlight how these results contribute to the extension of the classical view on memory formation, and lead to an innovative neural model of emotional memory.

12.3 An extended view of memory formation and neural model of emotional memory

12.3.1 Extending the classical view on memory formation

The classical view on memory formation states that information is initially encoded and labile but stabilizes over time during a process known as consolidation. Subsequently, memories can be retrieved but are considered to remain stable and stored, insensitive to further alterations.
The work presented in this thesis contributes to an extension of this view on memory formation by showing that memories are flexible and adaptive (Figure 12.1A). Upon retrieval the reactivation of a memory trace may render it instable and require restabilization in accordance with the reconsolidation hypothesis ((50, 400, 413); Chapters 7 and 8). In addition, upon retrieval a novel memory may be formed that includes a combination of the retrieved memory and the incoming sensory information at the time of retrieval, known as secondary encoding ((319); Chapters 9 and 10). Furthermore, at the time of retrieval a novel memory trace may be formed due to new learning that can theoretically be fully independent of the retrieved information. Extinction may be such an instant where a novel independent memory trace is created that may come to compete with an earlier fear memory during retrieval at a later point in time ((541, 590); Chapters 10 and 11). The relative independence of an extinction or safety memory from a fear memory is evidenced by the finding that a safety memory trace may also be created prior to the formation of a fear memory, such as in latent inhibition (591, 592). We thus contribute to an extension of the classical view on memory formation by including the processes of reconsolidation, secondary encoding, and new learning that may occur upon memory retrieval. Next, an innovative neural model of emotional memory for the extended view on memory formation will be provided.

12.3.2 Extending the neural model of emotional memory

12.3.2.1 The dominant neural model: encoding and consolidation

The dominant model on emotional memory considers the amygdala to be a critical site of plasticity for associative fear learning at the time of encoding (Figure 12.1B). In addition, the amygdala and noradrenaline release form the locus coeruleus increase processing in sensory regions (34-36) and the hippocampus is important for the formation of contextual and episodic memory. Furthermore, the amygdala is considered to strengthen contextual and episodic memory by upregulating neural excitability and plasticity in the hippocampus through a beta-adrenergic
mechanism during consolidation (Figure 12.1C). Beyond this dominant model the neural mechanisms that contribute to additional memory processes of emotional memory are being uncovered.

Here a model neural network that supports emotional memory retrieval, reconsolidation, secondary encoding, and extinction will be presented.

12.3.2.2 Retrieval

First the neural model on emotional memory will be extended to include retrieval (Figure 12.1D). The amygdala has been found to be important to the retrieval of emotional memory (487, 489, 492, 502, 504-506, 546) and amygdala-hippocampal connectivity has been found to increase during retrieval (488, 503). In addition, the dorsal medial prefrontal cortex (dmPFC) is considered to be important to the retrieval and expression of emotional memory (535, 537, 593, 594). Furthermore, noradrenergic modulation originating from the locus coeruleus supports memory retrieval (484, 492). The work presented in this thesis contributes to our understanding of retrieval of emotional memories by showing that the dmPFC contributes to the retrieval of emotional memory and is subject to beta-adrenergic modulation most likely originating from the locus coeruleus (Chapters 9 and 10). The dmPFC is an optimal candidate to support a retrieval function. Amygdala firing rate responses to conditioned stimuli are short lived (~50 -500 msec) (595) and may initiate emotional responses but do not sustain the response. The dmPFC receives inputs and projects back to the amygdala, the hippocampus and midbrain regions (596-602), and shows sustained firing rate responses throughout
the entire presentation time of conditioned stimuli which is associated with emotional responses (594). The dmPFC thus appears to play a pivotal role in the retrieval and expression of emotional memory. Having identified regions involved in retrieval of emotional memories next we will turn to the neural network involved in reconsolidation.

12.3.2.3 Reconsolidation

Evidence for reconsolidation was first obtained by blocking protein synthesis upon memory reactivation in the amygdala (50, 400), indicating a role for the amygdala in reconsolidation (Figure 12.1E). Subsequently, the hippocampus was found to be involved in the reconsolidation of context-dependent memories (401, 603-607). In addition a role for the dmPFC in reconsolidation has been detected (608-610). Furthermore, blocking noradrenaline in the amygdala or hippocampus can disrupt memories following reactivation (401, 515, 542, 606, 609-611). This suggests that reconsolidation may occur in all these regions and is modulated by noradrenaline. Whether a brain region is involved in the reconsolidation of a particular memory is determined by the dependency of the memory on that region. In addition, differences between regions, and in the dependency of memories on interregional connectivity may determine sensitivity to reconsolidation disturbances (Chapter 6). The work presented in this thesis has contributed to advancement in the understanding of reconsolidation, and has suggested an interaction between the amygdala, hippocampus and noradrenergic system during reconsolidation (Chapter 6, 7, 8, and 9). Next, the neural network that supports secondary encoding of emotional memories will be discussed.

12.3.2.4 Secondary encoding

Secondary encoding is the process whereby during every retrieval episode, retrieved mnemonic information, together with incoming sensory information, is stored as a novel memory trace (319). This results in the accumulation of multiple, related, mnemonic traces of a single episode within a slightly altered neuronal and experiential context, which together
aid the retrieval of an episode (319). In Chapter 9 it has been suggested that blocking enhanced recall of emotional stimuli with propranolol may prevent formation of a novel emotional memory trace, expressed as absent emotional memory enhancement on the next day. Considering that the task is known to involve the amygdala and hippocampus at encoding, this suggests an interaction between these regions and noradrenaline during secondary encoding. Indeed, a subsequent imaging study using a similar design showed involvement of the amygdala and hippocampus to be modulated by noradrenaline at the time of secondary encoding (471). The idea that during secondary encoding the retrieved mnemonic information becomes stored together with the experiential context at the time of retrieval predicts the involvement of the hippocampus. Evidence for this is provided in Chapter 10, where safety learning in a novel context is accompanied by hippocampal responses to the context presentation itself during retrieval, and subsequently this hippocampal involvement is detected during presentation of conditioned stimuli one day later, which correlates with the absence of fear. Noradrenaline was found to modulate the hippocampal involvement, and eventually lead to altered responses in the dmPFC (Chapter 10). Alternatively, the hippocampal involvement may reflect new learning (see below). In summary, this suggests that the amygdala, hippocampus, and dmPFC are involved in secondary encoding and under modulation of noradrenaline (Figure 12.1F). Finally, the neural network that underlies extinction will be discussed.

12.3.2.5 New learning: Extinction

At the time of retrieval a novel memory trace may be formed due to new learning that is independent of the retrieved memory trace. Extinction may be such an instance where a novel safety memory is created that is stored independent of a fear memory but may compete with it for expression. Here the neural network that supports extinction memory is described (Figure 12.1G).

The amygdala projects to the hippocampus, ventral- and dorsal mPFC, and locus coeruleus and in turn receives afferent inputs from the same
regions (598, 612-616). The hippocampus projects to the amygdala, vmPFC and dmPFC where the projections to the dmPFC appear to be inhibitory (334, 613, 616-618). The dmPFC and vmPFC are reciprocally connected, but the projections from the vmPFC to the dmPFC are likely inhibitory (619). Further, the dmPFC and vmPFC both project to the hypothalamus and monoaminergic nuclei in the brainstem (34, 614, 620-623), these latter connections are significant in that the vmPFC and dmPFC are the only neocortical regions in a position to exert control over neuromodulatory systems in the brainstem. In addition the locus coeruleus projects to widespread regions of the neocortex, the hippocampus and amygdala (624). The connectivity of this neural network supports fear and safety learning.

Simple associative fear memories as formed during conditioning critically rely on plasticity within the amygdala (5, 625-628). Contextual information on the relative danger or safety of a stimulus in a given environment is co-provided by the hippocampus (12, 518, 541, 629). As discussed above, the dmPFC is thought to support the retrieval and sustained expression of fear memories (594). Upon extinction learning the vmPFC is considered to come to inhibit the output of the amygdala (12, 518, 602, 629, 630) by projecting to the GABA-ergic intercalated cells of the amygdala (631). The work presented in this thesis directly shows the involvement of the vmPFC, dmPFC, and hippocampus in extinction (Chapter 10, 11). Indeed the vmPFC appears to function as a safety signal inhibiting fear, the dmPFC to relate to the expression of fear, and the hippocampus to provide a contextual safety signal. The novelty of the presented work highlights a role for noradrenaline in modulating dmPFC and hippocampus involvement in fear and safety learning (Chapter 10). In addition the work in this thesis extends understanding of the neural network of fear and safety learning by determining a modulatory influence of serotonin on dmPFC, insula, and sensory cortical functioning (latter two not depicted in Figure 12.1G for simplicity). These findings add to earlier work showing modulation of the dmPFC and vmPFC by noradrenaline (514, 519, 542), and effects of serotonin on vmPFC-amygdala connectivity related to fear and safety learning (570).
To conclude, an extensive neural network supports fear and safety learning as revealed by studies on extinction, and this network is critically modulated by noradrenaline and serotonin. Yet, the neural network underlying emotional memory is more complex than so far described.

12.3.2.6 Conclusion on the extended neural model of emotional memory

Here an extension of the classical view on memory formation and an innovative neural model of emotional memory have been provided. In doing so hopes are to increase understanding of the neural mechanisms that underlie emotional memory. Yet some notes are required.

