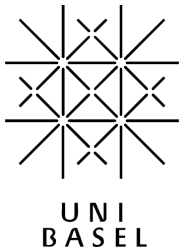


Hormonal and genetic modulation of memory processes in healthy humans: Focus on cortisol and *HDAC5*

Inauguraldissertation
zur
Erlangung der Würde einer Doktorin der Philosophie
vorgelegt der
Fakultät für Psychologie
der Universität Basel

von

Francina Ramona Hartmann

aus St.Moritz (GR), Schweiz

Basel, Mai 2015

Genehmigt von der Fakultät für Psychologie
auf Antrag von

Prof. Dr. Andreas Papassotiropoulos
Prof. Dr. Dominique J.-F. de Quervain

Basel, den 12.05.2015

_____

Prof. Dr. Roselind Lieb
Dekanin

Abstract

Individual differences in memory performance can be due to the influence of various hormones as well as genetic variations and epigenetic modifications. These complex molecular and genetic mechanisms can impact learning, memory consolidation and retrieval differentially. This thesis deals with the modulation of memory processes in healthy human subjects focusing on two viewpoints. Firstly, by addressing the influence of the stress hormone cortisol, as evidence from animal and human studies shows that cortisol can enhance memory consolidation and impair retrieval. Secondly, by analyzing genetic and epigenetic data to find a target associated with synaptic plasticity and memory formation.

To investigate if stress, induced by the cold pressor test, affects memory processes, a fear-conditioning paradigm was used. The stress group showed an increase in the cortisol level and reduced retrieval of the conditioned fear memory. In a further study, we investigated if inter-individual changes in basal cortisol levels affect episodic memory. Results showed an association between stronger decreases in cortisol levels during retrieval and a better recall performance.

In a large genetic study we focused on genetic polymorphisms tagging histone deacetylase 5 (*HDAC5*), a gene associated with synaptic plasticity and memory formation in animal models. We detected significant associations between these polymorphisms and episodic memory performance, especially for emotional information. Surprisingly, these polymorphisms were strongly associated with expression levels of a transcript in the vicinity of *HDAC5*.

These results may have implications for the understanding of the mechanisms underlying memory formation in healthy subjects and the interpretation of genetic data. Additionally, our results may have clinical implications for different neuropsychiatric disorders, such as depression, anxiety disorders or posttraumatic stress disorder, for which learning and memory play an important role.

Table of Contents

Figure Index

Acknowledgements

First of all, I would like to thank my supervisors Prof. Dr. med. Andreas Papassotiropoulos and Prof. Dr. med. Dominique de Quervain for their support and the possibility to be a part of their research team. During my time as PhD student they gave me the opportunity to acquire experiences in the interesting research field of cognitive and molecular neurosciences.

Thanks also to all my colleagues for their help and advice. Especially, I want to thank Dr. Angela Heck for her support, guidance and introduction to genetic analyses. I want as well mention Dr. Annette Milnik who introduced me in data analyses, and M.Sc. Bianca Auschra who gave me helpful suggestions.

Last but not least, I want to thank my parents for their support and the enabling of my studies as well as to my husband for his appreciation and support, especially during the time of writing this thesis.

Abbreviations

ACTH	adrenocorticotrophic hormone
AMPA	α-amino-3-hydroxy-5-methyl-4isoxazolepropionate
BLA	basolateral amygdala
CA^{2+}	calcium ion
CaMKII	CA^{2+}/calmodulin-dependent kinase II
cAMP	cyclic adenosine monophosphate
CBG	corticosteroid-binding globulin
CpG	cytosine-guanine dinucleotide
CPT	cold pressor test
CR	conditioned reaction
CREB	cyclic-AMP response element-binding protein
CRH	corticotropin-releasing hormone
CS	conditioned stimulus
DNA	deoxyribonucleic acid
FKBP5	FK506 binding protein 5
GC	glucocorticoid
GR	glucocorticoid receptor
GWAS	genome-wide association study
HAT	histone acetyltransferase
HDAC	histone deacetylase
HDAC5	histone deacetylase 5
HPA axis	hypothalamus-pituitary-adrenal axis
LTD	long-term depression
LTP	long-term potentiation
MAPK	mitogen-activated protein kinase
MG^{2+}	magnesium ion
MR	mineralocorticoid receptor
mRNA	messenger ribonucleic acid
MTL	medial temporal lobe
NA	noradrenaline
NMDA	N-methyl-D-aspartate

PKA	protein kinase A
PKC	protein kinase C
PP1	protein phosphatase 1
PTSD	posttraumatic stress disorder
SECPT	socially evaluated cold pressor test
SNP	single nucleotide polymorphism
TSST	Trier social stress test
US	unconditioned stimulus

1. Introduction

The ability to learn, store and retrieve information is important for the orientation in daily life. The skill to memorize allows us to acquire knowledge about the self and the environment, which enables us to act appropriately to different situational demands. The investigation of memory processes is therefore an important topic. To discover the mechanisms underlying different neurodegenerative and neuropsychiatric disorders that involve learning and memory processes as well as deficits, it is basic to first understand memory functions under normal conditions. Among others the transfer of acquired memories into long-term states.

This thesis deals with the modulation of memory processes and synaptic plasticity in healthy human subjects focusing on two viewpoints. First by addressing the hormonal modulation of memory processes, namely by dealing with the stress hormones glucocorticoids (GCs); cortisol when referred to humans. Evidence from human and animal studies shows that GCs on the one hand enhance memory consolidation and on the other impair memory retrieval (for reviews see de Quervain, Aerni, Schelling, & Roozendaal, 2009; Schwabe, Joels, Roozendaal, Wolf, & Oitzl, 2012; Wolf, 2009). GCs are also assumed to be involved in stress related disorders, as their release after a stressful experience is controlled by the hypothalamus-pituitary adrenal (HPA) axis (Schwabe, et al., 2012). In the first study (Bentz et al., 2013) included in this thesis we increased cortisol levels in healthy subjects using the cold pressor test (CPT) and linked these cortisol levels to the acquisition of a fear memory in an aversive differential conditioning paradigm. Cortisol is not only released in response to a stressful situation, its secretion also follows a circadian rhythm during the day (Lupien, Maheu, Tu, Fiocco, & Schramek, 2007). The effects of individual differences between cortisol levels, without experimental manipulation, are not well investigated. The focus of the second study (Ackermann, Hartmann, Papassotiropoulos, de Quervain, & Rasch, 2013a) included in this thesis was therefore on natural levels of cortisol. We investigated the influence of basal cortisol levels in reference to episodic memory consolidation and retrieval processes. Furthermore not only cortisol levels per se were investigated but also the changes in cortisol levels during recall performance.

The second aim of this thesis was to investigate the genetic modulation of memory processes in an episodic memory task by analyzing single nucleotide polymorphisms (SNPs) tagging the gene histone deacetylase 5 (*HDAC5*) (Hartmann et al., unpublished

manuscript); a gene involved in synaptic plasticity and memory formation in animal models (Agis-Balboa, Pavelka, Kerimoglu, & Fischer, 2013; Guan et al., 2002; Renthal et al., 2007). Moreover, HDAC5 is expressed in brain regions important for learning and memory, such as the hippocampus and the amygdala (Broide et al., 2007). Interestingly, Roozendaal et al. (2010) reported an interaction between GCs and histone modification mechanisms in memory consolidation. Rats treated with corticosterone after training displayed enhanced long-term memory performance in an object recognition task and increased histone acetylation levels in the hippocampus and the insular cortex. Additional treatment with a histone deacetylase (HDAC) inhibitor enhanced the effect of corticosterone on memory.

In our study we first used a candidate gene approach, analyzing SNPs situated in the *HDAC5* region. Additionally, in a second step, more dynamic processes were analyzed, i.e. the impact of *HDAC5* SNPs on DNA methylation as well as on mRNA expression levels.

This thesis includes the following three original research articles. I contributed to these papers by an involvement in data acquisition, data analyses and writing of the paper.

1. Influence of stress on fear memory processes in an aversive differential conditioning paradigm in humans.
   Bentz, D., Michael, T., Wilhelm, F.H., Hartmann, F.R., Kunz, S., von Rohr, I.R., & de Quervain, D.J.-F. (2013). *Psychoneuroendocrinology*, *38*(7), 1186-1197.

2. Associations between basal cortisol levels and memory retrieval in healthy young individuals.
   Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D.J.-F., & Rasch, B. (2013). *Journal of Cognitive Neuroscience*, *25*(11), 1896-1907.

3. Polymorphisms of HDAC5 are associated with episodic memory, DNA methylation and C17orf65 mRNA expression.
   Hartmann, F.R., Milnik, A., Auschra, B., Freytag, V., Spalek, K., Vogler, C., Vukojevic, V., de Quervain, D.J.-F, Papassotiropoulos, A., & Heck, A. *Unpublished manuscript.*

Additionally, I contributed to the following publications by an involvement in data acquisition and writing of the paper, which are included in the original research paper collection but not in this framework:

4. No associations between interindividual differences in sleep parameters and episodic memory consolidation.
   Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D.J.-F., & Rasch, B. (2014). *Sleep, in press.*

5. Sex-dependent dissociation between emotional appraisal and memory: A large-scale behavioral and fMRI study.
   Spalek, K., Fastenrath, M., Ackermann, S., Auschra, B., Coynel, D., Frey, J., Gschwind, L., Hartmann, F., van der Maarel, N., Papassotiropoulos, A., De Quervain, D., & Milnik, A. (2015). *Journal of Neuroscience, 35*(3), 920-935.

6. Hippocampal activation, memory performance in young and old, and the risk for sporadic Alzheimer's disease converge genetically to calcium signaling.
   Heck, A., Fastenrath, M., Coynel, D., Auschra, B., Bickel, H., Freytag, V., Gschwind, L., Hartmann, F., Jessen, F., Kaduszkiewicz, H., Maier, W., Milnik, A., Pentzek, M., Riedel-Heller, S.G., Spalek, K., Vogler, C., Wagner, M., Weyerer, S., Wolfsgruber, S., de Quervain, D.F.-J., & Papassotiropoulos, A. *Submitted.*

2. Theoretical Background

2.1 Memory: systems and neuronal background

Learning and memory are indispensable capacities to cope successfully with the demands of daily life. They allow flexible and adaptive behavior such as reaction to the environment. Disturbances however can have adverse consequences for our quality of life. Memory is the ability to encode, store and recall information over variable periods of time. On a temporal level, memory can be divided into short-term storage of information for seconds to minutes, and in the case of working memory the additional manipulation of the information, as well as into long-term storage of information for hours and up to years (Dickerson & Eichenbaum, 2010; Pause et al., 2013).

Furthermore the multiple forms of memory can be divided into declarative (explicit) and non-declarative (implicit) memory (Figure 1) (Kandel, Dudai, & Mayford, 2014).

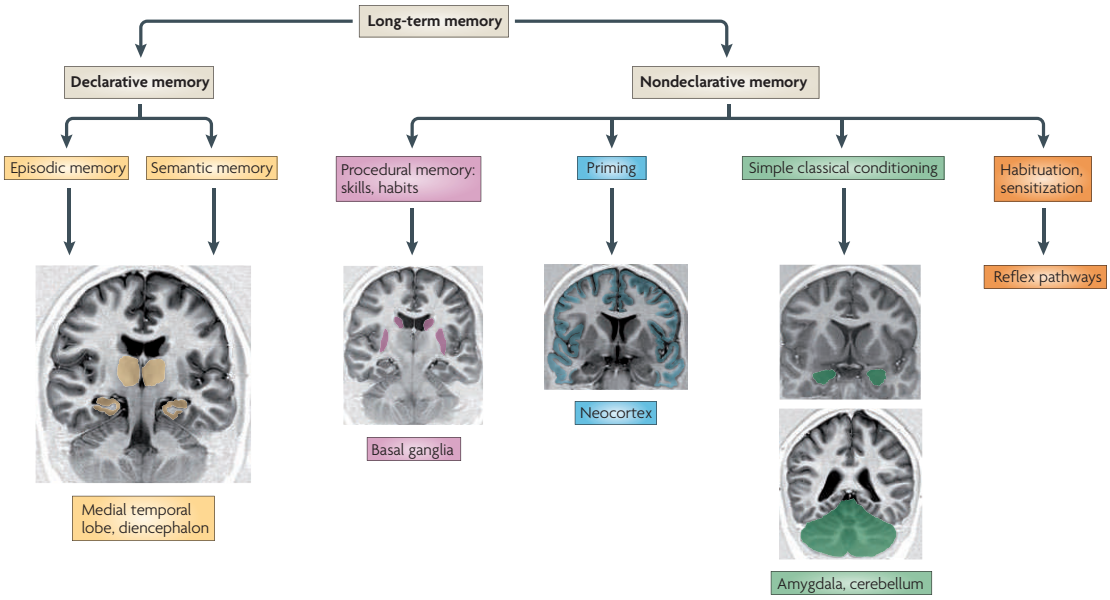


Figure 1. Memory systems of long-term memory and the brain regions involved (Henke, 2010).

Declarative memory is defined through the capacity for conscious recollection of facts and events. It allows the encoding of memories that build a relationship between multiple items and events as well as the ability to compare and contrast the remembered material.

Together these properties enable the modeling of the external world (Squire, 1992, 2004). Declarative memory can be subdivided into semantic memory (knowledge about facts) and episodic memory (knowledge about personal experiences and events of daily life) (Dickerson & Eichenbaum, 2010).

Non-declarative memory on the other hand is a heterogeneous collection of mainly non-conscious learning capacities that are expressed through performance rather than recollection. This includes priming processes, classical conditioning and memory for skills, the so-called procedural memory (Squire, 2004; Squire & Zola, 1996).

Both memory systems, declarative and non-declarative, are dependent on different brain regions and operate in parallel on a neuroanatomical level (Figure 1) (Squire, 2004). Due to the heterogeneous picture of non-declarative memory, various brain regions are implicated, for example the striatum in procedural memory, and the cerebellum, as well as the amygdala in classical conditioning, whereat the amygdala is especially important for the emotional responses (Kandel, et al., 2014; Squire & Zola, 1996). Declarative memory is dependent on the diencephalon and structures in the medial temporal lobe, where the hippocampus and adjacent cortical areas play an essential part (Squire, 2004; Squire & Zola, 1996).

2.1.1 Fear memory and classical fear conditioning

Fear learning and memory involve implicit but also explicit learning mechanisms. Classical conditioning described by Pavlov is a well-studied model of associative learning, including emotional and fear learning (LaBar & Cabeza, 2006; Maren, 2001). The principle of fear conditioning relies on a learned relationship between stimulus and response. An originally neutral stimulus (conditioned stimulus, CS) is associated with an aversive experience (unconditioned stimulus, US), leading to a fear response (conditioned reaction, CR) and to retrieval of the associated fear memory during later presentation with the CS. In animals for example, pairing of an initially neutral tone or light (CS) with a mild footshock (aversive US) leads to a conditioned fear response to the CS, expressed among other as freezing behavior (Bentz, Michael, de Quervain, & Wilhelm, 2010; LeDoux, Iwata, Cicchetti, & Reis, 1988; Myers & Davis, 2002). In humans, classical fear conditioning has been implicated in the mechanisms underlying memory processes involved in anxiety disorders as well as posttraumatic stress disorder (PTSD). CRs in humans during fear conditioning studies are measured as skin conductance responses or fear potentiated

eyeblink startle reflexes, both representing implicit aspects of fear memory (LaBar & Cabeza, 2006). Explicit memory processes operating in fear learning are expressed as expectations about the stimulus associations and as evaluation of the situation developed during fear acquisition (LaBar & Cabeza, 2006; Lovibond, 2006). In humans, they furthermore can be verbalized (Carter, O'Doherty, Seymour, Koch, & Dolan, 2006).

Repeatedly presentation of the CS in absence of the US can lead to a reduction of the CR. This extinction learning represents a new active learning process that is distinct from the initial fear learning acquisition and is accompanied by additional plasticity (Kitazawa, 2002; Myers & Davis, 2002). Extinction learning is the principle underlying exposure therapy in anxiety disorders (Bentz, et al., 2010).

Different brain regions are involved in explicit and implicit memory processes of fear conditioning, e.g. the hippocampus and the amygdala (Bechara et al., 1995; Knight, Cheng, Smith, Stein, & Helmstetter, 2004). Evidence from animal and human studies implicate that the amygdala plays an important role in fear response regulation (Delgado, Olsson, & Phelps, 2006; LeDoux, 2000). Findings from human studies show that lesions in the amygdala lead to deficits in fear conditioning (LaBar, LeDoux, Spencer, & Phelps, 1995) and furthermore amygdala activation in healthy subjects is correlated with the strength of the conditioned response (LaBar, Gatenby, Gore, LeDoux, & Phelps, 1998). The amygdala can be divided into several nuclei, which are associated with different functions (Jovanovic & Ressler, 2010). Inputs from diverse areas of the brain (e.g. thalamus, neocortex, olfactory cortex, hippocampus) enter the amygdala through the basolateral amygdala, where the formation of CS-US connections is believed to take place. The basolateral part has connections to the central nucleus of the amygdala, where the output of the information occurs through projections to autonomic and somatomotor structures, mediating fear responses (Delgado, et al., 2006; LeDoux, 1996). The hippocampus, however, is implicated primarily to the processing of contextual information (LaBar & Cabeza, 2006).

2.1.2 Episodic memory

Episodic memory refers to the ability to remember events or personal experiences both recent and past, including information about the spatial and temporal context. (Dere, Pause, & Pietrowsky, 2010; Tulving, 2002). It can be divided into short- and long-term memory, depending on the duration of the maintenance until retrieval occurs. Short-term memories

therefore are maintained only a few seconds up to hours whereas long-term memories can be stored for years or even for an unlimited duration (Dere, et al., 2010). Highly emotional events in particular can lead to an enhanced memory performance. Cahill, Gorski and Le (2003) for example could show that the degree of arousal at the time of encoding interacts with the effect of stress hormones on memory consolidation. Dolcos, LaBar and Cabeza (2005) furthermore could show that emotionally arousing pictures were better remembered than neutral ones and this effect even persisted one year after encoding.

Impairments of episodic memory have been observed in neurodegenerative disorders such as Alzheimer's Disease (Dere, et al., 2010; Dubois et al., 2007) or Parkinson's Disease (Williams-Gray, Foltynie, Lewis, & Barker, 2006) and in neuropsychiatric disorders including Schizophrenia and Major Depression (Dere, et al., 2010; Pause, et al., 2013). Important brain regions implicated in episodic memory processes are the medial temporal lobe (MTL), including key structures for memory such as hippocampus and amygdala, as well as the frontal cortex (Dickerson & Eichenbaum, 2010). Patients suffering from MTL damage, especially to the hippocampus, show impaired episodic memory. They are affected in their ability to learn new information (anterograde amnesia) but also in their memory for information that was learned before the damage (retrograde amnesia) (Bayley, Hopkins, & Squire, 2006; Bechara, et al., 1995; Scoville & Milner, 1957). Episodic memory impairments, however, can also occur in the course of healthy aging (Shing et al., 2010). Evidence from neuroimaging studies in healthy humans detected associations of MTL activation with encoding and retrieval of episodic memory content (Alkire, Haier, Fallon, & Cahill, 1998; Nyberg, McIntosh, Houle, Nilsson, & Tulving, 1996; Shing, et al., 2010; Squire et al., 1992).

Episodic memory additionally involves the frontal system, as has been shown e.g. in lesion studies (Davidson, Troyer, & Moscovitch, 2006). Especially the lateral prefrontal cortex is activated during strategic processes such as the use of strategies during encoding and other controlled processes during retrieval (Davidson, et al., 2006; Dickerson & Eichenbaum, 2010; Robin et al., 2015). The amygdala plays an additional role in memory, when it comes to emotionally arousing information. It has been shown that the amygdala is especially involved in the processing of the emotional part of the information but not of neutral information (Alkire, et al., 1998; Cahill et al., 1996; McGaugh, 2004). Furthermore, an imaging study demonstrated an emotional arousal driven increase in connectivity between the amygdala and the hippocampus during encoding of emotional information in relation to neutral information (Fastenrath et al., 2014).

2.2 Molecular mechanisms of memory formation and storage

The molecular biology underlying non-declarative and declarative memory processes has been studied extensively in the marine snail Aplysia and in the mammalian hippocampus, e.g. spatial memory in rodents representing hippocampus-dependent declarative memory (Bailey & Kandel, 2008; Bliss & Lomo, 1973; Kandel, 2001). It has been shown that encoding, consolidation and storage of information relies on activity-dependent modulation of synapses (Bliss, Collingridge, & Morris, 2003; Kandel, 2001). Learning-related synaptic plasticity can be divided into short-term and long-term memory processes with different underlying mechanisms (Bailey & Kandel, 2008). Short-term processes do not require the synthesis of new proteins and a single train of stimulation leads to modification of preexisting proteins and to the strengthening of existing connections (Kandel, 2001; Nguyen, Abel, & Kandel, 1994). In contrast, long-term synaptic changes after repeated trains of stimulation require transcription and translation of DNA, using cyclic adenosine monophosphate (cAMP)-dependent protein kinase, protein kinase A (PKA), mitogen-activated protein kinase (MAPK), and cAMP response element-binding protein (CREB). This leads to the formation of new synaptic connections and to longer lasting memories (Kandel, 2001; Milner, Squire, & Kandel, 1998; Nguyen, et al., 1994). Bliss and Lomo (1973) described a now well-known form of synaptic plasticity, the long-term potentiation (LTP) by discovering the activity-dependent plasticity in the hippocampus. High-frequency electrical stimulation of hippocampal regions facilitates chemical transmission by coincidentally activating pre- and postsynaptic elements and therefore inducing post-synaptic action potentials (Cooke & Bliss, 2006). The opposite of LTP is called long-term depression (LTD) and is characterized by a reduction in the efficiency of synaptic strength (Cortes-Mendoza, Diaz de Leon-Guerrero, Pedraza-Alva, & Perez-Martinez, 2013). LTP can be induced by a single train of electrical stimulation, is dependent on the quantity of repetitions and stable over time (Milner, et al., 1998). Furthermore, it is divided into an early phase lasting 2-3 hours and a more persistent long-lasting LTP of hours or even weeks, which parallels the short- and long-term processes in memory (Lynch, 2004).

An important point is that the molecular mechanisms underlying LTP vary between brain regions and different types of synapses (Cooke & Bliss, 2006; Milner, et al., 1998). LTP as it occurs at synapses in the hippocampus is induced via glutamate binding to postsynaptic α-amino-3-hydroxy-5-methyl-4-isoxazolepropionate (AMPA) receptors. The

following depolarization of the postsynaptic membrane activates *N*-methyl-D-aspartate (NMDA) receptors by removing the Mg^{2+} ion, which under resting conditions blocks ion flux (Figure 2) (Bliss & Collingridge, 1993; Cortes-Mendoza, et al., 2013; Kandel, 2001).

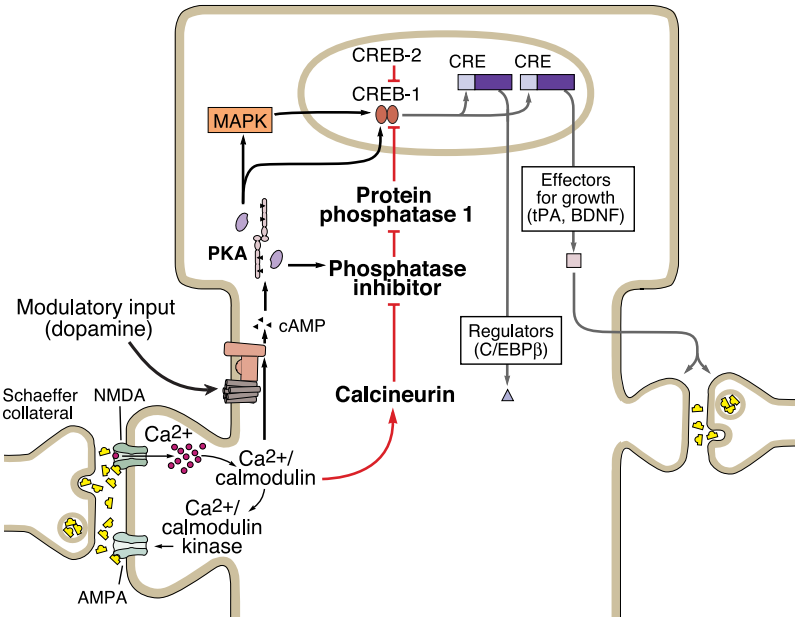


Figure 2. Long lasting long-term potentiation in the hippocampus (Kandel, 2001).

In the early LTP, a single train of action potentials leads then to CA^{2+} influx into the postsynaptic cell and to the activation of CA^{2+}/calmodulin-dependent kinase II (CaMKII), protein kinase C (PKC) and calcineurin (Abel & Lattal, 2001). During long-lasting LTP, when repeated trains of activations occur, CA^{2+} influx activates adenylyl cyclase, leading to increase of cAMP-dependent protein kinase levels, which in turn activates PKA (Abel & Lattal, 2001; Abel et al., 1997; Frey, Huang, & Kandel, 1993; Kandel, 2001). PKA recruits MAPK and both are then transported to the nucleus, where they phosphorylate CREB (Kandel, 2012; Martin et al., 1997). CREB furthermore activates targets (C/EBPB, EPA, BDNF) important for structural changes (Abel & Lattal, 2001; Kandel, 2001).

2.3 Hormonal modulation of memory processes: HPA axis, cortisol and noradrenaline

Among a variety of hormones supposed to affect memory processes (e.g. steroid hormones; (Ackermann et al., 2012)), GC effects on memory formation and retrieval have been demonstrated in various animal and human studies (for reviews see de Quervain, et al., 2009; Wolf, 2009). The stress hormones GCs (cortisol in humans, corticosterone in rodents) are secreted in a 24-h circadian rhythm under basal conditions. In humans, cortisol levels present a morning maximum, decline during the day and night and rise abruptly after the first hours of sleep (Lupien, et al., 2007).

Elevations of GC levels deviating from the normal daily rhythm occur as reaction to situations subjectively perceived as stressful (Krugers, Karst, & Joels, 2012). Two systems are activated by physiological or psychological stressors: The fast acting sympathetic nervous system including the release of the catecholamines adrenaline and noradrenaline (also known as epinephrine and norepinephrine) by the adrenal glands, and the slower HPA axis (de Kloet et al., 2006; Schwabe, et al., 2012). The activation of the HPA axis in response to a stressor is initiated by corticotropin-releasing hormone (CRH), released by the paraventricular nucleus of the hypothalamus. CRH in turn initiates the synthesis and secretion of the adrenocorticotrophic hormone (ACTH) from the anterior pituitary and its release into the bloodstream. ACTH is then transported to the adrenal cortex and induces there the secretion of GCs (Oakley & Cidlowski, 2013; Schwabe, Wolf, & Oitzl, 2010). A negative feedback loop regulates the stress reaction of the HPA axis. GCs inhibit the production and release of CRH and ACTH by a feedback to the hypothalamus and the anterior pituitary (Figure 3). These GC pulses occur in response to some, but not all stressors, resulting in a peak level approximately 20-40 min after the stressor followed by gradual return to baseline (Clements, 2013).

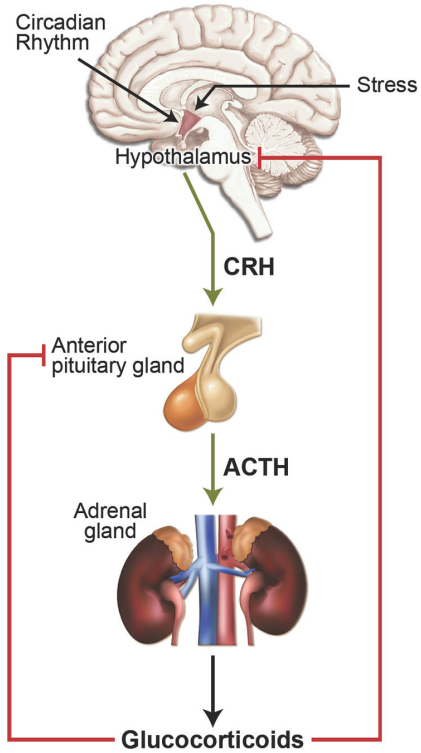


Figure 3. The hypothalamus-pituitary-adrenal (HPA) axis and the feedback loop through glucocorticoids (red line) (Oakley & Cidlowski, 2013).

After release from the adrenal cortex, GCs are able to cross the blood-brain barrier and therefore bind directly to mineralocorticoid receptors (MRs) and glucocorticoid receptors (GRs) in the brain. While MRs are observed most predominantly in limbic structures, GRs are distributed throughout the brain. Both receptors are expressed in the hippocampus, amygdala and prefrontal cortex, regions important for memory processes as mentioned before (Lupien & McEwen, 1997; Reul & de Kloet, 1985; van Ast et al., 2013). GRs have a low affinity to bind GCs and therefore become only occupied when GC levels are high, during stress or at the circadian peak. In contrast, MRs are almost entirely occupied under basal conditions as they have a 10-fold higher affinity for GCs (de Kloet, Oitzl, & Joels, 1999; Reul & de Kloet, 1985). It is supposed that due to their affinity for GCs, MRs are more implicated in the onset of the stress response and GRs are involved in the termination of the stress response and facilitate recovery. Furthermore, GRs are supposed to promote memory storage (de Kloet, Joels, & Holsboer, 2005). Stress and stress hormones have been shown to affect hippocampal plasticity by impairing LTP and facilitate LTD (Artola et al., 2006; Krugers, Goltstein, van der Linden, & Joels, 2006). On the other hand, hormones released during stress can enhance hippocampal synaptic efficiency and thus memory

performance by affecting AMPA receptor functioning (for review see Krugers & Hoogenraad, 2009). Additionally, it has been shown that low to mediate levels of GCs stimulate MRs and enhance hippocampal plasticity. The additional activation of GRs when the perceived stress is high and therefore GC levels increase, results in inhibitory effects on plasticity (de Kloet, et al., 2005; for review see Kim & Diamond, 2002).

As mentioned before, stress not only leads to the release of GCs, but also to secretion of adrenaline and noradrenaline (NA). Evidence shows that emotional arousal leading to NA release is essential for the modulation of stress hormone effects on memory processes (Krugers, et al., 2012). Contrary to GCs, adrenaline is not able to directly cross the blood-brain barrier. Emotional arousing experiences result in a release of adrenaline from the adrenal gland, which then activates vagal afferents to the nucleus of the solitary tract (de Quervain, et al., 2009; Schwabe, et al., 2012). Noradrenergic neurons in the nucleus of the solitary tract then induce the release of NA in the basolateral amygdala (BLA) by projecting directly to the BLA or indirectly via the locus coeruleus. The interaction of GC and NA in the BLA modulates memory processes in the prefrontal cortex, hippocampus, caudate nucleus and in other brain regions (Figure 4) (de Quervain, et al., 2009; Roozendaal, McEwen, & Chattarji, 2009; Schwabe, et al., 2012).

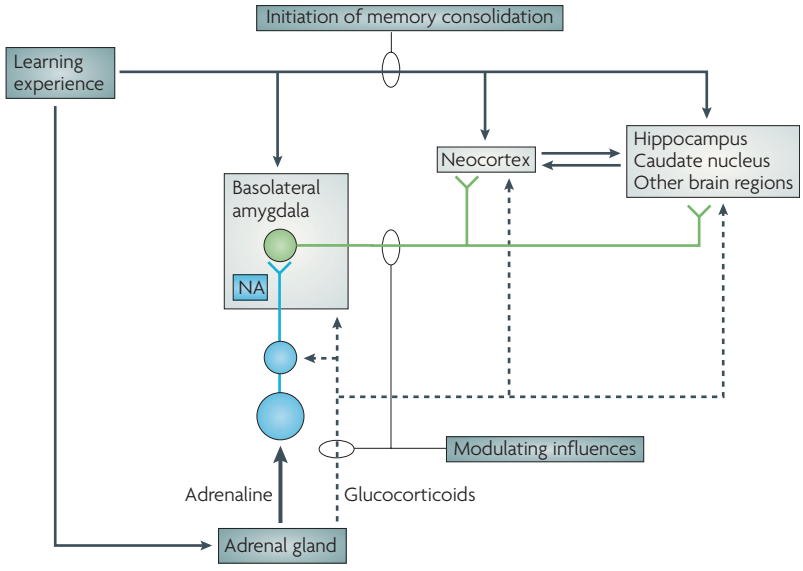


Figure 4. Memory modulation via glucocorticoids and noradrenaline (Roozendaal, et al., 2009).

Taken together it has been shown that stress effects on memory consolidation and retrieval require glucocorticoid and noradrenergic activation in the BLA (Schwabe, et al., 2012).

However it should be taken into account that the time of learning and memory testing in relation to the appearance of the stressor has an important influence on the memory outcome (Figure 5). This could also be an explanation for time-dependent differences in study results. It has been supposed that if learning takes place immediately after or during the exposure to a stressor, high NA levels and GC effects facilitate encoding by facilitating attention and other processes (Buchanan & Lovallo, 2001; Cahill, et al., 2003). In contrast, learning some time after the stressor occured, when hormone levels returned to baseline, supports memory consolidation via delayed, genomic GC actions by suppressing learning of new information (Joels, Pu, Wiegert, Oitzl, & Krugers, 2006; Schwabe, et al., 2012; Zoladz et al., 2011).

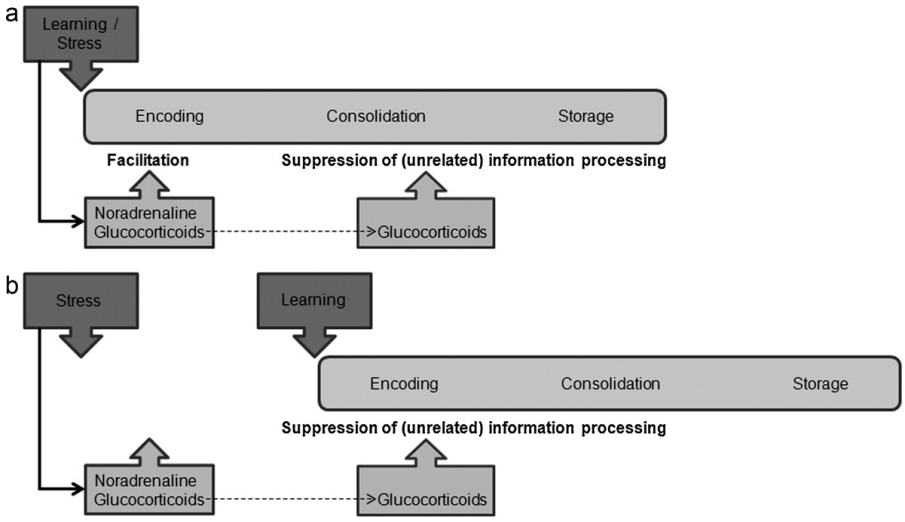


Figure 5. Time-dependent effects of stress on encoding and consolidation, where a) illustrates the case of stress induction shortly before or during learning and b) the case when stress is induced some time before learning occurs (Schwabe, et al., 2012).

2.3.1 Stress, Cortisol and fear memory processes

Fear memory is a form of network that contains information about the feared stimulus itself, about reactions to (verbal, physiological, behavioral) and evaluation of the stimulus (Bentz, et al., 2010; Foa & Kozak, 1986). One of the underlying mechanism of fear acquisition is classical conditioning (Myers & Davis, 2002). Once a fear memory is built, it can be strengthened by retrieval. This reactivation leads to reconsolidation of the memory and thus strengthens the aversive memory trace (Bentz, et al., 2010; Sara, 2000). It is assumed that these mechanisms play as well a part in the development of anxiety disorders

and PTSD (Bouton, Mineka, & Barlow, 2001; Mahan & Ressler, 2012; Mineka & Oehlberg, 2008).

Evidence from animal and human studies using different methods for stress induction or GC administration demonstrate that GCs enhance memory consolidation (Buchanan & Lovallo, 2001; Cahill, et al., 2003; Roozendaal, Okuda, de Quervain, & McGaugh, 2006) and impair long-term memory retrieval (de Quervain, Roozendaal, & McGaugh, 1998; de Quervain, Roozendaal, Nitsch, McGaugh, & Hock, 2000; Kuhlmann, Kirschbaum, & Wolf, 2005; Kuhlmann, Piel, & Wolf, 2005; Smeets, 2011). These findings have been applied to the therapy of anxiety disorders, as GCs could improve the effect of extinction therapy (de Quervain et al., 2011; Soravia et al., 2006; Soravia et al., 2014). GCs are supposed to inhibit retrieval of fear memories and thus interrupt the memory cycle by reducing reconsolidation. On the other side GCs then support the building of an extinction memory by enhancing the consolidation of the newly formed memory (Figure 6) (Bentz, et al., 2010; de Quervain, et al., 2009).

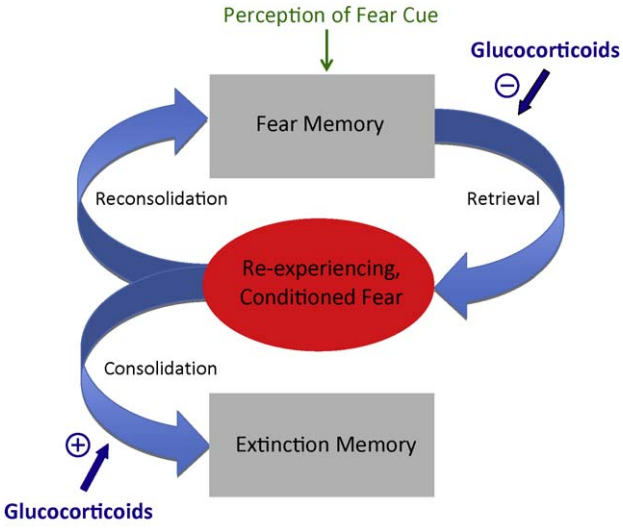


Figure 6. The role of glucocorticoids in the process of fear and extinction memory building (Bentz, et al., 2010).

A few studies investigated GC effects on fear or extinction learning in healthy humans and they reported gender-dependent differences for stress-induced cortisol levels and their effect on fear conditioning processes (Jackson, Payne, Nadel, & Jacobs, 2006; Merz et al., 2010; Zorawski, Cook, Kuhn, & LaBar, 2005). In the study "Influence of stress on fear memory processes in an aversive differential conditioning paradigm in humans" (Bentz, et

al., 2013) we used an aversive conditioning paradigm to induce fear memory in healthy humans and enhanced cortisol levels by stress induction using the CPT. Additionally, we were interested in gender differences.

2.3.2 Basal cortisol levels and episodic memory

Effects of GCs on memory processes depend on several factors as it has been shown before in this thesis, such as the different memory stages (acquisition, consolidation, retrieval). Furthermore they depend on whether the GC elevations are acute or chronic. Acute elevations enhance consolidation and impair retrieval (for reviews see de Quervain, et al., 2009; Schwabe, et al., 2012; Wolf, 2009). Chronically elevated GC levels are usually associated with impaired cognitive performance (de Quervain, et al., 2009; McEwen, 2001; Sapolsky, 2000).

Besides the effect of stress and enhanced GC levels, already basal levels of cortisol can have an impact on memory processes. GC levels are altered or reduced in psychiatric diseases, often accompanied by changes in cognition (Belanoff, Gross, Yager, & Schatzberg, 2001). Cortisol for example has been associated with cognitive deficits in depression (Hinkelmann et al., 2009) and with PTSD, which is characterized by an enhancement of memory retrieval (Mason, Giller, Kosten, Ostroff, & Podd, 1986; Yehuda, 2002; Yehuda & LeDoux, 2007). Only few studies have investigated the effect of natural circadian variation (basal cortisol) on memory in healthy subjects and have not found consistent results (Preuss, Schoofs, & Wolf, 2009; Putman, Van Honk, Kessels, Mulder, & Koppeschaar, 2004; Van Honk et al., 2003). In the study "Associations between basal cortisol levels and memory retrieval in healthy young individuals" (Ackermann, et al., 2013a), we investigated the effect of basal cortisol levels during encoding and retrieval in an episodic memory task. Furthermore, we analyzed in detail whether changes in basal cortisol levels during recall had an influence on memory performance. It has been shown previously that changes in cortisol levels during the time course of the study are associated with cognitive performance (Lee et al., 2007).

2.4 Genetic modulation of memory processes

The increasing knowledge about genetic and epigenetic information can encourage a better understanding of the biological mechanisms that underlie memory processes. In the study "Polymorphisms of HDAC5 are associated with episodic memory, DNA methylation and C17orf65 mRNA expression" (Hartmann et al., unpublished manuscript) we focused on the role that genetic polymorphisms of the *HDAC5* gene may play in episodic memory formation in healthy humans. Episodic memory is a genetically complex trait and twin studies reported heritability values up to 60% (Alarcon, Plomin, Fulker, Corley, & DeFries, 1998; Bouchard, Lykken, McGue, Segal, & Tellegen, 1990; Panizzon et al., 2011; Wilson et al., 2011; for review see Papassotiropoulos & de Quervain, 2011). Johansson et al. (1999) for example included 7 different memory measures and found the highest heritability values for the digit span backward test (49%) and a picture memory test (47%) in a sample of twins aged 80 and older. Varying results of genetic influences on memory can partly be explained by the memory tests used, measuring different components of episodic memory. Furthermore, short-delay recall and long-delay recall, representing temporal components, revealed overlapping as well as distinct genetic influences (Papassotiropoulos & de Quervain, 2011). Swan et al. (1999) report differential contributions of genetic and environmental influences to memory components in a sample of elderly twins. Whereas no evidence of a genetic influence on learning strategy or recognition memory could be found, verbal learning and delayed recall measures loaded on the same factor and revealed a heritability value of 56%. Furthermore, Panizzon et al. (2011) could find a heritability value of 36% for a factor influencing verbal learning, short- and long-delay free recall. However it is unclear whether they share the same underlying genetic factors (Panizzon, et al., 2011). These results of phenotypic variability attributed to heritable factors support the use of genetic studies to identify specific molecules and molecular pathways related to components of episodic memory (Papassotiropoulos & de Quervain, 2011).

Additionally to genetic influences on memory formation, the dynamic interplay between genes and experience is an important issue to examine. These dynamic epigenetic processes, such as DNA methylation and regulation of chromatin structure via histone modifications, affect gene expression independent of the DNA sequence (Mathews & Janusek, 2011; Sweatt, 2013). Several studies demonstrate the important role that DNA methylation and demethylation play in learning and memory formation (for review see

Day & Sweatt, 2011). DNA methylation levels, for example, are rapidly and dynamically regulated in the hippocampus of rats undergoing contextual fear conditioning. Increased methylation of protein phosphatase 1 (PP1) acutely silenced the gene and decreased methylation of reelin resulted in increased production of reelin mRNA. This shows that memory formation is dependent on the activation of some genes and silencing of others, as the reelin gene product promotes synaptic plasticity and long-term memory formation (Weeber et al., 2002) as well as memory consolidation in concert with suppression of PP1 (Miller & Sweatt, 2007). PP1 and reelin are together with bdnf, arc and calcineurin examples of memory-related genes that are epigenetically modified in response to experience (Day & Sweatt, 2011). Furthermore, chromatin modifications have been shown to play an important role in cognitive processes (Penney & Tsai, 2014). Levenson et al. (2004) could show that the formation of long-term fear memories in a contextual fear conditioning paradigm involves increased acetylation of histone H3 in the hippocampus of rats. Latent inhibition, a different form of long-term memory, however, was associated with increased acetylation of histone H4. These results point to the assumption that there might be specific patterns of histone modification for specific types of memory (Levenson & Sweatt, 2005).

Taken together, changes in DNA methylation and histone modifications occur in association with memory formation and in response to neuronal activity. This supports the involvement of an epigenetic code in learning and memory processes (Day & Sweatt, 2011; Miller & Sweatt, 2007) and the potential of epigenetic mechanisms to establish, limit and control neuronal function (Lattal & Wood, 2013).