In reality, the processes that support memory are not easily captured in boxes and psychological terminology such as in Figure 12.1A. Memory arises from the dynamic and complex system that is the brain (Chapter 6) and mnemonic processes will occur iteratively in high temporal resolution, and distinct process may even occur simultaneously in different regions of the brain. A more full depiction of the so far described neural network underlying emotional memory is provided in Figure 12.1H. Different memory functions will arise through dynamic interactions between these regions and neuromodulators may affect the state of the network and govern emotional memory. A model such as Figure 12.1H is necessarily a simplification and in this case excludes regions important for emotional memory that have been outside of the scope of this discussion. It will be a significant challenge for future research to come to a more complete description and understanding of the full dynamics of the neural network that supports emotional memory.

In conclusion, regardless of possible limitations, the models described in Figure 12.1 could aid in understanding the relative contributions of different regions within the neural network of emotional memory and provide insight into the neural mechanisms that underlie affective disorders.
12.4 Overlap brain regions

Investigating overlap between the neural network underlying emotional memory and the brain abnormalities in affective disorders could provide insight into the mechanism that contributes to the aetiology and persistence of these disorders.

Striking overlap exists in brain structures (Figure 12.2) that display structural abnormalities in PTSD and depression (Chapter 2), where neural responses are affected by propranolol administration prior to extinction learning (Chapter 10), affected by serotonin enhancement during instructed fear (Chapter 5), and determined by 5-HTTLPR genotype during conditioning-extinction (Chapter 11). The overlap pertains particularly to the dmPFC and the bilateral insula. Both regions have been reported to display altered functioning in patients with affective disorders that can restore their functionality following effective treatment (3). The dmPFC has been associated with fear memory storage and expression (535-538), with integration of cognitive information and motor output (205), and with cognitive control over emotional responses (14). Similarly, the insula has been implicated in pain perception (16), the anticipation of aversive stimuli (17), with the introceptive awareness of visceral responses (18, 19), and subjective feeling states (20). Despite initial investigations into the mechanistic processes supported by the dmPFC (537, 538), to date little is known about the functional role of this region in fear and safety learning, and its contribution to affective disorders. Even less is known about the mechanistic role of the insula. Increasing understanding of the contributions of these regions to fear and safety learning is critical and could contribute to understanding individual differences in symptoms and sensitivity to specific treatments.

To conclude, investigating the overlap between the neural network of emotional memory (Figure 12.1) and the brain abnormalities observed in affective disorders highlight contributions of the dmPFC and insula in the aetiology and persistence of affective disorders. Increasing understanding
of the contribution of these regions to emotional memory and affective disorders will be of considerable interest for future studies.

12.5 Integrating neuromodulatory function.

In this thesis the functional role of several neuromodulators has been studied, namely noradrenaline (Chapter 9, 10), serotonin (Chapter 5, 11), and to some extend dopamine (Chapter 5). Unlike direct synaptic transmission (through the neurotransmitters glutamate, y-aminobutyric acid, and glycine) where neurotransmitter release from one presynaptic site targets one postsynaptic site, neuromodulators are released by a small group of cells and diffuse through larger areas affecting multiple neurons, and ‘modulate’ the direct synaptic transmission. Here theories on the functional role of noradrenaline, and a theory on the functional interactions between serotonin and dopamine will briefly be presented, and a proposal of integration between these neuromodulators will be put forward.

Classically noradrenaline has been associated with arousal and been proposed to set the gain of activation in the neural network (632-634). However, these theories do not account for the observation that intermediate levels of noradrenaline are optimal for behaviour,
according to the Yerkes-Dodson law, and do not reflect stimulus-specific phasic responses of locus coeruleus neurons (34, 635). A prominent theory of the function of noradrenaline suggests that phasic and tonic responses by locus coeruleus neurons serve different, yet interacting, functions largely based on studies on stimulus detection (34). Phasic locus coeruleus responses are driven by outcome of task-related decision processes and facilitate ensuing behaviour. The idea here is that as evidence for decision options are collected in a layer of the neural network, and activity within a layer of the network crosses a critical threshold, locus coeruleus responses are triggered that drive neural processing in downstream layers towards the direction of the decision (34). Thus, phasic locus coeruleus responses may upregulate processing of relevant stimuli. In concordance with this suggestion we observe that processing of salient stimuli and memory formation are affected by propranolol which would have its greatest effect on postsynaptic phasic responses (Chapter 9 and 10) Alternatively, during periods of high tonic state of locus coeruleus activity a reduction in the ability to discriminate stimuli is observed reflecting a disengagement from the task and a search for alternative behavioural options (34). Together this suggests that tonic locus coeruleus activity serves as a gain function where phasic responses can facilitate task related performance, but increased tonic activation causes the system to become hyper-responsive to any stimulus initiating switching behaviour (636). Thus, interactions between the tonic and phasic responses of locus coeruleus neurons may determine whether the neural system exploits optimal processing of stimuli of high utility, or explores alternative options when utility of task-relevant stimuli are low (34, 624, 637, 638) Indeed locus coeruleus responses are generally observed prior to behavioural switches, and track changes in contingencies prior to behavioural adaptation (624). Building on the proposal that noradrenaline plays a role in optimizing neural decision processing according to stimulus utility, noradrenergic locus coeruleus activity may drive gross network reorganization and promote rapid network plasticity (175, 624), possibly driving the formation of novel neural ensembles associated with behavioural adaptation. In concordance, we observe noradrenergic dependent
changes in multiple regions that are part of the neural network of fear and safety learning (Chapter 10). Interestingly, locus coeruleus neurons do not appear to reflect stimulus valance or respond to the relative outcome of situations, and hence, the relative evaluation of the stimulus utility changes over time are likely determined by different systems. One possibility than is that locus coeruleus responses are governed by prefrontal cortical inputs where reward and cost outcomes of on-going behaviour are evaluated (34, 624), that may depend on an integration of serotonergic and dopaminergic signals which will be discussed next.

Interestingly, a recent theory integrating serotonin and dopamine function also emphasizes different contributions of the relative tonic and phasic responses of these neuromodulators (191). Classically dopamine has been associated with positive affect (reward) and behavioural activation (vigour), whereas serotonin has been associated with negative affect (aversion) and behavioural inhibition (inactivation). Cools and co-workers (2011) extent upon an earlier proposal on the role of dopamine that posits that the control of behavioural activation is a problem of trading off the cost (energetic) and benefits (faster reward gathering) over time (639). Thus, when rewards are frequently available, behaviour is sped up, but if rewards are infrequent behaviour is slowed down as the energetic cost to obtain the rewards becomes too high (639). Here, the tonic level of dopamine sets the average rate of reward against which phasic dopamine responses originating from the ventral tegmental area reflect the relative reward value (639). Cools and co-workers (2011) extent this by suggesting a similar function for serotonin responses of the dorsal raphe nucleus in signalling relative aversion and an interaction between dopamine and serotonin. Thus, when aversive outcomes are frequently available, behaviour is inhibited, but if aversive outcomes are infrequent behaviour is sped up (191). In a complex environment where rewarding and aversive outcomes will co-occur, phasic reward and aversive outcomes will be relative to a running weighted average of the tonic dopaminergic and serotonergic signals (191). Indeed, raising tonic serotonin levels modulates neural activity in a task that is known to activate the dopamine system, and neural functioning changes in a brain
region where an integration between serotonergic and dopaminergic projections occur (Chapter 5). Further, the medial frontal cortex is optimally situated to evaluate this integrative signal as it receives projections from both the ventral tegmental area and the dorsal raphe nucleus, and in turn projects back to both regions and the locus coeruleus (34, 640-642). The medial prefrontal cortex may thus track the relative value of a stimulus in a specific context based on dopaminergic and serotonergic signals and subsequently drive locus coeruleus responses.

When attempting to integrate the roles of noradrenaline, dopamine, and serotonin it is important to realize that noradrenergic responses are associated with stimulus onset and not the outcome of a situation (i.e. valance unspecific), whereas dopamine and serotonin responses are thought to reflect outcomes. Thus, during initial phases of a task dopamine and serotonin may signal the relative outcome or value of a stimulus and an integrated evaluation of the situation within the prefrontal cortex drives tonic locus coeruleus activity to an intermediate level so that subsequent phasic locus coeruleus responses enhance processing of salient stimuli. If the utility of a stimulus unexpectedly changes and becomes inadequate locus coeruleus activity is shifted to a higher tonic level, favouring enhanced processing of any incoming stimulus though gross network reorganization resulting in switching behaviour. This occurs because the gain amplification of stimuli with a previously sub-threshold value for selection in the decision process now cross the threshold. This could explain the finding in Chapter 10 that subjects during extinction learning do not show a general reduction of fear responses, but an inability to differentiate responses when given a noradrenergic antagonist. It is conceivable that dopamine and serotonin are able to provide a relative affective value signal, but the gain is too low due to the noradrenergic blockade that expression of this affective signal is hindered. Of interest here is also the functional role of the neuromodulator acetylcholine (not under direct investigation in this thesis). Both noradrenaline and acetylcholine neurons increase responses following contingency changes, but acetylcholine responses stay increased over a longer period, whereas noradrenaline responses
habituate quickly, which has led to the suggestion that acetylcholine is important for expected uncertainty (637, 638). Thus, where noradrenaline is important for unexpected uncertainty, increases bottom up processing, and triggers network reorganization; acetylcholine might be important for expected uncertainty, may maintain top-down processing, and considering its generally inhibitory function in the central nervous system might achieve so by preventing network reorganization.

In conclusion, neuromodulators influence neural processing indirectly and appear critical in the reorganization of network plasticity through effects on neural plasticity (624). The precise effects of neuromodulators likely depend on a critical interaction between the tonic level of activity and relative phasic responses. In addition, whereas dopamine and serotonin appear to provide a relative valance signal, allowing an integrated evaluation of the value of stimuli given a overall situation, noradrenaline and acetylcholine may be involved in modulating processing of stimuli themselves, and dependent on unexpected or expected changes in contingency trigger or inhibit a reorganization of neural network state. This view on the roles of different neuromodulators has implications for the interpretation of the studies presented in this thesis. Specifically, the realization that tonic and phasic activity of neuromodulators serve different functional roles has important implication for the interpretation of the pharmacological intervention studies in this thesis. The use of food targeting the serotonergic system (Chapter 5) likely increased central tonic serotonin levels. Thus the effects reported should be interpreted as effects of an increased running average of cost against which phasic (task evoked) signals are weighted. In contrast the studies manipulating noradrenaline (Chapter 9, 10) employed the β1β2-adrenoceptor antagonists propranolol. Considering that beta-adrenergic receptors are post-synaptic receptors this manipulation likely affected phasic noradrenergic responses and not the tonic state which is under control of alpha-adrenergic receptors, effects in these studies probably reflect the gain of stimulus processing. This interpretation of phasic versus tonic neuromodulatory signals is interesting but has not been directly investigated in the studies presented in this thesis limiting the
conclusions that can be drawn from the results. Next several limitations of the studies presented in this thesis will be discussed.

12.6 Limitations and outstanding questions

The studies presented in this thesis often raise questions that remain unanswered, have had unexpected results, and result in limitations of the conclusion that can be drawn. In addition, the results often call for further investigation in the future, and highlight topics that have so far been unexplored. Here several of these issues and future interests will be discussed.

12.6.1 Finding the amygdala isn’t peanuts.

Considering the prominent role of the amygdala in emotional memory (see also Figure 12.1) it is striking that we do not find amygdala involvement as revealed by BOLD-fMRI in the majority of the studies in this thesis. Both methodological considerations and the intrinsic neural function and architecture of the amygdala could provide an explanation for this discrepancy.

Animal studies have clearly revealed a critical role for the amygdala in emotional memory and particularly in the acquisition of learned fear (5, 9). Although several seminal papers have also revealed amygdala involvement in learned fear in humans using fMRI (6, 10, 643, 644), the detection of amygdala involvement in fear conditioning using fMRI in humans is far from ubiquitous (645). One explanation for an absence of amygdala detection is a dropout or distortion of the fMRI signal in this region of the brain (156). However, in the studies presented in this thesis signal was in fact obtained from the amygdala and specific effort was made to minimize distortion through the use of e.g. temporal lobe optimized sequences, multi-echo imaging, and fieldmaps. In addition, clear task-related amygdala activation was found during an emotional processing task (Chapter 5) indicating that the employed imaging methods are principally able to detect amygdala involvement in cognitive tasks. A
second explanation pertains to the observation that amygdala responses to salient stimuli in terms of increased firing rates are short lived (595). In contrast to the brief increases in firing rates of amygdala neurons, the BOLD-fMRI signal is slow and reflects an integrated average over longer periods of time. In addition, the use of box-car modulated regressors including the entire duration of stimulus presentation (Chapter 10 and 11), would limit the detection of short-lived neural responses and favour responses of longer duration. However, specifying regressors as stick-functions also did not result in the detection of task-related amygdala involvement (results not presented). A third explanation may be provided by the functional and architectural organization of the amygdala and the limited spatial resolution of fMRI. The spatial resolution of fMRI is in the millimetre order and the human amygdala is only few cubic millimetres in size. In addition, the amygdala is not a unitary nucleus, but consists of several subnuclei that serve different roles in the expression, acquisition, and extinction of fear (5, 518, 646, 647). Furthermore, the glutamatergic neurons of the amygdala subnuclei are interspersed with gaba-ergic intercalated cells (618, 631). Finally, recent studies indicate that even within the subnuclei of the amygdala a certain class of neurons increase their firing rates with the acquisition of fear, whereas others show a decrease and an increase in firing rates when fear is extinguished (598). Summating a neural signal over multiple subnuclei, over neurons with different neurotransmitters, and over functional subsets of cells could explain the absence of a net difference between experimental conditions in fMRI studies.

In conclusion, although the absence of amygdala effects in the studies presented in this thesis may appear striking this finding is not uncommon, and could be explained by the limitations of the imaging method used in interaction with the intrinsic neural function and architecture of the amygdala.
12.6.2 The vmPFC regulation of fear, and the default mode network.

The exact functional contribution of the vmPFC to fear and safety learning in humans as detected by BOLD-fMRI is obscured by its involvement in the so-called default mode network.

Rodent work has consistently indicated that the infralimbic cortex inhibits the output of the amygdala and in doing so regulates the expression of fear following extinction training (12, 513, 630, 648-651). The vmPFC is considered the human analogue of the rodent infralimbic cortex and in human BOLD-fMRI studies on extinction learning has consistently shown a relative deactivation that is greater in response to the CS+ than CS- (10, 524, 526, 528). In addition, this relative deactivation correlates negatively with the differential physiological fear responses (10). This has lead to the suggestion that the vmPFC response in humans reflects an inhibitory control signal over the amygdala, and idea supported by findings that the vmPFC and amygdala BOLD fMRI signal correlate negatively (10, 577). In Chapter 10 and 11 we replicate these findings. However, the relative deactivation of the vmPFC already becomes apparent during conditioning, and in our study is even greater following conditioning than extinction (Chapter 11). This could be the result of dynamic interactions between the amygdala and the vmPFC, where the vmPFC has a gating function over the amygdala but will reduce its gating signal if the amygdala drive is strong enough. Alternatively, it is of interest to note that the vmPFC is part of the so-called ‘default mode network’ (127, 128, 652-657). The default mode network is a set of brain regions that shows strong correlated activity under resting-state conditions, and activation of this network has been associated with internal processing (658). In contrast, the brain regions found to activate during conditioning and extinction (Chapter 10 and 11) also show correlated activity under resting-state conditions and are generally considered to be part of a ‘salience’ network (175, 658). The deactivation of the the vmPFC observed during conditioning and extinction could thus reflect a deactivation of the default mode network and an activation of the salience network, shifting
attention away from internal monitoring to salient external stimuli. In
support of this argument, we also observe a relative deactivation of the
posterior cingulate cortex/precuneus during conditioning and extinction
(Chapter 10 and 11), another region considered to be part of the default
mode network. Finally, it is of interest to note that the deactivation of
the default mode network has been associated with optimal learning and
memory formation (659, 660) a necessity for efficient conditioning and
extinction.

In conclusion, although human imaging studies have consistently
detected the involvement of the vmPFC in extinction learning and have
associated relative activity in this region with reduced fear responses, the
determination of its causal role in safety learning in humans is obscured
by its role in the default mode network.

12.6.3 Reconsidering reconsolidation.

Reconsolidation potentially allows the permanent alteration of memory
and has been discussed extensively in this thesis. Yet, it is important to
point out several limitations and outstanding questions of reconsolidation
research.

Initial studies on reconsolidation showed that reactivation of memory
traces renders these susceptible to disruption by treatments similar
to those that impair initial consolidation (50). A critical problem to the
demonstration of both consolidation and reconsolidation by disturbance
of memory is that this results in the absence of behaviour and thus
concords with instead of refutes the null-hypothesis. Hence, the absence
of a behavioural indication of memory could be due to many alternative
explanations, e.g. an insensitivity of the measuring method to detect
the expression of memory (400, 437, 661). For this reason an important
criteria that has been generated to demonstrate is reconsolidation is that
the memory impairment should not be due to an temporary inability to
access memory traces (448), hence should not return after longer delays.
The literature on the permanence of memory disturbance following
reactivation has been mixed with some studies finding disturbances for long periods of time (401, 406, 409, 424), whereas others have found the return of memories, specifically when memories are older or stronger (414-417). In addition memory enhancement following reactivation has been found (427, 476), and have led to the suggestion that reconsolidation serves to update memory (407, 444, 456). Together this raises the question whether reconsolidation is a generic process affecting all memories, and/or only parts of certain memories. In Chapter 6 the hypothesis has been postulated during reconsolidation synapses destabilize putting memories that depend on a limited number of synaptic weight changes or connections at risk of disruption, whereas memories that rely on re-wiring changes in cortico-cortical connections would be less sensitive to disturbance (466). Thus, simple associative conditioned fear memories that potentially depend on plasticity between a limited number of synaptic connection in the amygdala might particularly be sensitive to alteration following reactivation (420, 662), whereas episodic memory that is considered to depend on a more extensive neural network would be less sensitive to disturbance. As in Chapter 7 and 8 reconsolidation of episodic memory has been studied this raises several questions.

Memory performance as described in Chapter 7 and 8 is relatively low. This raises the question whether a similar disruption of reactivated memory had occurred if memories had been stronger. In addition, consolidation and reconsolidation impairments have been described as cue-dependent amnesia (46, 661, 663, 664), where the idea is that the absence of memory expression is dependent on the strength of the retrieval-cue. Thus, providing a stronger cue might trigger the expression of memory. In Chapter 7 during the study session subjects were shown all stimuli three times. Further, a stem-completion task was used to probe memory retrieval and subjects were instructed only to respond when certain of the to be recalled word in order to reduce the single-to-noise ratio of the task. Similarly, in Chapter 8 subjects were shown the stimuli only once, and received a multiple-choice recognition test, but were not shown the original stimulus material (except for the first slide). One may
wonder whether disturbances of memory would have been observed if memory encoding had been performed more often resulting in stronger memories, or would have been observed if stronger cueing paradigms had been used. Nevertheless, even if under such conditions memory expression would be evident this would still indicate that memory traces could be significantly weakened following reactivation.