2.4.1 Histone deacetylase

DNA is wrapped around a histone octamer containing the four core histones (H3, H4, H2A, H2B) and thus forms the nucleosome, a fundamental unit of the chromatin (Figure 7). Histone proteins are composed of a N-terminal tail, which possesses a large number of modification possibilities (Kouzarides, 2007; Penney & Tsai, 2014). Among others, acetylation, methylation and phosphorylation of histones are frequently named (Kouzarides, 2007) and play together with DNA methylation the most important role in chromatin remodeling (Cortes-Mendoza, et al., 2013). The combination of this histone modifications produce changes in gene expression that in turn influence memory formation (Day & Sweatt, 2011).

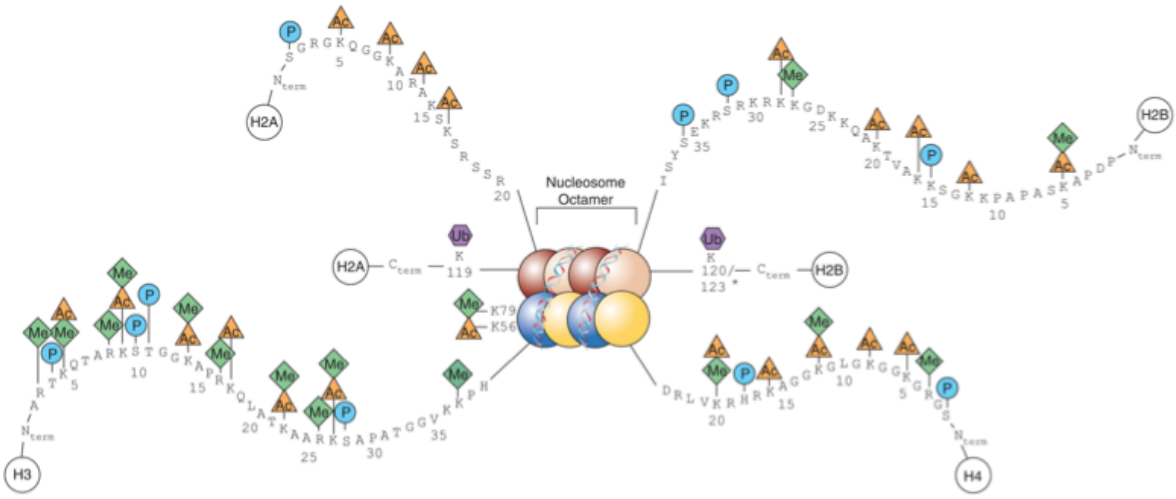


Figure 7. Histone modifications. Nucleosome octamer containing the histones H3, H4, H2A, H2B including acetylation (orange triangles), methylation (green diamonds) and phosphorylation (blue circles) (Keppler & Archer, 2008).

We will focus now on the process of acetylation. Histone acetyltransferases (HATs) add acetyl groups to histone proteins, making the chromatin accessible for transcription factors by relaxing the binding between histone tail and DNA. The more accessible chromatin leads to increased gene expression. Histone deacetylases (HDACs) otherwise remove acetyl groups and therefore are able to suppress gene expression (Day & Sweatt, 2011; Mahgoub & Monteggia, 2014). Acetylation of histones has been shown to play an important role in cognitive processes and in multiple neurological disorders (for review see Penney & Tsai, 2014), as well as in synaptic plasticity (Alarcon et al., 2004; Levenson, et al., 2004; Vecsey et al., 2007). Levenson et al. (2004) for example observed that elevated levels of histone acetylation by the use of HDAC inhibitors enhanced induction of LTP in area CA1 of the hippocampus. Additionally, they observed changes in acetylation of histone H3 in the same area one hour after contextual fear conditioning in rats. Finally, Haettig et al. (2011) could show that HDAC inhibition in mice modulates long-term memory in the hippocampus by using CREB.

2.4.2 Role for HDAC5 in learning and memory

HDAC5 is a class II HDAC and has been associated with synaptic plasticity and memory formation in animal studies (Guan, et al., 2002; Otsuki et al., 2010; Renthal, et al., 2007). Additionally, HDAC5 has been related to behavioral adaptations to emotional stimuli (cocaine addiction and chronic stress) (Renthal, et al., 2007) and to the pathogenesis of

depression in mice (Ookubo, Kanai, Aoki, & Yamada, 2013; Tsankova et al., 2006). Furthermore, HDAC5 expression is observed in brain regions that are important for learning and memory (i.e. hippocampus and amygdala) (Broide, et al., 2007). Importantly, class II HDACs, especially HDAC4 and HDAC5 can modify non-histone proteins by shuttle out of the nucleus in response to neuronal stimuli such as synaptic activity and CA^{2+} influx (Chawla, Vanhoutte, Arnold, Huang, & Bading, 2003; Mahgoub & Monteggia, 2014).

In our study (Hartmann et al., unpublished manuscript) we aimed at investigating the role that HDAC5 plays in episodic memory formation in healthy humans, as up to now the role of HDAC5 in humans is unexplored. In a first step, we conducted a candidate gene approach by analyzing SNPs lying in the *HDAC5* region. Additionally we analyzed the impact of these SNPs on DNA methylation levels in the promoter region of *HDAC5* as well as local and distant effects on mRNA expression levels. Epigenetic mechanisms that directly modulate chromatin structure to regulate gene expression such as DNA methylation have been implicated in neuronal plasticity and memory processes as well (Lattal & Wood, 2013).

3. Methods

3.1 Salivary cortisol

Measuring salivary cortisol is a common method to assess HPA axis function in psychoneuroendocrine research that has also been employed in our two studies (Ackermann, et al., 2013a; Bentz, et al., 2013). Cortisol is an end product of the HPA axis and circulates in blood. After secretion, a majority of the cortisol is bound to corticosteroid-binding globulin (CBG) and albumin. Only a small fraction is unbound and this freely circulating fraction is assumed to be the biological active form and to elicit the glucocorticoid effects on physiology and behavior. In blood, the fraction of bound and free cortisol can be measured. Free cortisol can enter saliva due to its low molecular weight through passive diffusion and therefore only the free fraction appears in saliva (Clements, 2013; Kirschbaum & Hellhammer, 1994; Levine, Zagoory-Sharon, Feldman, Lewis, & Weller, 2007).

Some advantages of measuring cortisol in saliva are that it is non-invasive, easy to collect and the compliance of the subjects is higher than for blood drawing. Furthermore, cortisol has been shown to be stable at room temperature. One advantage especially in view to stress research is that it causes no additional stress, as it could be the case in blood drawing (Clements & Parker, 1998; Kirschbaum & Hellhammer, 1994; Levine, et al., 2007).

Cortisol levels can be influenced by different factors and these should be taken into account when planning a study. Among others, age, gender (menstrual cycle, hormonal contraceptive use), diurnal rhythm of cortisol levels and habituation to the test situation can influence outcome measures (Clements, 2013; Kudielka, Hellhammer, & Wust, 2009; Kudielka & Kirschbaum, 2005). In our studies, we controlled for gender and use of hormonal contraceptives (Ackermann, et al., 2013a) or did our calculations for both sexes separately and included only women taking oral contraceptives, to avoid effects of the menstrual cycle (Bentz, et al., 2013). It has been shown that adult males usually show a larger increase in cortisol after stress induction compared to women (Kudielka, et al., 2009) this was likewise detected in our study (Bentz, et al., 2013).

3.2 Stress induction: the cold pressor test

Different methods have been developed to experimentally manipulate cortisol levels in subjects. The advantage of experimental stress induction in comparison to naturally occurring stress is that the situations can be standardized and there is the possibility to control for confounding variables (Kudielka, et al., 2009). Beside pharmacological manipulation (e.g. cortisone administration)(de Quervain, et al., 2000), various stress tests that stimulate the HPA axis and the sympathetic nervous system can be used to induce stress and therefore enhance cortisol levels. Two frequently used tests are the Trier social stress test (TSST) (Kirschbaum, Pirke, & Hellhammer, 1993) and the CPT (Lovallo, 1975). The TSST possesses important components for a situation to be perceived as stressful, such as high degree of ego involvement and anticipation of negative consequences. Subjects have to give a pretended job interview in front of a committee, which shows no emotional reaction to the subject (Kirschbaum, et al., 1993).

In our study (Bentz, et al., 2013) we used the CPT for stress induction. During the CPT, subject were instructed to immerse their arm up to the elbow into cold water (0-4 °C) for 3 min, or into warm water in the control condition (37-40 °C). The manipulation with the cold water evokes endogenous cortisol release and stimulates sympathetic activation (Lovallo, 1975). The sympathetic activation can be measured for example in an elevation of blood pressure (al'Absi, Petersen, & Wittmers, 2002) or skin conductance response (Buchanan, Tranel, & Adolphs, 2006).

For both methods the response magnitude can be variable between individuals as well as genders. This is also the case in the study from Bentz et al. (2013), where we could find clear effects only in males and we had also to exclude subjects who didn't show cortisol enhancement after the CPT. Among others, the evaluation of the stressor by the subject is important. Further features to turn a situation into a stressor are the novelty of the situation, as well as the feeling of uncontrollability (Dickerson & Kemeny, 2004). Schwabe et al. (2008) added to the CPT a socially evaluative component to enhance the subjective emotional component and therefore the HPA axis reaction. The socially evaluated cold pressor test (SECPT) expands the CPT by the component that subjects are watched by an investigator and videotaped during hand immersion (Schwabe, et al., 2008).

3.3 Behavioral genetics and epigenetics

Episodic memory is a complex polygenic trait (Papassotiropoulos & de Quervain, 2011). Genome-wide association studies (GWAS) and candidate gene studies implicated several genes to be associated with episodic memory such as *KIBRA* (Milnik et al., 2012; Papassotiropoulos et al., 2006), *CTNNBL1* (Papassotiropoulos et al., 2013b), *CPEB3* (Vogler et al., 2009), and *PKCα* (de Quervain et al., 2012).

Performing a GWAS is a useful method to discover novel genes and molecular pathways involved in polygenic traits as it allows the identification of associations between phenotypes and millions of genetic markers over the entire genome (Papassotiropoulos & de Quervain, 2011). One limitation of this method, however, is that due to the need for correction for multiple testing some genes may not reach genome-wide significance, although they may play a role in the generation of the phenotype (Frazer, Murray, Schork, & Topol, 2009). Knowledge about biological mechanisms underlying a trait and existing findings in animal and human studies allow the use of a candidate gene approach instead of a GWAS (Papassotiropoulos & de Quervain, 2011). Based on prior findings in animal studies, we used in our study (Hartmann et al., unpublished manuscript) the hypothesis confirming candidate gene approach by analyzing associations between SNPs tagging *HDAC5* and episodic memory. Besides SNPs, where one base at a locus in the DNA is changed, copy number variations can as well contribute to the genetic basis of phenotypic variation (Vogler et al., 2010).

In a second step, we analyzed more dynamic processes, namely DNA methylation and mRNA expression. Epigenetic changes, such as DNA methylation are enduring changes in gene expression without involvement of changes in the DNA sequence (Strachan & Read, 2011). They are supposed to explain additional phenotypic variance complementing the fraction attributed to genetic polymorphisms. Epigenetic information is susceptible to stress and other environmental influences and thus it is a result of the interaction between genome and environment (Levenson & Sweatt, 2005; Mathews & Janusek, 2011). Furthermore it can be passed on to subsequent generations (Bohacek & Mansuy, 2013).

DNA methylation occurs on cytosine residues and most often on cytosine-guanine dinucleotides (CpG) by addition of a methyl group (Griffiths & Hunter, 2014; Kristensen, Mikeska, Krypuy, & Dobrovic, 2008). This leads to structural changes of chromatin and to

downregulation of transcription and thus to reduced gene expression (Levenson & Sweatt, 2005; Miller & Sweatt, 2007).

Epigenetic mechanisms have been shown to be involved in learning and memory and to be important for the regulation of learning-dependent synaptic plasticity (Jarome & Lubin, 2014; Miller & Sweatt, 2007).

4. Original Research Papers

4.1. Influence of stress on fear memory processes in an aversive differential conditioning paradigm in humans

Bentz, D., Michael, T., Wilhelm, F.H., Hartmann, F.R., Kunz, S., von Rohr, I.R., & de Quervain, D.J.-F. (2013). *Psychoneuroendocrinology, 38*(7), 1186-1197.

Psychoneuroendocrinology (2013) 38, 1186—1197

Available online at www.sciencedirect.com

SciVerse ScienceDirect

journal homepage: www.elsevier.com/locate/psyneuen

Influence of stress on fear memory processes in an aversive differential conditioning paradigm in humans

Dorothée Bentz [a,*], Tanja Michael [b], Frank H. Wilhelm [c], Francina R. Hartmann [d], Sabrina Kunz [a], Isabelle R. Rudolf von Rohr [a], Dominique J.-F. de Quervain [a,e]

[a] University of Basel, Department of Psychology, Division of Cognitive Neuroscience, Birmannsgasse 8, 4055 Basel, Switzerland
[b] Saarland University, Department of Clinical Psychology und Psychotherapy, Campus, Gebäude A1 3, 66123 Saarbrücken, Germany
[c] University of Salzburg, Institute of Psychology, Division of Clinical Psychology, Psychotherapy, and Health Psychology, Hellbrunner Strasse 34, 5020 Salzburg, Austria
[d] University of Basel, Department of Psychology, Division of Molecular Neuroscience, Birmannsgasse 8, 4055 Basel, Switzerland
[e] University of Basel, Psychiatric University Clinic, Wilhelm Klein-Strasse 27, 4055 Basel, Switzerland

Received 18 June 2012; received in revised form 14 December 2012; accepted 19 December 2012

KEYWORDS
Fear conditioning;
Retrieval;
Emotional memory;
Emotional learning;
Acute stress;
Explicit memory;
Glucocorticoids

Summary It is widely assumed that learning and memory processes play an important role in the pathogenesis, expression, maintenance and therapy of anxiety disorders, such as phobias or post-traumatic stress disorder (PTSD). Memory retrieval is involved in symptom expression and maintenance of these disorders, while memory extinction is believed to be the underlying mechanism of behavioral exposure therapy of anxiety disorders. There is abundant evidence that stress and stress hormones can reduce memory retrieval of emotional information, whereas they enhance memory consolidation of extinction training. In this study we aimed at investigating if stress affects these memory processes in a fear conditioning paradigm in healthy human subjects. On day 1, fear memory was acquired through a standard differential fear conditioning procedure. On day 2 (24 h after fear acquisition), participants either underwent a stressful cold pressor test (CPT) or a control condition, 20 min before memory retrieval testing and extinction training. Possible prolonged effects of the stress manipulation were investigated on day 3 (48 h after fear acquisition), when memory retrieval and extinction were tested again. On day 2, men in the stress group showed a robust cortisol response to stress and showed lower unconditioned stimulus (US) expectancy ratings than men in the control group. This reduction in fear memory retrieval was maintained on day 3. In women, who showed a significantly smaller cortisol response to stress than men, no stress effects on fear memory retrieval were observed. No group differences were observed with respect to extinction. In conclusion, the present study provides evidence that stress can reduce memory retrieval of conditioned fear in men. Our findings may contribute to the understanding of the effects of stress and glucocorticoids on fear symptoms in anxiety disorders and suggest that such effects may be sex-specific.

* Corresponding author. Tel.: +41 61 267 02 28.
E-mail address: dorothee.bentz@unibas.ch (D. Bentz).

0306-4530/$ — see front matter
http://dx.doi.org/10.1016/j.psyneuen.2012.12.018

1. Introduction

Classical fear conditioning is a well-established paradigm to investigate emotional learning and memory processes thought to be involved in the pathogenesis and symptomatology of anxiety disorders (LaBar and Cabeza, 2006). Within the conditioning framework, anxiety disorders like panic disorder, phobias or post-traumatic stress disorder (PTSD) are considered as manifestations of conditioned fears (Bouton et al., 2001; Mineka and Oehlberg, 2008; Mahan and Ressler, 2012). Fear conditioning refers to an association of an originally neutral stimulus (conditioned stimulus, CS) with an aversive or traumatic experience (unconditioned stimulus, US). After fear acquisition, a later encounter with the CS leads to retrieval of the associated fear memory and to subsequent fear response (conditioned reaction, CR) and to further stabilization of the fear memory trace (Myers and Davis, 2002; Phelps and LeDoux, 2005; Bentz et al., 2010). Fear conditioning is usually sensitive to extinction, a form of learning characterized by a decrease in the amplitude and frequency of the CR when the CS is repeatedly presented without the US (Myers and Davis, 2002; Hermans et al., 2006). The inability to extinguish or inhibit maladaptive fear responses is a characteristic for most anxiety disorders (Hermans et al., 2006; Bentz et al., 2010). Experimental studies indicate that patients with anxiety disorders have deficits in extinction learning compared to healthy controls (Blechert et al., 2007; Michael et al., 2007).

Most fear conditioning studies in healthy humans and patients with anxiety disorders typically measure skin conductance responses (SCR) or fear potentiated startle to indicate fear learning and extinction. Both measures quantify implicit (non-conscious, unintentional) aspects of fear memory rather than explicit (conscious, intentional) memory processes that also operate in fear learning and expression (LaBar and Cabeza, 2006). During fear acquisition participants continuously develop expectancies about stimulus associations (Lovibond, 2006). Moreover as a result of repeated US—CS pairings, an affective valence is transferred from the US to the CS (evaluative conditioning) (de Houwer et al., 2001). Expectancy and evaluative learning have emotional components, but are also associated with explicit memory processes and can be verbalized and measured online in humans throughout conditioning paradigms (Carter et al., 2006). Neuropsychological findings from human lesion- and imaging studies indicate that explicit and implicit memory processes involved in fear conditioning are neurally dissociable and can be assigned to different brain regions, the hippocampus and the amygdala, respectively (Bechara et al., 1995; Knight et al., 2004). Further, there is evidence that patients with anxiety disorders have deficits in fear extinction of conditioned responses indicated by measures associated with both hippocampal and amygdaloidal memory processes (Blechert et al., 2007; Michael et al., 2007).

Evidence from animal and human studies indicates that stress and stress hormones influence learning and memory processes. The effect of glucocorticoids (GCs) on memory processes depends on several factors, such as the stage of memory process (acquisition, consolidation, or retrieval), emotional characteristics of the material and sex (de Quervain et al., 2009). There is abundant evidence that acute administration of GCs enhances memory consolidation (Roozendaal, 2000; Buchanan and Lovallo, 2001). Additionally to the enhancing effects of GCs on memory consolidation, GCs impair long-term memory retrieval processes (de Quervain et al., 1998, 2000; Roozendaal et al., 2004; Het et al., 2005; Kuhlmann et al., 2005a,b; Kuhlmann and Wolf, 2005; Buchanan et al., 2006). Especially emotionally arousing memory contents seem to be sensitive to the retrieval-impairing effects of GCs (Kuhlmann et al., 2005a,b; de Quervain et al., 2007). Therefore GCs might enhance extinction learning by inhibiting fear retrieval processes and promoting consolidation of extinction learning. There is preliminary evidence from studies with patients with different anxiety disorders (social phobia, specific phobia, PTSD) for the extinction enhancing effect of GCs. These studies found, besides lower stimulus-associated fear under GC treatment, prolonged effects outlasting the treatment period that may indicate enhanced extinction of fear (Aerni et al., 2004; Soravia et al., 2006; de Quervain et al., 2011). This is in line with animal studies that showed that GCs facilitate the consolidation of extinction memory, whereas suppression of GCs impairs extinction processes (Cai et al., 2006; Yang et al., 2006; Blundell et al., 2011). So far only a few studies investigated the effect of GCs or stress on fear and extinction learning in healthy humans (Zorawski et al., 2005; Grillon et al., 2006; Jackson et al., 2006; Stark et al., 2006; Zorawski et al., 2006; Nees et al., 2008; Luethi et al., 2009; Wolf et al., 2009; Kuehl et al., 2010; Merz et al., 2010, 2012a,b, 2013; Tabbert et al., 2010). Most studies used exclusively implicit measurements such as SCR (Zorawski et al., 2005; Jackson et al., 2006; Stark et al., 2006; Merz et al., 2010, 2012a,b, 2013; Tabbert et al., 2010), fear potentiated startle (Grillon et al., 2006; Nees et al., 2008; Wolf et al., 2009; Kuehl et al., 2010) and BOLD (blood oxygenation level dependent)-contrasts (Stark et al., 2006; Merz et al., 2010, 2012a,b, 2013; Tabbert et al., 2010) to identify fear acquisition and extinction. These studies indicate that stress-induced elevations of cortisol levels affect fear conditioning processes with different effects in men and women (interaction between stress and sex). Some studies utilizing SCR as indicator for fear acquisition show that high endogenous (Zorawski et al., 2005) and stress-induced cortisol levels (Jackson et al., 2006) seem to be associated with enhanced fear acquisition in men, but not women. fMRI studies that measured the influence of exogenously administered cortisol on neural correlates of fear acquisition show the opposite pattern with enhanced fear acquisition in women and impaired fear acquisition in men (Stark et al., 2006; Merz et al., 2010). However, there is no study so far that investigated the specific effects of stress on memory retrieval and extinction processes of conditioned fear (i.e. without affecting initial memory acquisition or consolidation processes).

In the present study we aimed at investigating the effects of stress on memory retrieval and extinction processes in a fear conditioning paradigm in healthy male and female participants. After fear memory has been acquired in a differential conditioning paradigm on day 1, cold pressor test (CPT) was used to induce stress before retrieval (as measured with the first extinction trial on day 2) and extinction training on day 2 (24 h after fear acquisition). Possible prolonged effects of the stress manipulation were investigated on day 3 (48 h after fear acquisition). CPT consisted of

arm immersion in cold water, and the control condition of arm immersion in warm water. Fear conditioning, retrieval and extinction were quantified by measures of implicit memory and explicit memory processes (SCR, US-expectancy and valence). Our hypothesis was that the stress manipulation on day 2 would lead to a reduction in memory retrieval (measured with the first extinction trial on day 2) and an enhancement in extinction consolidation processes. Based on previous findings for stress effects on fear learning and memory, we expected that sex may play a role in our paradigm as well (Zorawski et al., 2005; Jackson et al., 2006; Stark et al., 2006; Merz et al., 2010, 2012a,b, 2013).

2. Methods

2.1. Subjects

Participants were aged between 18 and 38 years and physically and mentally healthy. Exclusion criteria were: pregnancy, lactation, cardiovascular disorders, skin diseases, current psychopathology, substance abuse (including habitual smoking), and treatment with psychotropic medication, beta-blockers or steroids. We screened study eligibility over the phone via an interview including a demographic and physical health section and a section based on the Structured Clinical Interview for DSM-IV (Wittchen et al., 1997) to assess psychopathology. We measured all exclusion criteria again by means of questionnaires before study entry on day 1. At that time point we assessed depressive symptoms with a standardized depression scale in German (*Allgemeine Depressionsskala*/ADS, Hautzinger and Bailer, 1993); questionnaire scores had to be within a non-clinical range ($\leq$23). We used several questionnaires to characterize the study population concerning demography, trait and state anxiety (*State-Trait-Anxiety-Inventory*/STAI, German version, Laux et al., 1981), anxiety sensitivity (*Anxiety Sensitivity Index*/ASI, German version; Ehlers and Margraf., 1993) and trait and state emotionality (*Positive Affect and Negative Affect Schedule/PANAS*, German version, Krohne et al., 1996). Additionally the state version of the STAI and the PANAS were given once again on day 2 before start of testing. To avoid effects of menstrual cycle on outcome measures, only women taking monophasic oral contraceptives (OCs) with an ethinylestradiol component for at least the last three months were included; testing days were scheduled during the OC-intake phase. Pregnancy was excluded via a pregnancy test before study entry.

108 participants (51 men, 57 women) first entered the study (after the initial phone screening) and were randomized within sex to study groups. 14 participants (9 men, 5 women) had to be excluded during the experiment, because of technical problems (e.g. computer crash during stimulus presentation on one of the testing days) and 12 participants (6 men, 6 women) because of other reasons (e.g. substance abuse, high baseline blood pressure), resulting in 82 participants (36 men, 46 women). To analyze CPT-induced stress effects only those participants who showed a CPT-induced cortisol rise on day 2 (CPT-induced cortisol rise on day 2 was defined as cortisol level after CPT day 2 $-$ cortisol baseline level before CPT day 2 $\geq$ 0) were included in the stress group. In the control group only those participants who did not show any cortisol rise after the CPT control condition (CPT-C) (no CPT-induced cortisol rise on day 2 was defined as cortisol level after CPT-C day 2 $-$ cortisol baseline level before CPT-C day 2 $<$ 0) were included. 29 participants were excluded (6 men and 4 women from the stress group, 6 men and 13 women from the control group) resulting in 35 participants (13 men, 22 women) in the stress group and 18 participants (11 men, 7 women) in the control group. Participants excluded from the stress group did not differ from participants included in the stress group concerning any participant characteristics, demographic or psychometric measures (all *p*s $>$ .29). Participants excluded from the control group did show higher trait anxiety scores measured with the STAI-trait version, experienced the CPT-C as slightly more unpleasant and were slightly younger compared to participants included in the control group (all *p*s $<$ .04).

2.2. Conditioning procedure

The conditioning procedure builds on a conditioning paradigm from Michael et al. (2007), but the different conditioning phases are distributed over three days with habituation (HAB) and acquisition (ACQ) on day 1 and two consecutive extinction phases on day 2 (EXT I) and day 3 (EXT II). Each phase consisted of 5 CS-PLUS (neutral stimulus paired with US during ACQ) and 5 CS-MINUS (neutral stimulus not paired with US) trials. The CS-PLUS trials were only reinforced during ACQ on day 1, each CS-PLUS offset was immediately followed by US presentation. EXT I and EXT II consisted of a series of 5 unreinforced CS-PLUS and 5 CS-MINUS trials. CS duration was 8 s, during ACQ, the intertrial interval (ITI) was 18 $\pm$ 3 s (determined at random, counterbalanced over all trials). Four different pseudorandomized stimulus orders existed comprising the following restrictions: no more than two consecutive presentations of the same CS. ACQ had to start with a CS-PLUS in every stimulus order.

HAB on day 1 started with the instruction that two pictures (CS-PLUS, CS-MINUS) would be presented repeatedly, and one of the pictures would be occasionally accompanied by an electrical stimulation (US). The HAB phase was directly followed by the ACQ phase without further instructions concerning pairing of CS and US. No further instructions concerning pairing of CS and US were given on day 2 or day 3.

Two simple geometric figures (a triangle and a square) were used as CS. Both figures were gray in color and had an identical luminescence. The CS' sizes were 14 cm $\times$ 14 cm, with a resolution of 640 $\times$ 480 pixel; CS were presented in the midst of a white screen, 36.5 cm wide and 27.3 cm high. The triangle was used as CS-PLUS for half of the participants and CS-MINUS for the other half (counterbalanced over all participants).

An electrical stimulation was used as US (applied for 500 ms). On day 1, the participants adjusted the intensity of the electrical stimulation together with the experimenter to a level (four possible levels from 1 to 10 mA) that they described as being ''unpleasant and demanding some effort to tolerate''. The subjective unpleasantness of the US was measured on day 1 before HAB by means of a 100 mm visual analog scale (VAS) presented on the screen (anchors: ''slightly unpleasant'' to ''very unpleasant'', 0–100). Selected US-intensity on day 1 was maintained throughout the 3 study days.

US-expectancy (EXP) ratings were obtained immediately after each CS-PLUS/CS-MINUS presentation by means of a 100 mm VAS (''How much do you believe that the picture will be followed by an electrical stimulation?'', anchors: ''not at all'' and ''very much'', 0—100) presented on the monitor, resulting in altogether 20 EXP ratings for CS-PLUS/CS-MINUS respectively. The VAS appeared below the CS with concomitant CS presentation 8 s after CS-onset. At trial 1, 3, and 5 of each conditioning phase valence (VAL) and anxiety (ANX) ratings were obtained for CS-PLUS/CS-MINUS, resulting in 12 VAL and 12 ANX ratings for CS-PLUS/CS-MINUS respectively. A 100 mm VAS for VAL (''Please indicate how you perceived the last picture'', anchors: ''pleasant'' and ''unpleasant'', 0—100) and one for ANX (''The picture made me feel in the following way:'', anchors: ''not at all anxious'' and ''anxious'', 0—100) were presented on the monitor during the ITI following the CS-PLUS/CS-MINUS presentation. Participants had 5 s to answer each VAS, before the conditioning procedure proceeded automatically. Completion before the end of the 5 s commenced the ITI automatically. The participants were accustomed to the rating procedure during an initial trial run before beginning of HAB on day 1. At the end of ACQ on day 1 contingency awareness was assessed by presentation of CS-PLUS, CS-MINUS and a distractor figure (a gray circle of CS-size) with the question which one of the three pictures had been paired with the ''electrical stimulation'' on day 1.

2.3. Cold pressor test (CPT)

The CPT is a widely used, reliable, low-risk technique to stimulate sympathetic activation and endogenous GC release (Lovallo, 1975). During the CPT participants were requested to immerse their right arm up to the elbow in cold water (0—4 °C) for 3 min. Participants were instructed to keep their arm in the cold water as long as possible. They were allowed to remove their arm, if the pain became intolerable for them. Participants were told that the procedure implied no health risk, but is expected to be uncomfortable to painful. During the test, the experimenter measured the time and instructed the participants to remove their arm after 3 min arm immersion. After arm removal, each participant rested for 3 min whilst his/her arm was covered with a blanket. Afterwards participants were requested to indicate on a VAS how painful the CPT had been (anchors: ''not very unpleasant'', ''extremely unpleasant'', 0—100). The procedure of the CPT-C was the same except that the participants were requested to immerse their hand in warm water (37—40 °C) instead of cold water for 3 min. Additionally they were told that no health risk was implied and that no pain was expected.

Before starting the CPT we measured the participants baseline blood pressure (BP) with a non-invasive method, a CRITIKON Dinamap 1846 SX Vital Signs Monitor (220 V 50/60 Hz). Anyone with a baseline BP higher than 140/90 mmHg was excluded from the study. During the experiment BP was monitored 30 s, 3 min and 8 min after the beginning of the CPT. BP above 180/110 mmHg or below 100/60 mmHg, or a pulse below 50 during CPT lead to an abortion of the CPT.

2.4. Experimental procedure

The experiment took place in a temperature-controlled (21—23 °C), sound-attenuated room that was connected to an adjoining control room of the psychophysiological laboratory of the University of Basel. During the experiment participants were seated in a comfortable armchair placed one meter in front of a 19-in. monitor. To control for the diurnal cycle of cortisol all experiments were carried out between 1000 h and 1800 h. All participants were instructed to eat a small meal before coming to the laboratory. Additionally, all participants received a sweet drink (consisting of 120 ml syrup solved in water equaling 100 g sugar) at the beginning of each testing in an attempt to equalize glucose levels of all participants. Each participant was tested at the same time for three consecutive days to reduce the impact of diurnal variation on stress hormones (Het et al., 2005). The conditioning paradigm was distributed over 3 days with HAB and ACQ on day 1 and 24-h-delay-EXT (EXT I) on day 2 (with or without stress manipulation before EXT I according to group assignment) and 48-h-delay-EXT (EXT II) on day 3 (without any stress manipulation for either study group) (see Table 1).

On day 1 participants gave informed consent, filled out questionnaires, and baseline saliva samples and BP measurements were taken. Then electrodes for skin conductance measurements as well as for US application were attached and the US intensity was adjusted. After US intensity adjustment participants were instructed that from now on all instructions concerning the conditioning experiment would be presented in writing on the computer screen and the experimenter left the experimental room. On all three days the conditioning experiment began with a physiological baseline measurement while participants sat quietly looking at images of different landscapes (10 s each). On day 1 the conditioning procedure was directly followed by the CPT-C. All participants had to pass through the CPT-C because the experiment was part of a larger study that also looked at stress effects on fear memory consolidation (paper in preparation). On day 2 participants gave a saliva sample, filled out questionnaires and baseline BP was measured before they started with the CPT/CPT-C. EXT I began exactly 20 min after

Table 1 Overview of the conditioning paradigm distributed over three days.

	Phase	Stimuli
Day 1	Habituation	5 CS-PLUS (no US)/ 5 CS-MINUS (no US)
	Acquisition	5 CS-PLUS (paired with US)/ 5 CS-MINUS (no US)
	CPT control condition	
Day 2	CPT/CPT control condition	
	Extinction I	5 CS-PLUS (no US)/ 5 CS-MINUS (no US)
Day 3	Extinction II	5 CS-PLUS (no US)/ 5 CS-MINUS (no US)

US, unconditioned stimulus.

the start of the CPT/CPT-C, when peak cortisol responses were expected (Kirschbaum et al., 1999; Buchanan et al., 2006; Dickerson and Kemeny, 2004). Between the end of CPT/CPT-C and the beginning of EXT I the electrodes for skin conductance measurements and for US application were attached. Despite no electrical stimulation on day 2 and 3 US-intensity was adjusted to the intensity selected on day 1. On day 3 there was no stress manipulation. Participants started directly with EXT II after they gave their saliva baseline sample, filled out questionnaires, BP was measured and the electrodes for skin conductance measurements and for US application were attached.

2.5. Saliva measurements

Saliva was collected with Salivette (Sarstedt, Rommelsdorf, Germany). On day 1, saliva samples were taken directly before HAB/ACQ (baseline), after HAB/ACQ before the CPT-C and 20 min after beginning of the CPT-C. On day 2, saliva samples were taken directly before CPT/CPT-C (baseline), 20 min after beginning of CPT/CPT-C and after the end of the EXT I. On day 3, saliva samples were taken directly before EXT II (baseline) and after the end of EXT II. The saliva samples were stored at $-20\,^{\circ}C$ until biochemical analysis. Free cortisol in saliva was analyzed by using commercially available immunoassay. The inter- and intraassay coefficients of variation were $<10\%$. To reduce error variance caused by imprecision of the intraassay, all samples of one subject were analyzed in the same run.

2.6. Apparatus and physiological recordings

An electrical stimulator (constant current unit between 1 mA and 10 mA, STMISOC, BIOPAC Systems, Inc., Goleta, CA, USA) was used to deliver the US via 18-mm inner diameter Ag/AgCl electrodes (Red Dot™) on the right lower arm. Stimulus delivery, and physiological data acquisition were controlled by two PCs running E-Prime 2.0 (Psychology Software Tools, Inc., Pittsburgh, PA, USA) and Acq*Knowledge*® software (Version 3.81, Biopac Systems, Inc., Goleta, CA, USA). Subjective measurements were recorded by means of E-Prime. Physiological channels were recorded at a rate of 1000 Hz in continuous mode using Biopac MP150 system. From the ECG lead II, heart period (HP) was calculated as the interval in milliseconds between successive R-waves. For ease of interpretation HP was for converted to heart rate (HR). Skin conductance was obtained using 8-mm inner diameter Ag/AgCl electrodes (Red Dot™) filled with isotonic electrode paste (EDA-Paste TD-246, PAR Medizintechnik GmbH) (Fowles et al., 1981). Electrodes were placed on the middle phalanx of the index and middle fingers of the left hand and secured with normal tape and double-sided adhesive collars, which also served the purpose of helping to control the size of the skin area that comes in contact with the electrode paste as this affects the conductance value (Dawson et al., 2007). The constant voltage between the electrodes was 0.5 V. Two channels were obtained as control measures: (1) body movement was assessed using an accelerometer attached to the right shoulder and (2) respiration patterns were recorded using a pneumographic belt placed at the lower thorax/upper abdomen, since movement and respiration irregularities may trigger spurious SCRs.

2.7. Data preparation

Questionnaire data was entered, physiological recordings scored and subjective ratings were exported from E-Prime into SPSS statistics package for Macintosh (IBM SPSS, 20.0) by research assistants blind to condition. The skin conductance data was first corrected for any artifacts with the computer program ANSLAB 2.5 (Autonomic Nervous System Laboratory; Wilhelm and Peyk, 2005). This included careful visual inspection and the manual exclusion of SCRs that appeared to be influenced by movement, deep breaths, coughs, or sighs. SCR was calculated by subtracting the average skin conductance level (SCL) for the two seconds immediately before CS onset from the maximum SCL recorded between four seconds after CS onset and CS offset (four seconds duration, second interval response). Conditioning effects were based on the second interval response, as it is relatively unaffected by non-associative processes like dishabituation of the orienting response and is considered to represent the signal value of the CS (Lovibond, 1992; Michael et al., 2007). To approximate a normalized distribution, a Square Root (SQRT)-transformation was conducted. Since negative values cannot be SQRT-transformed, 1.5 was added to each SCR (transformed value = SQRT (SCR + 1.5)) (Dawson et al., 2007). For all outcome measurements of the differential conditioning procedure difference scores (DIFF $X_{phase,\ trial\ x}$) between the respective CS-PLUS and the CS-MINUS for each trial (DIFF $X_{phase,\ trial\ x}$; X specifies the respective outcome measure, $_{phase}$ the respective phase of the experiment, $_{x}$ the trial of the respective phase) were calculated with DIFF $X_{phase,\ trial\ x} > 0$ indicating conditioning effects.

2.8. Statistical analysis

Group differences in demographic and other participant and treatment characteristics (CPT/CPT-C temperature, US intensity) were analyzed with univariate ANOVAs with sex and group as between subject factors or chi-square tests. All analyses from day 2 on were calculated a priori within sex groups, because we hypothesized that the influence of stress on fear acquisition, retrieval and extinction would be modulated by sex (Zorawski et al., 2005; Jackson et al., 2006; Stark et al., 2006; Merz et al., 2010, 2012a,b, 2013). Additionally, all analyses were calculated with time of testing as a covariate to exclude the influence of diurnal variation on our results structure. Divergent results will be reported.

Furthermore cortisol measurements were calculated within sex groups with univariate ANOVAs with group as between subject factor, because sex differences in GC stress response are often reported in humans (Kajantie and Phillips, 2006; Cornelisse et al., 2011) with men showing higher cortisol stress responses (Kirschbaum et al., 1992, 1999), especially when women are taking OCs (Kirschbaum et al., 1999; Cornelisse et al., 2011). To quantify the CPT-induced cortisol rise on day 2, a difference score by means of the formula $CPT_{difference\ score}$ = cortisol level 20 min after CPT day 2 − cortisol level before CPT day 2 was calculated.

Separate analyses were conducted for each outcome measure (subjective measures: EXP, VAL, ANX; physiological measures: SCR, HR) and each conditioning phase (HAB, ACQ, EXT I and EXT II).

HAB/ACQ: To analyze if habituation and acquisition occurred and to control for pre-experimental group-differences repeated measures ANOVAs with trial (DIFF $X_{HAB,\ trial\ 5}$ and DIFF $X_{ACQ,\ trial\ 5}$) as repeated measures factor and group (stress, control) and sex (male, female) as between subject factor were calculated.

Retrieval I: To analyze CPT-induced stress effects on fear memory retrieval at EXT I univariate ANCOVAs with DIFF $X_{EXT\ I,\ trial\ 1}$ as dependent variable, DIFF $X_{ACQ\ I,\ trial\ 5}$ as covariate and group (stress, control) as between subject factor were calculated.

EXT I: To analyze if extinction occurred and to analyze CPT-induced stress effects on extinction repeated measures ANOVAs with trial (DIFF $X_{EXT\ I,\ trial\ 1}$ and DIFF $X_{EXT\ I,\ trial\ 5}$) as repeated measures factor and group (stress, control) as between subject factor were calculated.

Retrieval II: To analyze prolonged CPT-induced stress effects on fear memory retrieval at EXT I univariate ANCOVAs with DIFF $X_{EXT\ II,\ trial\ 1}$ as dependent variable and DIFF $X_{EXT\ I,\ trial\ 5}$ as covariate and group (stress, control) as between subject factor were calculated.

EXT II: To analyze if extinction occurred and to analyze prolonged CPT-induced stress effects on extinction repeated measures ANOVAs with trial (DIFF $X_{EXT\ II,\ trial\ 1}$ and DIFF $X_{EXT\ II,\ trial\ 5}$) as repeated measures factor and group (stress, control) as between subject factor were calculated.

All analyses were done with SPSS statistics package for Macintosh (IBM SPSS, 20.0). In case of violation of the sphericity assumption Greenhouse—Geisser correction was applied. All tests were two-tailed and a $p < .05$ was considered statistically significant.

Table 2 Demographic and participant characteristics and baseline measurements at day 1 before conditioning, day 2 before CPT and day 3 before extinction II.

	Control group	Stress group	Significance (p)		
			Sex	Group	Sex × group
Sex (males/females)[a]	11/7	13/22			
Age	25.08 (1.08)	23.72 (.78)	.19	.31	.07
BMI	22.96 (.59)	21.87 (.43)	.07	.14	.70
ADS	7.97 (1.12)	7.63 (.81)	.83	.81	.72
ASI	10.35 (1.34)	10.39 (.97)	.35	.98	.71
STAI trait	31.93 (1.43)	32.52 (1.04)	.11	.74	.59
STAI state day 1	32.69 (1.61)	34.97 (1.17)	.53	.26	.34
STAI state day 2	31.83 (1.57)	34.84 (1.14)	.86	.13	.11
STAI state day 3	31.65 (1.37)	34.49 (.99)	.48	.10	.08
PANAS PA trait	22.61 (1.48)	21.11 (1.07)	.42	.41	.28
PANAS NA trait	1.74 (.62)	2.21 (.45)	.46	.55	.16
PANAS PA state day 1	24.90 (1.18)	24.81 (.85)	.40	.95	.43
PANAS NA state day 1	3.46 (.83)	4.24 (.60)	.51	.45	.35
PANAS PA state day 2	21.90 (1.35)	21.54 (.98)	.68	.83	.89
PANAS NA state day 2	1.27 (.76)	2.41 (.55)	.67	.23	.05
PANAS PA state day 3	23.01 (1.43)	20.52 (1.04)	.88	.17	.34
PANAS NA state day 3	.58 (.32)	1.29 (.24)	.29	.09	.09
CPT-C evaluation day 1	5.64 (4.39)	10.41 (3.14)	.35	.38	.51
CPT/CPT-C evaluation day 2	1.67 (3.99)	83.30 (2.80)	.53	.01*	.62
CPT-C temperature day 1	38.48 (.24)	38.76 (.17)	.93	.34	.74
CPT/CPT-C temperature day 2	38.81 (.18)	2.03 (.13)	.12	.01*	.70
CPT/CPT-C duration day 2	180.00 (6.17)	168.70 (4.46)	.57	.14	.57
US unpleasantness rating	65.24 (5.03)	61.98 (3.64)	.45	.60	.68
UR (mean SCR ACQ trial 1—5)	.83 (.13)	77 (.10)	.62	.71	.90
US intensity (1—10 mA)[a,b]	1 mA: 5.6%	1 mA: 17.1%			
	2 mA: 16.7%	2 mA: 40.0%			
	5 mA: 50.0%	5 mA: 28.6%			
	10 mA: 27.8%	10 mA: 14.3%			

Data presented as mean (SEM). BMI, Body Mass Index; ADS, General Depression Scale; ASI, Anxiety Sensitivity Index; STAI, State-Trait Anxiety Inventory; PANAS, Positive Affect and Negative Affect Schedule (see text for details); CPT, cold pressor test; CPT-C, CPT control condition; US, unconditioned stimulus; UR, unconditioned reaction; SCR, skin conductance response; ACQ, acquisition; Significant effects $p < 0.05$ are identified by an asterisk (*).

[a] Pearson's Chi squared test for group $p > 0.05$.

[b] Pearson's Chi squared test for sex $p < 0.05$*.