Another concern pertains to the permanence of the disturbance of reactivated memory traces. Especially in the view of possible psychotherapeutical applications targeting the reconsolidation phenomena it is paramount that the alterations of memory are long-lasting. In Chapter 7 the disruption of reactivated memory persists at least for one week. A second interesting observation reported in Chapter 7 is that the correct recall of reactivated memory traces soon after an amnestic manipulation prevents the emergence of impaired reconsolidation. This finding either reflects that retrieval of a destabilized memory prior to the completion of reconsolidation restabilizes the memory trace, or that secondary encoding at the time of retrieval forms a trace that can be retrieved at subsequent testing. It would be of considerable interest to further examine the permanence and possible restabilization by retrieval effects of reactivated memories at risk of disturbance.

An interesting issue is whether the return of disturbed reactivated memories by definition excludes reconsolidation having occurred. Biomolecular studies have convincingly shown that synapses do destabilize upon reactivation, and that the restabilization of synapses can be disturbed resulting in alteration of memory (418-420, 422). It is however conceivable that only part of the connections of the neural ensemble that constitutes a memory trace are destabilized and disturbed, especially in the case of episodic memory that is thought to rely on extensive long-range neural networks. Further, the re-emergence of a functional neural memory ensemble may be initiated based on activity in only a subset of the neurons (270). One can phantom that in the case of a partial memory trace disruption, reiterative reactivation of the partial memory trace may
through synaptic plasticity re-recruit the connections and neurons that originally (or partially by replacement) constituted a neural ensemble that enables memory expression. In this dynamic view on memory the return of a behavioural expression of memory does not exclude the possibility that a reconsolidation mechanisms has occurred. This idea touches on an important concept of what a memory trace actually is and will be one of the most exciting questions to address in neuroscience in the near future.

Two other issues related to the studies on reconsolidation in this thesis are worth mentioning. One criticism on studies of reconsolidation has been that the disturbance of memory might be attributable to a heightened state of arousal at the time of memory reactivation evoked by the reminder rather than a general memory process (665). The disturbances of memory would than result from a state difference at the time of encoding and reactivation and test. The studies reported in Chapter 7 and 8 both contain arousing elements, either the application of ECT, or the occurrence of emotional items during encoding. However, especially for the results in Chapter 7 it is hard to phantom why only a subset of items would suffer from memory impairment whilst all items are reactivated and retrieved in the same state. A related issue concerns the idea that possibly only amygdala dependent memories show the reconsolidation phenomena. Although reconsolidation has predominantly been studies in task that are known to depend on the amygdala (400), disturbances of memories that are generally considered to be amygdala independent (403, 407), and disturbances by manipulations of brain regions outside the amygdala (401, 552, 666) have been observed. However, the amygdala may modulate neural activity in other areas and therefore it remains possible that only an amygdala contribution is disturbed in these studies. It would be of interest to study reconsolidation in the absence of amygdala influences, e.g. following lesions or in patients with Urbach-Wiethe disease.

A final limitation is specific to the results presented in Chapter 8. Electroconvulsive therapy was used to investigate the disturbance of reactivated memories. Two elements of ECT could contribute to the
observed memory impairment: the applied electrical stimulation itself, or the anaesthetics. We find electrode placement to be related to general memory performance suggesting that the electrical stimulation itself may cause the observed disruption of reactivated memory traces. In contrast, we do not observe a relationship between the anaesthetic dose and memory performance. Nevertheless it could be that the memory impairment is attributable to GABA-ergic medication.

Regardless of the issues discussed here, the results in Chapter 7 and 8 satisfy the criteria of the reconsolidation hypothesis, and thus provide evidence for the reconsolidation of episodic emotional memories in humans. In particular, the effects are found to be time-dependent, a critical criteria for a demonstration of reconsolidation that is often neglected. As evidenced in Chapter 9 and 10, the loss of memory following reactivation may also be attributable to new learning and effects at the time of retrieval. Therefore, any study on reconsolidation lacking a time-dependent demonstration should be treated with caution. The studies presented in Chapters 7-10 only differ on details in terms of memory reactivation yet instigate different memory processes. For future studies it is of considerable interest to determine the boundary conditions under which reconsolidation takes place.

Finally, several specific issues raised by studies in this thesis will be discussed.

12.6.4 Beta-blockers and skin conductance responses.

A possible confound in the study presented in Chapter 10 is the combined use of skin conductance responses and propranolol, which will be discussed here.

The study presented in Chapter 10 used skin conductance responses to parameterize fear responses, and employed the beta-blocker propranolol to alter fear memory. This raises the possibility that propranolol may have directly affected the peripheral sweat response obscuring the assessment
of fear memory. Historically there has been much debate about direct peripheral effects of noradrenaline on the skin conductance response (667), however eventually it became clear that myoepithelial cells responsible for the contractions of the secretory segment of sweat glands react to cholinergic stimulation only (667, 668). In addition, although arterial blood pressure and heart responses to stress are reduced by the beta-adrenergic antagonists nadolol, which does not cross the blood brain barrier, and therefore only has peripheral effects, it does not affect skin conductance responses (669). Also, local application of propranolol on the skin does not affect the sweat response during exercise (636). In line with this, in the study presented in Chapter 10 no effect of propranolol on baseline skin conductance responses to context presentations, or habituation trials are detected. In addition, the responses to the conditioned stimuli are far greater than the responses to these control measures, and therefore the obtained results are unlikely to be due a peripheral effect of propranolol on the sweat response.

In conclusion, the effects of propranolol on fear responses as described in Chapter 10 do not appear to be confounded by peripheral effects.

12.7 Understanding dynamic interactions within fear neural network

As summarized in Figure 12.1 and revealed by studies presented in this thesis (Chapter 5, 10, and 11), a large interactive neural network supports memory for emotional experiences. To date studies have large focussed on revealing the functional role of specific regions, however much less is known about the dynamic interactions between regions.

Beyond understanding function of specific regions a need exists to gain insight into the dynamic interactions within the neural network that gives rise to fear and safety learning. Recent methodological advancements have been developed to study dynamic interactions in brain networks based on BOLD-fMRI data. These methods include psychophysiological interaction (PPI) analyses and dynamic causal modelling (DCM) (670-672).
These methods allow one to study network interactions dependent on task performance. However, these methods are necessarily limited by the temporal sluggishness and spatial precision of the BOLD-fMRI signal. In addition, in practice DCM and PPI analyses require strong consistent statistical effects on single-subject level, and work most optimally for a restricted number of regions, and relative simple cognitive tasks. Thus, a true understanding of dynamic neural interactions in the fear and safety network may require invasive experiments.

Only a limited number of studies have invasively investigated dynamic neural interactions and provide a wealth of insight into fear and safety network functioning (175, 503, 518, 570, 630, 673-675). However, most of these studies measure changes in oscillatory coherence between regions within the neural network, but oscillatory coherence does not necessitate effective connectivity in terms of information processing (309). Therefore, a critical question to future research will be how oscillatory coherence between regions affect local spike rates, and drives behavioural expression. In addition such studies could provide critical information on the emergence and dissipation of neural ensembles that form a memory trace and reveal the nature of memory.

A need exists for more systems level studies to truly understand the interactions between brain regions that collectively give rise to emotional memory, and hence allow a more specific understanding of the ability to alter memories. Such studies would undoubtedly contribute to understanding dynamic nature of memory and the foundation of altering memories.

12.8 Offline periods

This thesis has focussed on the ability to alter memory for emotional experiences and has investigated the neural activity at the time of manipulation and at the time of outcome, i.e. retrieval. However, it is eminent that most alterations of memory occur in between these points of assessment.
The ability to alter memory often depended on processes that occur after the moment of learning of reactivation during so-called ‘off-line’ periods, e.g. consolidation, and reconsolidation. Neural activity patterns reflecting learned experiences are replayed during sleep and post-learning wake periods resulting in strengthening or weakening of memory (348, 369, 370). Although it is attractive to assume that similar processes occur following memory reactivation, and underlie reconsolidation effects, this remains not yet investigated. Understanding ‘off-line’ processing of memory following reactivation might provide further insight into the ability to alter memory for emotional experiences.

Future studies aimed at investigating the off-line replay of memory following reactivation would be of considerable interest and could provide fundamental insight into the mechanisms that allow alteration of memory.

12.9 Thoughts on the hippocampus, context, and memory systems.

Throughout this thesis studies have been presented that provide possibilities to alter memory for emotional experiences resulting in a loss of fear. An important outstanding question is under which circumstances fear may return. Of specific interest here is the role of contextual safety memory, its dependence on the hippocampus, and interactions between simple-associative and episodic memory.

Learning that a stimulus is safe is highly dependent on context, and the hippocampus is considered to be critical for contextual safety learning. Over time memories can undergo systems consolidation (248, 316, 318, 321), referring to the process whereby memories that are initially hippocampus dependent become incorporated within the neocortex and less dependent on the hippocampus. With systems consolidation, memories’ contextual information may be lost resulting in a return of fear, but could also lead to a generalization of safety to other contexts, yet little is known about these processes. In Chapter 10 hippocampal activity was found to be associated with less fear following contextual safety learning. For future studies it would be of considerable interest to test
whether fear would return if subjects were retested in the original context in which fear was acquired, or if the inhibition of fear also generalizes to a new context. Such studies would have important clinical implications, as it might indicate that optimal psychotherapy would require safety training in either the original fearful context, or require safety training over multiple contexts to obtain generalization of fear inhibition.

Most animal studies on memory alteration following reactivation have investigated simple associative learning that may rely on plasticity of several synapses within the amygdala. However, whereas the expression of fear can be disrupted following reactivation (Chapter 8 and 9), and emotional enhancement of memory may be abolished at reactivation (Chapter 9), memory is not fully lost, and moreover explicit episodic knowledge of a fearful event still remains (Chapter 10). In line with idea of the existence of multiple memory systems (229, 230, 233), episodic memory (memory for specific experiences) has been proposed to depend on the hippocampus, emotional expression of fear on the amygdala, and the emotional enhancement of episodic memory on up-regulation of hippocampal processing by the amygdala. Recent work on animals suggests that fear memories for episodic experiences that are hippocampus dependent return following disturbances at reactivation (476, 515). To which extent episodic memories for emotional experiences can be permanently altered, and to which extent episodic memory can contribute to the reinstatement of fear responses remains to be investigated. Here it is interesting to notice that explicit knowledge of the association between a stimulus and an aversive stimulus modulated the acquisition and persistence of fear (479). Following this issue is the question of the nature of memory for emotional experiences that contribute to fear and anxiety disorders. Are these memories simple associative and intrinsically amygdala centred, dependent on up-regulation of the hippocampus by the amygdala, or truly episodic? The answer to this question will determine the direction of future research investigating the ability to alter memory for emotional experiences aimed to optimize treatments.
In sum, to further increase understanding of the circumstances under which fear may return require investigation of the role of contextual safety learning and the role of the hippocampus. In addition, different elements of a memory trace, or memory systems, may interact in emotional memory, and altering one element of a fear memory but not another runs the risk that fear may return, another issue that requires further investigation.

12.10 Clinical implications.

The ability to alter memories for emotional experiences holds the promise to develop novel strategies to optimize treatment of affective disorders. Here the clinical implication of the work presented in this thesis will be discussed.

The relationship between structural brain abnormalities in posttraumatic stress disorder and depression were under investigation in Chapter 2 and 3. Several regions showed altered volume in both disorders, most notably the dmPFC and insula, and flashback re-experiences were found to be associated with regional volume reductions in posttraumatic stress patients. However, many of the detected regions that show volumetric abnormalities fall outside the neural network of emotional memory as depicted in Figure 12.1 This might be due to the specific patient population under investigation, but could alternatively reflect regions that contribute to the aetiology and persistence of affective disorders. Future studies will have to reveal if these results hold up, and if so, what the contribution of these regions is.

In addition, in Chapter 4 neural activity between the retrieval of normal emotional episodic memory and memories that evoke flashbacks were compared. Several brain regions were found to display greater activity to flashback evoking items, and notably regions in the medial temporal lobe were not found to be more active or even showed reductions in activity. These results can be taken to support the dual-processing theory (132, 133, 676), which states that flashback memories are vivid
memories but lack temporal context, and are supported by the dorsal visual stream that processes egocentric representations. However, there are several limitations to this study that could also explain the absence of the detection of involvement of medial temporal structures such as the hippocampus and amygdala in flashbacks. A first limitation is the relative small number of posttraumatic stress disorder patients that could be included in the final analyses. Thus, the study may not have had enough power to detect involvement of medial temporal regions. In addition, the definition of flashback evoking events versus normal emotional episodic memories was troublesome. Following the creating of their trauma script, posttraumatic stress patients indicated which items evoked flashback experiences. They also indicated when items evoked flashback experiences in the scanner. These indications overlapped, but not perfectly, and patients even indicated having had flashback experiences to items that were not from their own traumatic script. All items that patients indicated as having elicited flashback experiences were assigned flashback items for maximal power in the analyses, which may not have been an optimal definition. In addition, the normal emotional episodic items still came from subjects own trauma script, and are thus likely to be highly arousing as well. It could be that both conditions evoked such strong responses in regions like the amygdala that detection of involvement in flashbacks was obscured. Future studies will have to further address these concerns.

Several studies presented in this thesis indicate that altering memory for emotional experiences is possible at the time of retrieval or reactivation (Chapter 7-11). These findings serve as a proof-of-principle that altering memory is possible, findings that hold a promise for clinical application. An intriguing question is whether the methods that allow one to alter memory in well-controlled experimental settings translate to clinical practice.

Initially successful efforts to translate basic science findings on learning and memory into clinical practice have attempted to enhance extinction or exposure based therapy in clinical settings (517). The discovery that the
NMDA agonist D-cycloserine could enhance extinction learning in animals has led to applications in humans in combination with exposure therapy and shown promise in clinical studies for patients with posttraumatic stress disorder, obsessive-compulsive disorder, phobia, and social anxiety disorder (666, 677-681). Similarly, administration of the noradrenergic agonist yohimbine prior to exposure therapy has been found to decrease fear across exposures in claustrophobic patients and results in reduced fear in a follow-up rug free session (682, 683). Thus, initial translations from basic science to clinical application aimed to enhance extinction learning and exposure therapy have had promising results.

Clinical studies have started to build on basic science findings that memories can be altered by the pharmacological administration at the time of memory reactivation (405, 411, 507). The reactivation of traumatic memory through a trauma script driven reactivation in the presence of propranolol was found to reduce physiological responses at re-exposure to the trauma script in PTSD patients one week later (404). However, although propranolol and the benzodiazepine alprazolam were found to decrease negative affect scores during brief exposure therapy session, no sustained reductions were detected after a 48h wash-out period in four patients (684). Yet, reactivation of drug-related words in the presence of propranolol in abstinent heroin addicts reduced subsequent retrieval of these words (685). Hence, initial studies suggest that the combined reactivation of memory with propranolol administration may be beneficial in clinical populations.

The most exciting translation of basic science on the ability to alter memory to clinical practice exploits the memory reactivation-extinction paradigm (406, 454, 686). In a truly translational study, rats with a heroin dependence after daily retrieval of drug-associated memories prior to extinction training showed attenuated drug-induced reinstatement, spontaneous recovery, and renewal of drug conditioned effects and drug seeking. In parallel, reactivation of drug-associated memories in heroin addicts prior to extinction training resulted in attenuated drug cravings up to 130 days following the procedure (686). Hence, the reactivation-
extinction procedure promises to be of great clinical potential. Beyond applied techniques, knowledge of the neural mechanism underlying emotional memory might also contribute to optimized treatments.

The identification of the neural networks that underlie emotional memory and an increased understanding of neuromodulatory function could serve as a template to optimize treatments to the personal needs of patients. The identification of abnormal brain structure of brain function in particular regions could contribute to a specification of the cause of psychiatric disorders for a particular patient. For example one patient with posttraumatic stress disorder may have a hyper-reactive amygdala, triggering exaggerated fear responses; another suffer from distorted vmPFC function, resulting in a lack of inhibitory control over fear responses; or hyperbolic dmPFC functioning, resulting in increased sustained fear responses; yet another from abnormal insula functioning, resulting in hypersensitive cognitive awareness of fear responses; or over activity of the locus coeruleus resulting in increased general arousal states, or hippocampal hypoactivity causing an inability to inhibit fear in safe contextual situations. It is easy to imagine that dependent on the specific brain abnormality specific treatments may be more optimal. For example, extinction training specifically optimizes vmPFC regulated inhibitory fear regulation, thus patients with abnormal vmPFC functioning might benefit most (or least) from extinction-based treatments, such as exposure therapy. Similarly, the studies presented in Chapter 10 and 11 suggest that patients suffering from abnormal dmPFC or hippocampus functioning might benefit from propranolol treatment, whereas serotonergic treatment might have positive clinical outcomes in patients with altered dmPFC or insula activity. In addition, the theory on neuromodulators set out in this discussion suggest that propranolol might be effecting in patients that show hyperactivity in regions receiving input from the locus coeruleus, whereas patients with hyperactive locus coeruleus functioning itself might benefit more from clonidine treatment. Increased understanding of the neural mechanism that support emotional memory, and its malfunctioning in patients could thus contribute to optimize treatments. Of particular interest is here the development of
evidence-based personalized medicine. Apart from requiring progress in scientific understanding and an ability to characterize neural functioning on an individual basis, the development of personalized medicine in psychiatry will require a better specification of the individual symptoms of patients and a move away from the clustering of symptoms under a single disease label.

One final limitation worth mentioning is that treatment of affective disorders by altering memories may be beneficial, but is not a magic cure. After the many years patients struggle with psychiatric disease, patients social and professional lives are often highly disrupted, may suffer from unhealthy living styles, or have come to personify with their disease state, fully solving these situation will require active social counselling and/or cognitive and behavioural restructuring therapy. Altering memories for emotional experiences is just a first step.

In conclusion, basic scientific studies towards brain abnormalities in affective disorders and the ability to alter memories for emotional experiences holds the promise to develop novel strategies to optimize treatment of affective disorders. The first promising clinical studies targeting enhancement of extinction, or alteration of memory at reactivation have been conducted.