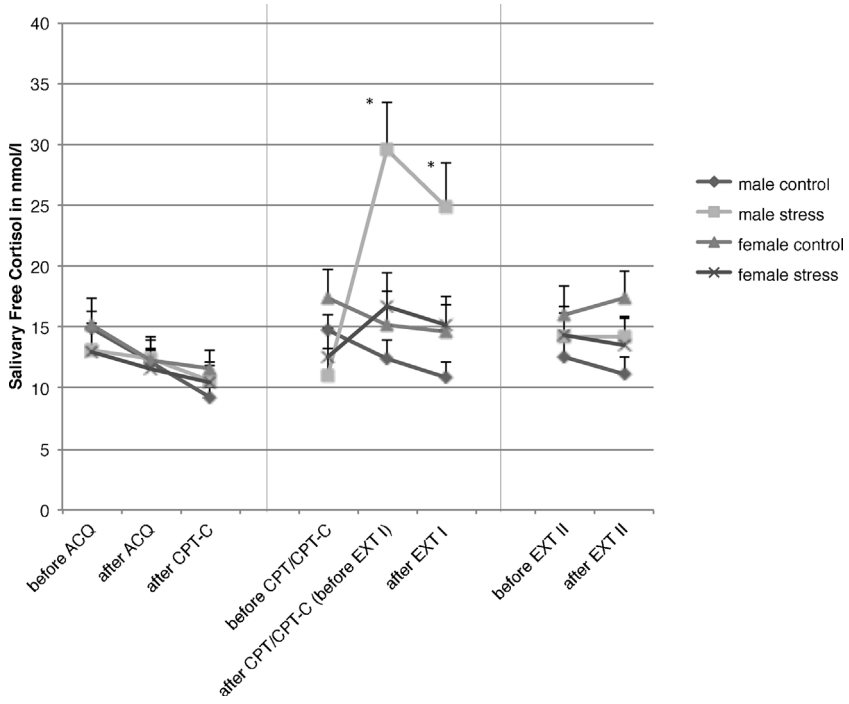


Figure 1 Salivary free cortisol (in nmol/l) for men and women in the cold pressor test (CPT) stress and the CPT control condition (CPT-C) control group before acquisition (ACQ), after ACQ and after CPT-C on day 1, before CPT/CPT-C, 20 min after CPT/CPT-C directly before extinction I (EXT I) and after EXT I on day 2 as well as before extinction II (EXT II) and after EXT II (days were separated visually by vertical lines). Values are depicted as means and SEM. Asterisks (*$p < 0.05$) indicate significant difference between men and women in the CPT stress group at indicated time-points (see text for details).

3. Results

3.1. Demographic, control, and psychometric variables

The groups did not differ significantly in demographic, participant characteristics and psychometric baseline measurements on day 1. Men chose a higher US-intensity (independent of group assignment) than women, but subjective US-unpleasantness did not differ by sex (Table 2).

3.2. Cortisol

We did not find group differences in cortisol levels before and after ACQ as well as after CPT-C on day 1 and before CPT or CPT-C on day 2 in men (all $ps \geq .26$) and women (all $ps \geq .08$). CPT/CPT-$C_{difference\ score}$ revealed a significant group effect in women ($F_{1,29} = 15.94$, $p < .01$, partial eta$^2 = .37$) and men ($F_{1,24} = 31.52$, $p < .01$, partial eta$^2 = .59$) indicating a significantly higher cortisol increase in the stress group than in the control group in both sexes. An univariate ANOVA further analyzing the CPT-induced cortisol rise with CPT/CPT-$C_{difference\ score}$ as dependent variable and sex (male, female) and group (stress, control) as between subject factors revealed, additionally to a main effect group ($F_{1,49} = 46.94$, $p < .01$, partial eta$^2 = .49$), an interaction effect for sex $\times$ group ($F_{1,49} = 13.21$, $p < .01$, partial eta$^2 = .21$) as well as a main effect for sex ($F_{1,39} = 10.42$, $p < .01$, partial eta$^2 = .21$). A post hoc univariate ANOVA within the stress group revealed a significantly higher $CPT_{difference\ score}$ in men compared to women ($F_{1,33} = 26.01$, $p < .01$, partial eta$^2 = .44$), with men showing approximately a four-times higher mean cortisol rise than women (Fig. 1). Men maintained heightened cortisol levels in the stress group compared to the control group after EXT I on day 2 ($F_{1,22} = 6.68$, $p = .02$, partial eta$^2 = .23$), whereas there was no group difference in women ($F_{1,27} = .06$, $p = .81$). No differences in cortisol levels before and after extinction on day 3 were found in men ($p \geq .23$) or women ($p \geq .14$).

3.3. Subjective measures

Habituation/acquisition, day 1

The repeated measures ANOVAs with trial (DIFF $X_{HAB,\ trial\ 5}$ and DIFF $X_{ACQ,\ trial\ 5}$) as repeated measures factor and group (stress, control) and sex (male, female) as between subject factors showed a significant effect for trial (all $ps < .01$, partial eta$^2 = .61$ (EXP), .52 (VAL), .27 (ANX)) with greater DIFF $X_{ACQ,\ trial\ 5}$ compared to DIFF $X_{HAB,\ trial\ 5}$ and no effect for group (all $ps \geq .32$) for any of the subjective measurements. Additionally for EXP a significant effect for sex ($F_{1,44} = 4.52$, $p = .04$, partial eta$^2 = .09$) was found with higher difference scores in women.

3.3.1. US-expectancy

Retrieval I, day 2

Men. The univariate ANCOVA with DIFF $EXP_{EXT\ I,\ trial\ 1}$ as dependent variable, DIFF $EXP_{ACQ,\ trial\ 5}$ as covariate and group (stress, control) as between subject factor found a significant

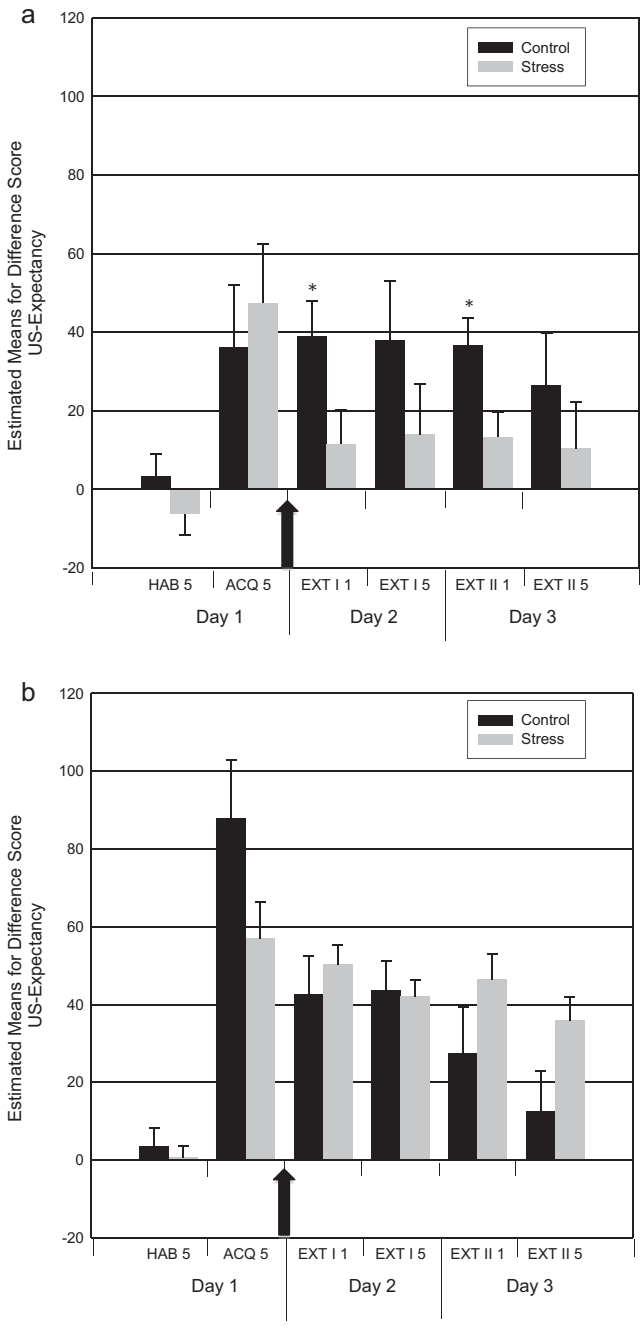


Figure 2 Difference score US-expectancy (EXP) at habituation (HAB) trial 5 (Diff $EXP_{HAB,\ trial\ 5}$), acquisition (ACQ) trial 5 (Diff $EXP_{ACQ,\ trial\ 5}$), extinction I (EXT I) trial 1 (Diff $EXP_{EXT\ I,\ trial\ 1}$), EXT I trial 5 (Diff $EXP_{EXT\ I,\ trial\ 5}$), extinction II (EXT II) Trial 1 (Diff $EXP_{EXT\ II,\ Trial\ 1}$) and extinction II (EXT II) trial 5 (Diff $EXP_{EXT\ II,\ trial\ 5}$) for men (a) and women (b) in the control and stress group. Memory retrieval was measured with the first extinction trial on day 2 (Diff $EXP_{EXT\ I,\ trial\ 1}$) and the first extinction trial on day 3 (Diff $EXP_{EXT\ II,\ trial\ 1}$). CPT was performed 20 min before EXT I on day 2 (indicated by arrow). Difference scores EXP for every time point were calculated by substracting the respective CS-MINUS from CS-PLUS at the trial under consideration. Values are depicted as estimated means and SEM of EXP difference scores at respective trial. The asterisk ($*p < 0.05$) indicates significant differences between the control and stress group at the respective trial (see text for details and *F*- and *p*-values).

group effect ($F_{1,20} = 4.94$, $p = .04$, partial $eta^2 = .20$) with the stress group showing a significant smaller DIFF $EXP_{EXT\ I,\ trial\ 1}$ score indicating that fear memory retrieval was influenced by CPT stress (Fig. 2a).

Women. The univariate ANCOVA with DIFF $EXP_{EXT\ I,\ trial\ 1}$ as dependent variable, DIFF $EXP_{ACQ,\ trial\ 5}$ as covariate and group (CPT stress, control) as between subject factor did not find a significant group effect ($F_{1,24} = .45$, $p = .51$) indicating that fear memory retrieval was not influenced by CPT stress (Fig. 2b).

Extinction I, day 2

Men/women. The repeated measures ANOVA with trial (DIFF $EXP_{EXT\ I,\ trial\ 1}$ and DIFF $EXP_{EXT\ I,\ trial\ 5}$) as repeated measures factor and group (stress, control) as between subject factor did not find a significant effect for trial in men ($F_{1,18} = .09$, $p = .77$) and women ($F_{1,25} = .60$, $p = .45$), group in men ($F_{1,18} = 1.59$, $p = .22$) and women ($F_{1,25} = 2.12$, $p = .16$) and no trial $\times$ group interaction in men ($F_{1,18} = .12$, $p = .74$) and women ($F_{1,25} = .00$, $p = .99$) indicating that extinction training per se did not lead to fear extinction and that there was no influence of stress on extinction performance.

Retrieval II, day 3

Men. The univariate ANOVA with DIFF $EXP_{EXT\ II,\ trial\ 1}$ as dependent variable, DIFF $EXP_{EXT\ I,\ trial\ 5}$ as covariate and group (stress, control) as between subject factor did find a significant group effect ($F_{1,18} = 6.27$, $p = .02$, partial $eta^2 = .27$) with smaller DIFF $EXP_{EXT\ II,\ trial\ 1}$ in the stress compared to the control group indicating that there was a prolonged effect of stress on fear memory retrieval.

Women. The univariate ANOVA with DIFF $EXP_{EXT\ II,\ trial\ 1}$ as dependent variable, DIFF $EXP_{EXT\ I,\ trial\ 5}$ as covariate and group (stress, control) as between subject factor did not find a significant group effect ($F_{1,23} = 1.91$, $p = .18$) indicating that there was no prolonged effect of stress on fear memory retrieval.

Extinction II, day 3

Men/women. The repeated measures ANOVA with trial (DIFF $EXP_{EXT\ II,\ trial\ 1}$ and DIFF $EXP_{EXT\ II,\ trial\ 5}$) as repeated measures factor and group (stress, control) as between subject factor did not find a significant effect for trial in men ($F_{1,19} = 1.43$, $p = .25$) and women ($F_{1,24} = 3.02$, $p = .10$) and group in men ($F_{1,19} = 3.10$, $p = .09$) and women ($F_{1,24} = 1.16$, $p = .29$) and no trial $\times$ group interaction in men ($F_{1,19} = 1.01$, $p = .33$) and women ($F_{1,24} = .95$, $p = .76$).

3.3.2. Valence, anxiety

Men/women. No effects for VAL and ANX were found for retrieval and extinction I on day 2 as well as for retrieval and extinction II on day 3 (all $ps \geq .09$).

3.4. Physiological measures

3.4.1. Skin conductance response

Acquisition/habituation. The repeated measures ANOVAs with trial (DIFF $SCR_{HAB,\ trial\ 5}$ and DIFF $SCR_{ACQ,\ trial\ 5}$) as repeated measures factor and group (stress, control) and sex (male, female) as between subject factor did not show a significant trial effect (all $ps \geq .13$) for SCR which might be due to the low number of acquisition trials. A significant

effect trial with higher DIFF $SCR_{ACQ,\ trial\ 5}$ compared to DIFF $SCR_{HAB,\ trial\ 5}$ indicating successful acquisition was the pre-condition to examine stress effects on retrieval and extinction processes within our conditioning paradigm. Therefore no further analyses for Retrieval I, EXT I, EXT II and Retrieval II were calculated.

3.4.2. Heart rate

Acquisition/habituation. The repeated measures ANOVAs with trial (DIFF $HR_{HAB,\ trial\ 5}$ and DIFF $HR_{ACQ,\ trial\ 5}$) as repeated measures factor and group (stress, control) and sex (male, female) as between subject factor showed no significant effect for trial (all $ps \geq .63$) for HR. No further analyses for Retrieval I, EXT I, EXT II and Retrieval II were calculated because the pre-condition, a significant effect trial with higher DIFF $HR_{ACQ,\ trial\ 5}$ compared to DIFF $HR_{HAB,\ trial\ 5}$ indicating successful acquisition, was not fulfilled.

4. Discussion

In the present study we aimed at investigating the effects of stress on memory retrieval and extinction processes in a fear conditioning paradigm in healthy male and female subjects. We found a stress-induced impairment in memory retrieval as measured with subjective US-expectancy ratings 24 h after acquisition in men who showed a robust cortisol response to the CPT. In women, who showed a significantly smaller cortisol response to stress, no stress effects on US-expectancy ratings were observed.

The finding of stress-induced impairment in fear memory retrieval is in line with several studies investigating memory retrieval within different learning paradigms in animals and humans (de Quervain et al., 1998, 2000; Roozendaal et al., 2004; Het et al., 2005; Kuhlmann et al., 2005a,b; Kuhlmann and Wolf, 2005; Buchanan et al., 2006). In contrast to the aforementioned studies the present results arise not from a classical declarative memory task that measures explicit memory processes, but from US-expectancy ratings in a fear conditioning paradigm. There is evidence that patients with anxiety disorders have problems to extinct fear expectancies (Blechert et al., 2007; Michael et al., 2007), indicating that the behavioral measure of the present study is of relevance for clinical conditions (see also Boddez et al., 2013).

In the present study, extinction training did not lead to a decrease in subjective measures. This lack of extinction in subjective measures is in line with results from other studies (de Houwer et al., 2001; Baeyens et al., 2005; Blechert et al., 2008) and might be due to low number of extinction trials in our conditioning paradigm. Additionally we did not observe a stress effect on extinction training. Thus, this study cannot make any conclusion with regard to stress effects on acquisition or consolidation of fear extinction. However, we found a prolonged effect of the stress intervention on memory retrieval, as the observed reduction of US-expectancy induced by stress on day 2, was also observed on day 3 (Fig. 2).

In women we did not find such an influence of stress on fear memory retrieval. This is in line with another study that found an effect of stress on fear memory retrieval in men, but not in women (Cornelisse et al., 2011). One explanation for the sexual dimorph stress effects on fear memory retrieval in our study could be that the cortisol increase in response to CPT stress in women was too small to unfold an impairing effect on memory retrieval. In our study, men showed an almost four times higher cortisol response than women, so we can not rule out that we would have seen the same effects in women with stronger cortisol stress responses.

It can be argued that the CPT is not an appropriate stress test to induce sufficient cortisol release in women compared to men. It is known that men show a better cortisol response to achievement challenges, whereas women show greater cortisol responses to social rejection challenges (e.g. the Yale Interpersonal Stressor, YIPS) (Stroud et al., 2000, 2002). The CPT stressor in our study was conceptualized as an achievement challenge rather than a social rejection challenge compared to other studies using the CPT (Schwabe et al., 2008).

Another explanation for the smaller cortisol response to CPT stress in women could be that only women taking oral contraceptives (OCs) took part in our study. Other studies investigating fear memory retrieval processes in women using OCs were also unable to detect stress effects on fear memory retrieval in these women (Kuhlmann and Wolf, 2005; Cornelisse et al., 2011). There is evidence that women taking OCs have a lower stress response measured in saliva cortisol compared to free cycling women in the follicular or luteal phase and men (Kirschbaum et al., 1999). Additionally a recent study found that women taking OCs show a blunted cortisol response to CPT stress compared to women not taking OCs (Nielsen et al., 2012). Saliva cortisol measures unbound cortisol (Kirschbaum and Hellhammer, 1989). A major part of circulating cortisol is bound to the corticosteroid-binding globulin (CBG) and only the unbound part is believed to be biologically active. Artificial and endogenous hormones influence the CBG levels with OCs having CBG increasing effects (Moore et al., 1978; Wiegratz et al., 2003). Another study applying the CPT in men as well as naturally cycling women found comparable CPT induced cortisol responses in men and women (Felmingham et al., 2012). This study further indicates that the blunted cortisol response in women in our study might be due to OC intake. For this reason it would be interesting to investigate women not taking OCs in a comparable study paradigm as ours. Additionally, it would be of value to investigate the influence of the natural fluctuations of sex hormones across the menstrual cycle on stress effects during fear memory retrieval and extinction processes.

The application of CPT to induce stress does not equal exogenous GC elevation as acute stress leads to further physiological changes besides cortisol release. Both catecholamines and cortisol are an integral part of the stress response and the CPT is known to engage both GC and noradrenergic activation (Lovallo, 1975). Recent animal studies and human studies indicate that concomitant noradrenergic release (e.g. by emotional arousal) is necessary for the memory modulating effect of GCs (Roozendaal et al., 2004, 2006; Buchanan et al., 2006; de Quervain et al., 2007). Therefore, it is possible that effects of CPT on memory retrieval are mediated by catecholamines or by their orchestrated actions with GCs.

Our findings may contribute to the understanding of effects of stress and GCs on fear symptoms in anxiety disorders. Anxiety disorders can be characterized as disorders involving disturbed memory processes responsible for the symptomatology and maintenance of these disorders.

Especially patients with PTSD are impaired by unwanted, uncontrollable memories regarding the traumatic experience (intrusions). A possibility to weaken aversive fear memory retrieval holds the potential to ameliorate the burden of the illness (Michael et al., 2005; de Quervain et al., 2009; Bentz et al., 2010). In a clinical study investigating the effects of cortisol treatment in a small number of patients with chronic PTSD we found evidence for such a beneficial effect. In this study acute low-dose cortisol administration reduced trauma retrieval and symptoms associated with traumatic memories without causing adverse side effects (Aerni et al., 2004). Additional evidence comes from a study reporting beneficial effects of GC administration on combat-related PTSD symptoms (Suris et al., 2010). GCs may not only reduce retrieval of traumatic memories in patients with PTSD but also retrieval of fear memory in patients with phobia and thereby reduce stimulus-induced fear. In a study in patients with social phobia we reported that cortisone administration reduced fear ratings during the anticipation, exposure, and recovery phases of a socio-evaluative stressor compared to placebo. Moreover, the stress-induced release of cortisol in placebo-treated subjects correlated negatively with fear ratings, suggesting that endogenously released cortisol in the context of a phobic situation buffers fear symptoms (Soravia et al., 2006). No such anxiolytic effect of GC administration was found in healthy humans exposed to the same paradigm suggesting that GCs do not have a general anxiolytic effect, but rather unfold their fear-reducing properties by influencing pathological fear memory processes (Soravia et al., 2009). In a study in participants with spider phobia, repeated oral administration of cortisol, but not placebo, 1 h before exposure to a spider photograph induced a progressive reduction of stimulus-induced fear (Soravia et al., 2006). Additionally, in a double-blind, randomized, placebo-controlled study in patients with height phobia we found that the administration of cortisol can enhance extinction-based psychotherapy (de Quervain et al., 2011). Together, the preclinical studies suggest that GC administration has acute effects on clinical symptoms by reducing the retrieval of aversive memories. Furthermore, there is evidence that GC administration facilitates fear extinction processes.

The present findings contribute to the understanding of the effects of stress and stress hormones on fear symptoms. Specifically, we found evidence that stress can induce a reduction of explicit fear memory in a conditioning paradigm in men but not women. This effect seems to be mediated through a reduction of fear memory retrieval. More studies are needed to investigate possible effects of stress and stress hormones on implicit forms of fear memory and on fear memory extinction processes in both sexes.

Role of funding source

Supported by grants from the Swiss National Science Foundation to D.Q. (PP00P3-123391 and CRSI33_130080/1).

Conflict of interest

None of the authors of the above manuscript has declared any actual or potential conflict of interest, which may arise from being named as an author of the manuscript.

Acknowledgements

Supported by grants from the Swiss National Science Foundation to D.Q. (PP00P3-123391 and CRSI33_130080/1).

References

Aerni, A., Traber, R., Hock, C., Roozendaal, B., Schelling, G., Papassotiropoulos, A., Nitsch, R.M., Schnyder, U., de Quervain, D.J., 2004. Low-dose cortisol for symptoms of posttraumatic stress disorder. Am. J. Psychiatry 161, 1488—1490.

Baeyens, F., Diaz, E., Ruiz, G., 2005. Resistance to extinction of human evaluative conditioning using a between-subjects design. Cogn. Emot. 19, 245—268.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., Damasio, A.R., 1995. Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. Science 269, 1115—1118.

Bentz, D., Michael, T., de Quervain, D.J., Wilhelm, F.H., 2010. Enhancing exposure therapy for anxiety disorders with glucocorticoids: from basic mechanisms of emotional learning to clinical applications. J. Anxiety Disord. 24, 223—230.

Blechert, J., Michael, T., Vriends, N., Margraf, J., Wilhelm, F.H., 2007. Fear conditioning in posttraumatic stress disorder: evidence for delayed extinction of autonomic, experiential, and behavioural responses. Behav. Res. Ther. 45, 2019—2033.

Blechert, J., Michael, T., Williams, S.L., Purkis, H.M., Wilhelm, F.H., 2008. When two paradigms meet: does evaluative learning extinguish in differential fear conditioning? Learn. Motiv. 39, 58—70.

Blundell, J., Blaiss, C.A., Lagace, D.C., Eisch, A.J., Powell, C.M., 2011. Block of glucocorticoid synthesis during re-activation inhibits extinction of an established fear memory. Neurobiol. Learn. Mem. 95, 453—460.

Boddez, Y., Baeyens, F., Luyten, L., Vansteenwegen, D., Hermans, D., Beckers, T., 2013. Rating data are underated: Vailidity of US expectancy in human fear conditioning. J. Behav. Ther. Exp. Psychiatry. 44, 201—206.

Bouton, M.E., Mineka, S., Barlow, D.H., 2001. A modern learning theory perspective on the etiology of panic disorder. Psychol. Rev. 108, 4—32.

Buchanan, T.W., Lovallo, W.R., 2001. Enhanced memory for emotional material following stress-level cortisol treatment in humans. Psychoneuroendocrinology 26, 307—317.

Buchanan, T.W., Tranel, D., Adolphs, R., 2006. Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. Learn. Mem. 13, 382—387.

Cai, W.H., Blundell, J., Han, J., Greene, R.W., Powell, C.M., 2006. Postreactivation glucocorticoids impair recall of established fear memory. J. Neurosci. 26, 9560—9566.

Carter, R.M., O'Doherty, J.P., Seymour, B., Koch, C., Dolan, R.J., 2006. Contingency awareness in human aversive conditioning involves the middle frontal gyrus. Neuroimage 29, 1007—1012.

Cornelisse, S., van Stegeren, A.H., Joëls, M., 2011. Implications of psychosocial stress on memory formation in a typical male versus female student sample. Psychoneuroendocrinology 36, 569—578.

Dawson, M.E., Schell, A.M., Filion, D.L., 2007. The electrodermal system. In: Cacioppo, J.T., Tassinary, L.G., Berntson, G.G. (Eds.), Handbook of Psychophysiology. 3rd ed. Cambridge University Press, New York, pp. 159—181.

de Houwer, J., Thomas, S., Baeyens, F., 2001. Associative learning of likes and dislikes: a review of 25 years of research on human evaluative conditioning. Psychol. Bull. 127, 853—869.

de Quervain, D.J., Aerni, A., Roozendaal, B., 2007. Preventive effect of {beta}-adrenoceptor blockade on glucocorticoid-induced memory retrieval deficits. Am. J. Psychiatry 164, 967—969.

de Quervain, D.J., Aerni, A., Schelling, G., Roozendaal, B., 2009. Glucocorticoids and the regulation of memory in health and disease. Front. Neuroendocrinol. 30, 358—370.

de Quervain, D.J., Bentz, D., Michael, T., Bolt, O.C., Wiederhold, B.K., Margraf, J., Wilhelm, F.H., 2011. Glucocorticoids enhance extinction-based psychotherapy. Proc. Natl. Acad. Sci. U. S. A. 108, 6621—6625.

de Quervain, D.J., Roozendaal, B., McGaugh, J.L., 1998. Stress and glucocorticoids impair retrieval of long-term spatial memory. Nature 394, 787—790.

de Quervain, D.J., Roozendaal, B., Nitsch, R.M., McGaugh, J.L., Hock, C., 2000. Acute cortisone administration impairs retrieval of long-term declarative memory in humans. Nat. Neurosci. 3, 313—314.

Dickerson, S.S., Kemeny, M.E., 2004. Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. Psychol. Bull. 130, 355—391.

Ehlers, A., Margraf, J., 1993. Angst vor der Angst: Ein neues Konzept in der Diagnostik der Angststörungen (''Fear of fear'': a new concept in the assessment of anxiety disorders). Verhaltenstherapie 3, 14—24.

Felmingham, K.L., Tran, T.P., Fong, W.C., Bryant, R.A., 2012. Sex differences in emotional memory consolidation: the effect of stress-induced salivary alpha-amylase and cortisol. Biol. Psychol. 89, 539—544.

Fowles, D.C., Christie, M.J., Edelberg, R., Grings, W.W., Lykken, D.T., Venables, P.H., 1981. Committee report. Publication recommendations for electrodermal measurements. Psychophysiology 18, 232—239.

Grillon, C., Pine, D.S., Baas, J.M., Lawley, M., Ellis, V., Charney, D.S., 2006. Cortisol and DHEA-S are associated with startle potentiation during aversive conditioning in humans. Psychopharmacology 186, 434—441.

Hautzinger, M., Bailer, M., 1993. Allgemeine Depressionsskala. Beltz Test GmbH, Goettingen.

Hermans, D., Craske, M.G., Mineka, S., Lovibond, P.F., 2006. Extinction in human fear conditioning. Biol. Psychiatry 60, 361—368.

Het, S., Ramlow, G., Wolf, O.T., 2005. A meta-analytic review of the effects of acute cortisol administration on human memory. Psychoneuroendocrinology 30, 771—784.

Jackson, E.D., Payne, J.D., Nadel, L., Jacobs, W.J., 2006. Stress differentially modulates fear conditioning in healthy men and women. Biol. Psychiatry 59, 516—522.

Kajantie, E., Phillips, D.I., 2006. The effects of sex and hormonal status on the physiological response to acute psychosocial stress. Psychoneuroendocrinology 31, 151—178.

Kirschbaum, C., Hellhammer, D.H., 1989. Salivary cortisol in psychobiological research: an overview. Neuropsychopharmocology 22, 150—169.

Kirschbaum, C., Wuest, S., Hellhammer, D., 1992. Consistent sex differences in cortisol responses to psychological stress. Psychosom. Med. 54., 648—657.

Kirschbaum, C., Kudielka, B.M., Gaab, J., Schommer, N.C., Hellhammer, D.H., 1999. Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus—pituitary—adrenal axis. Psychosom. Med. 61, 154—162.

Knight, D.C., Cheng, D.T., Smith, C.N., Stein, E.A., Helmstetter, F.J., 2004. Neural substrates mediating human delay and trace fear conditioning. J. Neurosci. 24, 218—228.

Krohne, H.W., Egloff, B., Kohlmann, C.-W., Tausch, A., 1996. Untersuchung mit einer deutschen Form der Positive and Negative Affective Schedule (PANAS). Diagnostica 42, 139—156.

Kuehl, L.K., Lass-Hennemann, J., Richter, S., Blumenthal, T.D., Oitzl, M., Schachinger, H., 2010. Accelerated trace eyeblink conditioning after cortisol IV-infusion. Neurobiol. Learn. Mem. 94, 547—553.

Kuhlmann, S., Kirschbaum, C., Wolf, O.T., 2005a. Effects of oral cortisol treatment in healthy young women on memory retrieval of negative and neutral words. Neurobiol. Learn. Mem. 83, 158—162.

Kuhlmann, S., Piel, M., Wolf, O.T., 2005b. Impaired memory retrieval after psychosocial stress in healthy young men. J. Neurosci. 25, 2977—2982.

Kuhlmann, S., Wolf, O.T., 2005. Cortisol and memory retrieval in women: influence of menstrual cycle and oral contraceptives. Psychopharmacology (Berl) 183, 65—71.

LaBar, K.S., Cabeza, R., 2006. Cognitive neuroscience of emotional memory. Nat. Rev. Neurosci. 7, 54—64.

Laux, L., Glanzmann, P., Schaffner, P., Spielberger, C.D., 1981. State-Trait-Angstinventar. Theoretische Grundlagen und Handanweisung. Beltz TestGmbH, Weinheim.

Lovallo, W.R., 1975. The cold pressor test and autonomic function: a review and integration. Psychophysiology 12, 268—282.

Lovibond, P.F., 1992. Tonic and phasic electrodermal measures of human aversive conditioning with long duration stimuli. Psychophysiology 29, 621—632.

Lovibond, P.F., 2006. Fear and avoidance: an integrated expectancy model. In: Craske, M.G., Hermans, D., Vansteenwegen, D. (Eds.), Fear and Learning. From Basic Processes to Clinical Implications. American Psychological Association, Washington, DC, pp. 117—132.

Luethi, M., Meier, B., Sandi, C., 2009. Stress effects on working memory, explicit memory, and implicit memory for neutral and emotional stimuli in healthy men. Front. Behav. Neurosci. 2, 5.

Mahan, A.L., Ressler, K.J., 2012. Fear conditioning, synaptic plasticity and the amygdala: implications for posttraumatic stress disorder. Trends Neurosci. 35, 24—35.

Merz, C.J., Tabbert, K., Schweckendiek, J., Klucken, T., Vaitl, D., Stark, R., Wolf, O.T., 2012a. Oral contraceptive usage alters the effects of cortisol on implicit fear learning. Horm. Behav. 62, 531—538.

Merz, C.J., Tabbert, K., Schweckendiek, J., Klucken, T., Vaitl, D., Stark, R., Wolf, O.T., 2012b. Neuronal correlates of extinction learning are modulated by sex hormones. Soc. Cogn. Affect. Neurosci. 7, 819—830.

Merz, C.J., Stark, R., Vaitl, D., Tabbert, K., Wolf, O.T., 2013. Stress hormones are associated with the neuronal correlates of instructed fear conditioning. Biol. Psychol. 92, 82—89.

Merz, C.J., Tabbert, K., Schweckendiek, J., Klucken, T., Vaitl, D., Stark, R., Wolf, O.T., 2010. Investigating the impact of sex and cortisol on implicit fear conditioning with fMRI. Psychoneuroendocrinology 35, 33—46.

Michael, T., Blechert, J., Vriends, N., Margraf, J., Wilhelm, F.H., 2007. Fear conditioning in panic disorder: enhanced resistance to extinction. J. Abnorm. Psychol. 116, 612—617.

Michael, T., Ehlers, A., Halligan, S.L., Clark, D.M., 2005. Unwanted memories of assault: what intrusion characteristics are associated with posttraumatic stress disorder? Behav. Res. Ther. 43, 613—628.

Mineka, S., Oehlberg, K., 2008. The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. Acta Psychol. (Amst.) 127, 567—580.

Moore, D.E., Kawagoe, S., Davajan, V., Nakamura, R.M., Mishell, D.R., 1978. An in vivo system in man for quantitation of estrogenicity. II. Pharmacologic changes in binding capacity of serum corticosteroid-binding globulin induced by conjugated estrogens, mestranol, and ethinyl estradiol. Am. J. Obstet. Gynecol. 130, 482—486.

Myers, K.M., Davis, M., 2002. Behavioral and neural analysis of extinction. Neuron 36, 567—584.

Nees, F., Richter, S., Lass-Hennemann, J., Blumenthal, T.D., Schachinger, H., 2008. Inhibition of cortisol production by metyrapone enhances trace, but not delay, eyeblink conditioning. Psychopharmacology 199, 183—190.

Nielsen, S.E., Segal, S.K., Worden, I.V., Yim, I.S., Cahill, L., 2012. Hormonal contraception use alters stress responses and emotional memory. Biol. Psychol. 92, 257—266.

Phelps, E.A., LeDoux, J.E., 2005. Contributions of the amygdala to emotion processing: from animal models to human behavior. Neuron 48, 175—187.

Roozendaal, B., 2000. 1999 Curt P, Richter award. Glucocorticoids and the regulation of memory consolidation. Psychoneuroendocrinology 25, 213—238.

Roozendaal, B., Hahn, E.L., Nathan, S.V., de Quervain, D.J., McGaugh, J.L., 2004. Glucocorticoid effects on memory retrieval require concurrent noradrenergic activity in the hippocampus and basolateral amygdala. J. Neurosci. 24, 8161—8169.

Roozendaal, B., Okuda, S., de Quervain, D.J., McGaugh, J.L., 2006. Glucocorticoids interact with emotion-induced noradrenergic activation in influencing different memory functions. Neuroscience 138, 901—910.

Soravia, L.M., de Quervain, D.J., Heinrichs, M., 2009. Glucocorticoids do not reduce subjective fear in healthy subjects exposed to social stress. Biol. Psychol. 81, 184—188.

Soravia, L.M., Heinrichs, M., Aerni, A., Maroni, C., Schelling, G., Ehlert, U., Roozendaal, B., de Quervain, D.J., 2006. Glucocorticoids reduce phobic fear in humans. Proc. Natl. Acad. Sci. U. S. A. 103, 5585—5590.

Schwabe, L., Bohringer, A., Chatterjee, M., Schachinger, H., 2008. Effects of pre-learning stress on memory for neutral, positive and negative words: different roles of cortisol and autonomic arousal. Neurobiol. Learn. Mem. 90, 44—53.

Stark, R., Wolf, O.T., Tabbert, K., Kagerer, S., Zimmermann, M., Kirsch, P., Schienle, A., Vaitl, D., 2006. Influence of the stress hormone cortisol on fear conditioning in humans: evidence for sex differences in the response of the prefrontal cortex. Neuroimage 32, 1290—1298.

Stroud, L.R., Salovey, P., Epel, E.S., 2002. Sex differences in stress responses: social rejection versus achievement stress. Biol. Psychiatry 52, 318—327.

Stroud, L.R., Tanofsky-Kraff, M., Wilfley, D.E., Salovey, P., 2000. The Yale Interpersonal Stressor (YIPS): affective, physiological, and behavioral responses to a novel interpersonal rejection paradigm. Ann. Behav. Med. 22, 204—213.

Suris, A., North, C., Adinoff, B., Powell, C.M., Greene, R., 2010. Effects of glucocorticoid on combat-related PTSD symptoms. Ann. Clin. Psychiatry 22, 274—279.

Tabbert, K., Merz, C.J., Klucken, T., Schweckendiek, J., Vaitl, D., Wolf, O.T., Stark, R., 2010. Cortisol enhances neural differentiation during fear acquisition and extinction in contingency aware young women. Neurobiol. Learn. Mem. 94, 392—401.

Wiegratz, I., Kutschera, E., Lee, J.H., Moore, C., Mellinger, U., Winkler, U.H., Kuhl, H., 2003. Effect of four different oral contraceptives on various sex hormones and serum-binding globulins. Contraception 67, 25—32.

Wilhelm, F.H., Peyk, P., 2005. ANSLAB: Autonomic Nervous System Laboratory (Version 2.5). Available at the SPR Software Resposity: http://www.sprweb.org.

Wittchen, H.U., Zaudig, M., Fydrich, T., 1997. SKID Strukturiertes Klinisches Interview für DSM-IV Achse I und II. Hogrefe, Goettingen.

Wolf, O.T., Minnebusch, D., Daum, I., 2009. Stress impairs acquisition of delay eyeblink conditioning in men and women. Neurobiol. Learn. Mem. 91, 431—436.

Yang, Y.L., Chao, P.K., Lu, K.T., 2006. Systemic and intra-amygdala administration of glucocorticoid agonist and antagonist modulate extinction of conditioned fear. Neuropsychopharmacology 31, 912—924.

Zorawski, M., Cook, C.A., Kuhn, C.M., LaBar, K.S., 2005. Sex, stress, and fear: individual differences in conditioned learning. Cogn. Affect. Behav. Neurosci. 5, 191—201.

Zorawski, M., Blanding, N.Q., Kuhn, C.M., LaBar, K.S., 2006. Effects of stress and sex on acquisition and consolidation of human fear conditioning. Learn. Mem. 13, 441—450.

4.2 Associations between basal cortisol levels and memory retrieval in healthy young individuals

Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D.J.-F., & Rasch, B. (2013). *Journal of Cognitive Neuroscience*, *25*(11), 1896-1907.

Associations between Basal Cortisol Levels and Memory Retrieval in Healthy Young Individuals

Sandra Ackermann[1,2,3], Francina Hartmann[1], Andreas Papassotiropoulos[1], Dominique J.-F. de Quervain[1], and Björn Rasch[2]

Abstract

■ Cortisol is known to affect memory processes. On the one hand, stress-induced or pharmacologically induced elevations of cortisol levels enhance memory consolidation. On the other hand, such experimentally induced elevations of cortisol levels have been shown to impair memory retrieval. However, the effects of individual differences in basal cortisol levels on memory processes remain largely unknown. Here we tested whether individual differences in cortisol levels predict picture learning and recall in a large sample. A total of 1225 healthy young women and men viewed two different sets of emotional and neutral pictures on two consecutive days. Both sets were recalled after a short delay (10 min). On Day 2, the pictures seen on Day 1 were additionally recalled, resulting in a long-delay (20 hr) recall condition. Cortisol levels were measured three times on Days 1 and 2 via saliva samples before encoding, between encoding and recall as well as after recall testing. We show that stronger decreases in cortisol levels during retrieval testing were associated with better recall performance of pictures, regardless of emotional valence of the pictures or length of the retention interval (i.e., 10 min vs. 20 hr). In contrast, average cortisol levels during retrieval were not related to picture recall. Remarkably during encoding, individual differences in average cortisol levels as well as changes in cortisol did not predict memory recall. Our results support previous findings indicating that higher cortisol levels during retrieval testing hinders recall of episodic memories and extend this view onto interindividual changes in basal cortisol levels. ■

INTRODUCTION

Glucocorticoids have a modulatory influence on memory processes. The effect of cortisol on memory strongly depends on the stage of memory consolidation (Schwabe, Joels, Roozendaal, Wolf, & Oitzl, 2012; de Quervain, Aerni, Schelling, & Roozendaal, 2009; Wolf, 2009). During memory formation, experimentally increased cortisol, pharmacologically or by stress induction, improves memory, in particular memory for emotionally arousing events (e.g., Cahill, Gorski, & Le, 2003; for a review, see Wolf, 2009). In contrast, cortisol impairs the retrieval of long-term memories. In rats, stress or systemic corticosterone administration before recall impairs recall of spatial memory of a water maze task acquired 24 hr earlier (de Quervain, Roozendaal, & McGaugh, 1998). In humans as well, administration of cortisone before retrieval testing impairs memory recall (Smeets, 2011; Tollenaar, Elzinga, Spinhoven, & Everaerd, 2008, 2009; de Quervain et al., 2003; de Quervain, Roozendaal, Nitsch, McGaugh, & Hock, 2000). Several studies show that the impairing influence of cortisol on retrieval of long-term memories is particularly pronounced for emotionally arousing material (Buchanan, Tranel, & Adolphs, 2006; Kuhlmann, Kirschbaum, & Wolf, 2005; Kuhlmann, Piel, & Wolf, 2005).

Effects of glucocorticoids on memory consolidation and retrieval depend on noradrenergic coactivation within the brain. Blockade of noradrenergic receptors in the amygdala diminishes cortisol-related memory enhancements (Roozendaal, Okuda, de Quervain, & McGaugh, 2006; Roozendaal, Okuda, Van der Zee, & McGaugh, 2006; van Stegeren et al., 2005; Quirarte, Roozendaal, & McGaugh, 1997); on the other hand, cortisol-induced retrieval impairments are blocked by concurrent administration of the adrenergic antagonist propranolol (de Quervain, Aerni, & Roozendaal, 2007; Roozendaal, Hahn, Nathan, de Quervain, & McGaugh, 2004; for a review, see Krugers, Karst, & Joels, 2012).

Considering biological mechanisms underlying reactivity and feedback processes of the HPA axis, mineralocorticoid (MR) and glucocorticoid receptors (GR) play an important role in mediating glucocorticoid effects in the brain. They are highly expressed in the limbic system (hippocampus and amygdala), regions important for emotion and cognition (Lupien & McEwen, 1997). MRs have a higher affinity for glucocorticoids than GRs and are almost saturated under basal levels. GRs become occupied under stress or when circadian glucocorticoid levels are high (Roozendaal, Okuda, de Quervain, et al., 2006; Reul & de Kloet, 1985).

[1]University of Basel, [2]University of Zurich, [3]Psychiatric University Hospital Zurich

Journal of Cognitive Neuroscience 25:11, pp. 1896–1907
doi:10.1162/jocn_a_00440

It has been suggested that MRs are implicated in the maintenance of basal activities of the stress system. On the other side, GRs, in interplay with MRs, seem to be implicated in the recovery from a stress response, hence the suppression of the HPA axis. The balance between MRs and GRs is important for HPA activity as well as for neuronal excitability, stress responsiveness, and behavioral adaptation (de Kloet, Vreugdenhil, Oitzl, & Joels, 1998). Furthermore, besides the intracellular effects of MR and GR, also a membrane-bound MR (Joëls, Karst, DeRijk, & De Kloet, 2008) and GR (Roozendaal et al., 2010) have been observed, which could be involved in rapid nongenomic effects on memory processes.

Although the effects of experimentally increased glucocorticoid levels on memory are well established, the relationship between natural circadian variation of cortisol (basal cortisol) and memory has received less attention. Basal cortisol levels follow a circadian rhythm (e.g., Kirschbaum & Hellhammer, 1989) and strongly differ between individuals (Kudielka, Hellhammer, & Wust, 2009). Furthermore, glucocorticoid levels are altered in psychiatric diseases such as depression and posttraumatic stress disorders (PTSD), which are often accompanied by cognitive deficits (Yehuda, 2002; Belanoff, Gross, Yager, & Schatzberg, 2001). The few studies that have investigated the effects of basal cortisol on memory formation in healthy individuals do not show consistent results; positive as well as negative relations between basal cortisol levels and memory for emotional information have been reported (Preuss, Schoofs, & Wolf, 2009; Putman, Van Honk, Kessels, Mulder, & Koppeschaar, 2004; Van Honk et al., 2003). Furthermore, It has been shown that changes in cortisol levels over the study visit are associated with cognitive performance (Lee et al., 2007). In contrast to basal cortisol levels during encoding, to our knowledge, the relation between basal cortisol levels or changes in cortisol levels during retrieval testing and memory recall in healthy young individuals is still unknown.

In this study, we aimed at investigating whether basal cortisol levels as well as changes in basal cortisol levels during recall are related to memory performance in a short-delay and a long-delay episodic memory task in a large population (n = 1225) of healthy young individuals. In addition, we were interested whether we could replicate previous findings of basal cortisol during encoding and memory performance.

METHODS

Participants

We had complete data from 1253 participants. Twenty-eight participants had to be excluded because their cortisol measures exceeded our outlier criterion (4 *SD*s from group mean). Data from 1225 healthy young women and men (812 women, 413 men) between 18 and 35 years (mean age = 22.49 years, *SD* = 3.59 years) were included in the analyses. Of the 812 women participating in the study, 429 women were taking hormonal contraceptives, and 383 women were not taking hormonal contraceptives. Participants were students or employees from the Basel area and were paid for their participation. They did not take any medication (except hormonal contraceptives) and reported no neurological or mental illness. The study was approved by the local ethics committee and all participants gave written informed consent before participation.

Procedure

The experiments were conducted on two consecutive days (Figure 1). On Day 1, participants received instructions and were trained on the tasks. After training, participants viewed emotional and neutral pictures of the picture memory task (Set 1). Afterwards participants performed on a working memory task (n-back). This task was followed by an unannounced free recall test of the previously seen

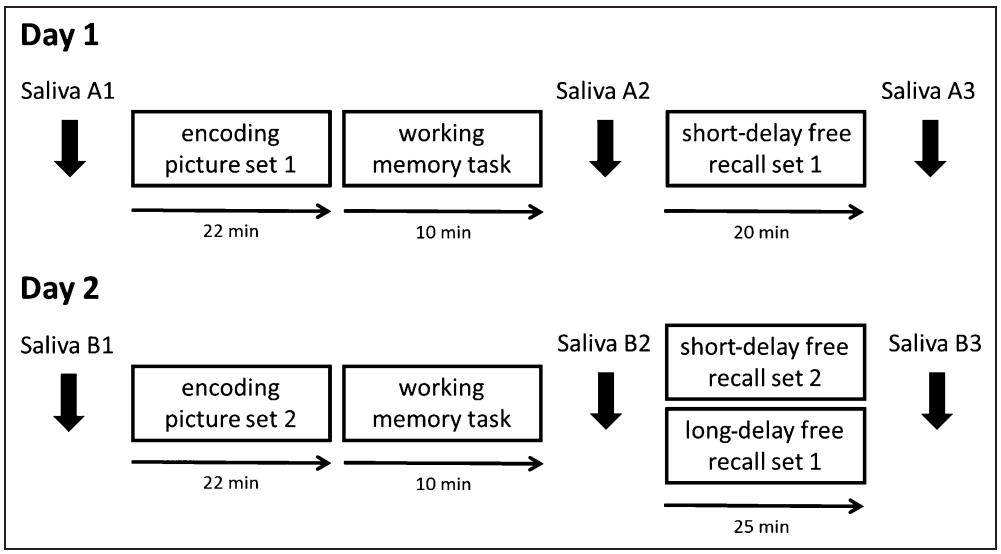


Figure 1. Study design and experimental procedure including point in time of cortisol measurements.

pictures (short-delay recall Day 1). Testing on Day 1 always occurred between 4:00 and 7:00 p.m. On Day 2 testing occurred between 1:00 and 3:00 p.m. Participants completed the same tasks again, although they saw a different set of emotional and neutral pictures (Set 2). On Day 2, participants were asked to freely recall all pictures seen 10 min earlier on the same day (short-delay recall Day 2) and the pictures seen 20 hr earlier on Day 1 (long-delay recall). On both days, saliva samples for cortisol determination were collected three times: before picture encoding, between picture encoding and picture recall, as well as after recall testing.

Picture Memory Task

The picture memory task consisted of 72 pictures taken from the International Affective Picture System (Lang, Bradley, & Cuthbert, 2008) as well as from in-house standardized picture sets. Stimuli consisted of two sets (Set 1 and Set 2) of 24 positive, 24 negative, and 24 neutral pictures interleaved with 24 scrambled pictures. Additionally, four pictures showing neutral objects were presented to control for primacy and recency effects (two pictures were shown in the beginning of the presentation, the other two at the end). These pictures were not included in the analysis. Set 1 was presented on Day 1, and Set 2 was presented on Day 2. The two sets were counterbalanced for ratings of arousal and valence as well as for visual complexity and presence of humans.

The pictures were presented in a quasi-randomized order so that a maximum of four pictures of the same category followed consecutively. A fixation-cross appeared for 500 msec before each picture. Then the picture was presented for 2.5 sec. After presentation of each picture, participants rated the presented picture according to its emotional valence (negative = 1, neutral = 2, positive = 3) and arousal (low = 1, medium = 2, high = 3) on a 3-point scale. Trials were separated by variable intertrial periods (9–12 sec). Participants were not told to memorize the pictures (incidental encoding).

For the free recall task, participants had to write down a short description of each picture. The participants were instructed to recall as many pictures as possible. There was no time limit for this task. Participants were not told how many pictures they saw during picture presentation; therefore, no expectation of the amount of the to-be-recalled pictures was mentioned to the participants. Two independent and blind raters analyzed the recalled pictures and decided for each picture whether it could be recognized as one of the presented pictures. The interrater reliability added up to .96 (Cronbach's alpha). Afterwards a third independent and blind rater decided on pictures, which were rated differently.

Working Memory Task

Between picture presentation and recall, participants performed on the 0- and 2-back versions of the *n*-back working memory task (Gevins & Cutillo, 1993). In this task, letters are presented successively in the center of the screen. In the 0-back condition, participants had to respond to the occurrence of the letter "x," which is a baseline measure of general attention, concentration, and RT. The 2-back task requires participants to respond to a letter repetition with one intervening letter (g - S - f - s). The latter condition required both the maintenance of the last two letters in memory, and updating of these remembered stimuli as each new stimulus was presented. The difference in accuracy between the 2-back and the 0-back condition represents a reliable measure of working memory. *n*-Back data were available for 1100 participants.

Saliva Samples

Cortisol was measured via saliva samples using Salivette collection tubes (Sarstedt, Germany). On both days, saliva samples were taken before picture presentation (Figure 1 and Table 1; Day 1: Sample A1; Day 2: Sample B1) between picture presentation and picture recall (Day 1: Sample A2; Day 2: Sample B2) as well as after recall testing (Day 1: Sample A3; Day 2: Sample B3).

We were interested in the relation between cortisol during retrieval testing or picture encoding, respectively, and recall success and therefore investigated associations in relation to average cortisol levels during encoding (A1 and A2) and retrieval (A2 and A3) on Day 1 as well as on Day 2 (B1 and B2 as well as B2 and B3, respectively). Finally, we examined the relationship between changes in cortisol levels during encoding (A2 minus A1) and during retrieval (A3 minus A2) on Day 1 and Day 2 (B2 minus B1 and B3 minus B2, respectively). Because of circadian rhythm, cortisol levels generally showed a decrease during the experimental sessions. Therefore, the change in cortisol was mostly negative and can be seen as a measure of decrease of cortisol during the tasks.

Cortisol levels were analyzed by the Technical University of Dresden, Germany. For cortisol analysis, saliva samples were centrifuged at 3000 rpm for 3 min after thawing. Concentrations of salivary free cortisol were measured using a commercially available chemiluminescence immunoassay (IBL, Hamburg, Germany) with intra- and interassay precision of 2.5% and 4.7%, respectively.

Statistical Analysis and Data Reduction

Data were analyzed with bivariate Pearson's correlations, partial correlations, repeated-measure ANOVAs, and *t* tests (SPSS Statistics 20.0, 2011). Statistical comparison of correlation coefficients was performed using the software "R" (R Development Core Team, 2012). Recalled pictures are presented as percentage of presented pictures. p values of $< .05$ were considered significant; for correlation analyses, we applied correction for multiple testing: We first calculated the correlations independent of emotional valence, resulting in 16 correlations (eight for average

Table 1. Descriptives of Cortisol Levels, Mean Cortisol, Change in Cortisol, and Memory Performance (n = 1225)

Cortisol Levels Day 1 (nmol/L)			*Cortisol Levels Day 2 (nmol/L)*		
A1	*A2*	*A3*	*B1*	*B2*	*B3*
7.51 ± 4.36	5.32 ± 2.79	4.39 ± 2.22	12.03 ± 6.63	8.91 ± 4.25	6.76 ± 2.96

Mean Cortisol Levels Day 1 (Log-transformed)		*Mean Cortisol Levels Day 2 (Log-transformed)*	
Encoding (Mean A1A2)	*Recall (Mean A2A3)*	*Encoding (Mean B1B2)*	*Recall (Mean B2B3)*
1.72 ± 0.47	1.47 ± 0.45	2.23 ± 0.45	1.96 ± 0.41

Change in Cortisol Day 1 (Log-transformed)		*Change in Cortisol Day 2 (Log-transformed)*	
Encoding (Change A2-A1)	*Recall (Change A3-A2)*	*Encoding (Change B2-B1)*	*Recall (Change B3-B2)*
−0.33 ± 0.23	−0.19 ± 0.18	−0.27 ± 0.24	−0.26 ± 0.21

Short Delay Memory Recall Day 1		*Short Delay Memory Recall Day 2*	
Positive pictures	47.66 ± 14.55%	Positive pictures	47.92 ± 15.71%
Negative pictures	44.29 ± 13.72%	Negative pictures	47.47 ± 15.74%
Neutral pictures	26.79 ± 12.71%	Neutral pictures	33.44 ± 15.86%

	Long Delay Memory Recall Day 2	
	Positive pictures	33.69 ± 15.00%
	Negative pictures	30.99 ± 14.31%
	Neutral pictures	18.87 ± 11.77%

Reported are mean ± *SD*.

cortisol levels and eight for change in cortisol levels). Using Bonferroni correction to correct for multiple testing, a p value of <.003 (i.e., $p < .05/16$) was considered significant. In case of significant correlation after correction for multiple testing, we analyzed correlations for the different valences separately and investigated whether the correlation coefficients of the different emotional valences differed significantly. For exploratory purposes, we also report all correlation coefficients for all valences (Tables 3 and 4). Where not stated differently, values are presented as mean ± *SEM*. Because of the known sex differences in memory recall and cortisol levels, we conducted additional analyses controlling for the influence of sex and use of hormonal contraceptives.

RESULTS

Salivary Cortisol

Because cortisol data (Table 1) were not normally distributed, we used log-transformed data for all analyses. Cortisol levels showed a significant point in time of Cortisol Measurement × Day interaction as well as significant main effects for Point in Time of Cortisol Measurement and Day (all $p < .001$). On both days, cortisol levels decreased over the three measurement points, and on average, cortisol levels were lower on Day 1 (4:00 to 7:00 p.m.; mean = 5.74 ± 0.08 nmol/L) as compared with Day 2 (1:00 to 3:00 p.m.; mean = 9.23 ± 0.12 nmol/L). These findings are in accordance with the well-known circadian variation of cortisol levels.

With respect to change in cortisol levels (decrease in cortisol; lower values indicate larger decrease) during encoding (Day 1: A2-A1; Day 2: B2-B1) and recall (Day 1: A3-A2; Day 2: B3-B2; Table 1), data showed a significant Point in Time of Cortisol Measurement × Day interaction as well as a significant main effect for Point in Time of Cortisol Measurement (all $p < .001$). On Day 1, decrease in cortisol was significantly larger during encoding than during recall (all $p < .001$). On Day 2, change in cortisol during encoding was not different from change during recall ($p = .21$). Decrease in cortisol levels during encoding was larger on Day 1 as compared with Day 2 ($p < .001$), whereas decrease in cortisol levels during recall was larger on Day 2 as compared with Day 1 ($p < .001$).

Given the possible influence of sex and use of hormonal contraceptives in respect to cortisol levels, we additionally compared cortisol levels of women taking

hormonal contraceptives, women not taking hormonal contraceptives, and men. The groups differed with respect to the decrease measures as well as the mean cortisol levels during encoding and retrieval on both days (all $p < .05$; for single comparisons, see Table 2). Women taking hormonal contraceptives generally showed less decrease in cortisol than the other two groups.

Picture Recall

In the short-delay recall conditions we found a significant Picture Valence × Day interaction, as well as significant main effects for Valence and Day (all $p < .001$; Table 1). On both days, participants recalled more emotional than neutral pictures (all $p < .001$). On Day 1, participants recalled more positive pictures than negative pictures ($p < .001$), whereas on Day 2 recall of positive and negative pictures did not differ ($p = .25$). On Day 2, participants recalled more negative and neutral pictures than on Day 1 (both $p < .001$), recall of positive pictures did not differ ($p = .52$).

In the long-delay condition, we found a significant main effect of Picture Valence ($p < .001$). Participants recalled more emotional pictures than neutral pictures (both $p < .001$); furthermore, positive pictures were better recalled than negative pictures ($p < .001$).

Cortisol during Picture Recall

Mean Cortisol Levels

None of the average levels of cortisol during picture recall (mean of A2 and A3, respectively, mean of B2 and B3) was significantly associated with recall performance, neither in relation to long-delay recall nor short-delay recall (all $p_{(uncorrected)} \geq .12$; Table 3). To get a more complete overview, we additionally checked whether single cortisol levels are associated with memory recall. However, none of the correlations reached significance after correction for multiple testing.

Change in Cortisol Levels

We observed a significant association between the decrease in cortisol levels during recall on Day 2 (B3-B2) and long-delay recall performance ($r = -0.13$, $R^2 = 1.69\%$; $p_{(uncorrected)} = .00001$; $p_{(Bonferroni\text{-}corrected)} = .0002$; Figure 2). Because it has previously been found that cortisol effects on memory are particularly pronounced for emotional stimuli, we compared correlations of decrease in cortisol levels with the different picture valences separately. After correction for multiple testing, we found significant correlations for positive and negative picture valences (Table 4); however, the correlation coefficients for the different picture valences did not significantly differ (all $p \geq .15$).

On Day 2, decrease in cortisol levels during recall (B3-B2) also correlated with short-delay recall of pictures on Day 2: $r = -0.11$, $R^2 = 1.21\%$; $p_{(uncorrected)} = .00006$; $p_{(Bonferroni\text{-}corrected)} = .001$; Figure 2). The correlation between cortisol levels during recall on Day 1 (A3-A2) and short-delay recall of pictures reached nominal significance but did not withstand correction for multiple testing (Day 1: $r = -0.08$, $p = .006$; $p_{(Bonferroni\text{-}corrected)} = .10$). Comparing correlations of cortisol levels with the different picture valences separately, on both days we found the highest correlation with short-delay recall of negative pictures, the lowest correlation with positive pictures and correlation with recall of neutral pictures in between (Table 4). On Day 2, the correlation coefficients for short delay of negative pictures and

Table 2. Comparison of Cortisol Levels between Women Taking Hormonal Contraceptives (whc; $n = 429$), Women Not Using Hormonal Contraceptives (wnc; $n = 383$) and Men ($n = 413$)

	whc	*wnc*	*Men*	*Single Comparisons*
Mean Cortisol Levels				
Encoding day 1 (mean A1A2)	1.75 ± 0.42	1.60 ± 0.51	1.81 ± 0.45	wnc < whc; wnc < men
Recall day 1 (mean A2A3)	1.55 ± 0.42	1.30 ± 0.48	1.53 ± 0.42	wnc < whc; wnc < men
Encoding day 2 (mean B1B2)	2.17 ± 0.42	2.24 ± 0.47	2.26 ± 0.47	whc < men
Recall day 2 (mean B2B3)	2.00 ± 0.39	1.92 ± 0.41	1.96 ± 0.43	wnc < whc
Change in Cortisol				
Encoding day 1 (change A2-A1)	−0.24 ± 0.17	−0.37 ± 0.22	−0.37 ± 0.27	wnc < whc; men < whc
Recall day 1 (change A3-A2)	−0.16 ± 0.14	−0.21 ± 0.18	−0.20 ± 0.22	wnc < whc; men < whc
Encoding day 2 (change B2-B1)	−0.17 ± 0.16	−0.33 ± 0.25	−0.33 ± 0.26	wnc < whc; men < whc
Recall day 2 (change B3-B2)	−0.19 ± 0.15	−0.32 ± 0.22	−0.29 ± 0.24	wnc < whc; men < whc

Reported are mean ± *SD*. Cortisol levels are log-transformed. Reported are significant post hoc comparisons (Bonferroni corrected). p values of <.05 are considered significant.

Table 3. Correlations between Average Cortisol Levels and Memory Recall (n = 1225)

	Cortisol Levels Day 1		*Cortisol Levels Day 2*	
	Encoding (Mean A1A2)	*Recall (Mean A2A3)*	*Encoding (Mean B1B2)*	*Recall (Mean B2B3)*
Short-delay Recall Day 1				
Positive				
r	−.00	−.01		
p	.91	.86		
Negative				
r	−.00	−.01		
p	.94	.63		
Neutral				
r	.02	.01		
p	.58	.81		
Short-delay Recall Day 2				
Positive				
r			.00	.01
p			.93	.84
Negative				
r			.05	.02
p			.10	.40
Neutral				
r			−.01	−.03
p			.65	.24
Long-delay Recall				
Positive				
r	−.02	−.01	−.02	−.04
p	.44	.64	.53	.22
Negative				
r	−.01	.00	−.03	−.05
p	.74	.99	.35	.07
Neutral				
r	−.01	−.01	.00	−.02
p	.76	.62	.98	.45

None of the correlations reached significance: Bonferroni-corrected p values ($p < .002$; i.e., 0.05/24) are considered significant.

short delay of positive pictures were significantly different ($t = 2.06$; $p = .04$). All other correlation coefficients for the different pictures valences did not differ (all $p \geq .17$).

To rule out the possibility, that the decrease measure (B3-B2) is influenced by the size of the first cortisol measure B2 (i.e., a larger first measure could lead to a larger decrease), we included B2 as covariate in an additional analysis. Including B2 as covariate did not alter the result pattern (association with long-delay recall performance: $r = -0.14, R^2 = 1.96\%; p_{(\text{uncorrected})} = .000001$; association with short-delay recall performance: $r = -0.11$, $R^2 = 1.21\%$; $p_{(\text{uncorrected})} = .00007$).

Given the differences in responses to stress effects on memory in women using hormonal contraceptives and women not using hormonal contraceptives (e.g., Nielsen, Segal, Worden, Yim, & Cahill, 2013), we additionally conducted separate analyses in women taking hormonal contraceptives and women not using hormonal contraceptives. Descriptively, associations were stronger for women not using hormonal contraceptives ($r = -0.19$, $R^2 = 3.61\%$; $p_{(\text{uncorrected})} = .0003$) than for women taking hormonal contraceptives ($r = -0.10$, $R^2 = 1.0\%$; $p_{(\text{uncorrected})} = .04$); for association between decrease in cortisol levels during recall on Day 2 (B3-B2) and long-delay recall performance. Effects in men were similar to the effects in women not using hormonal contraceptives ($r = -0.17$, $R^2 = 2.89\%$; $p_{(\text{uncorrected})} = .0004$). However, statistically, the correlation coefficients were not significantly different (all $p \geq .22$).

With respect to the association between the decrease in cortisol levels during recall on Day 2 (B3-B2) and short-delay recall performance, correlations were descriptively strongest in men ($r = -0.23$, $R^2 = 5.29\%$; $p_{(\text{uncorrected})} = .000004$), followed by women taking hormonal contraceptives ($r = -0.14$, $R^2 = 1.96\%$; $p_{(\text{uncorrected})} = .003$) and women not using hormonal contraceptives ($r = -0.12$, $R^2 = 1.44\%$; $p_{(\text{uncorrected})} = .02$). However statistically, the correlation coefficients did not significantly differ (all $p \geq .14$).

Cortisol during Encoding

Mean Cortisol Levels

None of the mean levels of cortisol during picture encoding (mean A1 A2 and mean B1 B2, respectively) were significantly associated with picture recall, neither in relation to long-delay recall nor in relation to short-delay recall (all $p_{(\text{uncorrected})} \geq .51$; Table 3).

Change in Cortisol Levels

In contrast to decrease in cortisol during recall, decrease in cortisol during encoding of pictures (A2-A1) did not predict long-delay recall ($r = 0.04$, $p_{(\text{uncorrected})} = .15$; $p_{(\text{Bonferroni-corrected})} > .99$; Table 4). In respect to short-delay recall, we did not find any significant correlations with decrease in cortisol levels during encoding (A2-A1 and B2-B1, respectively; both $p \geq .11$; $p_{(\text{Bonferroni-corrected})} > .99$; Table 4).

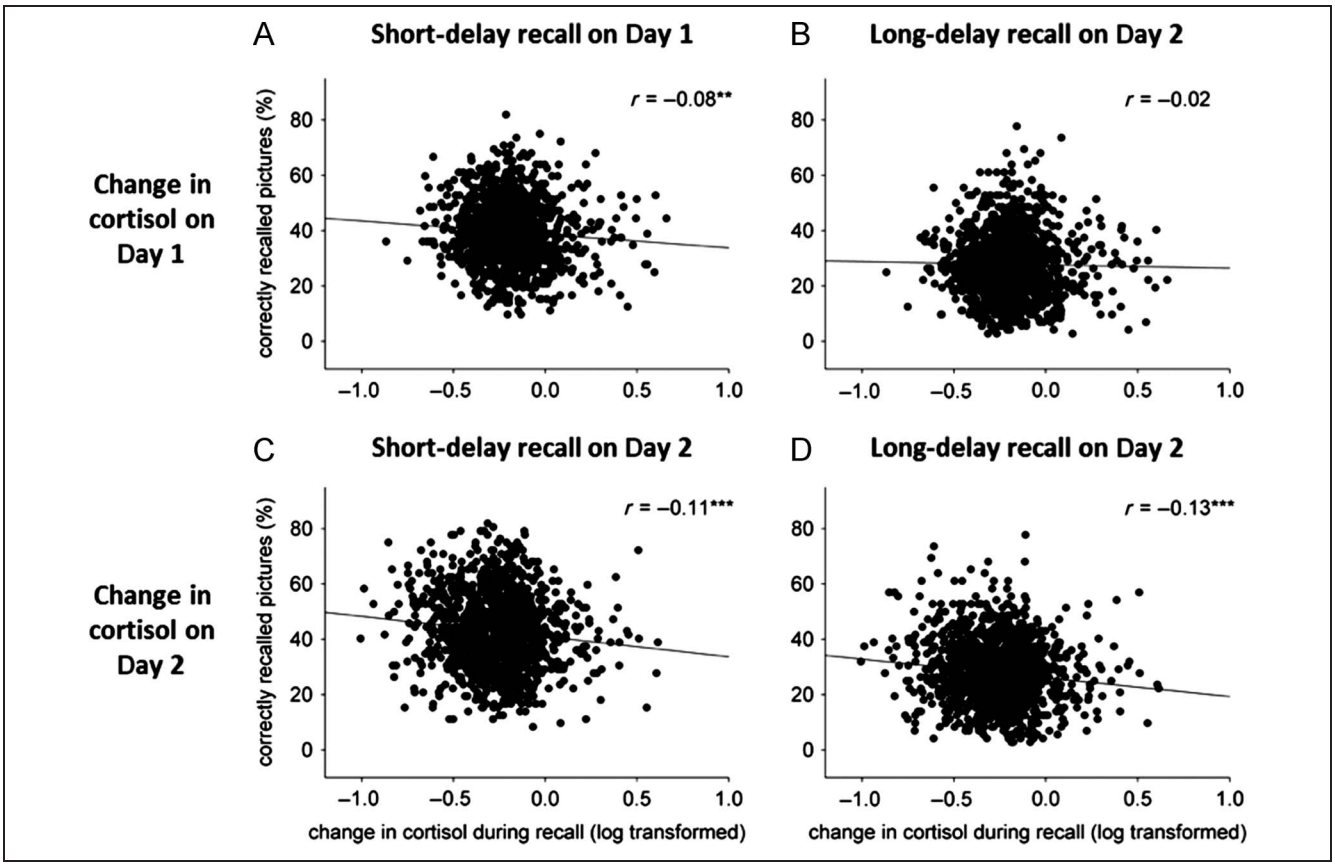


Figure 2. Associations between changes in cortisol levels during picture recall and recalled percentage of presented pictures (independent of valence). (A) Change in cortisol on Day 1 (A3-A2) and short-delay recall on Day 1. (B) Change in cortisol on Day 1 (A3-A2) and long-delay recall on Day 2. (C) Change in cortisol on Day 2 (B3-B2) and short-delay recall on Day 2. (D) Change in cortisol on Day 2 (B3-B2) and long-delay recall on Day 2. **p (uncorrected) < .01; ***p (uncorrected) < .001.

Table 4. Correlations between Change in Cortisol and Memory Recall (n = 1225)

	Cortisol Levels Day 1		*Cortisol Levels Day 2*	
	Encoding (Change A2-A1)	*Recall (Change A3-A2)*	*Encoding (Change B2-B1)*	*Recall (Change B3-B2)*
Short-delay Recall Day 1				
Positive				
r	.04	−.05		
p	.22	.07		
Negative				
r	.02	−.09		
p	.44	.003		
Neutral				
r	.01	−.06		
p	.74	.04		
Short Delay Recall Day 2				
Positive				
r			.08	−.07
p			.008	.01
Negative				
r			.02	**−.12***
p			.58	**<.001**
Neutral				
r			.03	**−.10***
p			.33	**<.001**
Long Delay Recall				
Positive				
r	.05	−.01	.04	**−.10***
p	.11	.71	.16	**<.001**
Negative				
r	.05	−.01	.04	**−.13***
p	.09	.73	.22	**<.001**
Neutral				
r	.00	−.03	−.00	−.09
p	.90	.31	.94	.003

Bonferroni-corrected p values ($p < .002$; i.e., 0.05/24) are considered significant. **Bold** font indicates significant results.

*Significant p values after Bonferroni correction.

In additional analyses controlling for possible effects of gender and use of hormonal contraceptives, results stayed similar; associations generally increased. Results did not change when correcting for valence or arousal ratings.

Working Memory

No significant correlations were found between working memory performance and basal cortisol levels or change in cortisol levels (all $p > .36$).

DISCUSSION

In this study, we investigated possible associations of naturally varying cortisol levels with free recall of emotional and neutral pictures. Stronger decreases in cortisol levels during recall testing predicted better memory recall in the long-delay as well as in the short-delay condition of Day 2 (the correlation with short-delay recall on Day 1 did not reach significance after correcting for multiple testing), independent of picture valence. We did not find any significant results for the average of cortisol during retrieval and recall performance. To have a more complete picture, we additionally investigated correlations of single cortisol levels with recall performance in an exploratory analysis; however, none of the results reached significance after correction for multiple testing. Furthermore, neither average cortisol levels during encoding nor changes in cortisol levels from baseline to encoding predicted memory performance in the short- or long-delay conditions. In respect to the natural variation of cortisol, these results point to an involvement of decrease of cortisol levels in the process of retrieving memories, rather than in memory acquisition. It is to note that the strength of the observed associations—although statistically highly significant—is rather small and explains only roughly 1–2% of the variation in memory performance. However, considering that we did not induce stress but investigated subtle variations in circadian cortisol levels during performance of the tasks, we in fact did expect small rather than large effects.

Our results are in line with previous studies examining glucocorticoid effects on memory retrieval. Administration of glucocorticoids before retrieval testing impaired memories acquired on the day before in animals and humans while leaving immediate recall unaffected (de Quervain et al., 1998, 2000). Increasing glucocorticoid levels by stress induction before retrieval lead to similar effects (de Quervain et al., 2009; Wolf, 2009, for reviews). In studies using different methods to induce an elevation of glucocorticoids, the impairing influence of cortisol on retrieval is particularly pronounced for emotionally arousing stimuli (Smeets, 2011; Smeets, Otgaar, Candel, & Wolf, 2008; Tollenaar et al., 2008; Buchanan et al., 2006; Kuhlmann, Kirschbaum, et al., 2005; Kuhlmann, Piel, et al., 2005). In our study, the association between change in cortisol during recall and recall was independent of picture valence, although the effect was most pronounced for negative pictures. In several previous studies, effects have also been found with respect to recall of neutral stimuli (Smeets, 2011; de Quervain et al., 2000, 2003). Besides differences in the method used to induce an elevation of glucocorticoids, studies also greatly differed in respect to the memory task, mode of recall (e.g., recognition vs. free recall), and time elapsed between encoding and recall.

Considering neuronal activity, the reduction in brain activity in medial-temporal lobe regions was predictive for the degree of memory impairment induced by preretrieval administration of corticosterone (de Quervain et al., 2003). In addition, cortisol effects on memory retrieval depend on concurrent noradrenergic activation of the amygdala, a brain region highly involved in emotional processing that has rich reciprocal projections with hippocampal brain regions (McGaugh, 2004). In human imaging studies, interaction between the hippocampus and the amygdala is greater during retrieval of emotional as compared with neutral information (Smith, Stephan, Rugg, & Dolan, 2006; Dolcos, LaBar, & Cabeza, 2005). Taken together, noradrenergic coactivation and amygdala–hippocampal interactions appear to be a prerequisite for cortisol-induced retrieval impairments, which may also underlie the effects of cortisol on recall performance in our study. However, with respect to change in cortisol levels during picture recall, we not only found an effect on recall of emotional pictures but also on recall of neutral pictures.

The encoding of emotional pictures might have induced an increase in emotional arousal and noradrenergic activity across the encoding of emotional and neutral pictures. This might in part explain why we found an effect not only on recall of emotional pictures but also on recall of neutral pictures. Yet, we did not find an association of cortisol levels with arousal ratings during picture viewing.

Furthermore, the effects of glucocorticoids on memory are mediated by binding to GR and MR. In respect to memory, the ratio of occupation of GR and MR is important (de Kloet et al., 1998). Therefore, we speculate that less cortisol decrease during the task might point to a ratio of GR/MR occupancy that is less sustentative for recall of memory.

Although change in basal cortisol measures sampled during retrieval testing predicted recall of memories, we did not observe a significant association with change in basal cortisol levels measured during encoding of these memories on the previous day. Stress- or pharmacological-induced cortisol elevations during memory formation are typically beneficial for storing emotional memories (Payne et al., 2007; Roozendaal, Okuda, de Quervain, et al., 2006), whereas reports on basal cortisol levels during encoding and memory have been inconsistent (Putman et al., 2004; Van Honk et al., 2003). In a similar experimental approach as in the current study, Preuss and colleagues (2009) reported a positive association between basal cortisol during encoding and free recall of emotional stimuli only when participants knew that their memory would be tested one day later (intentional encoding). The authors reported no association when participants were unaware that they had to remember the pictures later (incidental encoding), which is consistent with our findings, as in our study picture encoding and picture recall were incidental.

In addition to cortisol, the noradrenergic system is critically involved in memory formation (Roozendaal, Okuda, de Quervain, et al., 2006; Roozendaal, Okuda, Van der Zee, et al., 2006; van Stegeren et al., 2005;

Quirarte et al., 1997) and possibly also played a central role in memory formation in this study. As we did not measure noradrenergic activity, future studies need to further examine this important point.

Previous studies have reported substantial gender differences for the relationship between cortisol and memory (Andreano & Cahill, 2006; Jackson, Payne, Nadel, & Jacobs, 2006; Stark et al., 2006; Zorawski, Blanding, Kuhn, & LaBar, 2006; Wolf, Schommer, Hellhammer, McEwen, & Kirschbaum, 2001). We did not find substantial differences between men and women. However, a limitation of the current study is the missing information of women's cycle. There might be differences in the association of cortisol and memory in women in different stages of the cycle. We additionally conducted separate analyses for women taking hormonal contraceptives and women not using hormonal contraceptives given the reported differences in responses to stress effects on memory (e.g., Nielsen, Segal, Worden, Yim, & Cahill, 2013). However, we did not find any significant differences between correlation coefficients between these groups.

Furthermore, age might have an influence on the association between cortisol and memory. In our sample of young healthy individuals with a relatively narrow age range (18–35 years), we did not find an influence of age on the associations between cortisol and memory measures. However, it is possible that the picture might be different when investigating participants across a broader age span. In elderly participants, chronic elevation of cortisol over several years has been associated with worse declarative memory performance (e.g., Lupien et al., 2005). Therefore, it would be interesting to investigate whether chronically elevated cortisol levels in younger individuals are related to later memory complaints at an older age. It has previously been shown that in elderly participants elevated cortisol levels over several years lead to deficits in hippocampus-dependent memory and reduced hippocampal volume (Lupien et al., 1998).

Our results might have some clinical implications. Reduced basal cortisol levels have been observed in patients with PTSD (for a review, see Yehuda, 2002). On the background of the current findings, larger decrease in cortisol during memory recall might be related to facilitated recall of traumatic memories and could therefore influence disease status. Compatible with this notion, attempts to treat PTSD patients with cortisol lead to reduction of PTSD symptoms (de Quervain, 2006; Aerni et al., 2004). Furthermore, people with lower basal cortisol levels are at higher risk of developing PTSD after a traumatic event than people with higher basal cortisol levels (de Quervain et al., 2009). Our findings that less decrease in cortisol during retrieval hinders memory recall also in healthy participants adds to the notion that basal cortisol appears to be an important modulator for the accessibility and resistance of memories. This may be an implication for the development of new treatment options of PTSD. Our study strengthens and extends previous findings of glucocorticoids on memory by showing that also without any drastic experimental manipulation, less reduction in cortisol levels during memory retrieval is related to reduced recall of memories.

Acknowledgments

S. A., D. Q., A. P., and B. R. designed research; S. A., F. H., A. P., D. Q., and B. R. wrote the paper; S. A., F. H., and B. R. performed research and analyzed the data. We thank Stefanie Bartocha, Désirée Bruttin, Tobias Egli, Katrin Meier, Ursina Moor, and Cedric Zeindler for their help in conducting the experiment. This study was supported by grants from the University of Basle to B. R., the Swiss National Science Foundation to D. Q. (PP00P3-123391; CRSIK0_122691) and A. P. (PP00P3-114813; CRSIK0_122691), and the European Science Foundation to D. Q. and A. P. (EUROStress).

Reprint requests should be sent to Sandra Ackermann, Division of Biopsychology, Department of Psychology, University of Zurich, Binzmuehlestrasse 14 Box 05, 8050 Zurich, Switzerland, or via e-mail: sandra.ackermann@psychologie.uzh.ch.

REFERENCES

Aerni, A., Traber, R., Hock, C., Roozendaal, B., Schelling, G., Papassotiropoulos, A., et al. (2004). Low-dose cortisol for symptoms of posttraumatic stress disorder. *American Journal of Psychiatry, 161,* 1488–1490.

Andreano, J. M., & Cahill, L. (2006). Glucocorticoid release and memory consolidation in men and women. *Psychological Science, 17,* 466–470.

Belanoff, J. K., Gross, K., Yager, A., & Schatzberg, A. F. (2001). Corticosteroids and cognition. *Journal of Psychiatric Research, 35,* 127–145.

Buchanan, T. W., Tranel, D., & Adolphs, R. (2006). Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learning & Memory, 13,* 382–387.

Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: Interaction with the degree of arousal at encoding. *Learning & Memory, 10,* 270–274.

de Kloet, E. R., Vreugdenhil, E., Oitzl, M. S., & Joels, M. (1998). Brain corticosteroid receptor balance in health and disease. *Endocrine Reviews, 19,* 269–301.

de Quervain, D. J. (2006). Glucocorticoid-induced inhibition of memory retrieval: Implications for posttraumatic stress disorder. *Annals of the New York Academy of Sciences, 1071,* 216–220.

de Quervain, D. J., Aerni, A., & Roozendaal, B. (2007). Preventive effect of beta-adrenoceptor blockade on glucocorticoid-induced memory retrieval deficits. *American Journal of Psychiatry, 164,* 967–969.

de Quervain, D. J., Aerni, A., Schelling, G., & Roozendaal, B. (2009). Glucocorticoids and the regulation of memory in health and disease. *Frontiers in Neuroendocrinology, 30,* 358–370.

de Quervain, D. J., Henke, K., Aerni, A., Treyer, V., McGaugh, J. L., Berthold, T., et al. (2003). Glucocorticoid-induced impairment of declarative memory retrieval is associated with reduced blood flow in the medial temporal lobe. *European Journal of Neuroscience, 17,* 1296–1302.

de Quervain, D. J., Roozendaal, B., & McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature, 394,* 787–790.

de Quervain, D. J., Roozendaal, B., Nitsch, R. M., McGaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature Neuroscience, 3,* 313–314.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2005). Remembering one year later: Role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences, U.S.A., 102,* 2626–2631.

Gevins, A., & Cutillo, B. (1993). Spatiotemporal dynamics of component processes in human working memory. *Electroencephalography and Clinical Neurophysiology, 87,* 128–143.

Jackson, E. D., Payne, J. D., Nadel, L., & Jacobs, W. J. (2006). Stress differentially modulates fear conditioning in healthy men and women. *Biological Psychiatry, 59,* 516–522.

Joëls, M., Karst, H., DeRijk, R., & De Kloet, E. R. (2008). The coming out of the brain mineralocorticoid receptor. *Trends in Neurosciences, 31,* 1–7.

Kirschbaum, C., & Hellhammer, D. H. (1989). Salivary cortisol in psychobiological research: An overview. *Neuropsychobiology, 22,* 150–169.

Krugers, H. J., Karst, H., & Joels, M. (2012). Interactions between noradrenaline and corticosteroids in the brain: From electrical activity to cognitive performance. *Frontiers in Cellular Neuroscience, 6,* 15.

Kudielka, B. M., Hellhammer, D. H., & Wust, S. (2009). Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology, 34,* 2–18.

Kuhlmann, S., Kirschbaum, C., & Wolf, O. T. (2005). Effects of oral cortisol treatment in healthy young women on memory retrieval of negative and neutral words. *Neurobiology of Learning and Memory, 83,* 158–162.

Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *Journal of Neuroscience, 25,* 2977–2982.

Lang, P. J., Bradley, M. M., & Cuthbert, B. N. (2008). *International Affective Picture System (IAPS): Affective ratings of pictures and instruction manual*. Gainesville: University of Florida.

Lee, B. K., Glass, T. A., McAtee, M. J., Wand, G. S., Bandeen-Roche, K., Bolla, K. I., et al. (2007). Associations of salivary cortisol with cognitive function in the Baltimore memory study. *Archives of General Psychiatry, 64,* 810–818.

Lupien, S. J., De Leon, M., De Santi, S., Convit, A., Tarshish, C., Nair, N. P., et al. (1998). Cortisol levels during human aging predict hippocampal atrophy and memory deficits. *Nature Neuroscience, 1,* 69–73.

Lupien, S. J., Fiocco, A., Wan, N., Maheu, F., Lord, C., Schramek, T., et al. (2005). Stress hormones and human memory function across the lifespan. *Psychoneuroendocrinology, 30,* 225–242.

Lupien, S. J., & McEwen, B. S. (1997). The acute effects of corticosteroids on cognition: Integration of animal and human model studies. *Brain Research Reviews, 24,* 1–27.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual Review of Neuroscience, 27,* 1–28.

Nielsen, S. E., Segal, S. K., Worden, I. V., Yim, I. S., & Cahill, L. (2013). Hormonal contraception use alters stress responses and emotional memory. *Biological Psychology, 92,* 257–266.

Payne, J. D., Jackson, E. D., Hoscheidt, S., Ryan, L., Jacobs, W. J., & Nadel, L. (2007). Stress administered prior to encoding impairs neutral but enhances emotional long-term episodic memories. *Learning & Memory, 14,* 861–868.

Preuss, D., Schoofs, D., & Wolf, O. T. (2009). Associations between endogenous cortisol levels and emotional memory in young women: Influence of encoding instructions. *Stress, 12,* 379–387.

Putman, P., Van Honk, J., Kessels, R. P., Mulder, M., & Koppeschaar, H. P. (2004). Salivary cortisol and short and long-term memory for emotional faces in healthy young women. *Psychoneuroendocrinology, 29,* 953–960.

Quirarte, G. L., Roozendaal, B., & McGaugh, J. L. (1997). Glucocorticoid enhancement of memory storage involves noradrenergic activation in the basolateral amygdala. *Proceedings of the National Academy of Sciences, U.S.A., 94,* 14048–14053.

R Development Core Team. (2012). *R: A language and environment for statistical computing*. Vienna: R Foundation for Statistical Computing.

Reul, J. M., & de Kloet, E. R. (1985). Two receptor systems for corticosterone in rat brain: Microdistribution and differential occupation. *Endocrinology, 117,* 2505–2511.

Roozendaal, B., Hahn, E. L., Nathan, S. V., de Quervain, D. J., & McGaugh, J. L. (2004). Glucocorticoid effects on memory retrieval require concurrent noradrenergic activity in the hippocampus and basolateral amygdala. *Journal of Neuroscience, 24,* 8161–8169.

Roozendaal, B., Hernandez, A., Cabrera, S. M., Hagewoud, R., Malvaez, M., Stefanko, D. P., et al. (2010). Membrane-associated glucocorticoid activity is necessary for modulation of long-term memory via chromatin modification. *Journal of Neuroscience, 30,* 5037–5046.

Roozendaal, B., Okuda, S., de Quervain, D. J., & McGaugh, J. L. (2006). Glucocorticoids interact with emotion-induced noradrenergic activation in influencing different memory functions. *Neuroscience, 138,* 901–910.

Roozendaal, B., Okuda, S., Van der Zee, E. A., & McGaugh, J. L. (2006). Glucocorticoid enhancement of memory requires arousal-induced noradrenergic activation in the basolateral amygdala. *Proceedings of the National Academy of Sciences, U.S.A., 103,* 6741–6746.

Schwabe, L., Joels, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: An update and integration. *Neuroscience and Biobehavioral Reviews, 36,* 1740–1749.

Smeets, T. (2011). Acute stress impairs memory retrieval independent of time of day. *Psychoneuroendocrinology, 36,* 495–501.

Smeets, T., Otgaar, H., Candel, I., & Wolf, O. T. (2008). True or false? Memory is differentially affected by stress-induced cortisol elevations and sympathetic activity at consolidation and retrieval. *Psychoneuroendocrinology, 33,* 1378–1386.

Smith, A. P., Stephan, K. E., Rugg, M. D., & Dolan, R. J. (2006). Task and content modulate amygdala-hippocampal connectivity in emotional retrieval. *Neuron, 49,* 631–638.

Stark, R., Wolf, O. T., Tabbert, K., Kagerer, S., Zimmermann, M., Kirsch, P., et al. (2006). Influence of the stress hormone cortisol on fear conditioning in humans: Evidence for sex differences in the response of the prefrontal cortex. *Neuroimage, 32,* 1290–1298.

Tollenaar, M. S., Elzinga, B. M., Spinhoven, P., & Everaerd, W. A. (2008). The effects of cortisol increase on long-term memory retrieval during and after acute psychosocial stress. *Acta Psychologica, 127,* 542–552.

Tollenaar, M. S., Elzinga, B. M., Spinhoven, P., & Everaerd, W. (2009). Immediate and prolonged effects of cortisol, but not propranolol, on memory retrieval in healthy young men. *Neurobiology of Learning and Memory, 91,* 23–31.

Van Honk, J., Kessels, R. P., Putman, P., Jager, G., Koppeschaar, H. P., & Postma, A. (2003). Attentionally modulated effects of cortisol and mood on memory for emotional faces in healthy young males. *Psychoneuroendocrinology, 28,* 941–948.

van Stegeren, A. H., Goekoop, R., Everaerd, W., Scheltens, P., Barkhof, F., Kuijer, J. P., et al. (2005). Noradrenaline mediates amygdala activation in men and women during encoding of emotional material. *Neuroimage, 24,* 898–909.

Wolf, O. T. (2009). Stress and memory in humans: Twelve years of progress? *Brain Research, 1293,* 142–154.

Wolf, O. T., Schommer, N. C., Hellhammer, D. H., McEwen, B. S., & Kirschbaum, C. (2001). The relationship between stress induced cortisol levels and memory differs between men and women. *Psychoneuroendocrinology, 26,* 711–720.