12.11 Conclusion

The work presented in this thesis has contributed to increased understanding of the ability to alter memories for emotional experiences. By presenting an extended view on memory formation, a model of the neural network that supports emotional memory, identifying the overlap between this neural network and brain abnormalities in patients with affective disorders, and putting forward an integrated view on neuromodulatory function, hopes are that this work will both have increased fundamental understanding of the mechanism that drive behaviour and contribute to the development to optimize treatments for affective disorders.
“When you wake up we will have months, maybe years to piece together the broken fragments of your past; better yet we can invent memories that fit your fantasies. For the time being, I will tell you about myself and the other members of this family we both belong to, but don’t ask me to be precise, because inevitably errors will creep in. I have forgotten a lot, and some of the facts are twisted. There are places, dates, and names I don’t remember; on the other hand, I never forget a good story”

-Paula-

Isabel Allende, 1995
THE STORY COMES TO AN END.

KABOOM! Glass splinters. Swiftly we turn our heads. ‘The principles of neuroscience’ has fallen of the bookshelf and shattered a vase on its way down. I look at Ben, he smiles: “This would have brought me back in the past, made me experience it all over again, but now, nothing.” Anna smiles and says: “I am happy the treatment has worked so well for him.”

Let’s hope that fundamental science can contribute to make this a clinical reality for patients like Ben in the near future.
Emotionele gebeurtenissen worden vaak goed onthouden en helpen ons bij het aanpassen aan onze omgeving, maar kunnen ook leiden tot traumatische herinneringen zoals in het geval van posttraumatische stressstoornissen (PTSD). Marijn Kroes onderzocht wat er in het brein van patiënten met PTSD veranderd, wat er gebeurt als zij zich traumatische gebeurtenissen herinneren, en of het mogelijk is om geheugen voor emotionele gebeurtenissen permanent te veranderen of zelfs uit te wissen. Zijn onderzoek toont aan dat zowel de structuur als het functioneren van het brein bij PTSD patiënt is veranderd. Verder laat zijn werk zien dat het mogelijk is om herinneringen aan emotionele gebeurtenissen uit te wissen door de elektrische activiteit van het brein te verstoren na geheugen heractivatie. Hiernaast blijken emotionele herinneringen permanent te kunnen worden veranderd door het blokkeren dan wel activeren van de stoffen noradrenaline en serotonine in het brein tijdens het herinneren van emotionele gebeurtenissen of het aflernen van angstreacties. Deze aanpassingen van emotioneel geheugen gaan gepaard met veranderingen in hersengebieden die belangrijk zijn voor de regulatie van emoties en deze gebieden overlappen met de abnormale brein functies in patiënten met PTSD. Marijn Kroes bevindingen geven aan dat geheugen voor emotionele gebeurtenissen permanent kunnen worden veranderd en kunnen een belangrijke bijdrage leveren aan de ontwikkeling van nieuwe psychiatrische behandelingen.

NEDERLANDSE SAMENVATTING.
REFERENCES.


181. Ingram KM, Mickes L, Wixted JT (2011): Recollection can be weak and familiarity can be strong. Manuscript submitted for publication.


425. Gordon WC, Spear NE (1973): The effects of strychnine on recently acquired and reactivated passive avoidance memories. Physiol Behav. 10:1071.


635. Yerkes RM, Dodson JD (1908): The relation of strength of stimulus to rapidity of habit formation. Journal of Comparative Neurology and Psychology. 18:459-482.


Sara SJ (2010): Reactivation, retrieval, replay and reconsolidation in and out of sleep: connecting the dots. Front Behav Neurosci. 5.


32. van Dijk, J.P. (2010). On the Number of Motor Units. Radboud University Nijmegen, Nijmegen, the Netherlands.


38. Grootens, K.P. (2010). Cognitive dysfunction and effects of antipsychotics in schizophrenia and borderline personality disorder. Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands.


45. Timmer, N.M. (2011). The interaction of heparan sulfate proteoglycans with the amyloid b protein. Radboud University Nijmegen, Nijmegen, the Netherlands.


47. van Grootel, T.J. (2011). On the role of eye and head position in spatial localisation behaviour. Radboud University Nijmegen, Nijmegen, the Netherlands.


57. van der Linden, M.H. (2011). Experience-based cortical plasticity in object category representation. Radboud University Nijmegen, Nijmegen, the Netherlands.


61. van Leeuwen, T.M. (2011). 'How one can see what is not there': Neural mechanisms of grapheme-colour synaesthesia. Radboud University Nijmegen, Nijmegen, the Netherlands.


64. Voermans, N. (2011). Neuromuscular features of Ehlers-Danlos syndrome and Marfan syndrome; expanding the phenotype of inherited connective tissue disorders and investigating the role of the extracellular matrix in muscle. Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands.


78. van Barneveld, D.C.P.B.M. (2012). Integration of exteroceptive and interoceptive cues in spatial localization. Radboud University Nijmegen, Nijmegen, the Netherlands.
86. Verhagen, L. (2012). How to grasp a ripe tomato. Utrecht University, Utrecht, the Netherlands.


89. Xiang, HD. (2012). The language networks of the brain. Radboud University Nijmegen, Nijmegen, the Netherlands.


107. Van Eijndhoven, P. (2012). State and trait characteristics of early course major depressive disorder. Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands.


122. Lagro, J. (2013). Cardiovascular and cerebrovascular physiological measurements in clinical practice and prognostics in geriatric patients. Radboud University Nijmegen Medical Centre, Nijmegen, the Netherlands.
I could not have wished for a more wonderful PhD supervisor. Guillen I can’t begin to describe the importance of your supervision, guidance, mentoring, and ‘parental management’ to the completion of my PhD. I vividly remember (falsely perhaps?) being crash-landed into your office during my first visit to the Donders, discussing a possible PhD position in your lab, the work I had done previously, and science in general in a very pleasant and intellectually stimulating atmosphere. This atmosphere has been characteristic for the entire length of my PhD period and one I’ve greatly enjoyed. Your management style has created a lab that is structured, organized, but highly supportive, interactive, social, and intellectually challenging, a place and style I’ve much enjoyed and appreciated over the years. Everyone is different, me, highly chaotic, all over the place, hyper-active, etc., yet you’ve somehow found a way to supervise me that allowed me to work independently, stay enthusiastic about science and our projects, whilst slowly but steadily teaching me to "FOCUSS!", learn brevity, and comprehensiveness, and be put in place when necessary (yes sometimes this was necessary). It is an amazing quality if a supervisor can adjust its supervising style to a student so that he most optimally functions and develops, and I’ve much appreciated you for it. I’ve greatly enjoyed our scientific discussion whilst writing our review, often extending to discussion that allowed me to have a peak into what it means to have a research group; our visit to Edinburgh, and our discussions on how to implement the ERC programme; interactions that were not only scientifically interesting but very pleasant on a personal level. Exemplary for your efforts to allow your PhD students to develop most optimally is your decision to allow to me work in Francesco’s lab for the last remainder of my PhD period so that I could learn a branch of neuroscience that I was very interested in. Which employer allows an employee to go work for someone else? Who does that? Guillen Fernandez, that’s who, and I am very grateful for it. I fondly hope that we will be able to work together, and collaborate in the future and to continue our very amiable interactions. Guillen I would like to thank you, not just for allowing me to develop as a scientist, but, more important, as a person.

Big Barrie, Uncle Buck, Yoda, Dude, Mate, Bryan, where to start? Sometimes you meet people in your life that you just click with. What began as an interview
for an internship position at the FIL soon ended up on a terrace over coffee, and in a bar over beers (and peanuts of course), culminating in an unstoppable flow of research ideas, critical discussions and loads of laughs. One research project became two research projects became x-research projects, as too many questions needed to be investigated, and results needed to be scrutinized. Late nights following hospital hours, weekends outside shifts, interactions full of enthusiasm, fun, and an absolute love for science you’ve infected me with. One can learn a lot from books, but someone has to show you the ropes. You are my scientific mentor and I can only say that I can’t phantom anyone better. Your incredible intelligence, scientific knowledge, complete dedication to train a student to the best of your abilities, are characteristics I much appreciate. All that amidst your ever cheerful perseverance issues: "Well done young skywalker", or more often "You Muppet!", "reeeeeconsolidation", make it incredibly stimulation, informative, and just good fun to have worked with you over the years. This thesis would have never existed if not for you driving the Barry Beamer across a grass field in Nijmegen, just one example of your incredible support and I can only thank you for it. It’s an absolute privilege to have been able to keep working together during my PhD and to have you as a promotor. The lines between supervisor and friend have long become blurred, you’ve become so much more, a friend, and at times a bigger brother perhaps. We’ll now continue to work together in your lab in Madrid, and hopefully keep our collaborations going even after that, and I’m sure we’ll keep up the amazingly fun times we’ve had in London, during your visits to Amsterdam, mine to Madrid and Marbella. "On to the mattresses", and lets introduce Madrid to the ‘... Dr. Strange’!

**Guido.** Barely done with your own PhD you managed to get me acquainted with the Donders and up and running with my PhD in no time at all. Interaction with you has been smooth sailing from the start. We’ve worked together as colleagues, you’ve given me incredible independence, but we’re always there whenever I needed advise or help. The road towards a PhD is not always easy, there are bumps and potholes to stumble over, and it’s incredible helpful and pleasant to have someone around whom has just conquered them all and shows you the way. You’ve been an incredible support and have always had my back when needed, which has made this PhD a much more enjoyable journey, and I
am grateful of you for it. We’ve had long discussion on our research projects, its progress, how to deal with the issues facing us, which I have much enjoyed and have contributed greatly to completion of this thesis. You’ve been an incredible source of advice and help even after you left the Donders, a quality and effort I am thankful for. You’ve now successfully gone on to start your own lab in Amsterdam, and will even be following up on the work we’ve done together. I can only hope to follow your example one day. Hopefully we can continue to work and socially interact together in the future as I’ve much enjoyed and appreciated you over the last years.