Yehuda, R. (2002). Post-traumatic stress disorder. *The New England Journal of Medicine, 346,* 108–114.

Zorawski, M., Blanding, N. Q., Kuhn, C. M., & LaBar, K. S. (2006). Effects of stress and sex on acquisition and consolidation of human fear conditioning. *Learning & Memory, 13,* 441–450.

4.3 Polymorphisms of HDAC5 are associated with episodic memory, DNA methylation and C17orf65 mRNA expression

Hartmann, F.R., Milnik, A., Auschra, B., Freytag, V., Spalek, K., Vogler, C., Vukojevic, V., de Quervain, D.J.-F., Papassotiropoulos, A., & Heck, A. *Unpublished manuscript*.

Title: Polymorphisms of HDAC5 are associated with episodic memory, DNA methylation and C17orf65 mRNA expression.

Authors: Francina R. Hartmann[1], Annette Milnik[1,2], Bianca Auschra[1], Virginie Freytag[1], Klara Spalek[3], Christian Vogler[1,2], Vanja Vukojevic[1,4], Dominique J.-F. de Quervain[2,3], Andreas Papassotiropoulos[1,2,4], Angela Heck[1,2]

Affiliations:

[1] Division of Molecular Neuroscience, Department of Psychology, University of Basel, Basel, Switzerland

[2] University of Basel, Psychiatric University Clinics, Basel, Switzerland

[3] Division of Cognitive Neuroscience, Department of Psychology, University of Basel, Basel, Switzerland

[4] Life Sciences Training Facility, Department Biozentrum, University of Basel, Basel, Switzerland

Running title: HDAC5 and human memory performance

Abstract

Histone deacetylation modulates gene expression by removing acetyl groups from the histone tail. Histone deacetylases (HDACs) have been associated with synaptic plasticity and memory formation in animal models. In this study we investigated the association between polymorphisms of the histone deacetylase 5 (*HDAC5*) gene, a member of class II HDACs, and episodic memory performance in two independent samples of healthy humans (screening sample n=1155, replication sample n=869). Two single nucleotide polymorphisms (SNPs) rs184478 and rs11079983 were significantly associated with free recall performance particularly for emotional stimuli in both samples. Furthermore, genotypes of SNP rs184478 were associated with methylation levels in the promoter region of *HDAC5*. No significant association between rs184478 and peripheral mRNA expression levels of HDAC5 in a subsample of the replication sample (n=358) could be found. However rs184478 was significantly associated with expression levels of C17orf65, lying in the vicinity of *HDAC5*. Additional analysis of cortical mRNA expression levels in human post-mortem brain tissue in three publicly available databases revealed associations of HDAC5 cortical expression levels with SNP rs184478. Although our data cannot support *HDAC5* as candidate, these findings show that SNP rs184478 may play a role in episodic memory performance. This study emphasizes the importance of using expression analyses in addition to genotype data to avoid wrong conclusions about the involved genes.

Introduction

Histone acetylation is a key epigenetic mechanism that modulates gene transcription by adding acetyl groups to the lysine residue within the N-terminal of histones. Histone deacetylases (HDACs) remove acetyl groups, thus leading to a less accessible chromatin structure and a suppressed gene expression.[1] HDACs are categorized in two classes (class I and class II) based on sequence similarities.[2] Compared with class I HDACs, which are universally expressed in the central nervous system and the periphery, class II HDACs show more specific expression patterns: In particular, expression patterns of HDAC5 and HDAC4 affect brain regions that are crucial for learning and memory (i.e. hippocampus and amygdala).[3]

Even though the exact mechanisms are not known yet, epigenetic changes impact on synaptic plasticity and memory formation (for review, see[1]). Evidence from rodent models supports a role for HDAC5 in learning and memory: Several studies found that pharmacological inhibition of Hdac goes along with enhanced memory performance in mice[4-6] and HDAC inhibitors have been shown to ameliorate memory deficits in neurodegenerative disorders rodents models.[7-9] However, also the opposite direction of effect was observed: E.g., mice lacking Hdac5 have poorer spatial and associative memory in several behavioral tasks, e.g. the open field test, elevated-plus maze and the Morris water maze.[10] Furthermore, phencyclidine-treated mice showing a schizophrenia-like endophenotype were less impaired in an object recognition task when prefrontal expression of Hdac5 was enhanced.[11] Moreover, mice that were exposed to chronic emotional stimuli (cocaine or/and stress) had lower Hdac5 expression in the nucleus accumbens, and *Hdac5* knockout mice showed a hypersensitivity to cocaine and chronic stress, pointing to a general role of Hdac5 in reward-dependent synaptic plasticity.[12]

Additionally, HDAC5 has been involved both in axon regeneration[13] and in the regulation of adult neurogenesis, controlling maturation and survival of newborn neurons.[14] However, evidence for implication of HDAC5 in learning and memory is ambiguous, since impairments have not been observed in hippocampus-dependent learning and memory in *Hdac5* knockout mice.[15]

Besides a possible role in learning and memory, accumulating evidence from pharmacological research indicates that HDAC5 is involved in the pathogenesis of depression. Hippocampal Hdac5 expression is decreased by the antidepressant imipramine and Hdac5 overexpression in the same area inhibits the antidepressant effect of imipramine on depression-like behavior in mice.[16] Recently it has been shown that Hdac5 expression in the amygdala, the key brain structure for the formation of emotional memory, is increased by antidepressants.[17]

Taken together, different lines of evidence from animal models point to a role of HDAC5 in both memory formation and modulation of emotional content. Up to now no study exists addressing the impact of genetic variants of *HDAC5* on human memory. Against this background we investigated common polymorphisms of the *HDAC5* gene for associations with episodic memory performance as well as for memory of emotional stimuli in healthy humans. We also investigated if the associated polymorphisms modulate the methylation level of CpG sites in the promoter region of *HDAC5.* Furthermore we tested both local and distant effects of associated polymorphisms with mRNA expression levels derived from blood and the brain.

Materials and methods

Screening sample

We included 1155 healthy young subjects from the Basel area (Switzerland) in our study (372 male, 783 female; mean age = 22.52 ± 3.55 (SD) years; range: 18-35 years).[18] All participants gave written informed consent prior to participation and filled in a health status questionnaire. They reported no medication intake and were free of any neurological or psychiatric condition at the time of the experiment.

Replication sample

The replication sample included 869 cognitively healthy young subjects from the Basel area, Switzerland (355 male, 514 female; mean age = 22.51 ± 3.34 (SD) years; range: 18-35 years).[18] Subjects reported to be free of any neurological or psychiatric condition and not to take medication at the time of the experiment and were independent from the subjects in the screening sample. In this sample, the encoding phase of the picture-based episodic memory task and the n-back task (both described in the following section) took place in an MRI scanner, whereas the unexpected free recall task was performed outside the scanner.

The ethics committee of the canton Basel approved the study protocols for both samples.

Episodic memory task: All subjects (screening and replication sample) performed an identical episodic memory task. They were presented 72 emotional and neutral pictures (24 negative, 24 positive and 24 neutral) taken from the International Affective Picture System (IAPS)[19] and from in-house standardized picture sets. Two additional neutral pictures were presented at the beginning and two at the end of the task to control for

primacy and recency effects. These pictures were not included in the analysis. Additionally, 24 scrambled pictures were included. Their background contained the color information of all pictures used in the experiment and was overlaid with a crystal and distortion filter (Adobe Photoshop CS3, Adobe Systems Inc., San Jose, CA, USA). On the foreground geometrical figures of varying shape, size and orientation were shown.

The pictures were presented for 2.5s in a quasi-randomized order. A maximum of four pictures of the same category (e.g. animals, humans, landscape) and valence occurred consecutively. Between the pictures a fixation-cross appeared on the screen for 500ms and the trials were separated by a variable inter-trial period of 9-12s. During this inter-trial period after each picture, subjects rated it according to valence (negative, positive, neutral) and arousal (high, medium, low) on a three-point scale (Self-assessment Manikin, SAM). Scrambled pictures were rated according to their shape (vertical, symmetric, horizontal) and size (large, medium, small). Subjects were not instructed to recall the pictures later (incidental recall).

Working memory: After the episodic memory task, the participants performed the 0- and 2-back versions of the n-back task.[20] In this task letters are presented successively in the center of the screen. In the 0-back version, subjects had to respond to the letter "x" by a button-press, representing a baseline measure of general attention, concentration, and reaction time. In the 2-back version, subject had to respond to a letter repetition with one intervening letter (g – S – f – s). This version required the maintenance of the last two letters in memory and updating of these remembered stimuli as each new stimulus was presented.

After the n-back task, participants were instructed to freely recall as many pictures as they remember by writing down a description of the pictures without any time limit. Two

independent investigators rated the picture descriptions and scored them as correctly recalled if they could assign the description to a presented picture. The inter-rater reliability was 0.96 in the screening and 0.99 in the replication sample (Cronbach's alpha).

Procedural memory (Screening sample only): Subjects performed at the end of the session a finger-tapping task to measure procedural memory.[21] Subjects had to press four numeric keys on the computer keyboard using their non-dominant hand. The sequence "4-2-3-1-4" was repeatedly shown during 30s and the subjects were instructed to repeat this sequence as quickly and accurately as possible. The numeric sequence was displayed in white letters on a black screen throughout the 30s, to hold working memory demand on a minimum. For each button press a white dot appeared on the screen instead of the pressed number. After each round subjects were informed about their performance (number of correct sequences; number of total sequences performed). Subjects completed 12 rounds with a resting period of 30s after each round.

Array-based single nucleotide polymorphism (SNP) genotyping

Saliva was collected for DNA isolation using the Oragene DNA sample collection kit (DNA Genotek Inc., ON, Canada). DNA isolation was done using standard procedures.

For genotyping purposes we used Affymetrix Genome-Wide Human SNP Array 6.0 (Affymetrix Inc., Santa Clara, CA, USA). Samples were processed using Affymetrix® SNP 6 Core Reagent Kit (Affymetrix), following manufacturer recommended standard procedure. The hybridized arrays were washed and stained using GeneChip® Fluidics Station 450 (Affymetrix). Following, the arrays were scanned with GeneChip Scanner (Affymetrix), according to the manufacturer's instructions, using Affymetrix GeneChip

Command Console (AGCC, version 3.0.1.1214). Generation of SNP calls and Array quality control were performed using the command line programs of the Affymetrix Power Tools package (version: apt-1.14.2). According to the manufacturer's recommendation, contrast QC was chosen as QC metric, using the default value of greater or equal than 0.4. Mean call rate for all samples averaged 98.5%. All samples passing QC criteria were subsequently genotyped using the Birdseed (v2) algorithm. Bayesian Clustering Algorithm[22] was applied on genome-wide summary statistics to identify and exclude atypical individual samples. Briefly, considering a combination of two summary statistics, the algorithm infers each sample's posterior probability to belong to the outliers' class. A first outlier assessment was based on genome-wide call rate and heterozygosity rates, for which extreme values may be indicative of a genotyping bias. The second assessment aimed at identifying subjects with unusual ancestry according to the majority of the samples. This was done by projecting the samples' genotypic data on the two first PCA components inferred from HapMap data using YRI, CEU and CHB-JPT populations. Thus we excluded non-Caucasian subjects (screening sample, n=120, replication sample, n=95) as well as subjects with genotyping bias (screening sample, n=19, replication sample, n=19), finally including n=1155 (screening sample) and n=869 (replication sample) in the analysis.

SNP selection

Samples were genotyped using Affymetrix Human SNP-array 6.0. 10 SNPs mapping on *HDAC5* (±20kb) were identified (Supplementary Table 1). Of the 10 SNPs, 5 SNPs did not fulfill our inclusion quality criteria: One SNP deviated from Hardy-Weinberg equilibrium (P(HWE)≤0.05), two SNPs had a minor allele frequency of <1% and four SNPs had a

genotype call rate <99% (Supplementary Table 1). We used the TAGGER software[23] implemented in HapMap to check the coverage of genetic variability of *HDAC5* provided by the Affymetrix SNPs. We used a forced choice method with a MAF threshold of 0.05 and r^2 threshold of 0.8 for the selection of tagging SNPs. Using the HapMap release 27, phase II+III, the five selected SNPs captured 93% of genetic variability of *HDAC5* with a mean r^2 of 0.97.

Statistical analysis

The number of correctly recalled pictures served as phenotype of interest. Gender and age were included as covariates into the analysis.

PLINK software version 1.07 (http://pngu.mgh.harvard.edu/purcell/plink/)[24] was used for genetic analysis. Genetic associations were run under the assumption of an additive model.

The nominal significance threshold was set to $P \leq 0.05$. Correction for multiple testing was done using Bonferroni correction for the number of tests ($\alpha = 0.05/5$). Non-genetic analyses and additional independent t-tests for comparisons of the mean memory performances between genotype groups were done using SPSS (IBM SPSS Statistics, version 20, 2011).

Additionally for both samples the genomic control inflation factor lambda (λ_{GC})[25] was calculated to assess genetic admixture using the WGA Viewer software (http://compute1.lsrc.duke.edu/softwares/WGAViewer/).[26] We calculated λ_{GC} values for both samples, resulting in a $\lambda_{GC}=1.007$ for the screening and $\lambda_{GC}=1.002$ for the replication sample, indicating the absence or only a minor admixture in our samples.

RNA isolation and microarray hybridization

Blood samples (n=380) were collected using PAXgene Blood RNA Tubes (PreAnalytix Qiagen/BD, Switzerland). Total RNA was isolated with the PAXgene Blood miRNA Kit (PreAnalytix, Switzerland). A second purification step was performed with the miRNeasy Micro Kit (Qiagen, Germany). The concentration and quality of the RNA was determined using Nanodrop 2000 (ThermoScientific, USA) and RNA Nano 6000 Kit on Bioanalyzer 2100 instrument (Agilent, USA).

Globin Removal

GLOBINclear™-Human Kit (Ambion, USA) was used for a non-enzymatic depletion of the alpha and beta globin mRNA starting from 1µg of total RNA preparations derived from whole blood, following a standard procedure. Concentration and quality of the GLOBINclear™ treated RNA was assessed as described above. Subsequently, the alpha and beta globin mRNA depletion was measured by qPCR. In brief: for reverse transcription, 350ng of total RNA was denatured for 8 min at 70°C followed by ice incubation in the presence of 25ng Anchored Oligo(dT)20 Primer (Invitrogen, USA) and 75ng Random Decamers Primers (Ambion, USA). In the RT reaction, cDNA was generated in 25µl reaction using the Super RT kit (HT Biotechnology, Santa Cruz, CA USA). Upon completion of the reaction, the volume was adjusted to 200µl in Lambda DNA solution (5ng/µl final concentration; Promega, Fitchburg, WI USA). The primers were designed against splice variants that contain the alpha-Globin gene: alpha-Globin Forward: 5'-GCACGCGCACAAGCT-3', and alpha-Globin Reverse: 5'- GGGTCACCAGCAGGCA-3' (Microsynth, Switzerland). The expression levels were normalized to *RPLPO* gene (human large ribosomal protein) using the following primers: RPLP0-Ex3-4_FW, 5'-

CTCTGGAGAAACTGCTGC-3' and RPLP0-Ex3-4_RV, 5'-CTGATCTCAGTGAGGTCC-3' (Sigma Aldrich, USA). qPCR was performed using the Power SYBR Green PCR Master Mix (Life Technologies, USA) according to standard recommendations, in 12µl final volume of reaction, using 2µl of cDNA template, on RotorGene 6000A instrument (Corbett Research Pty Ltd, Sydney Australia). Cycling conditions were as follows: 95°C, 60s – 40x (95°C, 3s - 56°C, 10s – 72°C, 4s). A melting curve analysis (61°C to 95°C, rising by 0.7°C / 3s) was used to check amplification specificity. Threshold cycles (crossing point) were determined using the Rotor-Gene software version 6.1 (Corbett Research, Australia). *RPLPO* was selected as reference gene for normalization. Expression levels were normalized using a geometric mean level of expression[27]. Fold differences were calculated using the delta-delta Ct method[28] using the qBasePlus software (Biogazelle, Gent, Belgium).

Target Synthesis, Labeling and Hybridization

Target synthesis was performed using Ambion® WT Expression Kit (Ambion, Life Technologies, USA) starting from 250ng of high-quality GLOBINclear™ treated RNA, following the standard procedure. 5.16µg of target cDNA was labeled and prepared for hybridization with the GeneChip® WT Terminal Labeling and Hybridization Kit (Affymetrix, USA). Samples were hybridized to the Affymetrix GeneChip Human Transcriptome Array 2.0 (Cat# 902162) for 16 hours (45°C, 60rpm) in the Hybridization oven 640 (Affymetrix, USA).

The arrays were washed and stained on the Fluidics Stations 450 (Affymetrix) using the Hybridization Wash and Stain Kit (Affymetrix, USA) according to protocol FS450_0001. Scanning was performed on the Affymetrix GeneChip Scanner 3000 7G (Affymetrix, USA).

DAT images and CEL files of the microarrays were generated using Affymetrix GeneChip Command Control software (Affymetrix, USA).

Microarray data analysis

In order to account for technical inter-array variation we performed a full quantile-normalization. Feature summarization was conducted using a median-polish on transcript-level according to the HTA 2.0. lib-set-version 0.3. (Affymetrix Power Tools version: 1.16.0). We calculated a principal component analysis (PCA) based on all resulting transcript values. Based on visual inspection of the graph depicting PC1 vs PC2, we excluded subjects showing the most extreme deviations regarding the principal component scores ($n_{outlier}$=10) and repeated the normalization and quantification procedure. Subsequently, a PCA comprising transcript-level gene expression data of n=370 individuals was calculated and expression values were corrected for PC1 to PC10, sex and age using linear regression.

HumanMethylation450 BeadChip data

Microarray-based DNA methylation analysis was performed at ServiceXS (ServiceXS B.V., Leiden, The Netherlands) on the HumanMethylation450 BeadChip (Illumina, Inc., San Diego, CA, USA). DNA was isolated from blood and the samples were randomized and processed in 6 plates. In brief, bisulfite conversion was performed with 500 ng genomic DNA input using the EZ DNA Methylation Gold Kit (Zymo Research, Irvine, CA, USA). Bisulfite conversion quality control on the samples was performed with DNA qPCR reaction and subsequent melting curve analysis.[29] Bisulfite-converted DNA was processed and hybridized on the HumanMethylation450 BeadChip (Illumina, Inc.), according to the

manufacturer's instructions. Data were extracted and analyzed from the idat files using the R package RnBeads.[30] During preprocessing, the background was subtracted using the R package methylumi (method „noob"[31]). For within-array normalization we used the subset-quantile normalization (SWAN) algorithm.[32] The following probe categories were excluded from the final dataset: non-CpG context probes, probes with a SNP overlapping the target CpG as well as probes with more than 2 overlapping SNPs upstream the CpG target (based on Illumina Infinium 450K annotation file, MAF>0.01), probes associated with sex chromosomes and non-specific probes. Additionally, Greedycut algorithm iteratively removed the dataset probes and samples of the highest impurity (beta values with detection *P*-value below set threshold of 0.05 were considered as unreliable).

Next, beta values per CpG were further step-wise post-processed: 1) Logit-transformation (R-package car[33]). 2) Z-transformation per plate (correcting for plate effects by mean-centering and setting the variance to one). 3) Regressing out the first 8 principal components, which resulted in a correction for technical biases and blood cell counts. The counts for the following cell types were retrieved from whole blood: white blood cell count in total, as well as lymphocyte, neutrophil and mixed counts separately; red blood cell count, number of thrombocytes. Subjects' white blood cell counts were significantly ($p < 6 \times 10^{-5}$) associated with the third, fourth and seventh PCA axis. 4) Regressing out gender and age. 5) Regressing out the imputed genotype probabilities of the 50-mer variant with the highest MAF: Sample-specific polymorphic variants in the 50-mer region were identified by using dbSNP (NCBI dbSNP Human Build 140) and the imputed data of our samples. For the imputed data, haplotype estimation was performed using SHAPEITv2[34] allowing a per individual and a per SNP missing rate for observed markers of max. 5%. After pre-phasing, genotype imputation was performed using IMPUTE v2.3.0,

which imputes missing genotypes using a multi-population reference panel.[35] The integrated variant callset of 1092 individuals from the 1000 Genomes Project (release v3 in NCBI build 37/hg19 coordinates, March 2012) served as panel data. Methylation data was available for 513 subjects of the replication sample. We included all CpG sites in the promoter and the intragenic region of *HDAC5* in the analyses. Pearson correlations were done in R (R Core Team, 2012) to test for genotype-dependent differences in methylation level. Calculations were based on the post-processed CpG signal.

The time interval between measuring the cognitive phenotypes and blood collection was on average 348 days (median 314 days).

Analysis of cortical gene expression data

Data are based on the survey of genetic human cortical gene expression published by Myers et al.[36] Gene expression studies of 193 samples from the cerebral cortex of neuropathologically normal brains were carried out with the Illumina HumanRefseq-8 Expression BeadChip (Illumina Inc., San Diego, CA, USA). For genome-wide genotyping, the Affymetrix GeneChip Human Mapping 500K Array Set was used. The complete data files were downloaded from http://labs.med.miami.edu/myers/. Expression levels of HDAC5 transcript probe GI_21237796 were used as the dependent variable. The genetic association analysis was run under the assumption of an additive model.

RegulomeDB was used to identify DNA features and regulatory elements in non-coding regions.[37]

A second data set of gene expression studies of brains from Late-onset Alzheimer disease patients (n=176) and healthy controls (n=191) was used, published by Webster et al.[38] Genome-wide genotyping was done using the Affymetrix GeneChip Human Mapping 500K

Array Set and gene expression analysis was done using the Illumina HumanRefseq-8 Expression BeadChip (Illumina Inc.).

Results

Associations with episodic memory performance: Screening and replication sample

In the screening sample (n=1155), two SNPs rs184478 and rs11079983 were associated with episodic memory performance, (both SNPs: P=0.004, $P_{Bonferroni}$=0.02; r^2=0.96; Table 1).

We then performed *post hoc* analyses for emotional and neutral pictures separately. Both SNPs were associated with memory for emotional pictures (both SNPs: P=0.002; Table 1), whereas neutral pictures showed no significant association (both SNPs: P>0.05; Table 1). Both negative and positive pictures contributed to the association observed for emotional pictures (rs184478: positive pictures: P=0.002; negative pictures: P=0.03; rs11079983: positive pictures: P=0.003; negative pictures: P=0.02; Table 1). Neither SNP was associated with arousal or valence rating of the pictures (all P-values >0.05).

Both SNPs were associated significantly with memory performance also in the replication sample of 869 healthy young subjects (both SNPs: P=0.03; Table 2). The direction of the effect was the same as in the screening sample, with carriers of the major allele (rs184478: G; rs11079983: C) showing better memory performance. Again, the association was significant for emotional pictures but not for neutral pictures (emotional pictures: both SNPs: P=0.03; neutral pictures: both SNPs: P>0.05; Table 2). We also addressed the specificity of the genetic association findings by testing further cognitive phenotypes. We did not find significant associations between the two SNPs and performance with further cognitive domains, i.e. working memory (n-back) or procedural

memory (Walker) (all *P*s>0.13).

Exclusion of participants with substance abuse as well as daily consumption of cannabis and alcohol improved the associations between episodic memory and SNPs rs184478 and rs11079983 in the screening sample (n=1110; P=0.003 and P=0.002, respectively) and stayed the same in the replication sample (n=864; P=0.025 and P=0.033, respectively).

Association of SNP rs184478 with DNA methylation

Methylation data was available for a subsample of the replication sample (n=513).

We analyzed the association between SNP rs184478 and 19 CpG sites in the promoter and gene region of *HDAC5*. There was a significant association between SNP rs184478 and CpG cg17842157 (*P*=0.0001; $P_{Bonferroni}$=0.002, corrected for 19 tests). cg17842157 resides within an *HDAC5* promoter-associated CpG island (Supplementary Figure 1), which generally shows low methylation levels. At this CpG site, methylation levels were highest in major allele (GG) homozygotes (Figure 1). In this subsample of 513 individuals, methylation level of cg17842157 was not correlated with memory performance (*r*=-0.02, *P*=0.71).

Putative functional relevance of SNP rs184478

A query of bioinformatic databases revealed that SNP rs184478 maps within a Transcription Factor ChIP-seq site, a DNaseI Hypersensitivity Cluster and a H3K4Me1 histone mark as identified by the ENCODE project.[39] Furthermore, SNP rs184478 scores „1b“on RegulomeDB, indicating that it is likely linked to the expression of a gene target C17orf65.

Association of SNP rs184478 with mRNA expression levels

Based on the reported link between SNP rs184478 and methylation, we have furthermore checked if SNP rs184478 has any local or distant effects on gene expression. This hypothesis was tested in a subgroup of the replication sample for which mRNA levels from peripheral blood were available (n=358). Although we could not find any genotype-dependent differences in the mRNA level of HDAC5 (*P*=0.59) we observed a strong association for mRNA expression levels of C17orf65 (*P*=3.305e-15), a transcript lying in the vicinity of HDAC5. This link is in complete concordance with data of the Brain eQTL Almanac (Braineac; http://www.braineac.org/) database, which report the association between SNP rs184478 and the transcript of C17orf65 as best hit (*P*=0.00086).

Next, we checked if the link between HDAC5 expression and genotypes of SNP rs184478 could be tissue-specific. This hypothesis was confirmed in three public available data sets: Firstly, linear regression analysis including gender, age at death, post mortem interval, and the transcript detection rate as covariates, revealed that SNP rs184478 was correlated with cortical HDAC5 expression in the Myer's data set (n = 193)[36], containing cortical tissue of brains from neuropathologically normal humans (*P*=0.018, Figure 2). Secondly, we observed an association between SNP rs184478 and cortical HDAC5 expression in a second database published by Webster et al.[38] (*P*=0.018) containing cortical tissue of brains from late-onset Alzheimer disease patients combined with healthy controls (n=363). Linear regression analysis was corrected for gender, APOE status, age at death, cortical region, day of expression hybridization, institute source of sample, postmortem interval and transcript detection rate. Of note, in this case-control study, the controls were the same used in Myer's data. Thirdly, SNP rs184478 was associated with HDAC5 expression in the occipital cortex in the Braineac data set

(*P*=0.0059). However, the direction of effect in this sample was contrary to the results observed in the datasets published by Myer's as well as Webster, as minor allele carriers showed the lowest expression pattern.

Discussion

Our study reports an association between polymorphisms of *HDAC5* and episodic memory performance but shows as well that the functional relevance of *HDAC5* cannot be supported by our expression data. In two independent samples of cognitively healthy young participants we identified and replicated two SNPs in high linkage disequilibrium (LD), rs184478 and rs11079983, which were associated with free recall performance in an episodic memory task. Major allele carriers showed a better performance compared to non-carriers.

DNA methylation data that were available for a subgroup of the replication sample also showed rs184478 genotype-dependent differences in methylation levels of cg17842157, a CpG site located within a CpG island in the *HDAC5* promoter region. Specifically, minor allele carriers showed lower methylation levels than non-carriers. Next we tested if the observed effect on methylation has regional or distant impacts on gene expression.

In absolute accordance with publicly available databases (Braineac database), we observed massive effects of rs184478 on peripheral gene expression of C17orf65, lying in close vicinity of HDAC5. To date nothing is known about the specific function of C17orf65. HDAC5 expression levels derived from blood did not show any genotype-dependent differences. We next tested tissue specificity of our finding: Converging evidence of HDAC5 expression in brain tissue was observed in three publicly available databases. In the Myers and Webster datasets, minor allele carriers of SNP rs184478 had higher HDAC5

cortical mRNA expression levels. However, in the Braineac database the direction of effect was reversed. This divergence in effect may be explained by the type of tissue in which mRNA were measured: In the Myers and Webster data, mRNA levels were measured in cortical tissue probes, whereas in the braineac database, RNA was extracted by smaller brain regions, i.e. the occipital lobe. Therefore it seems plausible that the tissue-specificity also applies to the particular tissue under investigation, which has been reported by Kang et al.[40] and Ramasamy et al.[41] Furthermore an explanation of the divergence direction of effect could be that the SNP data used in the Braineac database derived from imputed data, which provides only estimates of the genotypes and is therefore more noise.

Divergent results however have also been observed in animal studies. Focusing on the link between minor allele carriers showing lower performance in an episodic memory task parallels the results from several studies, which showed that pharmacological inhibition of Hdac goes along with enhanced memory performance in mice.[4-6] Fischer et al.[7] could show, that chronic injection of the HDAC inhibitor sodium butyrate led to memory recovery and elevated levels of synaptic marker proteins in mice that have developed synaptic and neuronal loss. Other studies, which observed lower performance in cognitive tasks associated with lower HDAC5 mRNA levels, are not in line with these findings.[10, 11]

Although we could find a robust SNP-phenotype association and there might be some tissue-specific effects, the heterogeneous mRNA expression results indicate that a ubiquitous link between SNP rs184478 and expression levels of HDAC5 is not possible and we cannot infer HDAC5 as the appropriate target in our samples. We furthermore hypothesis that the SNP rs184478 may play a role in episodic memory performance, but it seems to be connected to the expression of C17orf65 instead of HDAC5 to which the SNP

has been assigned. This assumption is furthermore supported by the results from RegulomeDB where SNP rs184478 is characterized to be likely linked to the expression of the gene target C17orf65. Although, several studies point to a role of Hdacs in general[6, 42] and particularly of HDAC5[12, 43, 44] for synaptic plasticity, we can only speculate that the analyzed region may have a functional relevance, but we are not able to make any conclusion about the involved mechanisms.

We would like to stress that our study relied on the concept of common tagging SNPs available on the utilized SNP array. Hence, the observed SNP association reflects an LD-dependent signal of a putative functional genetic variant. Providing further replication in independent samples, sequencing efforts could follow to identify the causal genetic variant.

A further limitation is that the methylation and expression data rely on blood, resulting in a proxy tissue approach. However, this issue is thoroughly discussed by Tylee et al.,[45] who state that DNA methylation signatures in genomic regions rich in cytosine-guanine dinucleotides (as it is the case with HDAC5 cg17842157, that is positioned in the CpG island) generally show stable epigenetic signatures across brain and non-brain tissues. Furthermore, several studies could show that peripheral measures of methylation and expression were related to memory performance as well as memory-related brain activity.[46-49]

To conclude, our data cannot support *HDAC5* as candidate for episodic memory. Even though we could find a robust behavioral finding and we assume that SNP rs184478 has some functional relevance. Our results may give some hints that the assignment of SNPs to the closest gene, which up to now has been done, must not necessarily be the best candidate, as SNPs can also influence expression of genes at a distance. This could even

open the possibility to find more functional variants, as up to now only a small fraction of all identified SNPs are proven functional candidates. Furthermore our findings emphasize the importance of adding expression analysis to candidate gene approaches to prevent wrong conclusions when it comes to the analysis of complex genetic traits.

References

1 Day JJ, Sweatt JD. Epigenetic mechanisms in cognition. *Neuron* 2011; **70:** 813-829.

2 Gregoretti IV, Lee YM, Goodson HV. Molecular evolution of the histone deacetylase family: functional implications of phylogenetic analysis. *Journal of molecular biology* 2004; **338:** 17-31.

3 Broide RS, Redwine JM, Aftahi N, Young W, Bloom FE, Winrow CJ. Distribution of histone deacetylases 1-11 in the rat brain. *Journal of molecular neuroscience : MN* 2007; **31:** 47-58.

4 Haettig J, Stefanko DP, Multani ML, Figueroa DX, McQuown SC, Wood MA. HDAC inhibition modulates hippocampus-dependent long-term memory for object location in a CBP-dependent manner. *Learn Mem* 2011; **18:** 71-79.

5 Stefanko DP, Barrett RM, Ly AR, Reolon GK, Wood MA. Modulation of long-term memory for object recognition via HDAC inhibition. *Proc Natl Acad Sci U S A* 2009; **106:** 9447-9452.

6 Vecsey CG, Hawk JD, Lattal KM, Stein JM, Fabian SA, Attner MA *et al.* Histone deacetylase inhibitors enhance memory and synaptic plasticity via CREB:CBP-dependent transcriptional activation. *J Neurosci* 2007; **27:** 6128-6140.

7 Fischer A, Sananbenesi F, Wang X, Dobbin M, Tsai LH. Recovery of learning and memory is associated with chromatin remodelling. *Nature* 2007; **447:** 178-182.

8 Mahgoub M, Monteggia LM. A role for histone deacetylases in the cellular and behavioral mechanisms underlying learning and memory. *Learn Mem* 2014; **21:** 564-568.

9 Penney J, Tsai LH. Histone deacetylases in memory and cognition. *Science signaling* 2014; **7:** re12.

10 Agis-Balboa RC, Pavelka Z, Kerimoglu C, Fischer A. Loss of HDAC5 impairs memory function: implications for Alzheimer's disease. *Journal of Alzheimer's disease : JAD* 2013; **33:** 35-44.

11 Koseki T, Mouri A, Mamiya T, Aoyama Y, Toriumi K, Suzuki S *et al.* Exposure to enriched environments during adolescence prevents abnormal behaviours associated with histone deacetylation in phencyclidine-treated mice. *Int J Neuropsychopharmacol* 2012; **15:** 1489-1501.

12 Renthal W, Maze I, Krishnan V, Covington HE 3rd, Xiao G, Kumar A *et al.* Histone deacetylase 5 epigenetically controls behavioral adaptations to chronic emotional stimuli. *Neuron* 2007; **56:** 517-529.

13 Cho Y, Sloutsky R, Naegle KM, Cavalli V. Injury-induced HDAC5 nuclear export is essential for axon regeneration. *Cell* 2013; **155:** 894-908.

14 Schneider JW, Gao Z, Li S, Farooqi M, Tang TS, Bezprozvanny I *et al.* Small-molecule activation of neuronal cell fate. *Nat Chem Biol* 2008; **4:** 408-410.

15 Kim MS, Akhtar MW, Adachi M, Mahgoub M, Bassel-Duby R, Kavalali ET *et al.* An essential role for histone deacetylase 4 in synaptic plasticity and memory formation. *J Neurosci* 2012; **32:** 10879-10886.

16 Tsankova NM, Berton O, Renthal W, Kumar A, Neve RL, Nestler EJ. Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nat Neurosci* 2006; **9:** 519-525.

17 Ookubo M, Kanai H, Aoki H, Yamada N. Antidepressants and mood stabilizers effects on histone deacetylase expression in C57BL/6 mice: Brain region specific changes. *J Psychiatr Res* 2013; **47:** 1204-1214.

18 Heck A, Fastenrath M, Ackermann S, Auschra B, Bickel H, Coynel D *et al.* Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. *Neuron* 2014; **81:** 1203-1213.

19 Lang PJ, Bradley MM, Cuthbert BN. *International Affective Picture System (IAPS): Affective Ratings of Pictures and Instructional Manual*. University of Florida: Gainesville, FL, 2008.

20 Gevins A, Cutillo B. Spatiotemporal dynamics of component processes in human working memory. *Electroencephalography and clinical neurophysiology* 1993; **87:** 128-143.

21 Walker MP, Brakefield T, Seidman J, Morgan A, Hobson JA, Stickgold R. Sleep and the time course of motor skill learning. *Learn Mem* 2003; **10:** 275-284.

22 Bellenguez C, Strange A, Freeman C, Donnelly P, Spencer CC. A robust clustering algorithm for identifying problematic samples in genome-wide association studies. *Bioinformatics* 2012; **28:** 134-135.

23 de Bakker PI, Yelensky R, Pe'er I, Gabriel SB, Daly MJ, Altshuler D. Efficiency and power in genetic association studies. *Nat Genet* 2005; **37:** 1217-1223.

24 Purcell S, Neale B, Todd-Brown K, Thomas L, Ferreira MAR, Bender D *et al.* PLINK: a toolset for whole-genome association and population-based linkage analysis. *Am J Hum Genet* 2007; **81:** 559-575.

25 Devlin B, Roeder K. Genomic control for association studies. *Biometrics* 1999; **55:** 997-1004.

26 Ge D, Zhang K, Need AC, Martin O, Fellay J, Urban TJ *et al.* WGAViewer: software for genomic annotation of whole genome association studies. *Genome research* 2008; **18:** 640-643.

27 Vandesompele J, De Preter K, Pattyn F, Poppe B, Van Roy N, De Paepe A *et al.* Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biol* 2002; **3:** RESEARCH0034.

28 Pfaffl MW. A new mathematical model for relative quantification in real-time RT-PCR. *Nucleic Acids Res* 2001; **29:** e45.

29 Kristensen LS, Mikeska T, Krypuy M, Dobrovic A. Sensitive Melting Analysis after Real Time- Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. *Nucleic Acids Res* 2008; **36:** e42.

30 Assenov Y, Muller F, Lutsik P, Walter J, Lengauer T, Bock C. Comprehensive analysis of DNA methylation data with RnBeads. *Nat Methods* 2014; **11:** 1138-1140.

31 Davis S, Du P, Bilke S, Triche TJ, Bootwalla M. methylumi: Handle Illumina methylation data. R package version 2.10.0. 2014.

32 Maksimovic J, Gordon L, Oshlack A. SWAN: Subset-quantile within array normalization for illumina infinium HumanMethylation450 BeadChips. *Genome Biol* 2012; **13:** R44.

33 Du P, Zhang X, Huang CC, Jafari N, Kibbe WA, Hou L *et al.* Comparison of Beta-value and M-value methods for quantifying methylation levels by microarray analysis. *BMC Bioinformatics* 2010; **11:** 587.

34 Delaneau O, Zagury JF, Marchini J. Improved whole-chromosome phasing for disease and population genetic studies. *Nat Methods* 2013; **10:** 5-6.

35 Howie BN, Donnelly P, Marchini J. A flexible and accurate genotype imputation method for the next generation of genome-wide association studies. *PLoS Genet* 2009; **5:** e1000529.

36 Myers AJ, Gibbs JR, Webster JA, Rohrer K, Zhao A, Marlowe L *et al.* A survey of genetic human cortical gene expression. *Nature genetics* 2007; **39:** 1494-1499.

37 Boyle AP, Hong EL, Hariharan M, Cheng Y, Schaub MA, Kasowski M *et al.* Annotation of functional variation in personal genomes using RegulomeDB. *Genome Research* 2012; **22:** 1790-1797.

38 Webster JA, Gibbs JR, Clarke J, Ray M, Zhang W, Holmans P *et al.* Genetic control of human brain transcript expression in Alzheimer disease. *American Journal of Human Genetics* 2009; **84:** 445-458.

39 Bernstein BE, Birney E, Dunham I, Green ED, Gunter C, Snyder M. An integrated encyclopedia of DNA elements in the human genome. *Nature* 2012; **489:** 57-74.

40 Kang HJ, Kawasawa YI, Cheng F, Zhu Y, Xu X, Li M *et al.* Spatio-temporal transcriptome of the human brain. *Nature* 2011; **478:** 483-489.

41 Ramasamy A, Trabzuni D, Guelfi S, Varghese V, Smith C, Walker R *et al.* Genetic variability in the regulation of gene expression in ten regions of the human brain. *Nat Neurosci* 2014; **17:** 1418-1428.

42 Haggarty SJ, Tsai LH. Probing the role of HDACs and mechanisms of chromatin-mediated neuroplasticity. *Neurobiology of learning and memory* 2011; **96:** 41-52.

43 Guan Z, Giustetto M, Lomvardas S, Kim JH, Miniaci MC, Schwartz JH *et al.* Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. *Cell* 2002; **111:** 483-493.

44 Otsuki K, Uchida S, Wakabayashi Y, Matsubara T, Hobara T, Funato H *et al.* Aberrant REST-mediated transcriptional regulation in major depressive disorder. *J Psychiatr Res* 2010; **44:** 378-384.

45 Tylee DS, Kawaguchi DM, Glatt SJ. On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *American journal of medical genetics Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics* 2013; **162B:** 595-603.

46 Klengel T, Mehta D, Anacker C, Rex-Haffner M, Pruessner JC, Pariante CM *et al.* Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nat Neurosci* 2013; **16:** 33-41.

47 Menke A, Klengel T, Rubel J, Bruckl T, Pfister H, Lucae S *et al.* Genetic variation in FKBP5 associated with the extent of stress hormone dysregulation in major depression. *Genes Brain Behav* 2013; **12:** 289-296.

48 Ursini G, Bollati V, Fazio L, Porcelli A, Iacovelli L, Catalani A *et al.* Stress-related methylation of the catechol-O-methyltransferase Val 158 allele predicts human prefrontal cognition and activity. *J Neurosci* 2011; **31:** 6692-6698.

49 Vukojevic V, Kolassa IT, Fastenrath M, Gschwind L, Spalek K, Milnik A *et al.* Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *J Neurosci* 2014; **34:** 10274-10284.

Figure legends

Figure 1 Association of SNP rs184478 genotypes with methylation level of cg17842157. The *HDAC5* SNP rs184478 is associated with methylation level of the CpG site cg17842157 in a subsample of the replication sample (n=513) that underwent array-based interrogation of methylation levels. Black bars indicate raw beta values of methylation levels, whereas higher values indicate higher methylation levels; error bars are s.e.m.

Figure 2 Association of SNP rs184478 genotypes with HDAC5 cortical expression levels. The *HDAC5* SNP rs184478 is associated with expression levels (rank-invariant normalized intensities) of the HDAC5 transcript GI_21237796 in the cortices of 193 non-demented deceased subjects. Black bars indicate mean expression levels of GI_21237796; error bars are s.e.m. Statistics were run under the assumption of an additive genetic model.

Figure 1

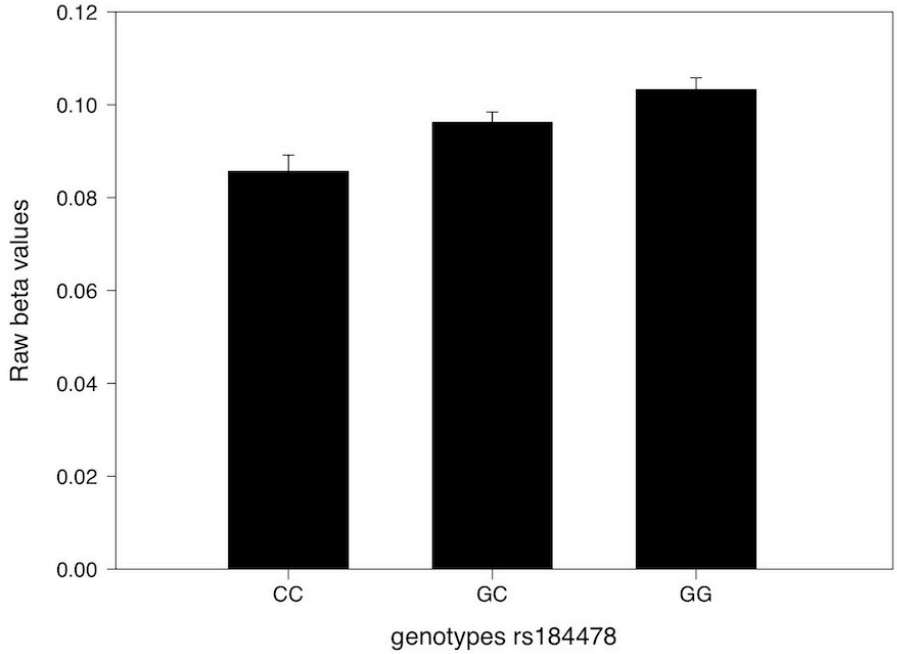


Figure 2

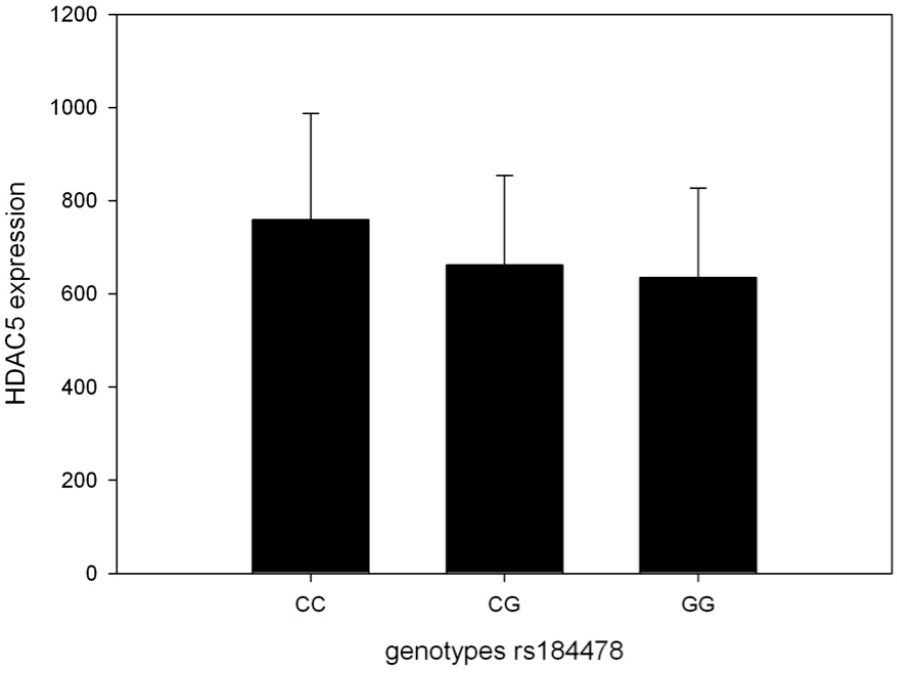


Table 1 Genotype-dependent memory performance in the screening sample[a]

Genotype	*Negative pictures*	*Positive pictures*	*Neutral pictures*	*Emotional pictures*	*All pictures*
rs184478	*mean ± s.e.m.*	*mean ± s.e.m.*	*mean ± s.e.m.*	*mean ± s.e.m.*	*mean ± s.e.m.*
CC (N=88)	10.76 ± 0.35	10.83 ± 0.35	6.19 ± 0.31	21.59 ± 0.61	27.78 ± 0.80
CG (N=432)	10.31 ± 0.15	11.37 ± 0.17	6.34 ± 0.15	21.69 ± 0.28	28.03 ± 0.38
GG (N=631)	10.93 ± 0.13	11.81 ± 0.14	6.58 ± 0.12	22.74 ± 0.24	29.31 ± 0.33
	P = 0.027	*P* = 0.002[b]	*P* = 0.113	*P* = 0.002[b]	*P* = 0.004[b]
rs11079983					
TT (N=91)	10.64 ± 0.35	10.81 ± 0.35	6.19 ± 0.31	21.45 ± 0.61	27.64 ± 0.81
TC (N=433)	10.36 ± 0.15	11.43 ± 0.16	6.33 ± 0.15	21.79 ± 0.28	28.12 ± 0.38
CC (N=631)	10.92 ± 0.13	11.77 ± 0.14	6.60 ± 0.12	22.69 ± 0.24	29.29 ± 0.33
	P = 0.021	*P* = 0.003[b]	*P* = 0.066	*P* = 0.002[b]	*P* = 0.004[b]

[a] Note: Nominal *P*-values of association tests assuming a linear additive model with age and gender as covariates

[b] *P*-values surviving Bonferroni correction

Table 2 Genotype-dependent memory performance in the replication sample[a]

Genotype	*Negative pictures*	*Positive pictures*	*Neutral pictures*	*Emotional pictures*	*All pictures*
rs184478	*mean ± s.e.m.*	*mean ± s.e.m.*	*mean ± s.e.m.*	*mean ± s.e.m.*	*mean ± s.e.m.*
CC (N=61)	11.16 ± 0.37	12.02 ± 0.44	6.89 ± 0.40	23.18 ± 0.72	30.07 ± 1.04
CG (N=330)	11.09 ± 0.18	11.87 ± 0.18	6.98 ± 0.17	22.96 ± 0.32	29.94 ± 0.43
GG (N=477)	11.63 ± 0.15	12.40 ± 0.16	7.33 ± 0.15	24.03 ± 0.27	31.37 ± 0.38
	P = 0.046	*P* = 0.058	*P* = 0.104	*P* = 0.026	*P* = 0.025
rs11079983					
TT (N=62)	11.16 ± 0.37	12.02 ± 0.44	6.95 ± 0.40	23.18 ± 0.71	30.13 ± 1.02
TC (N=330)	11.11 ± 0.18	11.87 ± 0.18	6.99 ± 0.17	22.98 ± 0.32	29.97 ± 0.43
CC (N=477)	11.61 ± 0.15	12.41 ± 0.16	7.32 ± 0.15	24.02 ± 0.27	31.34 ± 0.38
	P = 0.058	*P* = 0.057	*P* = 0.156	*P* = 0.030	*P* = 0.034

[a] Note: Nominal *P*-values of association tests assuming a linear additive model with age and gender as covariates.

Supplementary information

Supplementary Table 1 SNP-IDs, map positions[a], function, allele variants and *P*-values of the test for deviation from Hardy-Weinberg Equilibrium (HWE), minor allele frequency (MAF) and call rates in the screening sample

SNP-ID	Position on hg19	Location/Function	Alleles	HWE	MAF	Call rate
rs228762	42138392	intron (Gene LSM12)	C/T	0.0275	0.153	0.979
rs850857	42154011	downstream gene variant	G/A	0.7074	0.471	0.999
rs190144	42164584	intron	A/G	0.2583	0.254	0.996
rs615070	42176591	intron	G/A	0.3590	0.255	0.993
rs400460	42182770	intron	A/G	0.3678	0.264	0.986
rs228768	42191893	intron	C/A	0.9010	0.317	0.986
rs192028	42197429	intron	A/G	1	0.001	0.979
rs11079983	42198170	intron	T/C	0.1788	0.274	0.999
rs184478	42209751	unknown (UCSC's predicted function: upstream gene variant)	C/G	0.2521	0.274	0.996
rs9915408	42212006	unknown (UCSC's predicted function: upstream gene variant (Gen C17orf53))	C/T	1	0.0007	1

[a]Positions according to the February 2009 Human Reference Sequence (UCSC genome built version hg19)

Legend to Supplementary Figure

Supplementary Figure 1 Promoter and gene region of *HDAC5*. Plots represent gene (orange), promoter region (light blue) and CpG islands (grey) of *HDAC5*. Positions of gene, promoter region and CpG islands are annotated via RnBeadsData package (http://rnbeads.mpi-inf.mpg.de).[1]

A) Association between *HDAC5* SNPs (from left to right: rs850857, rs190144, rs615070, rs11079983, rs184478) and emotional picture recall in the screening (green) and replication (blue) samples. Vertical axis represents minus log-transformed *P*-value of the association test. Horizontal axis represents chromosomal position. The red dotted line indicates nominal significance level (P<0.05). The top associated SNP rs184478 is highlighted and the position is marked by the black dotted line.

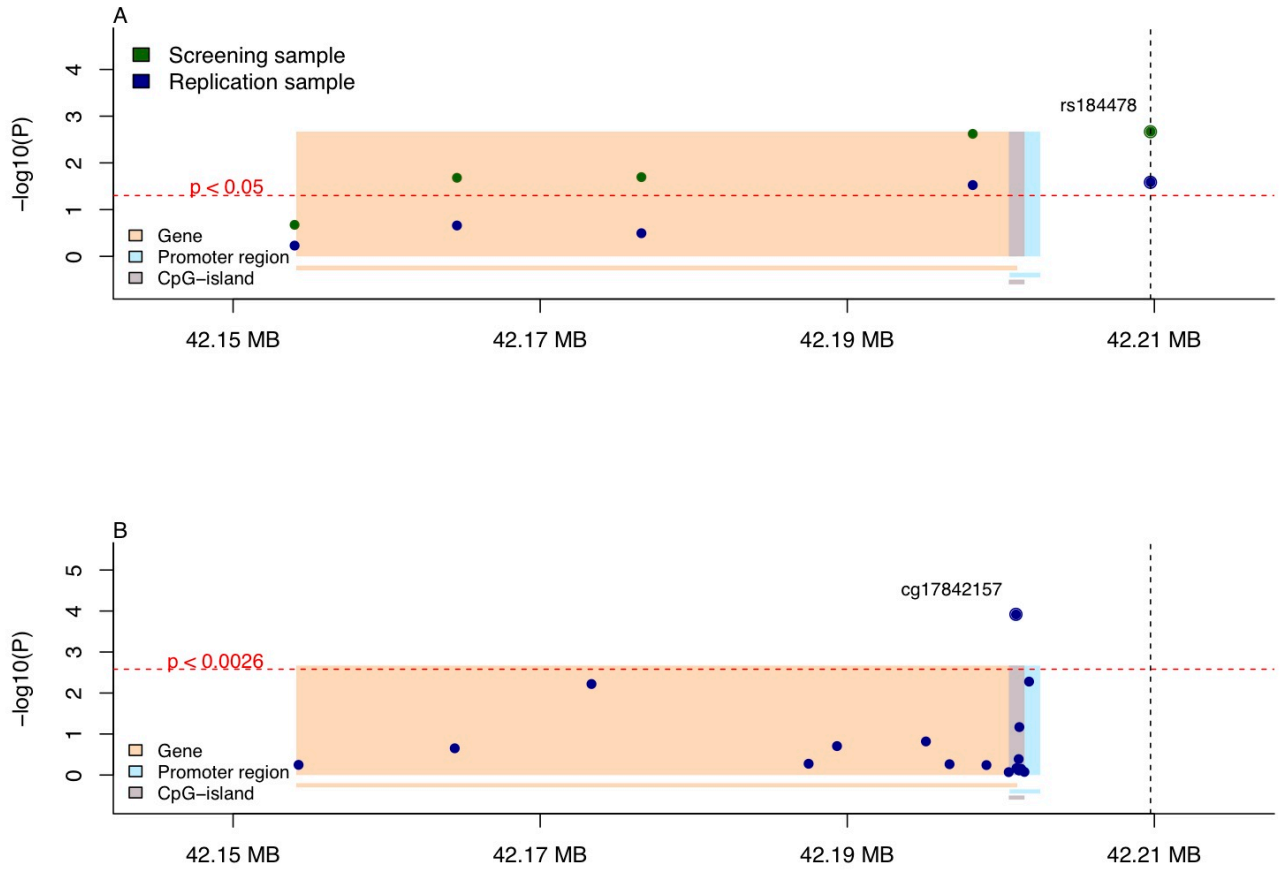
B) Association between SNP rs184478 and DNA methylation at 19 CpG sites. Vertical axis represents minus log-transformed *P*-value of the association test in a subset of the replication sample (N=513). Horizontal axis represents chromosomal position.

The black dotted vertical line indicates the position of rs184478. The red dotted horizontal line indicates Bonferroni corrected significance threshold.

Reference

1 Assenov Y, Müller F, Lutsik P, Walter J, Lengauer T, Bock C. *Compehensive Analysis of DNA Methylation Data with RnBeads* 2014.

Supplementary Figure 1



4.4 No associations between interindividual differences in sleep parameters and episodic memory consolidation

Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D.J.-F., & Rasch, B. (2014). *Sleep, in press*.

No Associations between Interindividual Differences in Sleep Parameters and Episodic Memory Consolidation

Sandra Ackermann, PhD[1,2,3]; Francina Hartmann, MSc[1]; Andreas Papassotiropoulos, MD[1,4,5]; Dominique J.-F. de Quervain, MD[5,6]; Björn Rasch, PhD[2,7]

[1]Department of Psychology, Division of Molecular Neuroscience, University of Basel, CH-4055 Basel, Switzerland; [2]Department of Psychology, Division of Biopsychology, University of Zurich, CH-8050 Zurich, Switzerland; [3]Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric Hospital, University of Zurich, CH-8032 Zurich, Switzerland; [4]Department Biozentrum, Life Sciences Training Facility, University of Basel, CH-4056 Basel, Switzerland; [5]Psychiatric University Clinics, University of Basel, CH-4012 Basel, Switzerland; [6]Department of Psychology, Division of Cognitive Neuroscience, University of Basel, CH-4055 Basel, Switzerland; [7]Department of Psychology, Cognitive Biopsychology and Methods, University of Fribourg, CH-1701 Fribourg, Switzerland

DISCLOSURE STATEMENT: This was not an industry supported study. Financial support was from grants of the University of Basel (to Dr. Rasch).; the Swiss National Science Foundation to Dr. de Quervain, (PP00P3-123391 to Dr. de Quervain, PP00P3-114813 to Dr. Papassotiropoulos, CRSIK0_122691 to Dr. Papassotiropoulos and Dr. de Quervain, and PP00P1_133685 to Dr. Rasch); the European Science Foundation (EUROStress to Dr.Papassotiropoulos and Dr. de Quervain) and from the University of Zurich, clinical research priority project “Sleep and Health” (to Dr. Rasch and Dr. Erich Seifritz). The authors have indicated no financial conflicts of interest.

Short title: Sleep and Episodic Memory in a Large Sample

Submitted for publication July, 2014
Submitted in final revised form September, 2014
Accepted for publication September, 2014

Address correspondence to: Björn Rasch, Department of Psychology, Cognitive Biopsychology and Methods, University of Fribourg, CH-1701 Fribourg, Switzerland; Tel: +41 26 300 76 37; Fax: +41 26 300 97 12; E-mail: bjoern.rasch@unifr.ch; and Sandra Ackermann, Department of Psychology, Division of Biopsychology, University of Zurich, CH-8050 Zurich, Switzerland; Tel: +41 44 635 71 75; E-Mail: sandra.ackermann@psychologie.uzh.ch

Acknowledgments: We thank Erich Seifritz for his support through the clinical research priority project “Sleep and Health.” We thank Sami Abdel Aziz, Stefanie Bartocha, Désirée Bruttin, Tobias Egli, Karen Hellhammer, Nathalie Holz, Marietta Jäger, Carmen Kohler, Katrin Meier, Ursina Moor, Julia Rihm, Sarah Schoch, Myriam Stebler, Manuela Steinauer, Valeria Suhl, Eva Vanova, Irina Wächter, and Cedric Zeindler for their help in conducting the experiments.

Abstract

Study Objectives: Sleep and memory are stable and heritable traits that strongly differ between individuals. Sleep benefits memory consolidation, and the amount of slow wave sleep, sleep spindles, and rapid eye movement sleep have been repeatedly identified as reliable predictors for the amount of declarative and/or emotional memories retrieved after a consolidation period filled with sleep. These studies typically encompass small sample sizes, increasing the probability of overestimating the real association strength. In a large sample we tested whether individual differences in sleep are predictive for individual differences in memory for emotional and neutral pictures.

Design: Between-subject design.

Setting: Cognitive testing took place at the University of Basel, Switzerland. Sleep was recorded at participants’ homes, using portable electroencephalograph-recording devices.

Participants: Nine hundred-twenty-nine healthy young participants (mean age 22.48 $\pm$ 3.60 y standard deviation).

Interventions: None.

Measurements and results: In striking contrast to our expectations as well as numerous previous findings, we did not find any significant correlations between sleep and memory consolidation for pictorial stimuli.

Conclusions: Our results indicate that individual differences in sleep are much less predictive for pictorial memory processes than previously assumed and suggest that

previous studies using small sample sizes might have overestimated the association strength between sleep stage duration and pictorial memory performance. Future studies need to determine whether intraindividual differences rather than interindividual differences in sleep stage duration might be more predictive for the consolidation of emotional and neutral pictures during sleep.

Key Words: declarative memory, rapid eye movement sleep, sample size, sleep EEG, slow wave sleep

Introduction

Sleep and memory are stable traits. Twin studies have revealed that genetic differences account for approximately 50% of the interindividual variance in memory performance,[1] as well as the duration of sleep stages including slow wave sleep (SWS), whereas genetic contribution to rapid eye movement (REM) sleep is less conclusive.[2] Highest heritability values have been revealed for electroencephalographic (EEG) spectral power density as well as REM density (h^2 = 64–96%).[2-4] In addition, sleep is remarkably stable in the same individual across multiple nights whereas sleep largely differs between individuals.[5] Importantly, these interindividual differences are remarkably robust against sleep disturbances, first night effects or prior sleep deprivation.

Memory consolidation profits from sleep after learning.[6,7] In particular, SWS and its characteristic slow oscillatory activity have been implicated in plastic processes underlying the consolidation of declarative memories of events and facts.[6] According to the active system consolidation account, slow oscillatory activity during SWS synchronizes spontaneously reactivated hippocampal memories with thalamocortical spindle activity, thereby facilitating plastic and integrative processes in cortical areas involved in long-term storage.[7] Alternatively, the synaptic homeostasis hypothesis implicates slow wave activity (SWA) during SWS in processes of synaptic downscaling, preparing the brain for new learning the next day.[8] In contrast to SWS, REM sleep and its associated theta activity have been mostly associated with the reprocessing of emotional memories.[9]

Although the mechanism underlying memory consolidation processes during sleep are increasingly understood, it is still an open question whether interindividual differences in sleep parameters are predictive for memory. Several studies have reported strong associations between the amount of nonrapid eye movement (NREM) sleep, SWS, or SWA

during NREM sleep and memory processes during sleep, ranging from r = 0.69 to r = 0.94.[10–12] In addition, sleep spindle number and density were highly correlated with overnight retention of memories as well as general learning ability and intelligence (r = 0.56 to r = 0.68).[13–15] Finally, the amount of REM sleep or REM-associated theta power was highly predictive for recall of emotional memories (r = 0.63 to r = 0.88).[16] However, these correlations are not consistently observed. More importantly, almost all of these studies used very small sample sizes (range of N in the aforementioned studies: N = 6 to N = 31), raising two main issues: (1) Studies using small sample sizes typically overestimate the real effect size, as only very large correlations can reach significance because of a lack of statistical power. The problem increases when multiple sleep and memory parameters are correlated in the same study. (2) Nonsignificant correlations cannot be interpreted because of a lack of statistical power and a high chance of false negatives. The low informative value of results using small sample sizes have been emphasized just recently,[17–19] calling for studies with larger sample sizes and sufficient statistical power to validate current findings.

In the current study, we rigorously investigated how interindividual differences in sleep are related to memory performance in 929 young healthy volunteers. Participants viewed emotional and neutral pictures in the evening and freely recalled them the next day. Sleep was recorded at home using a mobile EEG recording device. We predicted that overnight retention of neutral pictures is significantly associated with the amount of SWS, power of SWA, and sleep spindle density. In addition, we hypothesized that the retention of emotional pictures is positively correlated to the amount of REM sleep and theta activity during REM sleep.

Materials and Methods

Participants

We had complete data from 985 subjects. Fifty-six subjects had to be excluded because their measures (memory or/and sleep parameters) exceeded our outlier criterion (four standard deviations (SD) from group mean). Data from 929 healthy young women and men (633 women, 296 men) between 18 and 35 y (mean age 22.48 $\pm$ 3.60 y [SD]) were included in the analyses. EEG frequency data were available from 885 subjects; data from 44 subjects could not be analyzed because of EEG artefacts. Participants were students or employees from the Basel area and were paid for their participation. They did not take any medication (except hormonal contraceptives), and reported no neurological or mental illness. To get used to wearing a portable EEG recording device, subjects spend a night at home wearing a portable dummy EGG recording device before entering the study. The study was approved by the local ethics committee and all participants gave written informed consent prior to participation.

Procedure

The experiments were conducted on 2 consecutive days (Figure 1). On day 1, after participants had arrived electrodes were applied. Afterward, participants received instructions and were trained on the tasks. After training, participants viewed emotional and neutral pictures of the picture memory task (picture set 1). Afterward, they performed on a working memory task for 10 min (n-Back). This task was followed by an unannounced free recall test of the previously seen pictures (short-delay free recall day 1, picture set 1). The session ended with a finger-tapping task. Testing on day 1 always occurred between 15:30 and 20:00. After testing on day 1, participants spend the night at home wearing a portable

EEG recording device. On day 2, testing occurred between 12:30 and 15:00. Participants came back to the laboratory and viewed another set of emotional and neutral pictures (picture set 2). Previous studies have shown that interference learning before recall makes the effect of sleep on memory consolidation more discernable.[20] After 10 min performance on the working memory task, participants freely recalled pictures from both set 1 (long-delay free recall picture set 1) and set 2 (short-delay free recall day 2, picture set 2) and afterward performed on the recall phase of the finger sequence tapping task (see Figure 1 for a summary of the procedure). Participants were tested in groups of 1-6 individuals.

Picture Memory Task

The picture memory task consisted of 72 pictures taken from the International Affective Picture System (IAPS),[21] as well as from in-house standardized picture sets. Stimuli consisted of two sets (picture set 1 and picture set 2) of 24 positive, 24 negative, and 24 neutral pictures interleaved with 24 scrambled pictures. In addition, four pictures showing neutral objects were presented to control for primacy and recency effects (two pictures were shown in the beginning of the presentation, the other two at the end). These pictures were not included in the analysis. Picture set 1 was presented on day 1, picture set 2 was presented on day 2. The two sets were counterbalanced for ratings of arousal and valence as well as for visual complexity and presence of humans.

The pictures were presented in a quasirandomized order so that a maximum of four pictures of the same category followed consecutively. A fixation-cross appeared for 500 ms before each picture. Then the picture was presented for 2.5 sec. After presentation of each picture, subjects rated the presented picture according to its emotional valence (negative = 1, neutral = 2, positive = 3) and arousal (low = 1, medium = 2, high = 3) on a three-point scale.

Trials were separated by variable intertrial periods (9–12 sec). Participants were not told to memorize the pictures (incidental encoding).

For the free recall task, participants had to write down a short description of each picture. The participants were instructed to recall as many pictures as possible. There was no time limit for this task. Participants were not told how many pictures they saw during picture presentation; therefore, no expectation of the amount of pictures to be recalled was mentioned. Two independent and blind raters analyzed the recalled pictures and decided for each picture whether it could be recognized as one of the presented pictures. The interrater reliability added up to 0.96 (Cronbach α). Afterward, a third independent and blind rater decided on pictures, which were rated differently.

Participants recalled the pictures learned on day 1 (picture set 1) 10 min after encoding (short-delay free recall picture set 1 on day 1) as well as 20 h after encoding (long-delay free recall picture set 1 on day 2). Pictures learned on day 2 (picture set 2) were recalled 10 min after encoding (short delay free recall picture set 2 on day 2; see Figure 1). Overnight memory retention was calculated as relative retrieval performance of picture set 1 with learning performance before the retention interval (short delay recall picture set 1 on day 1) set to 100% (long delay free recall picture set 1 / short delay free recall picture set 1 * 100%).

Working Memory Task

Between picture presentation and recall, participants performed on the 0- and 2-back versions of the n-back working memory task.[22] In this task, letters are presented successively in the center of the screen. In the 0-back condition, participants had to respond to the occurrence of the letter ‘x’, which is a baseline measure of general attention, concentration, and reaction time. The 2-back task requires participants to respond to a letter repetition

with one intervening letter (g – S – f – s). The latter condition required both the maintenance of the last two letters in memory and updating of these remembered stimuli as each new stimulus was presented. We analyzed differences in accuracy between the 2-back and the 0-back condition as this variable represents a reliable measure of working memory.[23] Complete n-back data were available for 857 subjects.

Procedural Memory Task

After memory recall, procedural memory was measured with a motor learning task.[24] The participants worked on a sequential finger-tapping task using their nondominant hand (i.e., with the fingers of the left hand for a right hander and *vice versa*). They had to press four numeric keys on the computer keyboard, repeating the sequence "4-2-3-1-4" as quickly and accurately as possible for 30 sec, followed by a resting phase of 30 sec. The numeric sequence was displayed on the computer screen throughout the task to keep working memory demands at a minimum. Participants could only see a white star displayed on the black screen for every pressed key, but not the actual number pressed. On day 1, participants completed 12 rounds, each round lasted for 30 sec, and 30 sec rest between rounds. On day 2, they completed three rounds with the same sequence of numbers and three rounds with a new sequence of numbers. The key press responses were recorded and for each round the number of overall responses and the number of correct responses was scored. The following measures were calculated: baseline (mean of correct responses of rounds 9 to 12 [on day 1]), recall (mean of correct responses of rounds 1 to 3 [on day 2]), new sequence (mean of correct responses of rounds 4 to 6 [on day 2]), and learning (difference between recall and baseline). Data on procedural memory was available for 902 participants.

EEG Recordings, Sleep Analysis, and Spindle Count

Sleep was recorded at home using a mobile EEG recording device (Somnoscreen Neuro, Somnomedics, Germany). Six Ag-AgCl electrodes were placed according to the international 10–20 System (Fz, C3, Pz, Oz, left and right mastoid). Electrodes were physically referenced to Cz. Additionally, an electroocculogram (EOG), an electromyogram (EMG, chin) and an electrocardiogram (ECG) were recorded for standard polysomnography. Finally, an actimeter was used to monitor movements. EEG signals were recorded between 0.2 – 35 Hz, EOG between 0.2 – 35 Hz, and EMG and ECG between 1 – 128 Hz. The sampling rate for the EEG channels as well as for the EMG and the ECG channels was 256 Hz, and the sampling rate for the EOG channels was 128 Hz.

Sleep Scoring

Analysis of sleep data was restricted to the period between the lights-off and lights-on markers provided by the participants. If participants had forgotten these markers, sleep onset and offset were determined visually. For sleep stage analysis, data were referenced to the right mastoid. Sleep scoring of all sleep data was performed by an automatic algorithm (Somnolyzer 24 × 7) provided by the Siesta Group, Vienna, according to standard criteria.[25] Scoring accuracy of the algorithm has been validated in several studies.[26,27] For the total time in bed every 30-sec epoch was scored as NREM sleep stage 1, 2, 3, 4, or REM sleep with SWS defined by the sum of time spent in sleep stages 3 and 4. Sleep onset was defined by the first period in stage 1 sleep followed immediately by stage 2 sleep. SWS sleep latency and REM sleep latency was determined with reference to sleep onset.

Frequency Analysis

For frequency analysis, EEG data was rereferenced to the averaged mastoids. Data of Cz, which was used as physical reference during data acquisition, were reinstated during re-referencing. Fz, Cz, C3, Pz, and Oz electrodes were then referenced to the averaged mastoids for frequency analyses. Then, sleep scoring data were imported and used as segmentation markers. Data were segmented to 30-sec periods of wakefulness, stage 1 sleep, NREM sleep (consisting of sleep stages 2, 3, and 4) and REM sleep. For frequency analysis, equally sized segments of EEG data consisting of 1,024 datapoints (4 sec) with 100-point overlap were created. Data quality was controlled by using an automatic artifact rejection procedure: segments were kept for further analysis when (1) the maximal difference in EMG activity was < 150 muV, (2) the maximal voltage step in each EEG channel (Fz, Cz, Pz, Oz) was < 50 muV/ms, and (3) the maximal difference in each channel was below 300 muV (500 muV during NREM sleep). Power in each frequency band was calculated in each artifact-free segment for each EEG channel using a fast Fourier transform (FFT) with a 10% Hanning window (resolution 0.25 Hz). Then, power spectra were averaged over all segments. Averaged power was calculated for the following frequency bands: slow oscillation band (0.5 – 0.75 Hz); delta band (0.75 – 4.5Hz); theta band (4.5 – 8 Hz); alpha band (8 – 11 Hz), slow spindle band (11 – 13 Hz); fast spindle band (13 – 15 Hz); beta band (15 – 25 Hz). Because 50-Hz artifacts were greatest for Oz, this electrode was discarded from the analysis.

Spindle Analysis

Spindles (counts and density) during NREM sleep stages were analyzed because of their well-known relationship with overnight retention of memories.[15,28,29] Discrete spindles are a characteristic feature of sleep stage 2 and occur also in SWS, but are virtually absent during REM sleep. Slow (< 13 Hz) and fast spindles (> 13 Hz) were separately identified at the three selected EEG recording sites (Fz, Cz, Pz) during NREM sleep stages 2, 3, and 4, based on an algorithm adopted from previous studies.[15,30] In brief, frequency power was extracted in the frequency bands of interest (10–13 Hz; 13–15 Hz), and the events were counted for which the power signal exceeded a fixed threshold (± 10 muV) for an interval lasting 0.5–3 sec. Spindles were counted separately in each channel during EEG segments free of movement artifacts (maximal EMG difference < 150 muV). Mean spindle counts were calculated by averaging spindle counts of all three channels. To calculate mean spindle density, mean spindle counts were divided by the number of analyzed 30-sec epochs. The two separate spindle bands were chosen based on previous studies that demonstrated the presence of two kinds of spindles in humans possibly linked to different aspects of cognitive function, i.e., slow spindles that prevail over the frontal cortex and show greater topographical variability than the fast spindles that concentrate over the parietal cortex.[31,32]

REM Analysis

Average REM density was calculated by dividing the number of 1-sec periods during REM sleep that contained REM by the total number of 1-sec REM sleep epochs.[33] REM during REM sleep was detected automatically and defined as rapid signal changes in the EOG channel (> 0.8 mV/s) after movement artefact rejection and application of a 50-ms moving average.

Statistical Analysis and Data Reduction

Data were analyzed with bivariate Pearson correlations, partial correlations, regressions, repeated-measures analysis of variance, and *t* tests (SPSS Statistics 19.0). Statistical comparison of correlations coefficients was performed using the software 'R'.[34] Recalled pictures are presented as percentage of presented pictures. Overnight memory retention was calculated as relative retrieval performance, with learning performance before the retention interval set to 100%.

$P < 0.05$ was considered significant; where appropriate we corrected for multiple testing using Bonferroni correction. According to our 5 *a priori* hypotheses, we used a Bonferroni corrected significance level of $p_{Bonferroni} = 0.05$ ($p_{nominal} = 0.05 / 5 = 0.01$) for statistical testing of the critical correlation coefficients. At this significance level, the statistical power for detecting correlations coefficients as low as $r = 0.2$ was $1\text{-}\beta > 99\%$ in our study. Thus, in the case of nonsignificance, the nonexistence of correlations $r > 0.2$ can be inferred with a > 99% certainty. The significance level for all other correlations as well as for group comparisons was set to $P = 0.05$.

Unless indicated differently, values are presented as mean ± standard error of the mean. Because we found significant sex differences in memory recall and sleep measures, we conducted all analyses controlling for the influence of sex. We also detected a significant association of various sleep parameters with age; therefore, we additionally controlled for influences of age.

Results: Pilot Study

In a pilot study (n = 55; 18 men; mean age 24.73 y ± 3.55 [SD]), we first checked whether the used memory task is sleep dependent. Participants either encoded the pictures in the evening and recalled them after sleep (sleep group), or encoded the pictures in the morning and recalled the pictures in the evening (wake group). Sleep was not recorded and the participants did not complete any tasks other than encoding and recall of pictures. Memory retention was calculated as relative retrieval performance with learning performance before the retention interval (short-delay free recall of picture set 1) set to 100%. Participants who slept after picture encoding (n = 27) remembered significantly more pictures (75.69 ± 3.29 %) than participants being awake (n = 28; 59.68 ± 3.39 % $t(53) = 3.39$, $P = 0.001$). Importantly, the benefit of sleep on memory retention did not differ between the three valence categories (+13.78 %, +16.99 %, +19.67 %, for negative, positive, and neutral pictures, $P = 0.80$). In addition, we did not find any differences in learning performance of set 1 or 2 nor between learning performance of participants who slept after picture encoding and those being awake (short delay after 10 min, all $P > 0.52$, see Supplementary Table 1 for descriptive values), excluding possible time of day effects. Thus, memory consolidation in this experimental paradigm clearly benefits from sleep in a valence-independent manner. However, because we used a between-subject design in our pilot study and do not have a control group that stays awake during the retention period overnight, possible effects of circadian rhythm on long delay recall cannot entirely be excluded.

Results: Main Study

Picture Memory Performance

In general, emotional pictures were significantly better remembered than neutral pictures in both short- and the long-delay condition (all $P < 0.001$). However, overnight retention scores (with performance at learning set to 100%) did not differ between the three valence categories (negative pictures: 69.56 ± 0.74%, positive pictures: 69.22 ± 0.71%, neutral pictures: 68.99 ± 1.00%, $P = 0.82$). Compared to men, women had higher retention scores for positive pictures ($P < 0.001$) and marginally higher retention scores for negative pictures ($P < 0.05$). They did not differ with respect to the retention score for neutral pictures ($P = 0.23$). Age did not influence overnight retention scores (all $P > 0.12$). Memory variables conformed to a normal distribution (see Supplementary Figure 1). Subjective arousal ratings of the pictures were highest for negative pictures (set 1: 2.33 ± 0.01; set 2: 2.25 ± 0.01), medium for positive (set 1: 1.84 ± 0.01; set 2: 1.76 ± 0.01), and lowest for neutral pictures (set 1: 1.37 ± 0.01; set 2: 1.31 ± 0.01, $F(2,1830) = 4010.13$, $P < 0.001$). Pictures seen on day 1 were generally rated as more arousing than pictures seen on day 2 ($P < 0.001$).

Sleep Parameters

The distribution of %SWS and %REM as well as fast sleep spindle density conformed to a normal distribution. The distribution of power values of SWA during NREM as well as theta during REM sleep were asymmetric and were therefore log-transformed to conform to a normal distribution (see Supplementary Figure 2). Generally, women had a longer sleep duration than men ($P < 0.001$; Supplementary Table 2) and a trend for a higher percentage SWS ($P = 0.10$). Men and women did not differ in relation to percentage of REM sleep ($P =$

0.12). Furthermore, women showed higher SWA during NREM sleep and higher sleep spindle density than men (both P < 0.001; Supplementary Table 2). Despite the rather narrow age range (i.e., 18–35 y; mean age 22.48 ± 3.60 y [SD]), older participants had significantly lower percentage of SWS (r = -0.28, P < 0.001), lower SWA during NREM sleep (r = -0.40, P < 0.001), lower theta activity during REM sleep (r = -0.25, P < 0.001) as well as a lower spindle density during NREM sleep (r = -0.15, P < 0.001). Conversely, age correlated positively with percentage stage 2 sleep (r = 0.17, P < 0.001) and percentage REM sleep (r = 0.08, p = 0.01). Based on these results, all calculated associations between memory and sleep were corrected for age and sex.

Associations Between Overnight Memory Consolidation and Sleep Parameters

Episodic Memory and Sleep

Unexpectedly, we could not confirm our hypotheses regarding an association between the overnight retention of neutral pictures and sleep parameters of NREM sleep. Neither %SWS (r = -0.01, $p_{nominal}$ = 0.70, $p_{Bonferroni}$ > 0.99), SWA (r = -0.05, $p_{nominal}$ = 0.17, $p_{Bonferroni}$ = 0.85) nor spindle density (r = -0.01, $p_{nominal}$ = 0.81, $p_{Bonferroni}$ > 0.99) predicted overnight retention of neutral pictures. Similarly, no correlation was observed when using overnight retention of all pictures independently of their valence (all -0.05 ≤ r ≤ 0.01 , Figure 2; Table 1). In addition, we did not observe any correlation with memory for emotional pictures and overnight theta activity during REM sleep (r = -0.02, $p_{nominal}$ = 0.60, $p_{Bonferroni}$ > 0.99). We observed a nominally significant association between retention of emotional pictures and %REM sleep, which were unexpectedly negatively correlated (r = -0.60, $p_{nominal}$ = 0.03, $p_{Bonferroni}$ = 0.15). A similarly directed correlation was also observed for neutral pictures (r = -0.07, $p_{nominal}$ = 0.05, $p_{Bonferroni}$ = 0.25; Supplementary Table 3) and for all pictures

independent of their valence (r = -0.08, $p_{nominal}$ = 0.02, $p_{Bonferroni}$ = 0.10, Figure 2, Table 1). However, these correlations did not withstand Bonferroni correction. Thus, in striking contrast to our expectation, none of the expected correlations between sleep parameters (% SWS, SWA during NREM sleep, spindle density during NREM sleep, %REM sleep, and theta activity during REM sleep) and overnight memory retention reached significance.

In addition, we analyzed whether sleep parameters correlate with short-delay recall (Table 1). Correlations between short-delay recall and %REM sleep reached Bonferroni-corrected significance (short-delay recall day 1: r = 0.108, $p_{nominal}$ < 0.001, $p_{Bonferroni}$ < 0.005; short-delay recall day 2: r = 0.094, $p_{nominal}$ = 0.004, $p_{Bonferroni}$ = 0.02). Also, the correlation between short-delay recall on day 1 and theta activity during REM sleep reached significance (r = -0.085, $p_{nominal}$ = 0.01, $p_{Bonferroni}$ = 0.05).

When correcting for short-delay recall performance on day 1, the correlation between overnight memory retention and %REM sleep reached significance (r = -0.101, $p_{nominal}$ = 0.003, $p_{Bonferroni}$ = 0.02). However, correlations between overnight memory retention and other sleep parameters stayed nonsignificant also when correcting for short-delay recall performance on day 1.

In further exploratory analyses, neither total sleep time, nor %wake, nor %N1, nor %N2, nor SWS latency, nor REM latency, nor slow or fast spindle density during NREM sleep, nor REM density was associated with overnight memory consolidation (all $p_{Bonferroni} \geq 0.10$; Supplementary Table 3).

Control Measures

No correlation was observed between %SWS, %REM, %N2, SWA, theta activity during REM sleep, or spindle density during NREM sleep and overnight retention in the procedural

memory task or the overnight improvement (accuracy and reaction time) in the working memory task (all $-0.05 \leq r \leq 0.06$, all $p_{nominal} > 0.05$).

Discussion

In the current study we show that although the picture memory task we used is clearly sleep dependent, recall performance in this task is not significantly associated with time spent in SWS or REM sleep. In fact, no associations between sleep and pictorial memory parameters reached significance after correction for multiple testing, except for a small negative correlation (only explaining roughly 1% of variance) between overnight memory retention and REM sleep when additionally correcting for short-delay recall on day 1.

Our results do not contradict the notion that sleep plays an important role for consolidating memories, and the results of our pilot study clearly replicate the improved recall of pictures after a retention period of sleep as compared to wakefulness (in a valence-independent manner). The nonfindings of sleep-memory associations in our study cannot be attributed to missing sleep dependency of the memory task we used in the current study. The type of recall, i.e., free recall, also does not explain our null findings, as beneficial effects of sleep on free recall have been repeatedly shown.[7] Indeed, sleep effects on declarative memory are most pronounced for cued recall and free recall but are smaller for recognition tests.[35] Moreover, using a declarative as well as a procedural memory task should not diminish the effects, as in previous studies using both tasks a clear benefit of sleep for the declarative memory task has been shown.[36,37]

However, our null results concerning sleep and memory associations clearly show that interindividual differences in specific sleep parameters (i.e., time spent in SWS or REM sleep, SWA during NREM sleep, theta during REM sleep, sleep spindle density) do not reliably predict interindividual differences in memory retention of emotional and neutral pictures across sleep. Please note that we did not test whether intraindividual differences in the amount of sleep stages (i.e., variations in sleep architecture across multiple nights in the same subject) are related to intraindividual variations in memory performance. The possible relation between sleep and memory measured in multiple occasions in the same participant might possibly be much more important for memory processes occurring during sleep as compared to interindividual differences and remains to be determined in future studies.

In many previous studies on interindividual correlations between sleep and memory parameters, large correlation coefficients have been found in relatively small sample sizes. With respect to declarative memory, several studies have reported correlations between different hippocampus-dependent declarative memory tasks and NREM sleep[10–12] or spindle activity.[13,15,38] In addition, associations between REM sleep and emotional memory have been reported.[16] Furthermore, motor skill learning has been associated with stage N2 NREM sleep.[28] Associations of motor skill learning with sleep spindles have been found as well.[39,40] Often the reported correlations between sleep and memory parameters are large; in the aforementioned studies the reported r of significant correlations ranges from r = 0.5 and r = 0.9; explaining approximately 25% to 80% of the variance. All samples in which these correlations between memory parameters and sleep have been found are rather small (in the aforementioned studies sample sizes ranges between N = 6 and N = 31). However, in studies with small sample sizes, effects are often overestimated because of larger random variations in the sample from the true association in the population.[19] Overestimation is

even more likely when multiple testing is performed; for example, when comparing duration of different sleep stages (or night halves, night quarters, etc.) with several memory parameters, the likelihood of reporting false positives is increased. Unfortunately, an appropriate correction for multiple comparisons is also not always done in these studies. Finally, because of the small sample size, nonsignificant correlations cannot be reliably rejected because of a lack of statistical power.[17,18] These circumstances may call into question the value of the reported effects and may to some extent explain the partially inconsistent findings seen in the current sleep and memory literature concerning correlations between different sleep stages and memory consolidation parameters. It is important to note that in most of the aforementioned studies, the reported association between sleep and memory parameters is typically not the main result of these studies, possibly explaining less rigorous testing of correlations.

In our exploratory analysis we also investigated associations between sleep time and memory parameters and did not find any significant results. Likewise, previous studies show that sleep time is not related to measures of cognitive performance nor brain size across mammalian species.[41,42] As an example, guinea pigs and baboons have about the same sleep duration.[41] Another example shows that bats sleep about 18 to 20 h a day whereas elephants sleep only about 3 to 4 h a day. [42]

Our study has some important limitations that limit the generalizability of our results. First, encoding in the current study was incidental, and our participants were not informed that they had to recall the pictures again after sleep. Previous studies have shown that effects of sleep on intentional encoded memory are larger than on incidentally encoded memory.[35] Moreover, only if participants were told of the future relevance of the memories after incidental encoding, sleep benefited memory consolidation as compared to a retention

period filled with wakefulness, and only then memory was correlated with SWA and spindle count during SWS.[43] However, also in our study recall of pictures was improved after a retention interval filled with sleep as compared to wakefulness, excluding that the incidental encoding and lack of future relevance completely abolished the positive effect of sleep on memory in our study.

Second, our null findings are limited to memory for neutral and emotional pictures, and cannot necessarily be generalized to memory for associative stimuli (e.g. word-pairs) or spatial locations. However, free recall of pictures clearly is a declarative and episodic memory task involving hippocampal areas, and current theories on sleep and memory processes do not clearly separate episodic memory for pictures from other types of episodic memory. Future studies will need to test the existences of interindividual associations between sleep and memory parameters also for other learning tasks. Please note that we also did not find any correlations between sleep parameters and consolidation in a procedural sequential finger- tapping task.

Third, in contrast to many sleep and memory studies, study participants did not sleep right after learning. In our study, approximately 6 h passed between encoding and sleep onset, and during this time, participants were on their own and not in the laboratory. It might be possible that time spent in different sleep stages only predicts consolidation of memories that are encoded shortly before sleep. However, in our view, this would in fact strongly diminish the importance for memory consolidation processes occurring during sleep for memories encoded during the day.

Finally, we only recorded 1 night with polysomnography. Therefore, we cannot tell what effects are state effects and what effects are trait effects. However, sleep EEG is a stable marker,[5] and study participants underwent an adaptation night wearing a dummy EEG

recording device and therefore were used to wearing the portable EEG recording device. As mentioned previously, to be able to distinguish intraindividual and interindividual effects of sleep parameters on memory, several nights of sleep EEG should be recorded. Furthermore, a comparison of subjects showing normal sleep patterns with subjects having sleep disturbances might have shed more light on the association of sleep parameters and memory consolidation and the role of normal versus disturbed sleep.

In conclusion, we found that interindividual differences only contribute to a small extend to the effect of sleep on memory consolidation. The current results point to a stronger involvement of intraindividual differences in sleep in respect to overnight memory consolidation as compared to interindividual factors and lead to the question whether sleep in general is important for memory consolidation but not so much the various sleep stages *per se*.