To me neuroscience is fundamentally interesting, yet its relationship to clinical practice is a strong motivating factor to do research. Whilst in London I was offered the opportunity to explore the clinical side of neuroscience. Chris I am very grateful to have been involved in your research endeavours. You and Mat had put forward an incredible effort to study PTSD and depression patients within an MRI scanner, an absolutely monumental task, and allowed me hitch along for the ride. You’ve taught me great insight into the problems that faces these patients and enlightened me with the relationship between clinical abnormalities and theories of brain function. Your enthusiasm and friendly demeanour, wonderful interactions during my stay in London, return visits, and Skype conversations made it an absolute pleasure to have worked with you. I’m very happy that our three papers together have made it into this thesis. Thank you, and I hope we will soon meet up over a wonderful glass of wine.

Diner during the first Fernandez lab retreat I was at (and I’m paraphrasing here). Marijn: "I’ve been studying reconsolidation in humans, but the problems is that we can’t just use protein synthesis inhibitors or ECT or so to impair memory", Indira: "oh, but we have a lot of patients that receive ECT that participate in studies, we could try to work with them." This marked the start of what might be the most monumental study in this thesis, and certainly one of the most interesting research projects I have found myself to be involved in. Indira without you this study would have never some about. You helped me work out the study design, arranged contacts within the RUNMC and with the psychiatrists in Arnhem. Working with severely depressed patients is not always easy, but you’ve helped me out here, taught me about their condition, and have been
supportive throughout the course of this project. I much enjoyed our meetings, both scientific, and social, like during the annual BBQ at your home, and hope we will continue to do so in the future.

One email, an incredibly friendly response, and before I knew it I was standing in the OR witnessing an ECT session and trying to keep up with the high-speed train of uncanning enthusiasm that is Jeroen. When we met you had just started your own PhD and were setting up your first study, all that while running a full time psychiatric clinical job. You’ve truly made our research project see the light. Everything needed was arranged, everything explained, dealt with and solved in no time at all. It is a wonderful characteristic of you: never a problem, always a solution, deal-with-it-immediately attitude, all with great flair. You, and the other clinicians: Bas, Michael, Joep, Boudewijn, Lucas, made me feel right at home in Arnhem, welcomed me into your circle. Each visit, and every ECT session I was at was accompanied by fruitful discussions, essential questions, insightful explanations, and an incredible helpfulness from all of you. I’ve had an amazing time in Arnhem thanks to all of you. I felt honoured to be able to help out, even if only a little, on your own projects, and am intrigued to see how you and Guido will continue the work we’ve started. Jeroen thank you for everything, and I hope I will be able to defend my thesis as well as your outstanding performance.

One more person form Arnhem deserves to be mentioned separately. Oscar you have been the backbone of our ECT study. Kept an eye on possible candidates, helped me approach them, collect information, allowed me to crash your office, and been helpful in any way possible. On most visits we ended up talking about our daily lives, what we had been up to over the weekend, future plans, etc, etc. I’ve much enjoyed our talks and working with you, and would much like to thank you for the wonderful time in Arnhem.

Ray you allowed me into a lab few-to-none interns ever make it into. My time in your lab, and at the FIL in general have truly shaped me as a scientist, both in work attitude, scientific thinking, and goals to strive for. The FIL was an incredibly inspiring environment and one I have very fond memories of. I’ve much enjoyed working with you, learning from you, am honoured to have been
able work with one of the great scientists to have founded the field I now work in. Thank you.

**Joris, Jonas, Hasan,** working with you on our joint research project has been a very insightful and pleasant experience. I enjoyed our interactions, and am appreciative to have gained some insight into the workings of scientific projects within a commercial environment. I was quite amazed by the results of our study, hope it will see the light of publication soon, and would like to thank you for all your patients and support.

Several internship students have profoundly contributed to the work in this thesis. **Klodiana** running our propranolol study would not have been possible without you. You are an incredibly motivated and fun student, and I much enjoyed our time together testing all those subjects. I hope you are enjoying your new PhD position and will do well, and wish you all the best. **Sanne** its been a pleasure to work with you, and your rigidity, attention to detail, and hard work have led to great stimulus material, and together with **Martijn** a huge data set I still have to figure out what to do with, and eventually contributed greatly to our subsequent studies.

Thank you **Hans-Christiane Pape, Roshan Cools, and Sander Daselaar** for being on the evaluation committee of my thesis. I am sorry for making you read this much text, but I hope you found it interesting, and would like to thank you for the feedback on my thesis, which might not have been normal procedure but was very informative for me.

A special thank you goes out to all the **participants** of the studies presented in this thesis. Especially to the **patients** for whom participation is not always easy, but were willing to put forward their best efforts, and often shared their highly insightful stories.

The Fernandez lab is a delight to work in largely because of all the great colleagues that make the lab into a highly interaction, stimulation, social, and supportive environment. I’ve absolutely loved all the discussion during lab meetings, retreats, coffees, and lunches, and am very grateful to all your help
and support in shaping and reflecting upon the work in this thesis. **Atsuko, Carly, Daphne, Dirk, Eelco, Erno, Frauke, Floris, Hein, Isabella, Jan-Willem, Jenny, Linda, Lindsey, Marieke, Mariet, Marlieke, Marloes, Mark, Mirre, Nicole, Nils, Noortje, Olga, Ruud, Sarah, Susanne, Shaozheng, Yu, and all interns** (I hope I didn’t miss anyone, if so sorry! I mean you to), you are an incredibly bright, fun, and great bunch of people and I am very pleased to have worked with you all. Carly, Atsuko, Floris, Isabella, Mariet, and Susanne I feel privileged to continue to collaborate with you and look forward to work with you in the future. Erno, thank you for all advise, and on your help to find a post-doc position, and I hope we will run into each other in the Big Apple at some point. To all of you, I am sure we will continue to meet in the future and am looking much forward to it.

**Sabine** what would I have done without you? Two major drug studies, one clinical study, all associated ethical, safety, and organizational issues, health screenings, blood sampling, genetic sampling, could I have placed a bigger burden on you somehow? Yet you managed to keep everything in check and run smoothly. Your contribution to my PhD is invaluable and I am very pleased to have worked with you in an always enjoyable setting.

The Donders is an incredible place and houses some of the smartest people I’ve ever met, who work on cutting-edge science in a highly competitive field, yet are able to maintain a very social, helpful, supportive, and stimulating atmosphere. A big thank you to the **PIs** who’ve been able to construct such a wonderful institute, and especially **Roshan and Christian D** for their scientific discussions and support. Some people I would like to mention separately. **Hanneke** we spend time together in London at the FIL and continued to interact at the Donders, you are the example of a great promising young scientist and a lovely person and I’ve learn a great deal from you. Hope to keep running into you in the next city we may find ourselves in. **Stephen**, I’m happy to have been reunited with your ever-intriguing character and to have been able to find our way through the SCR analyses together. **Lennart** thank you for sharing all your knowledge on fMRI analyses and for letting me take your thesis as an example. **Niels** we shared offices only briefly but I had great fun every time we were there together.
Working at the Donders is especially pleasant and efficient due to its incredible organization and support that is provided by Tildie, Arthur, Sandra, Nicole, Marek, Sander, erik, uriel, rene, Edward, Paul, who cannot be appreciated enough.

Francesco thank you for allowing me to spend a "sabbatical" in your lab and subsequently taking me on as a post-doc. You’re work has been a great inspiration, and I’ve truly enjoyed the opportunity to acquire some insight into systems neuroscience of memory and electrophysiology from you. It’s been incredibly fun to work with you and learn from you. I highly appreciate all your efforts to explain all methods, techniques, and theories that exist in this field. Spending time in your lab has led me to a new level of nerdy-ness and I love it ("can't we use an arduino for that?"). You’ve been an incredibly hospitable and fun person inside and outside that lab, a very stimulating mind to be around, and I feel fortunate to have met you. Together with Esther and Tim we’re having an amazingly interesting time setting up a new lab, and despite sometimes frustrating slow progress you’ve maintained a great working atmosphere, so thank you all. I am sure great things will come out of the lab soon, and I hope that we will be able to continue our collaboration.

My latest endeavour into electrophysiology started at the UVA, and I would specifically like to thank Cyriel for allowing me to spend a wonderful time in your group. Conrado I've had a great time working with you and setting up our project, so thank you. And a further thank you goes out to everyone in the Cognitive and Systems Neuroscience group.

Benno, Piray, Erika, Areg, and the others from the Roozendaal lab, our recent interactions and all your help has been very informative and enjoyable so thank you.

Richard, Lisa, Dorothy, Tomonori, Szu-Han, Adrian, and the others at the Morris lab in Edinburgh, thank you for having been incredibly helpful, and I truly enjoyed our scientific discussions, joint meetings, and visits to Edinburgh that have been very inspirational.
Every now and again you get the opportunity to meet the who-is-who in science and at times they leave a mark on your scientific thinking. In my case some of the people to thank for this are Bruce McNaughton, Christian Buchel, Christina Alberini, Devon Mueller, Dominique deQuervain, Eleanor Maguire, Daniela Schiller, Georgy Buzsaki, Joe LeDoux, Karim Nader, Karl Friston, Liz Phelps (I highly anticipate working together in the future and would like to thank you for all the efforts you have made so far, its been interesting and flattering), Lynn Nadel, Edvar Moser, Mick Rugg, Paul Frankland, and Raphael Kalisch, thank you for some very interesting discussions.