Abbreviations

ECG, electrocardiogram

EEG, electroencephalography

EMG, electromyogram

EOG, electrooccculogram

IAPS, International Affective Picture System

N2, stage 2 sleep

REM, rapid eye movement

SWA, slow wave activity

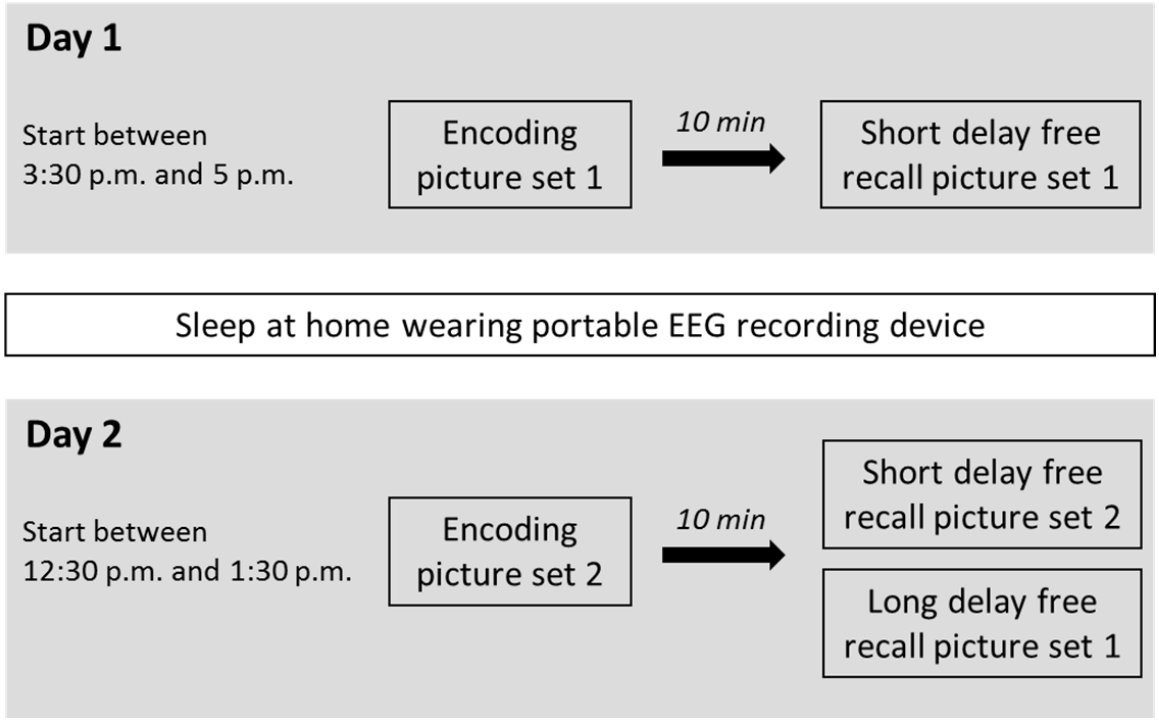
SWS, slow wave sleep

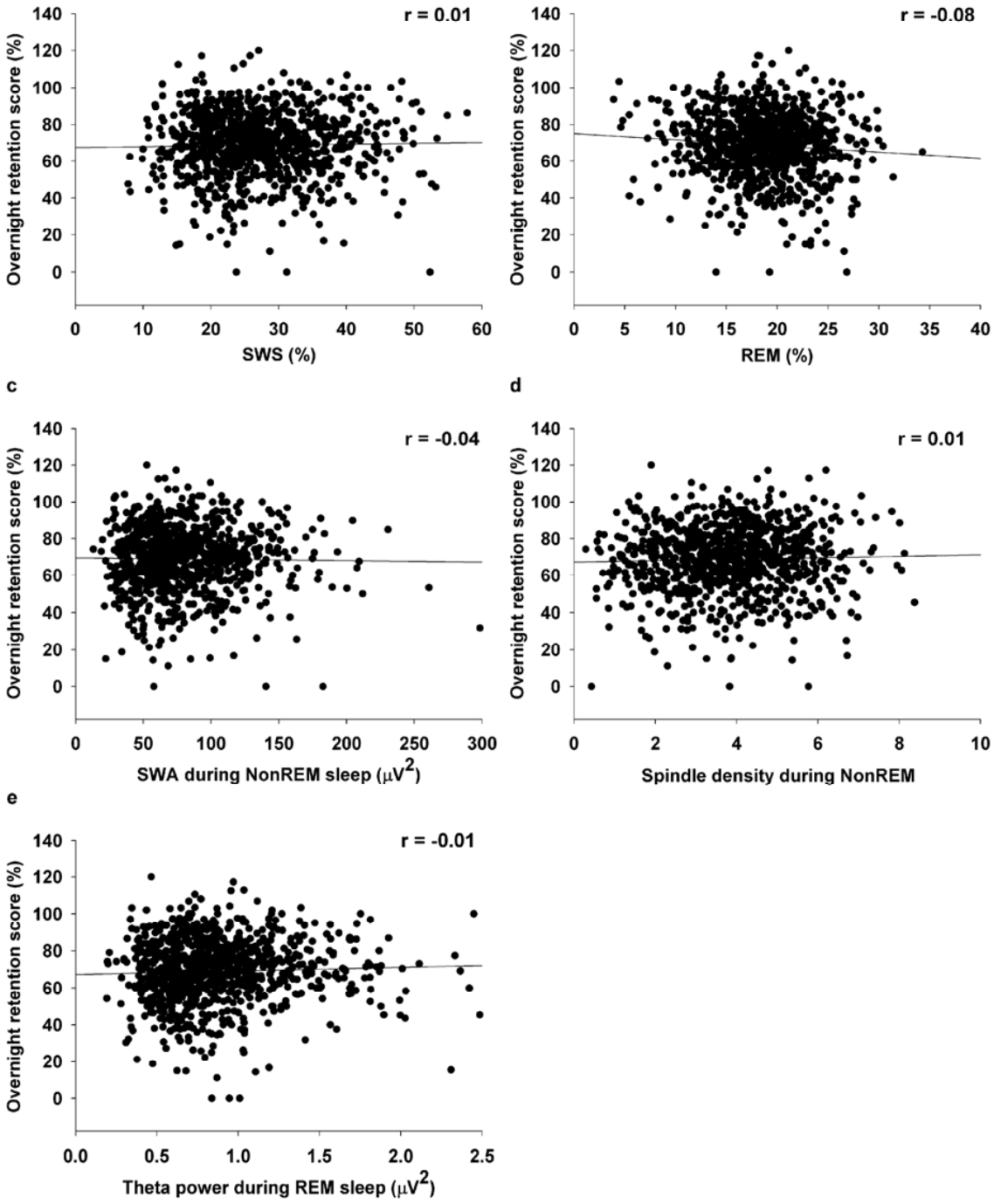
Figure Legends

Figure 1. Summary of the study design and procedure. On day 1 the participants arrived between 15:30 and 17:00. First, the electroencephalography (EEG) electrodes were positioned on the participants' heads. Afterward they viewed 72 negative, positive, and neutral pictures (set 1) and rated them for valence and arousal. After 10 min during which participants performed on a n-back working memory task, they freely recalled the pictures (short-delay free recall picture set 1). Then they performed on a finger sequence-tapping task. After the experimental session, participants slept at home while sleep EEG was recorded.

On day 2, participants came back to the laboratory (arrival between 12:30 and 13:30) and viewed another set of 72 pictures from the International Affective Picture System (set 2). After 10 min during which participants again performed on a n-back working memory task, participants freely recalled pictures from set 2 seen on day 2 (10 min short-delay recall free recall picture set 2) as well as from set 1 seen on day 1 (20 h long-delay free recall picture set 1). Finally, participants performed on the recall phase of the sequential finger-tapping task.

Figure 2. Associations between sleep parameters and overnight memory retention (relative retrieval performance with learning performance before the retention interval (short-delay free recall picture set 1) set to 100%), independent of picture valences. (A) Association between slow wave sleep (SWS) and overnight memory retention score. (B) Association between rapid eye movement (REM) sleep and overnight memory retention score. (C) Association between slow wave activity (SWA) during nonrapid eye movement (NREM) sleep and overnight memory retention score. (D) Association between spindle


Day 1
Start between
3:30 p.m. and 5 p.m.
Encoding
picture set 1
10 min
Short delay free
recall picture set 1
Sleep at home wearing portable EEG recording device
Day 2
Start between
12:30 p.m. and 1:30 p.m.
Encoding
picture set 2
10 min
Short delay free
recall picture set 2
Long delay free
recall picture set 1


r = 0.01
Overnight retention score (%)
SWS (%)
r = -0.08
REM (%)
c
r = -0.04
SWA during NonREM sleep (μV²)
d
r = 0.01
Spindle density during NonREM
e
r = -0.01
Theta power during REM sleep (μV²)

density during NREM sleep and overnight memory retention score. (E) association between theta activity during REM sleep and overnight memory retention score.

Table 1. Correlations between memory performance and sleep parameters; N = 929

		Overnight memory retention	Short-delay recall day 1	Short-delay recall day 2	Long-delay recall
SWS (%)	r	0.007	-0.059	-0.074	-0.035
	P	*0.83*	*0.72*	*0.03*	*0.28*
REM sleep (%)	r	-0.08	**0.108**	**0.094**	0.027
	P	*0.02*	***<0.001***	***0.004***	*0.41*
SWA during NREM sleep (μV^2)	r	-0.04	-0.041	-0.07	-0.048
	P	*0.23*	*0.23*	*0.04*	*0.15*
Theta activity during REM sleep (μV^2)	r	-0.011	**-0.085**	-0.043	-0.062
	P	*0.74*	***0.01***	*0.20*	*0.06*
Spindle density during NREM sleep	r	0.007	-0.017	0.047	0.037
	P	0.82	0.61	0.16	0.28

Reported are nominal P values. Note that Bonferroni corrected P values (corresponding to $p_{uncorrected} < 0.01$; i.e., 0.05/5) are considered significant. Corrected for the influences of sex and age. Results stayed similar when separately analyzing associations with emotional and neutral pictures. Bold typeface indicates significant correlations.

NREM = nonrapid eye movement; REM = rapid eye movement; SWA = slow wave activity; SWS = slow wave sleep.

References

1. Volk HE, McDermott KB, Roediger HL, Todd RD. Genetic influences on free and cued recall in long-term memory tasks. Twin Res Hum Genet 2006;9:623–31.

2. Linkowski P. EEG sleep patterns in twins. J Sleep Res 1999;8:11–3.

3. De Gennaro L, Marzano C, Fratello F, et al. The electroencephalographic fingerprint of sleep is genetically determined: a twin study. Ann Neurol 2008;64:455–60.

4. Ambrosius U, Lietzenmaier S, Wehrle R, et al. Heritability of sleep electroencephalogram. Biol Psychiatry 2008;64:344–8.

5. Buckelmüller J, Landolt H-P, Stassen HH, Achermann P. Trait-like individual differences in the human sleep electroencephalogram. Neuroscience 2006;138:351–6.

6. Diekelmann S, Born J. The memory function of sleep. Nat Rev Neurosci 2010;11:114–26.

7. Rasch B, Born J. About sleep's role in memory. Physiol Rev 2013;93:681–766.

8. Tononi G, Cirelli C. Sleep function and synaptic homeostasis. Sleep Med Rev 2006;10:49–62.

9. Walker MP, van der Helm E. Overnight therapy? The role of sleep in emotional brain processing. Psychol Bull 2009;135:731–48.

10. Wagner U, Kashyap N, Diekelmann S, Born J. The impact of post-learning sleep vs. wakefulness on recognition memory for faces with different facial expressions. Neurobiol Learn Mem 2007;87:679–87.

11. Peigneux P, Laureys S, Fuchs S, et al. Are spatial memories strengthened in the human hippocampus during slow wave sleep ? Neuron 2004;44:535–45.

12. Backhaus J, Junghanns K, Born J, Hohaus K, Faasch F, Hohagen F. Impaired declarative memory consolidation during sleep in patients with primary insomnia: Influence of sleep architecture and nocturnal cortisol release. Biol Psychiatry 2006;60:1324–30.

13. Cox R, Hofman WF, Talamini LM. Involvement of spindles in memory consolidation is slow wave sleep-specific. Learn Mem 2012;19:264–7.

14. Tamminen J, Payne JD, Stickgold R, Wamsley EJ, Gaskell MG. Sleep spindle activity is associated with the integration of new memories and existing knowledge. J Neurosci 2010;30:14356–60.

15. Gais S, Mölle M, Helms K, Born J. Learning-dependent increases in sleep spindle density. J Neurosci 2002;22:6830–4.

16. Nishida M, Pearsall J, Buckner RL, Walker MP. REM sleep, prefrontal theta, and the consolidation of human emotional memory. Cereb Cortex 2009;19:1158–66.

17. Button KS, Ioannidis JP, Mokrysz C, et al. Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 2013;14:365–76.

18. Fiedler K. Voodoo correlations are everywhere - not only in neuroscience. Perspect Psychol Sci 2011;6:163–71.

19. Thornton A, Lee P. Publication bias in meta-analysis: its causes and consequences. J Clin Epidemiol 2000;53:207–16.

20. Ellenbogen JM, Hulbert JC, Stickgold R, Dinges DF, Thompson-Schill SL. Interfering with theories of sleep and memory: sleep, declarative memory, and associative interference. Curr Biol 2006;16:1290–4.

21. Lang PJ, Bradley MM, Cuthbert BN. International affective picture system (IAPS): affective ratings of pictures and instruction manual. Gainesville, FL.: University of Florida, 2008.

22. Gevins A, Cutillo B. Spatiotemporal dynamics of component processes in human working memory. Electroencephalogr Clin Neurophysiol 1993;87:128–43.

23. Rasch B, Spalek K, Buholzer S, et al. A genetic variation of the noradrenergic system is related to differential amygdala activation during encoding of emotional memories. Proc Natl Acad Sci U S A 2009;106:19191–19196.

24. Walker MP, Brakefield T, Morgan A, Hobson JA, Stickgold R. Practice with sleep makes perfect: sleep-dependent motor skill learning. Neuron 2002;35:205–11.

25. Rechtschaffen A, Kales A. A manual of standardized terminology, techniques and scoring system for sleep stages of human sleep. Los Angeles: Brain Information Service. Brain Information Institute UCLA, 1968.

26. Anderer P, Gruber G, Parapatics S, Dorffner G. Automatic sleep classification according to Rechtschaffen and Kales. Conf Proc IEEE Eng Med Biol Soc 2007;2007:3994–7.

27. Anderer P, Gruber G, Parapatics S, et al. An E-health solution for automatic sleep classification according to Rechtschaffen and Kales: validation study of the Somnolyzer 24 x 7 utilizing the Siesta database. Neuropsychobiology 2005;51:115–33.

28. Nishida M, Walker MP. Daytime naps, motor memory consolidation and regionally specific sleep spindles. PLoS One 2007;2:e341.

29. Fogel SM, Smith CT, Cote KA. Dissociable learning-dependent changes in REM and non-REM sleep in declarative and procedural memory systems. Behav Brain Res 2007;180:48–61.

30. Schimicek P, Zeitlhofer J, Anderer P, Saletu B. Automatic sleep-spindle detection procedure: aspects of reliability and validity. Clin Electroencephalogr 1994;25:26–9.

31. Schabus M, Dang-Vu TT, Albouy G, Balteau E, Boly M, Carrier J, et al. Hemodynamic cerebral correlates of sleep spindles during human non-rapid eye movement sleep. Proc Natl Acad Sci U S A 2007;104:13164–9.

32. Zeitlhofer J, Gruber G, Anderer P, Asenbaum S, Schimicek P, Saletu B. Topographic distribution of sleep spindles in young healthy subjects. J Sleep Res 1997;6:149–55.

33. Ficca G, Scavelli S, Fagioli I, Gori S, Murri L, Salzarulo P. Rapid eye movement activity before spontaneous awakening in elderly subjects. J Sleep Res 2004;13:49–53.

34. R Development Core Team. R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing, 2012.

35. Diekelmann S, Wilhelm I, Born J. The whats and whens of sleep-dependent memory consolidation. Sleep Med Rev 2009;13:309–21.

36. Rasch B, Born J, Gais S. Combined blockade of cholinergic receptors shifts the brain from stimulus encoding to memory consolidation. J Cogn Neurosci 2006;18:793–802.

37. Gais S, Born J. Low acetylcholine during slow-wave sleep is critical for declarative memory consolidation. Proc Natl Acad Sci U S A 2004;101:2140–4.

38. Schabus M, Gruber G, Parapatics S, et al. Sleep spindles and their significance for declarative memory consolidation. Sleep 2004;27:1479–85.

39. Fogel SM, Smith CT. Learning-dependent changes in sleep spindles and Stage 2 sleep. J Sleep Res 2006;15:250–5.

40. Tamaki M, Matsuoka T, Nittono H, Hori T. Fast sleep spindle (13 – 15 Hz) activity correlates with sleep-dependent improvement in visuomotor performance. Sleep 2008;31:204–11.

41. Siegel, J. M. The REM sleep-memory consolidation hypothesis. Science 2001;294:1058–63.

42. Siegel, J. M. Clues to the functions of mammalian sleep. Nature 2005;437:1264–71.

43. Wilhelm I, Diekelmann S, Molzow I, Ayoub A, Mo M, Born J. Sleep selectively enhances memory expected to be of future relevance. J Neurosci 2011;31:1563–9.

Supplementary Information

No Associations Between Interindividual Differences in Sleep Pparameters and Episodic Memory Consolidation

Subtitle: Sleep and episodic memory performance in a large sample

Sandra Ackermann, PhD[1,2,3]; Francina Hartmann, MSc1; Andreas Papassotiropoulos, MD[1,4,5]; Dominique J.-F. de Quervain, MD[5,6]; Björn Rasch, PhD[2,7]

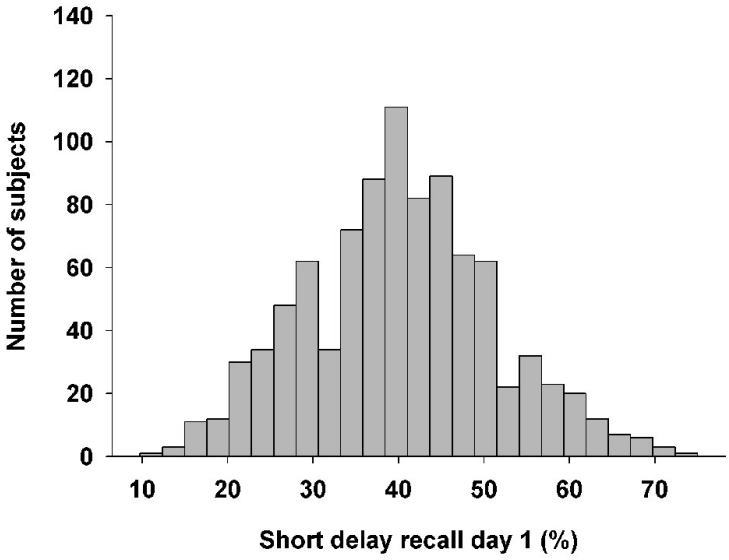
[1]Department of Psychology, Division of Molecular Neuroscience, University of Basel, CH-4055 Basel, Switzerland; [2]Department of Psychology, Division of Biopsychology, University of Zurich, CH-8050 Zurich, Switzerland; [3]Department of Psychiatry, Psychotherapy and Psychosomatics, Psychiatric Hospital, University of Zurich, CH-8032 Zurich, Switzerland; [4]Department Biozentrum, Life Sciences Training Facility, University of Basel, CH-4056 Basel, Switzerland; [5]Psychiatric University Clinics, University of Basel, CH-4012 Basel, Switzerland; [6]Department of Psychology, Division of Cognitive Neuroscience, University of Basel, CH-4055 Basel, Switzerland; [7]Department of Psychology, Cognitive Biopsychology and Methods, University of Fribourg, CH-1701 Fribourg, Switzerland

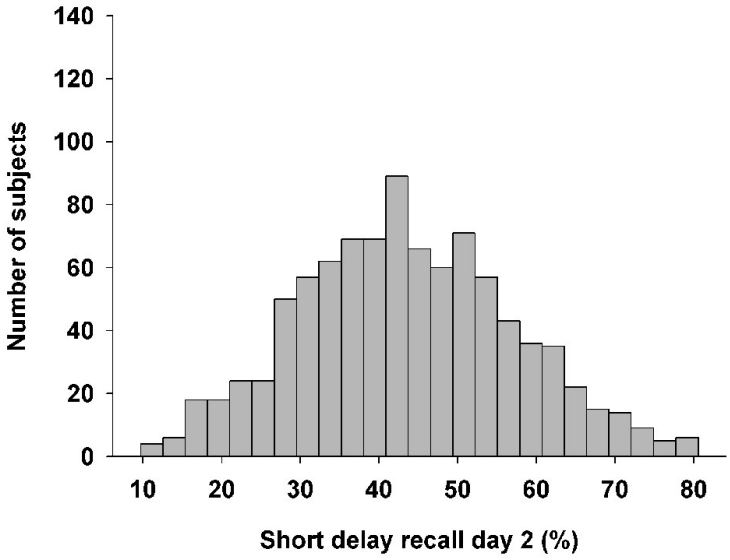
Corresponding authors: Björn Rasch, Department of Psychology, Cognitive Biopsychology and Methods, University of Fribourg, CH-1701 Fribourg, Switzerland. Phone: +41 26 300 76 37, Fax: +41 26 300 97 12, E-mail: bjoern.rasch@unifr.ch; and Sandra Ackermann, Department of Psychology, Division of Biopsychology, University of Zurich, CH-8050 Zurich, Switzerland. Phone: +41 44 635 71 75, E-Mail: sandra.ackermann@psychologie.uzh.ch

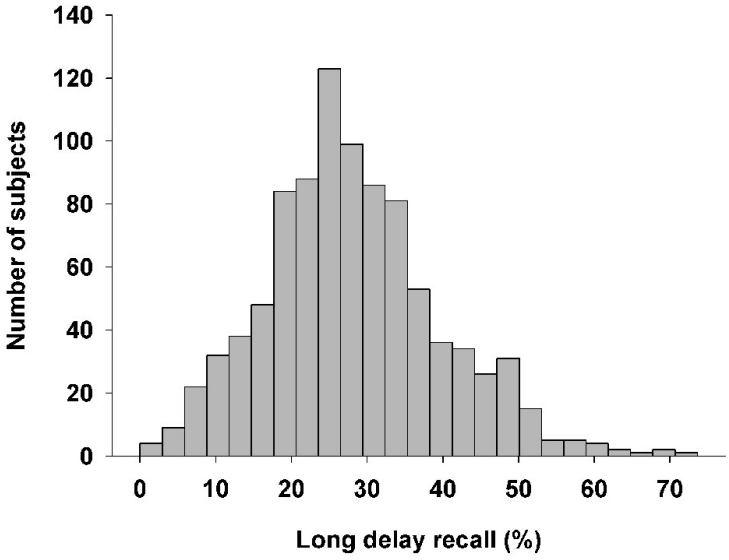
Figure Captions

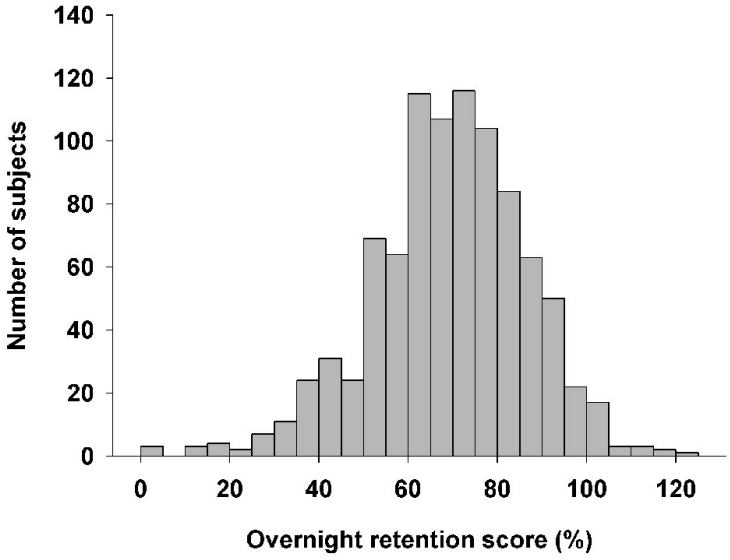
Supplementary Figure 1. Histograms of memory measures (N = 929). Short-delay free recall measures from day 1 and day 2 as well as the long-delay free recall measure are depicted as % recalled of all pictures presented. Overnight memory retention was calculated as relative retrieval performance with learning performance before the retention interval (short delay free recall day 1) set to 100%.

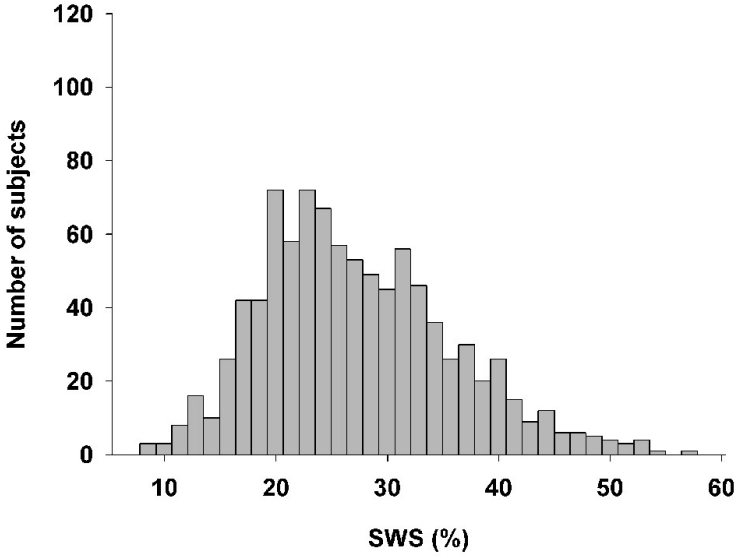
Supplementary Figure 2. Histograms of sleep stages (N = 929) and sleep electroencephalography frequencies (N = 885). Because the distributions of slow wave activity (SWA) during nonrapid eye movement (NREM) sleep and theta activity during rapid eye movement (REM) sleep were asymmetric, data from those two variables was log-transformed for all analyses.

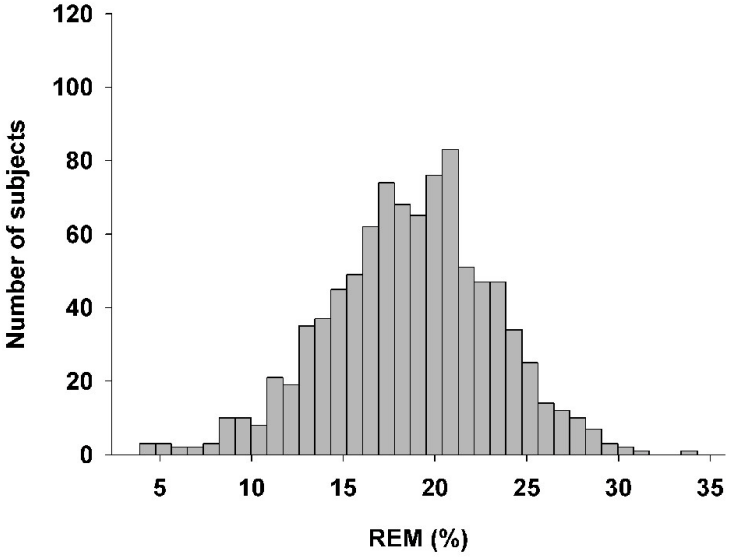

Number of subjects
140
120
100
80
60
40
20
0
10
20
30
40
50
60
70
Short delay recall day 1 (%)

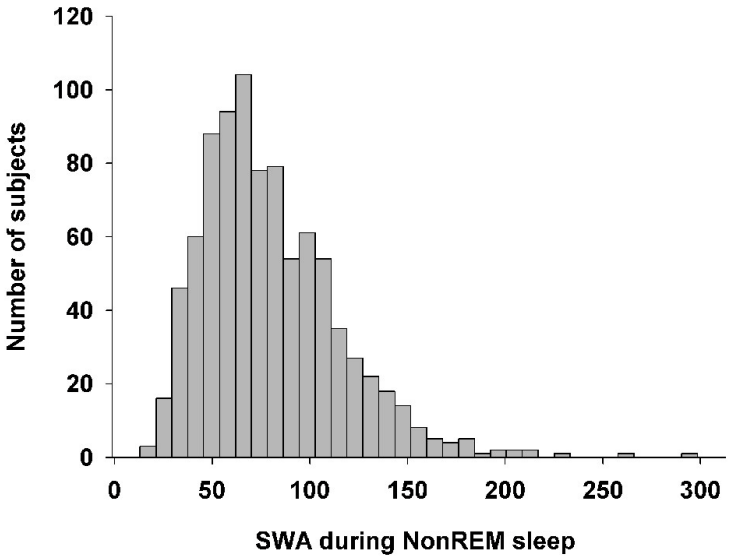

Number of subjects
140
120
100
80
60
40
20
0
10
20
30
40
50
60
70
80
Short delay recall day 2 (%)

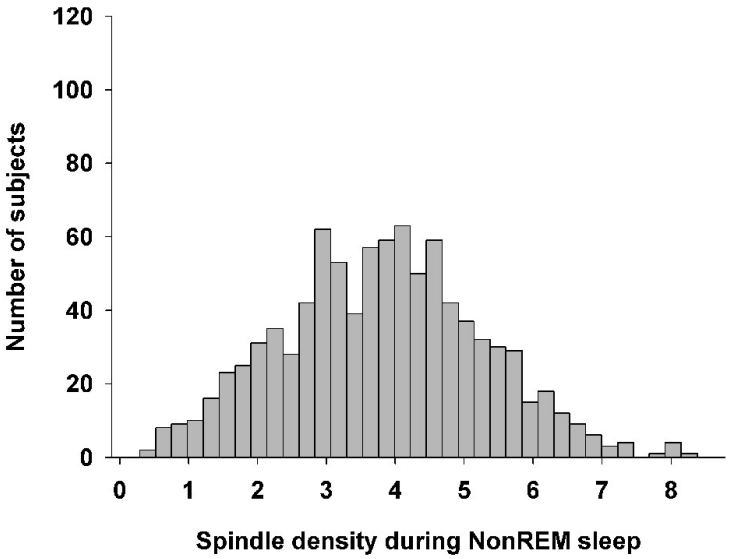

Number of subjects
140
120
100
80
60
40
20
0
0
10
20
30
40
50
60
70
Long delay recall (%)

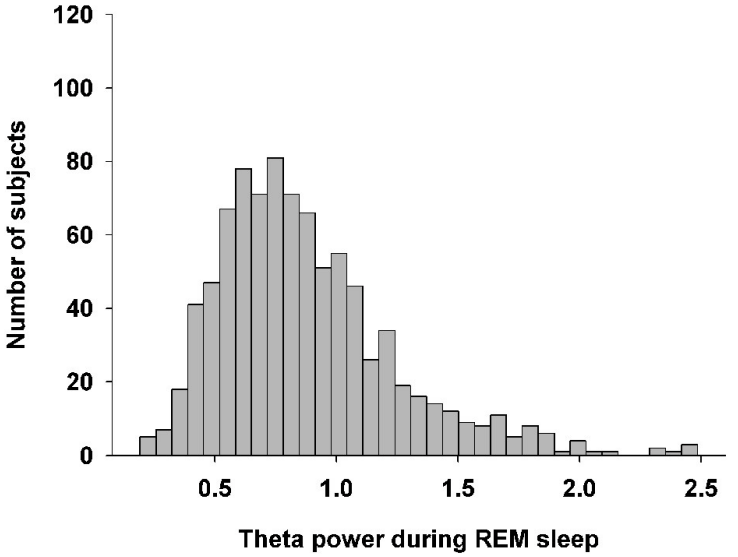

Number of subjects
140
120
100
80
60
40
20
0
0
20
40
60
80
100
120
Overnight retention score (%)


Number of subjects
120
100
80
60
40
20
0
10
20
30
40
50
60
SWS (%)


Number of subjects
120
100
80
60
40
20
0
5
10
15
20
25
30
35
REM (%)


Number of subjects
120
100
80
60
40
20
0
0
50
100
150
200
250
300
SWA during NonREM sleep


Number of subjects
120
100
80
60
40
20
0
0
1
2
3
4
5
6
7
8
Spindle density during NonREM sleep


Number of subjects
120
100
80
60
40
20
0
0.5
1.0
1.5
2.0
2.5
Theta power during REM sleep

Supplementary Table 1. Descriptives of memory performance in the pilot study; N = 55

Short-delay memory recall day 1

	Sleep group (n = 27)	Wake group (n = 28)	T (53)	P
Positive pictures	45.52 ± 2.36 %	40.62 ± 1.91 %	1.62	0.11
Negative pictures	39.66 ± 2.51 %	40.92 ± 2.63 %	-0.35	0.73
Neutral pictures	25.31 ± 2.29 %	23.96 ± 2.05 %	0.44	0.66
All pictures	36.83 ± 1.96 %	35.17 ± 1.75 %	0.63	0.53

Short-delay memory recall day 2

	Sleep group	Wake group		
Positive pictures	37.35 ± 2.56 %	39.58 ± 2.07 %	-0.68	0.50
Negative pictures	41.05 ± 2.92 %	38.54 ± 2.06 %	0.71	0.48
Neutral pictures	29.17 ± 3.14 %	24.85 ± 2.39 %	1.10	0.28
All pictures	35.85 ± 2.52 %	34.33 ± 1.61 %	0.52	0.61

Overnight retention score

	Sleep group	Wake group		
Positive pictures	76.03 ± 5.11 %	62.25 ± 4.25 %	**2.80**	**0.04**
Negative pictures	75.37 ± 3.94 %	58.38 ± 5.11 %	**2.62**	**0.01**
Neutral pictures	76.07 ± 7.24 %	56.40 ± 5.44 %	**2.18**	**0.03**
All pictures	75.69 ± 3.29 %	59.68 ± 3.39 %	**3.39**	**0.001**

Reported are means ± standard error of the mean of recalled pictures (short-delay recall: percentage of presented pictures; overnight retention score: relative retrieval performance with learning performance before sleep (short-delay free recall day 1) set to 100%. Bold typeface indicates significant differences.

Supplementary Table 2. Descriptives of sleep parameters in men and women; N = 929

	Men	Women	T	P
Total sleep time (min)	436.66 ± 4.38	480.47 ± 2.98	**-8.29**	**< 0.001**
Wake (%)	3.24 ± 0.20 %	2.80 ± 0.11 %	**2.11**	**0.04**
S1 (%)	3.66 ± 0.11 %	2.79 ± 0.06 %	**7.85**	**< 0.001**
S2 (%)	47.47 ± 0.44 %	48.26 ± 0.29 %	-1.51	0.13
SWS (%)	26.61 ± 0.49 %	27.60 ± 0.34 %	-1.65	0.10
REM (%)	19.01 ± 0.27 %	18.52 ± 0.18 %	1.55	0.12
SWA during NonREM sleep (μV^2)	67.15 ± 1.71	85.97 ± 1.50	**-7.57**	**< 0.001**
Theta activity during REM sleep (μV^2)	0.66 ± 0.01	0.98 ± 0.02	**-10.92**	**< 0.001**
SWS latency (min)	16.36 ± 0.41	15.94 ± 0.30	0.82	0.42
REM latency (min)	86.76 ± 2.15	85.81 ± 1.42	0.38	0.71
Spindle Density during NREM sleep	3.44 ± 0.08	3.99 ± 0.06	**-5.48**	**< 0.001**
Slow spindle density during NREM sleep	2.00 ± 0.07	2.00 ± 0.05	0.03	0.98
Fast spindle density during NREM sleep	1.92 ± 0.07	2.80 ± 0.05	**-10.76**	**< 0.001**
REM density	0.18 ± 0.01	0.20 ± 0.00	**-2.41**	**0.02**

Reported are means ± standard error of the mean. Bold typeface indicates significant differences.

NREM = nonrapid eye movement; REM = rapid eye movement; SWS = slow wave sleep.

Supplementary Table 3. Correlations between sleep parameters and memory recall, sleep stages: *n* = 929; frequency measures: *n* = 885

		Overnight memory retention			
Sleep parameters		All pictures	Positive pictures	Negative pictures	Neutral pictures
Planned analyses					
% SWS	r	0.007	0.012	0.025	-0.013
	P	*0.83*	*0.71*	*0.44*	*0.70*
SWA during NREM sleep (μV^2)	r	-0.040	-0.017	-0.013	-0.046
	P	*0.23*	*0.62*	*0.71*	*0.17*
Spindle density during NREM sleep	r	0.007	0.00	0.023	-0.008
	P	*0.82*	*> 0.99*	*0.49*	*0.81*
% REM	r	-0.080	-0.057	-0.058	-0.066
	P	*0.02*	*0.09*	*0.08*	*0.05*
Theta activity during REM sleep (μV^2)	r	-0.011	0.010	-0.029	0.022
	P	*0.74*	*0.76*	*0.39*	*0.51*
Exploratory analyses					
Total sleep time	r	-0.034	-0.037	-0.050	0.061
	P	*0.30*	*0.26*	*0.13*	*0.06*
% Wake	r	0.026	0.014	0.015	0.032
	P	*0.42*	*0.68*	*0.66*	*0.33*
% N1	r	0.040	0.006	0.040	0.029
	P	*0.23*	*0.86*	*0.22*	*0.38*
% N2	r	0.022	0.018	-0.011	0.036
	P	*0.50*	*0.59*	*0.73*	*0.28*
SWS latency	r	0.003	-0.008	-0.016	0.005
	P	*0.93*	*0.80*	*0.63*	*0.88*
REM latency	r	0.074	0.067	0.036	0.063
	P	*0.02*	*0.04*	*0.27*	*0.05*
Slow spindle density during NREM sleep	r	-0.006	-0.009	0.019	-0.014
	P	*0.87*	*0.78*	*0.56*	*0.68*
Fast spindle density during NREM sleep	r	0.009	0.004	0.019	-0.006
	P	*0.78*	*0.91*	*0.56*	*0.87*
REM density	r	-0.054	-0.036	-0.021	-0.069
	P	*0.11*	*0.28*	*0.52*	*0.04*

None of the correlations reached significance: Bonferroni-corrected P values.

($p_{nominal} < 0.01$; i.e. 0.05/5) are considered significant. Reported are nominal P values.

NREM = nonrapid eye movement; REM = rapid eye movement; SWS = slow wave sleep.

4.5. Sex-dependent dissociation between emotional appraisal and memory: A large-scale behavioral and fMRI study

Spalek, K., Fastenrath, M., Ackermann, S., Auschra, B., Coynel, D., Frey, J., Gschwind, L., Hartmann, F., van der Maarel, N., Papassotiropoulos, A., De Quervain, D.J.-F., & Milnik, A. (2015). *Journal of Neuroscience, 35*(3), 920-935.

Behavioral/Cognitive

Sex-Dependent Dissociation between Emotional Appraisal and Memory: A Large-Scale Behavioral and fMRI Study

Klara Spalek,[1] Matthias Fastenrath,[1] Sandra Ackermann,[2] Bianca Auschra,[2] David Coynel,[1] Julia Frey,[1] Leo Gschwind,[1] Francina Hartmann,[2] Nadine van der Maarel,[1] Andreas Papassotiropoulos,[2,3,4] Dominique de Quervain,[1,4] and Annette Milnik[2,4]

Department of Psychology, [1]Division of Cognitive Neuroscience and [2]Division of Molecular Neuroscience, [3]Life Sciences Training Facility, and [4]University Psychiatric Clinics, University of Basel, 4009 Basel, Switzerland

Extensive evidence indicates that women outperform men in episodic memory tasks. Furthermore, women are known to evaluate emotional stimuli as more arousing than men. Because emotional arousal typically increases episodic memory formation, the females' memory advantage might be more pronounced for emotionally arousing information than for neutral information. Here, we report behavioral data from 3398 subjects, who performed picture rating and memory tasks, and corresponding fMRI data from up to 696 subjects. We were interested in the interaction between sex and valence category on emotional appraisal, memory performances, and fMRI activity. The behavioral results showed that females evaluate in particular negative ($p < 10^{-16}$) and positive ($p = 2 \times 10^{-4}$), but not neutral pictures, as emotionally more arousing ($p_{\text{interaction}} < 10^{-16}$) than males. However, in the free recall females outperformed males not only in positive ($p < 10^{-16}$) and negative ($p < 5 \times 10^{-5}$), but also in neutral picture recall ($p < 3.4 \times 10^{-8}$), with a particular advantage for positive pictures ($p_{\text{interaction}} < 4.4 \times 10^{-10}$). Importantly, females' memory advantage during free recall was absent in a recognition setting. We identified activation differences in fMRI, which corresponded to the females' stronger appraisal of especially negative pictures, but no activation differences that reflected the interaction effect in the free recall memory task. In conclusion, females' valence-category-specific memory advantage is only observed in a free recall, but not a recognition setting and does not depend on females' higher emotional appraisal.

Key words: **arousal; episodic memory; picture task; sex differences; valence**

Introduction

Sex differences are observed for a wide range of parameters in human research, including biological markers, physiological measurements, behavior, neuropsychological traits, or neuropsychiatric disorders (Davis et al., 1999; Holden, 2005; Kudielka and Kirschbaum, 2005; McCarthy and Konkle, 2005; Cahill, 2006, 2014; Tolin and Foa, 2006; Andreano and Cahill, 2009; McLean and Anderson, 2009; Su et al., 2009; Jazin and Cahill, 2010; Miettunen and Jääskeläinen, 2010; Balliet et al., 2011; Bao and Swaab, 2011; Cross et al., 2011; Trent and Davies, 2012; Ingalhalikar et al., 2014). A person's sex is defined by genetic, as well as by gender identity, which includes psychological, behavioral, and social aspects (Egan and Perry, 2001; Meyer-Bahlburg, 2010).

Episodic memory is a complex polygenic behavioral trait, influenced by genetic and environmental factors along with their interactions (Read et al., 2006; Volk et al., 2006; Papassotiropoulos and de Quervain, 2011). An important modulating factor for episodic memory performance is the perceived emotionality of the learned material (Roozendaal and McGaugh, 2011). Specifically, the more information is perceived as arousing, the more likely it will be remembered (LaBar and Cabeza, 2006). This memory-enhancing effect of emotional arousal is partially mediated through activation of the amygdala (Cahill et al., 1996; McGaugh and Roozendaal, 2002; McGaugh, 2004).

There is evidence that men and women react differently to emotional material (Gard and Kring, 2007). Especially for aversive material, it has been shown that women rate emotional stimuli as more arousing compared with men and additionally have stronger reactions to aversive pictures, as measured by physiological responses like event-related potentials (ERPs), electromyography (EMG), and startle response (Bradley et al., 2001; Gard and Kring, 2007; Lithari et al., 2010). Furthermore, there is evidence that females outperform males in episodic memory tasks related to recall of verbal material, faces, and pictures (Herlitz et al., 1997, 2013; de Frias et al., 2006; Bloise and Johnson, 2007; Andreano and Cahill, 2009). This females' advantage can already be shown in childhood and puberty (Kramer et al., 1997; Herlitz et al., 2013) and is stable over time (de Frias et al., 2006). The question arises whether females' stronger perception of emotionally arousing information may lead to stronger encoding of emo-

Received June 12, 2014; revised Nov. 14, 2014; accepted Nov. 18, 2014.

Author contributions: K.S., A.P., D.d.Q., and A.M. designed research; K.S., S.A., B.A., J.F., L.G., F.H., and N.v.d.M. performed research; K.S., M.F., B.A., D.C., L.G., and A.M. analyzed data; K.S., M.F., B.A., D.C., J.F., L.G., F.H., N.v.d.M., A.P., D.d.Q., and A.M. wrote the paper.

This work was funded by the Swiss National Science Foundation, Sinergia Grant CRSI33_130080 to D.d.Q. and A.P.

The authors declare no competing financial interests.

Correspondence should be addressed to Annette Milnik, University of Basel, Division of Molecular Neuroscience, Birmannsgasse 8, CH-4009 Basel. E-mail: annette.milnik@unibas.ch.