**Eelke** we’ve been brothers since I was one year old, and we’ve been best friends ever since. Its quite incredible to always have someone around you can rely on, trust, talk to, play with, party with, see the world with, and discuss everything in life with. With you around I know we could face anything and always come out on top. Your one of the brightest, hard-working, and most hilarious persons I know, and I am proud to have you as a brother. It’s been amazingly interesting to discuss the overlap between business and science with you over the years, and our recent undertaking to develop talks for companies and begin our own start-up has been wonderful. Its impressive to see how much you’ve learned about neuroscience in a very short period, to the point that I think you would do a pretty good job at defending my thesis yourself. Hence, I could not wish for a better paranimph, but more importantly, I could not wish for a better brother.

**Dennis** I’ve known you since I was seven, looked up against you, the club champion boys till 10 is an idol at that age. At 12 we became friends, or maybe more accurate, you adopted me as a little brother (yes little I know), a difference of three years is an enormous gap at that age. It grew over the years; went to practice together, started skateboarding together, going out, playing music, went to gigs, throughout those years I had an example to follow, someone to hang out with, and a friend to call on. As we grew up our friendship equated, yet you remained the most dependable and trustworthy friend I have. It’s an incredible quality that comes natural to you. However far away, you were always close. I still have a pile of papers with emails we send each other during my year in the US, sharing everything that happened in life intermixed with the biggest pile of
nonsense that made roflmo a reality. Next, I moved to Amsterdam, you followed a little later, and we remained best friends over the years. I am incredibly happy to see that you seem to have found your place in life over the last couple of years, and wish you all the best. I may once again disappear to some far away place, but you’ve always been by my side, now as my paranimph, and I hope you’ll always be. Simply put, I am just incredibly fortunate to have you as a friend.

I’ve been incredibly privileged to have friends in high school that I still have today; **Elmer, Migh** (a special thank you for helping me design this thesis), **Robin, Jarrik**, many that I had known but truly became friends in Amsterdam; **Daan, Ernie, Gerben, Hans, Jeroen, Laurens, Kreits, Matthieu**; and others that became friends and joined the crew over the years; **Alex, Eli, Johanneke, Kaspar, Liza, Lotte, Maya, Oscar, Saskia, and Tim**. You are an incredible bunch of people and we’ve shared so many good times. A separate mention goes out to **Marlous and Barbara**; you’ve been the greatest photo-models, and possibly without you knowing it many people have remembered you well but subsequently forgotten you all in the name of science. **Thomas, Thomas, Jitse, and Frederic** I’ve had the best of times hanging out with you, going to concerts and festivals, could we please have another silly dancing competition? To all of you thank you.

This thesis was largely written whilst hanging out in bars with **Tara** 'biggest rock-star in neuroscience' till the wee hours of the night. I’d never thought I’d meet someone like you in one of the most peculiar places of the world. Neuroscience and music occupy a large part of my life and sharing a similar taste for both with someone is amazing. I’ve had a great time hanging out with you, going to concerts with you, sharing frustrations over elements of a PhD with you, and gaining insight on life with you. I hope to keep running into you wherever in the world, and keep sharing more awesome experiences, which inevitably happens with you around.

During a PhD, science can come to occupy ones mind completely, to an overwhelming point at times. Luckily there has always been one place where I could tune out, relax, and have fun: **Chip & Charge**. Throughout the years I’ve had a great time with so many people so thank you all. A special thanks to
the boys that have shown me every corner of the courts: Folkert, GP, Justin, Laurens, Luuk, the boys of HE1, and men of HE2; de 'matige finale spelers’ Frederiek, Irlen, Miriam, Nienke, you’re an incredible bunch and I sure hope you’ll come over for diner wherever in the world I may end up; the ‘party-gang’ for sharing the love, when are we going again?

Lotte "I don’t have time to explain the complexities of memory storage" you’re on.

KKK-GWL: Martijn, Steffan, and more recently Mark, roomies that you can play tennis with, go out with, have fun with, and that are simply just friends, what more can you wish for?

The "hele grote meneren" of JOS, this season was so much fun, so thank you, and I hope to join you for a few more matches this year. I’ve already heard talk of the ’k-word’.

Pa en Luus, het enige stukje opzettelijke Nederlands is aan jullie voorbehouden. Pap ik had hier nooit gestaan zonder jou. Biologisch gezien logisch, maar dat bedoel ik niet. Ik herinner me niet veel van mijn vroege jeugd, maar ik weet wel dat ik altijd een onvoorwaardelijke liefde heb gekend, altijd een veille en vertrouwde plek om naar terug te keren, tot de dag van vandaag. Twee kleine blagen alleen opvoeden kan niet eenvoudig zijn, maar jij deed het, gaf ons een normaal leven. Ik heb het je niet altijd makkelijk gemaakt, maar je hebt me altijd alle vertrouwen en ruimte gegeven en gesteund waar nodig. Het is geweldig om een vader te hebben waar je altijd bij terecht kunt, mee kunt praten, advies kunt krijgen, kunt plagen, en mee kunt lachen en ik hoop dit nog heel lang met je te kunnen doen. Straks sta ik daar op een podium voor al die professoren in toga en dat heb ik vooral aan jou te danken, dus dan mag je best wel een beetje trots zijn, ook op jezelf (maar zullen we deze keer de foto’s aan iemand anders overlaten?). Pap dank je wel. Luus je hebt Chris weer gelukkig gemaakt ik hoop dat je dit blijft doen, en daar ben ik je onnoemelijk dankbaar voor, maar ik ben je vanaf dag 1 gaan waarderen om jou. Je bent een onlosmakelijk deel van ons gezin, een geweldig lieve, slimme, ondersteunende vrouw, een bron van advies, en bied een plek om even tot rust te komen. Dank je wel, en ik hoop nog heel
lang van jou en je culinaire uitspattingen te genieten. Cobie en Harry er bestaat zoiets als een 'extended amygdala' en een 'extended family', die termen mogen jullie zelf opzoeken en zelf bedenken met welke ik jullie associeer. Jullie zijn de meest warme mensen die ik ken en ik ben blij dat jullie altijd in mijn leven zijn gebleven. Ik heb onnoemelijk veel met jullie gelachen, plezier gehad, en heel veel gedeeld, en hoop dit nog lang te blijven doen. En als ik nog een keer jullie herinneringen kan misbruiken voor mijn eigen wetenschappelijke gewin zal ik het zeker niet laten.
PUBLICATION LIST.

Peer reviewed international publication


Articles under review


13. L. Genzel, M. C. W. Kroes, M. Dresler, and F. P. Battaglia (under rev). Light sleep vs. slow wave sleep in memory consolidation: A question of global vs. local processes?

**Articles in submission or preparation**


17. A. Takashima, F. van der Ven, M. C. W. Kroes, and G. Fernandez (in prep). The fate of neutral memories encoded under arousing context: Neural correlates of sustaining memories over time.
CURRICULUM VITAE.

Marijn Kroes was born on the 5th of March, 1983 in Veenendaal, the Netherlands. After he obtained his high school diploma at the Rembrandt College in 2000, he completed a senior year at Dodge City High School, Kansas, USA, as an exchange student. Completing a ‘propedeuse’ in social work and counselling he went on to study psychonomics at the University of Amsterdam, and completed his bachelors thesis entitled "Stress and memory: generative en degenerative influences of corticosteroids on several neurophysiological memory mechanisms" under supervision of Dr. Harm Krugers. In 2005 he enrolled in the masters programme Neuroscience and Cognition at Utrecht University, where during his first internship he studied the reverse apparent motion effect using human psychophysiology and in vivo single cell recordings under supervision of Prof. Dr. Martin Lankheet, and Dr. Roger Bours. For his second masters internship he moved to the Wellcome Trust Centre for Neuroimaging at the University College London, UK. under the auspices of Prof. Dr. Raymond Dolan, and Prof. Dr. Bryan Strange, Marijn Kroes worked on research projects aimed at understanding the role of emotions on memory in humans employing psychophysiological, neuroimaging, pharmacological, and genetic approaches. This work culminated in an internship report, as well as a master thesis supervised by Prof. Dr. Bryan Strange entitled "functional segregation of the hippocampus". Prior to obtaining his masters degree in 2008, he started as a
research assistant at the department of Clinical Health Psychology at University College London investigating functional and structural brain abnormalities in patients with depression and/or posttraumatic stress disorder together with Prof. Dr. Chris Brewin, Prof. Dr. Michael Rugg, and Dr. Matthew Whalley. In 2008 Marijn Kroes returned to the Netherlands to start a PhD project with Prof. Dr. Guillen Fernandez and under supervision of Dr. Guido van Wingen. His PhD work investigated the ability to alter memories for emotional experiences using functional magnetic resonance imaging, psychophysiology, pharmacological, and genetic approached in both healthy human subjects and psychiatric patients, the results of which are described in this thesis. The last six months of his PhD period he spend on a ’sabbatical’ (sic) in the lab of Dr. Francesco Battaglia at the University of Amsterdam investigating schema consolidation using multi-unit electrophysiological recordings and optogenetics. Marijn Kroes currently lives in Amsterdam and works as a post-doctoral fellow in the lab of Dr. Francesco Battaglia at the Radboud University Nijmegen probing the neural networks of memory.