DOI:10.1523/JNEUROSCI.2384-14.2015

 0270-6474/15/350920-16$15.00/0

Table 1. Descriptive information for the included samples and tasks

	Sample 1	Sample 2	Sample 3	Sample 4	All
Females (%)	73	66	64	60	65
Mean age	21.2	22.4	24.1	22.4	22.3
Age range	18–28	18–35	18–38	18–35	18–38
Ongoing study	No	Yes	No	Yes	—
N_{max}	511	1638	104	1145	3398
Picture-rating task	9-point scale	3-point scale	3-point scale	3-point scale	
Valence rating *N*	503	1482	102	1131	3218
Valence rating reaction speed *N*	0	851	0	872	1723
Arousal rating *N*	503	1482	102	1131	3218
Valence rating reaction speed *N*	0	832	0	853	1685
Picture-memory task	3 × 10 pictures	3 × 24 pictures	3 × 24 pictures	3 × 24 pictures	
−10 min delayed recall *N*	510	1481	104	1137	3232
−20–24 h delayed recall *N*	501	1477	0	0	1978
Recognition *N*	0	0	101	1119	1220
Words short-delay memory task *N*	511	1430	0	0	1941
fMRI Encoding *N*	0	0	0	696	696
fMRI Recognition *N*	0	0	0	686	686

For the ongoing studies, the status of the samples is from April 2013. N, sample size.

tional stimuli, thereby inducing an extra advantage in emotional episodic memory performance.

Here we assessed the influence of sex on the emotional appraisal and the recollection of pictures with varying emotional content, as well as on the brain activity during encoding and recognition of these pictures. In the present study, we were particularly interested whether the valence category of the stimulus material (i.e., positive, neutral, and negative pictures) differentially influences the association between sex and a given phenotype, which can be studied with interaction analysis. The advantage of an interaction analysis is the gain in specificity, accompanied with the disadvantage of a greater model complexity and a reduced model stability (Blalock, 1966; Kreft et al., 1998). Due to the large sample sizes in the present study, we were able to analyze not only main effects of sex, but also interaction effects between sex and valence category (positive, neutral, and negative pictures), treating both as factors, with sex being a between-subjects factor and valence category a within-subjects factor. The behavioral data enabled us to disentangle two questions: first, whether valence-category-specific sex differences in the perceived emotionality of pictorial stimuli are linked to corresponding differences in memory performance. Second, whether the valence-category-specific females' memory advantage is memory task-independent and can be found in a free recall, as well as a recognition setting. By analyzing valence-category-specific sex differences in brain activity while encoding and while recognizing pictures, we aimed at identifying neuronal underpinnings of the sex and valence-category-specific differences in behavior.

Materials and Methods

Participants. We analyzed data of $N = 3398$ subjects from four different samples (Table 1). Overall, 65% of the subjects were female and the mean age was 22.3 years (range 18–38). Subjects were recruited from the areas of Zurich (Samples 1, 3) and Basel (Samples 2, 4) in Switzerland. Sampling strategy was to recruit large samples of healthy young adults, without further restrictions. Advertising was done mainly in the Universities of Zurich and Basel and in local newspapers. Subjects were free of any neurological or psychiatric illness, and did not take any medication (apart from oral contraception) at the time of the experiment. Women using different methods of hormonal contraceptives (e.g., oral, spiral, patch) and naturally cycling women were included in the study without restrictions. For the analyzed datasets (status April 2013) we have sufficient information regarding hormonal contraceptives only for Sample 4. Forty-three percent of the females were naturally cycling; for one subject information is missing. Of the females using hormonal contraceptives, 50% used oral contraception (not further characterized). The ethics committee of the Canton Basel and Zurich approved the experiments. Written informed consent was obtained from all subjects before participation. The fMRI analyses were based on Sample 4 only.

Behavioral tasks descriptions. Subjects performed three related tasks that were included in the main analyses, a picture-rating task ($N = 3218$ subjects) and two retrieval tasks: a free-recall task ($N_{max} = 3232$ subjects) and a recognition task ($N = 1220$ subjects). Table 1 gives an overview of all analyzed performances and number of subjects per sample who performed the task. The picture-rating task consisted of the presentation of $N_{max} = 24$ pictures per valence category (negative, neutral, and positive; see below, Description of the used pictures sets). Subjects rated the presented pictures according to valence (negative, neutral, positive) and arousal (low, middle, high) on a nine-point or three-point scale. Subjects of Samples 2–4 additionally encoded 24 scrambled pictures with a geometrical object in the foreground. The object had to be rated regarding its form (vertical, symmetric, horizontal) and size (small, medium, large). In the unannounced free recall picture memory task, subjects had to freely recall these pictures after 10 min (short delay, SD) and eventually additionally after 20–24 h (long delay, LD). Subjects were instructed to describe the pictures with short keywords, to note as much as they can remember related to the remembered pictures and to describe as many of the pictures as possible. Two independent and blinded raters scored these descriptions to identify the number of correctly recalled pictures (Cronbachs α was 91–98%). A third independent rater then decided for the pictures rated inconsistently. In the picture recognition task, 144 pictures were presented, 72 previously seen pictures from the picture-rating task (which already had to be freely recalled) and 72 completely new pictures (24 negative, 24 neutral, and 24 positive pictures). The subjects rated the pictures as remembered, familiar, or new. We used the correctly remembered previously seen pictures as recognition performance measurement.

Statistical analyses of the behavioral data. The rating scales (three- or nine-point scale) as well as the number of stimuli (3 × 10 or 3 × 24) differed between samples. Therefore, it was necessary for the overall analyses to *z*-transform the data. To standardize the output of the different analyses, we *z*-transformed all task performances for each sample separately. Hence, we corrected but could not test for differences between samples.

Ratings (valence and arousal) and memory performances (short-delay free recall, long-delay free recall, recognition) were analyzed by calculating five main (mixed) models with subject as random effect, and sex (female, male; between-factor), valence category (negative, positive and neutral; within-factor), and the interaction term between sex and valence category as contrasts of interest (fixed effects). The models were esti-

mated by REML (restricted maximum-likelihood estimation). Age was included as covariate in all models. Statistical tests for significance were done with *F* tests. *Post hoc* tests for the three different valence categories separately were done with linear models (*t* test), with sex as the variable of interest.

The following additional analyses were done to investigate the free recall memory performances more in depth: (1) short- and long-delay free recall performances were compared by calculating an overall model with time-point as an additional fixed-effect, and the three-way interaction between sex, valence category, and time-point. (2) To correct for the impact of ratings, reaction speed and verbal memory (words short-delay free recall) on the picture memory performances, we additionally included these variables (as main effects and as interaction term with valence category) as possible predictive variables of the picture memory performance in the mixed models, individually and in combination. These models were labeled as "full models." The main models including age, sex, valence category, and the interaction between sex and valence category were labeled as "reduced models." Estimation was done for these analyses with maximum-likelihood. Full and reduced models were compared with the log-likelihood test.

In case of group comparisons (males vs females) we estimated Cohen's *d* as effect size measurement. The estimate of *d* was based on the *t* value of the linear models, but not on the mean and standard deviation of the task performance. Therefore, *d* is corrected for the effects of all confounding variables included in the linear model. By convention, $d = 0.2$ is considered to be a small, $d = 0.5$ to be an intermediate and $d = 0.8$ to be a large effect (Cohen, 1992). Due to the factor coding in our analyses, a positive *d* means that females scored higher on a given phenotype compared with males. For the mixed models effects, which include a repeated measurement, we report the generalized η^2 (Bakeman, 2005). An $\eta^2 = 2\%$ is considered to be small, $\eta^2 = 15\%$ is considered to be intermediate, and $\eta^2 = 35\%$ to be a large effect (Cohen, 1992). Effect sizes calculated for repeated measurements of a factor are influenced by the correlation between the repeated measurements, and can therefore not easily be compared to effect sizes for factors, which are calculated between independent groups.

All calculations were done in R (R Development Core Team, 2011), the mixed model calculations were done with the nlme package (Pinheiro et al., 2011), calculations of the generalized η^2 were done with the ezANOVA package (Lawrence, 2012). All models were calculated with full datasets per subject, which results in an orthogonal design regarding factors with repeated measurements. All reported *p* values are nominal *p* values. To account for the fact that we calculated five main models for the five phenotypes (valence rating, arousal rating, picture short-delay free recall, picture long-delay free recall, and recognition), only results with a *p* value <0.01 will be called statistically significant; *p* values smaller than 1×10^{-16} were not expressed with exact values.

Study description Sample 1. The experiment took place on 2 consecutive days in lecture halls in groups of ~30 subjects. In the following, we describe the parts of the experiment that were relevant for our analyses. On day 1, subjects received information about the study and written informed consent was obtained. Afterward they viewed six series of five semantically unrelated nouns presented at a rate of one word per second with the instruction to learn the words for immediate free recall after each series. The words were taken from the collections of Hager and Hasselhorn (1994) and consisted of 10 neutral words such as "angle," 10 positive words such as "happiness," and 10 negative words such as "poverty." The order of words was pseudorandom, with each group of five words containing no more than three words per valence category. After a distraction task (D2 task), subjects underwent an unexpected delayed free-recall test of the learned words after ~5 min (words short-delay recall). The free recall of a word was considered successful only if it was spelled correctly or with a single letter typo that did not make it become a different word. Approximately 20 min later the picture-rating task during encoding started: participants were presented the pictures (3×10, Set 1 see below, Description of the used pictures sets) and had to rate every picture after its presentation according to valence and arousal on a nine-point scale (duration: 5 min). After a distraction task of 10 min subjects had to freely recall these pictures with a time limit of 6 min. The distraction task was a decision-making task known as the dilemma task. The subjects read six short descriptions (~100 words and 1 diagram each), detailing life-threatening scenarios and the choice between two suboptimal outcomes, one of which they had to choose. On the second day, ~8 min after arrival, subjects were asked to freely recall the pictures from day 1 (24 h delayed recall), again with a time limit of 6 min. The total length of the experimental procedure on day 1 was ~2.5 h, and on day 2 ~50 min. Participants received 70 CHF for their participation.

Study description Sample 2. The experiment took place on 3 d in groups of 1–7 subjects. The time interval between day 1 and 2 was on average 15 d, whereas days 2 and 3 took place on 2 consecutive days. Here we describe the parts of the experiment at days 1, 2, and 3 that were relevant for our analyses. On day 1, subjects received information about the study and written informed consent was obtained. After ~50 min, subjects performed the word-recall tasks as described in Sample 1. The only difference was the distraction tasks, here a free recall of a figural memory task (Rey visual design learning task) and the encoding of abstract figures (Kimura figures). On day 2 after ~1.5 h, the picture-related tasks started: participants received instructions and were trained on the picture-rating task and a working memory task (*N*-back). After training, participants performed the picture-rating task (20 min, 3×24 meaningful pictures, Set 2 see below, Description of the used pictures sets, 1×24 scrambled pictures). While viewing the pictures, subjects had to rate the perceived valence and arousal of each picture on two three-point scales. The working memory task (10 min) served as a distraction task. It was followed by the unannounced free recall test (no time limit) of the pictures. On day 3 after ~15 min, the second picture-task related block took place: participants completed again the picture-rating task (20 min) with a new set of emotional and neutral pictures (3×24 meaningful pictures, 1×24 scrambled pictures). They again rated the perceived valence and arousal of each picture on two three-point scales. Afterward they performed the working memory task (10 min). Participants were then asked to freely recall (no time limit) the pictures seen 10 min earlier and the pictures from day 2 (20 h delayed recall). The total length of the experimental procedure on day 1 was 1.5 h, on day 2 was ~3 h, and on day 3 2 h. Participants received 25 CHF/h for participation. This is an ongoing study.

Study description Samples 3 and 4. Study design and procedures were mostly identical between Samples 3 and 4, which were conducted in two different sites with two different MRI scanners. The study of Sample 3 was the prestudy of Sample 4 with slight differences in scanning procedures. After receiving general information about the study and giving their written informed consent, participants were instructed and then trained on the picture-rating task and a working memory task (*N*-back) they later performed in the MR scanner. After training, participants were positioned in the scanner. Subjects received earplugs and headphones to reduce scanner noise. Their head was fixed in the coil using small cushions and they were instructed not to move their heads. Pictures were presented in the scanner using MR-compatible LCD goggles (VisualSystem, NordicNeuroLab). Eye correction was used when necessary. Functional MR images were acquired during the picture-rating task (3×24 meaningful pictures, Set 2, see next paragraph, 1×24 scrambled pictures) and during the working memory task. Participants spent 30 min in the scanner (20 min picture-rating task, 10 min working memory task). After the presentation of each picture, subjects had to rate the perceived valence and arousal on two three-point scales. The working memory task served as distraction task. After completing the tasks, participants left the scanner for the unannounced free recall test of the pictures (no time limit). After finishing the free recall, subjects were instructed and trained on the recognition task outside the scanner. Following training subjects were again positioned in the MR scanner. In the first 20 min, they performed the recognition task (old pictures seen in the picture-rating task in combination with new pictures from Set 3, see next paragraph) and in the last 20 min structural scans were acquired. The total length of the experimental procedure was ~3–4.5 h. Participants received 25 CHF/h for participation. The study of Sample 4 is an ongoing study.

Description of the used pictures sets. On the basis of normative valence scores pictures from the International Affective Picture System (Lang et al., 1988) were assigned to emotionally negative, neutral and positive

picture groups (ranges for each set separately per valence; Set 1: negative: 1.5–3.7, neutral: 4.6–5.5, positive: 5.6–8.2; Set 2: negative: 1.4–3.5, neutral: 4.4–5.6, positive: 7.1–8.3; Set 3: negative: 1.8–3.6, neutral: 4.5–5.7, positive: 7.0–8.3). For Sets 2 and 3, neutral pictures (Set 2: 8 pictures; Set 3: 6 pictures) from in-house standardized pictures sets were selected to equate the picture sets for visual complexity and content (e.g., human presence).

(f)MRI data acquisition (Sample 4 only). Measurements were performed on a Siemens Magnetom Verio 3 T wholebody MR unit equipped with a 12-channel head coil. Functional time series were acquired with a single-shot echo-planar sequence using parallel imaging (GRAPPA). We used the following acquisition parameters: TE (echo time) = 35 ms, FOV (field-of-view) = 22 cm, acquisition matrix = 80 × 80, interpolated to 128 × 128, voxel size: 2.75 × 2.75 × 4 mm^3, GRAPPA acceleration factor $r = 2.0$. Using a midsagittal scout image, 32 contiguous axial slices placed along the anterior–posterior commissure plane covering the entire brain with a TR (repetition time) = 3000 ms ($\alpha = 82°$) were acquired using an ascending interleaved sequence. A high-resolution T1-weighted anatomical image was acquired using a magnetization prepared gradient echo sequence (MP-RAGE, TR = 2000 ms; TE = 3.37 ms; TI = 1000 ms; flip angle = 8°; 176 slices; FOV = 256 mm, voxel size = 1 × 1 × 1 mm^3).

MRI construction of a population-average anatomical probabilistic atlas. Automatic segmentation of the subjects' T1-weighted images was used to build a population-average probabilistic anatomical atlas. More precisely, each participant's T1-weighted image was first automatically segmented into cortical and subcortical structures using FreeSurfer (v4.5, http://surfer.nmr.mgh.harvard.edu/; Fischl et al., 2002). Labeling of the cortical gyri was based on the Desikan–Killiany Atlas (Desikan et al., 2006), yielding 35 regions per hemisphere. The segmented T1 image was then normalized to the study-specific anatomical template space using the subject's previously computed warp field, and affine-registered to the MNI (Montreal Neurological Institute) space (see below, fMRI preprocessing). Nearest-neighbor interpolation was applied, to preserve labeling of the different structures. The normalized segmentations were finally averaged across subjects, to create a population-average probabilistic atlas. Each voxel of the template could consequently be assigned a probability of belonging to a given anatomical structure, based on the individual information of $N = 612$ subjects.

Experimental design: fMRI picture-rating task. We used an event-related design consisting of 100 trials, including two primacy and two recency trials depicting neutral information, 24 scrambled pictures, and 24 pictures per valence category (positive, negative, neutral). The pictures were presented for 2.5 s in a quasi-randomized order so that a maximum of four pictures of the same category were shown consecutively. A fixation-cross appeared on the screen for 500 ms before each picture presentation. Trials were separated by a variable intertrial period (period between appearance of a picture and the next fixation cross) of 9–12 s (jitter). During the intertrial period, participants subjectively rated the meaningful pictures according to valence (positive, neutral, negative) and arousal (high, medium, low) on a three-point scale (Self Assessment Manikin) by pressing the button with the fingers of their dominant (right-handed: 97%; left-handed: 72%) or nondominant hand (right-handed: 3%; left-handed: 28%). For scrambled pictures, participants rated form (vertical, symmetric, horizontal) and size (small, medium, large) of the geometrical object in the foreground.

Experimental design: fMRI picture recognition task. We used an event-related design consisting of 144 trials. Per trial pictures from two different sets was presented. Each set contained 72 pictures (24 pictures for each stimulus category), one of the sets of stimuli was new (i.e., not presented before), the other old (i.e., presented during the picture-rating task). The pictures were presented for 1 s in a quasi-randomized order so that at most four pictures of the same category (i.e., negative new, negative old, neutral new, neutral old, positive new, positive old) were shown consecutively. A fixation-cross appeared on the screen for 500 ms before each picture presentation. Trials were separated by a variable intertrial period of 6–12 s (jitter) that was equally distributed for each stimulus category. During the intertrial period, participants subjectively rated the picture as remembered, familiar or new on a three-point scale by pressing a button with the fingers of their dominant or nondominant hand (see previous paragraph).

fMRI analyses software. Preprocessing and first level analyses were performed using SPM8 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging, London, UK; http://www.fil.ion.ucl.ac.uk/spm/) implemented in MATLAB R2011b (MathWorks). Second level analyses were done by using GLM Flex (Martinos Center and Mass General Hospital, Charlestown, MA; http://nmr.mgh.harvard.edu/harvardagingbrain/People/AaronSchultz/Aarons_Scripts.html) in MATLAB. GLM Flex is capable of dealing with missing values on group level. The region-of-interest (ROI) analyses were done in R (R Development Core Team, 2011), mixed model calculations were done with the nlme package (Pinheiro et al., 2011).

fMRI preprocessing. Volumes were slice-time corrected to the first slice and realigned using the "register to mean" option. A mean image was generated from the realigned series and coregistered to the structural image. The functional images and the structural images were spatially normalized by applying DARTEL, which leads to an improved registration between subjects. Normalization incorporated the following steps: (1) structural images of each subject were segmented using the "New Segment" procedure in SPM8. (2) The resulting gray and white matter images were used to derive a study-specific group template. The template was computed from a subpopulation of $N = 612$ subjects of this study (see above, MRI construction of a population-average anatomical probabilistic atlas). (3) An affine transformation was applied to map the group template to MNI space. (4) Subject-to-template and template-to-MNI transformations were combined to map the functional images to MNI space. The functional images were smoothed with an isotropic 8 mm full-width at half-maximum Gaussian filter.

fMRI first-level analyses and parameter estimation. Intrinsic autocorrelations were accounted for by AR(1) and low-frequency drifts were removed via high-pass filter (time constant 128 s). For each subject, evoked hemodynamic responses to event-types with zero duration were modeled with a delta function (e.g., button presses), whereas events with a nonzero duration (e.g., picture presentation) were modeled with a boxcar function. Each event was convolved with a canonical hemodynamic response function. Per general linear model the pictures of the three valence categories positive, neutral, and negative and the scrambled picture category were modeled separately. Activity during the picture-rating task was assessed in three different ways: (1) by contrasting activity during the presentation of meaningful pictures against activity during the presentation of scrambled pictures. (2) By contrasting activity during the presentation of later remembered pictures against activity during the nonremembered pictures. (3) By investigating a linear valence and arousal-dependent modulation of signal intensity using parametric analysis (Büchel et al., 1998). The parametric analyses were based on the subject-specific ratings per picture. Therefore, we had to exclude all subjects with monomorphic ratings within one valence category (number of excluded subjects per valence category for valence rating: positive $N = 14$, negative $N = 52$, neutral $N = 18$; number of excluded subjects per valence category for arousal rating: positive $N = 3$, negative $N = 2$, neutral $N = 29$). (4) The activity during the recognition of pictures was assessed by contrasting activity during the presentation of old pictures against activity during the presentation of new pictures. Button presses and rating scale presentation during the ratings were modeled separately. In addition, six movement parameters from spatial realigning were included as regressors of no interest.

fMRI group analyses. Subject-specific parameter estimates from the first-level analyses were entered in the second-level (group) analyses as dependent variables. The minimum number of subjects per voxel was set to be 150. The maximum number of subjects for analyses 1, 2, and 3 (encoding) was $N = 696$, and for recognition (4) $N = 686$. For three analyses, i.e., (1) picture-rating task meaningful versus scrambled pictures, (2) picture-rating task remembered versus nonremembered pictures, and (4) recognition old versus new pictures, we calculated an ANOVA with sex as between-factor (male, female), valence category as within-factor (positive, neutral, negative), and the interaction term between sex and valence category. Statistical tests of significance were done using F and t tests. The minimum cluster size was set to 5 voxels and we

Table 2. Sample-specific raw data of the analyzed task performances

Sample	Valence category	Sex	Picture arousal rating	Picture valence rating	Picture memory SD	Picture memory LD	Recognition correctly remembered	Recognition false alarm	Picture arousal rating reaction speed	Picture valence rating reaction speed	Words memory SD
Sample 1	Positive	Female	3.93 (1.37), 367	2.2 (0.77), 367	6.44 (1.61), 372	6.65 (1.62), 365					3.31 (1.46), 372
	Positive	Male	3.6 (1.31), 136	1.93 (0.79), 136	5.97 (1.63), 138	6.02 (1.79), 136					2.86 (1.31), 139
	Neutral	Female	1.44 (0.92), 367	0.7 (0.61), 367	4.83 (1.72), 372	4.86 (1.66), 365					2.76 (1.46), 372
	Neutral	Male	1.48 (0.87), 136	0.72 (0.63), 136	4.46 (1.6), 138	4.46 (1.74), 136					2.27 (1.44), 139
	Negative	Female	4.87 (1.38), 367	−2.45 (0.74), 367	6.3 (1.65), 372	6.4 (1.63), 365					3.05 (1.48), 372
	Negative	Male	4.26 (1.38), 136	−2.01 (0.76), 136	6.07 (1.61), 138	6.04 (1.57), 136					2.67 (1.3), 139
Sample 2	Positive	Female	0.86 (0.38), 989	0.75 (0.18), 989	12.07 (3.25), 989	8.61 (3.52), 987			0.8 (0.23), 566	0.77 (0.19), 571	2.97 (1.53), 954
	Positive	Male	0.81 (0.39), 493	0.72 (0.22), 493	10.04 (3.56), 492	6.73 (3.35), 490			0.82 (0.24), 266	0.79 (0.21), 280	2.47 (1.4), 476
	Neutral	Female	0.38 (0.29), 989	0.09 (0.16), 989	6.65 (3.03), 989	4.7 (2.8), 987			0.76 (0.22), 566	0.81 (0.21), 571	2.44 (1.38), 954
	Neutral	Male	0.36 (0.28), 493	0.1 (0.15), 493	5.78 (3.02), 492	3.91 (2.7), 490			0.76 (0.21), 266	0.84 (0.22), 280	2.05 (1.37), 476
	Negative	Female	1.37 (0.34), 989	−0.81 (0.18), 989	10.88 (3.22), 989	7.65 (3.39), 987			0.8 (0.21), 566	0.81 (0.21), 571	2.58 (1.43), 954
	Negative	Male	1.19 (0.39), 493	−0.7 (0.23), 493	9.96 (3.26), 492	6.78 (3.35), 490			0.89 (0.24), 266	0.88 (0.24), 280	2.24 (1.47), 476
Sample 3	Positive	Female	0.93 (0.39), 66	0.76 (0.19), 66	12.84 (3.81), 67		19.14 (3.85), 66	0.21 (0.57), 66			
	Positive	Male	0.86 (0.4), 36	0.7 (0.2), 36	11.76 (4.36), 37		19.6 (4.76), 35	0.23 (0.43), 35			
	Neutral	Female	0.42 (0.32), 66	0.1 (0.17), 66	7.18 (3.3), 67		18.18 (4.9), 66	0.41 (0.61), 66			
	Neutral	Male	0.39 (0.26), 36	0.06 (0.16), 36	6.92 (3.88), 37		18.57 (6.18), 35	0.23 (0.49), 35			
	Negative	Female	1.39 (0.32), 66	−0.82 (0.2), 66	11.61 (3.29), 67		19.8 (3.24), 66	0.18 (0.49), 66			
	Negative	Male	1.28 (0.36), 36	−0.75 (0.21), 36	10.54 (3.96), 37		21.06 (3.66), 35	0.17 (0.45), 35			
Sample 4	Positive	Female	0.95 (0.37), 679	0.77 (0.17), 679	12.63 (3.37), 684		19.06 (3.84), 671	0.26 (0.68), 671	0.79 (0.22), 516	0.73 (0.18), 522	
	Positive	Male	0.9 (0.36), 452	0.76 (0.19), 452	11.36 (3.34), 453		19.21 (4.1), 448	0.3 (0.67), 448	0.8 (0.24), 337	0.73 (0.19), 350	
	Neutral	Female	0.38 (0.28), 679	0.08 (0.16), 679	7.34 (3.2), 684		18.57 (4.79), 671	0.27 (0.55), 671	0.79 (0.22), 516	0.78 (0.2), 522	
	Neutral	Male	0.36 (0.26), 452	0.12 (0.17), 452	6.83 (3.08), 453		19.16 (4.77), 448	0.32 (0.64), 448	0.78 (0.23), 337	0.79 (0.22), 350	
	Negative	Female	1.42 (0.29), 679	−0.83 (0.17), 679	11.36 (3.34), 684		19.86 (3.55), 671	0.19 (0.48), 671	0.77 (0.21), 516	0.79 (0.19), 522	
	Negative	Male	1.27 (0.34), 452	−0.73 (0.22), 452	11.33 (3.34), 453		20.73 (3.43), 448	0.25 (0.67), 448	0.84 (0.23), 337	0.82 (0.21), 350	

Mean (standard deviation), and sample size for all analyzed task performances, separately for the three valence categories and sex. Data is additionally shown separately for the four included samples, because the rating scales and number of items per task differ between the samples (see Table 1). SD, short delay; LD, long delay.

applied a familywise error (FWE) correction for the significance threshold on whole-brain (WB) level of $P_{\text{FWE-WB}} < 0.05$ (meaningful vs scrambled: $F_{(2,2082)} \geq 12.77$, $t_{(2082)} \geq/\leq \pm 4.49$; remembered vs nonremembered: $F_{(2,2082)} \geq 12.80$, $t_{(2082)} \geq/\leq \pm 4.49$; old vs new: $F_{(2,2052)} \geq 13.03$, $t_{(2052)} \geq/\leq \pm 4.54$). In case of a significant interaction between sex and valence category, we further investigated the source of significant interaction with *post hoc* tests at the cluster level (see below, fMRI ROI analysis).

Due to the relevance of the medial temporal lobe (hippocampus, parahippocampal gyrus, and entorhinal cortex) and amygdala for (emotional) memory performance (Milner, 1972; Henke et al., 1999; Schacter and Wagner, 1999; Cabeza and Nyberg, 2000; de Quervain et al., 2003; Phelps, 2004) we performed *post hoc* additional small-volume corrected (SVC) analyses in the same way as done on WB level. By focusing on these regions we lowered the significance threshold to $P_{\text{FWE-SVC}} < 0.05$ (meaningful vs scrambled: $F_{(2,2082)} \geq 8.78$, $t_{(2082)} \geq/\leq \pm 3.56$; remembered vs nonremembered: $F_{(2,2082)} \geq 8.80$, $t_{(2082)} \geq/\leq \pm 3.57$; old vs new: $F_{(2,2052)} \geq 8.95$, $t_{(2052)} \geq/\leq \pm 3.60$).

Additionally, we identified brain regions associated with the subjective valence or arousal ratings for the three valence categories separately (analysis 3, linear relationship). Statistical tests of significance were done using *t* tests. Minimum cluster size was set to 5 voxels, the FWE correction on WB level to $P_{\text{FWE-WB}} < 0.05$ (arousal: positive pictures $t_{(692)} \geq/\leq \pm 4.64$, negative pictures $t_{(693)} \geq/\leq \pm 4.66$, neutral pictures $t_{(666)} \geq/\leq \pm 4.70$; valence: positive pictures $t_{(681)} \geq/\leq \pm 4.63$, negative pictures $t_{(643)} \geq/\leq \pm 4.68$, neutral pictures $t_{(677)} \geq/\leq \pm 4.61$). These analyses were done mainly for visualization purpose.

fMRI ROI analysis. From those voxel clusters showing a significant interaction effect between sex and valence category at the group-level for the contrast meaningful versus scrambled, we extracted the subject-specific parameters estimated in the first-level analysis. Next, we averaged the parameter estimates within each valence category and cluster for each subject (averaged first-level estimates per subject, valence, and ROI). All further analyses were done using linear (mixed) models in combination with ANOVA. The (averaged first-level) parameter estimates were again assigned as dependent variable. In case of mixed models, estimation was done by REML. Statistical tests of significance were done using *F* and *t* tests. Age was included as covariate in all models. Subjects were treated as random effect. Per ROI, we calculated two analyses:

The first analysis was performed to confirm and extend the results of the fMRI second-level ANOVA. Therefore, we included sex and valence category and the interaction term between sex and valence category as fixed effects. We performed *post hoc* tests to clarify the source of interaction, contrasting two of the three possible valence categories against each other (negative vs neutral, negative vs positive, positive vs neutral).

The next steps were done to further characterize all regions that showed a negative-specific sex effect. First, we identified all regions with a significant main effect of sex specifically for the negative picture category. Second, we investigated the linear relationship between the meaningful versus scrambled contrast parameters and the task performances (behavioral data: averaged ratings and memory performances), especially of the negative and negative against neutral valence categories. In these models task performance, sex, and valence category were assigned as fixed effects. All reported *p* values were nominal *p* values. The significance threshold was adapted to $p < 0.002$ to account for the number of extracted ROIs (encoding meaningful vs scrambled 25 ROIs).

Results

For the behavioral data, the mean and standard deviation of the task performances, separately for the four samples, the two sex groups, and the three valence categories are summarized in Table 2. Figure 1 depicts the task performances after *z*-transformation for all four samples combined, separately for the two sex groups and the three valence categories. The reported effect sizes were corrected for all covariates included in the analyses. Due to the factor coding of sex, a positive *d* means that females scored higher on a given phenotype than males.

Task 1: picture-rating task, valence and arousal ratings

Behavioral data

Across both sexes, subjects' averaged valence and arousal ratings showed substantial differences between valence category (valence

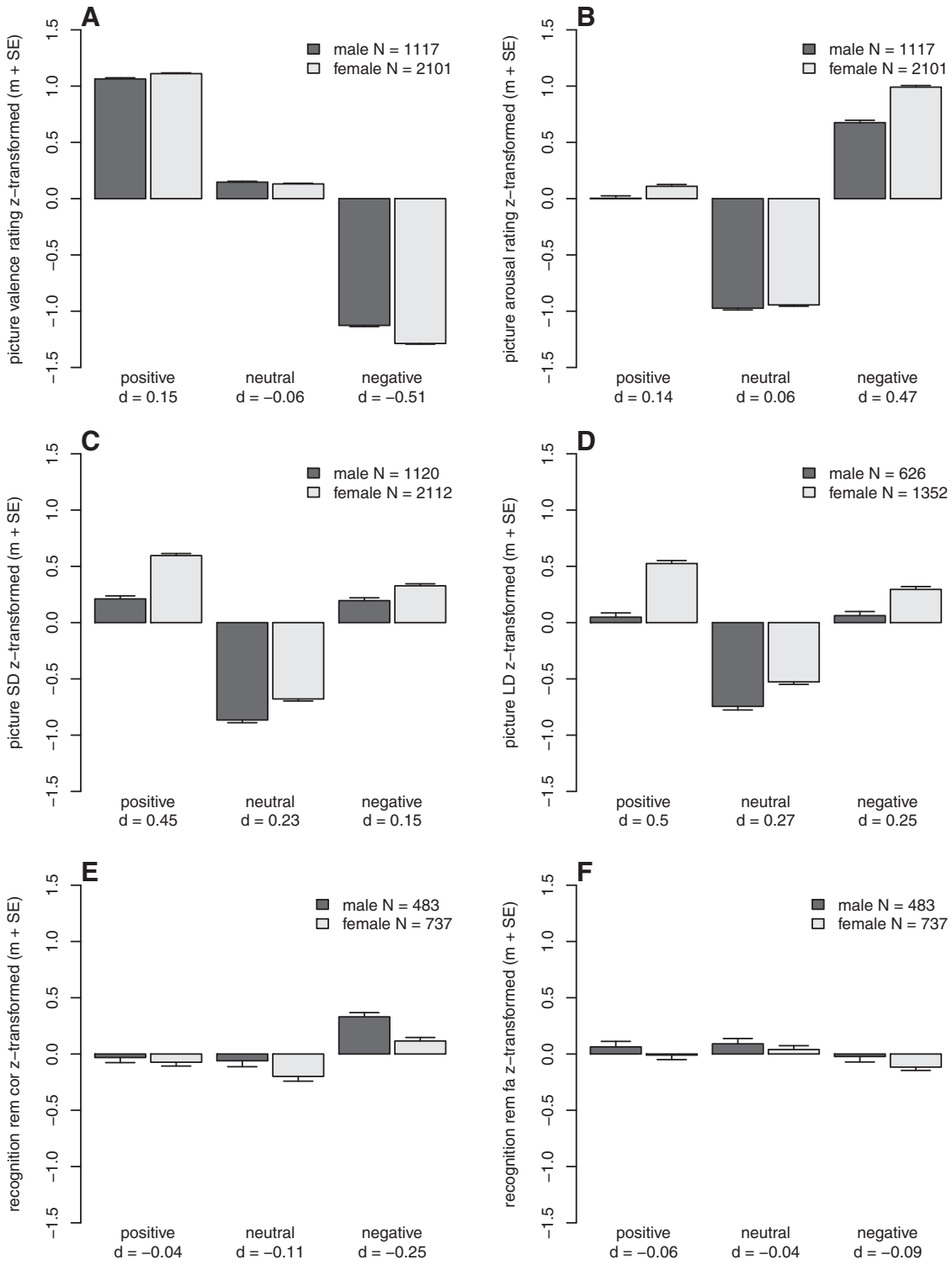


Figure 1. Results of the behavioral analyses. The task performances are *z*-transformed, therefore a negative task performance denotes that the performance in this group was lower than the average performance. ***A***, Picture valence rating. ***B***, Picture arousal rating. ***C***, Short delay (SD) memory performance. ***D***, Long delay (LD) memory performance. ***E***, Recognition performance, correctly remembered old pictures (rem cor). ***F***, Recognition performance, false alarm new pictures (rem fa). m + SE, mean and standard error of the mean; d, effect size.

rating main effect of valence category: $F_{(2,6432)} = 50{,}737.76$, $p < 1 \times 10^{-16}$, $\eta^2 = 91.32\%$; arousal rating main effect of valence category: $F_{(2,6432)} = 12{,}764.24$, $p < 1 \times 10^{-16}$, $\eta^2 = 56.96\%$). *Post hoc* tests showed that pictures from the emotional valence categories were significantly more extremely rated compared with the neutral pictures (valence rating positive vs neutral: $t_{(3217)} = -149.11$, $p < 1 \times 10^{-16}$, negative vs neutral: $t_{(3217)} = -190.14$, $p < 1 \times 10^{-16}$; arousal rating positive vs neutral: $t_{(3217)} = -93.24$, $p < 1 \times 10^{-16}$, negative vs neutral: $t_{(3217)} = 158.46$, $p < 1 \times 10^{-16}$; Fig. 1*A*,*B*).

There were significant interaction effects between sex and valence category on the valence rating ($F_{(2,6432)} = 95.32$, $p < 1 \times 10^{-16}$, $\eta^2 = 1.94\%$) and on the arousal rating ($F_{(2,6432)} = 75.08$, $p < 1 \times 10^{-16}$, $\eta^2 = 0.77\%$; Fig. 1). *Post hoc* tests showed that

Table 3. Results of the fMRI picture-rating task during encoding contrast meaningful versus scrambled pictures

Whole brain analyses results						ROI results based on the averaged estimates per cluster				
							Post hoc tests			
			Peak voxel MNI coordinates				Sex *x*-valence category analyses for different subsets of valence categories			
Region	H	F_{max}	*X*	*Y*	*Z*	*N*	Neg, neu, pos: *p*	Neg, neu: *p*	Neg, pos: *p*	Pos, neu: *p*
Frontal lobe										
Paracentral lobule*	L	19.7	−13.75	−30.25	40	49	8.9×10^{-10}*	3.4×10^{-7}*	5.2×10^{-9}*	0.86
Precental gyrus 1	L	18.39	−57.75	5.5	0	17	5.3×10^{-8}*	0.00011*	6.1×10^{-9}*	0.13
Precental gyrus 2*	L	14.26	−46.75	0	4	6	2.1×10^{-7}*	2.7×10^{-5}*	1.6×10^{-7}*	0.45
Precental gyrus 3*	L	19.48	−35.75	−13.75	48	51	1.8×10^{-8}*	0.00014*	6.7×10^{-10}*	0.084
Precental gyrus 4*	L	15.73	−16.5	−11	76	10	6.9×10^{-8}*	7.8×10^{-8}*	5×10^{-5}*	0.11
Precental gyrus 5	R	14.8	60.5	8.25	4	5	1.9×10^{-7}*	3.4×10^{-6}*	1.3×10^{-6}*	0.69
Precental gyrus 6 †	R	15.15	46.75	−2.75	52	6	6.1×10^{-7}*	0.0024	2.9×10^{-8}*	0.032
Superior frontal gyrus	R	19.19	8.25	2.75	64	47	3.6×10^{-9}*	5.1×10^{-7}*	5×10^{-8}*	0.98
Parietal lobe										
Inferior parietal cortex	L	14.31	−38.5	−82.5	28	9	1.3×10^{-7}*	1.1×10^{-6}*	2.6×10^{-6}*	0.4
Precuneus cortex	L	18.27	−8.25	−49.5	52	65	8.5×10^{-9}*	4.1×10^{-7}*	3.5×10^{-7}*	0.49
Superior parietal cortex	L	13.86	−19.25	−46.75	68	7	3.7×10^{-7}*	1.1×10^{-5}*	1.4×10^{-6}*	0.97
Supramarginal gyrus 1	L	14.11	−63.25	−22	16	5	9.3×10^{-7}*	0.00022*	2.9×10^{-7}*	0.23
Supramarginal gyrus 2*	L	15.37	−52.25	−27.5	20	16	7.7×10^{-8}*	2.1×10^{-5}*	1×10^{-8}*	0.56
Supramarginal gyrus 3	L	16.84	−60.5	−35.75	32	22	6.3×10^{-8}*	1.2×10^{-5}*	3×10^{-8}*	0.53
Supramarginal gyrus 4*	R	13.92	49.5	−27.5	28	11	4×10^{-7}*	3×10^{-5}*	3×10^{-7}*	0.71
Occipital lobe										
Cuneus cortex*	L	14.95	−13.75	−77	20	19	1.3×10^{-7}*	0.00012*	2.5×10^{-8}*	0.19
Lingual gyrus 1 †	L	16.36	−13.75	−55	−8	25	1×10^{-7}*	0.0067	6×10^{-9}*	0.0038
Lingual gyrus 2	R	14.61	24.75	−49.5	4	5	5.2×10^{-7}*	0.00049*	6.1×10^{-8}*	0.12
Cingulate cortex										
Cingulate cortex, caudal anterior division	L	15.07	0	16.5	28	10	1×10^{-7}*	3.4×10^{-7}*	0.00012*	0.041
Cingulate cortex, Posterior division	R	15.19	11	−27.5	40	16	7.2×10^{-8}*	1×10^{-6}*	3.2×10^{-6}*	0.32
Cerebellum										
Cerebellum cortex 1	L	21.77	−33	−57.75	−52	73	1.2×10^{-12}*	1.1×10^{-8}*	9.3×10^{-12}*	0.58
Cerebellum cortex 2*	R	14.92	24.75	−44	−28	11	1.8×10^{-8}*	1.7×10^{-6}*	6.7×10^{-8}*	0.91
Cerebellum white matter	R	14.63	24.75	−46.75	−48	5	2.1×10^{-7}*	1.6×10^{-5}*	1.7×10^{-7}*	0.69
Temporal lobe										
Superior temporal gyrus ††	L	19.23	−35.75	−2.75	−24	19	3.5×10^{-10}*	1.3×10^{-8}*	2.8×10^{-8}*	0.85

The table gives an overview about all brain regions that showed a significant interaction effect for sex and valence category on whole-brain level. The *post hoc* tests revealed, that all but two regions (marked with †, precental gyrus 6 and lingual gyrus 1) showed a significant ($p < 0.002$) negative specific sex *x*-valence interaction. Regions marked with an asterisk (*) additionally survived all filtering steps of the ROI analyses. In these regions, additionally to the sex *x*-valence interaction effect, there was a significant main effect of sex for negative pictures and a significant correlation with valence or arousal rating of negative pictures. For all clusters, except the left paracentral lobule, this correlation was significantly stronger for negative in comparison to the neutral picture category, at least for one of the two ratings. The relevant significant *p* values for the filtering are printed in bold. $p < 0.002$ are considered significant and marked with an asterix (*). H, Hemisphere; *N*, number of voxels; ME, main effect. ††Reported is the closest gray matter area identified manually. d,r, effect sizes.

females rated the valence and the arousal especially of negative emotional material more extreme than males, with medium effect sizes (valence: $t_{(3215)} = -13.83$, $p < 1 \times 10^{-16}$, $d = -0.51$; arousal: $t_{(3215)} = 12.57$, $p < 1 \times 10^{-16}$, $d = 0.47$). The ratings of positive material were also significantly more extreme in females (valence: $t_{(3215)} = 4.09$, $p = 4.4 \times 10^{-5}$, $d = 0.15$; arousal: $t_{(3215)} = 3.72$, $p = 2 \times 10^{-4}$, $d = 0.14$), but with small effect sizes. There were no significant differences between the two sexes for the ratings of neutral stimuli (valence: $t_{(3215)} = -1.5$, $p = 0.13$, $d = -0.06$; arousal: $t_{(3215)} = 1.53$, $p = 0.13$, $d = 0.06$).

fMRI data

Because we observed sex-specific differences in emotional ratings of negative and positive, but not neutral pictures (significant interaction effect between sex and valence category), we were interested whether we could identify a neuronal correlate explaining these sex- and valence-category-specific differences in rating. In the first-level analysis, activity during the picture-rating task was assessed by contrasting activity during the presentation of meaningful pictures against activity during the presentation of scrambled stimuli (positive vs scrambled, neutral vs scrambled, negative vs scrambled). In the (second-level) group analysis, we calculated an ANOVA with sex as between-factor (male, female), valence category as within-factor (positive, neutral, negative) and the interaction term between sex and valence category. We identified significant ($p_{FWE\text{-}WB} < 0.05$) clusters for the interaction effect between sex and valence category in several regions with an emphasis on motor-relevant regions in the frontal and parietal cortices, and in the cerebellum (Table 3; Fig. 2). No additional suprathreshold clusters were identified when applying SVC ($p_{FWE\text{-}SVC} < 0.05$) for bilateral medial temporal lobe regions (hippocampus, parahippocampal gyrus, entorhinal cortex, and amygdala) only. Figure 3*A*,*B* shows the results of the main effects sex and valence category.

In the ROI analysis, we first identified for all clusters the origin of the significant interaction between sex and valence category. These *post hoc* tests showed that in all but two regions within the precentral gyrus and the lingual gyrus (Table 3), the negative valence category drove the significant interaction effect between sex and valence category, meaning that the differences between negative and positive as well as negative and neutral pictures became significant, but not the difference between positive and neutral pictures.

In the next step, we identified all regions that additionally showed a significant main effect of sex for negative pictures only. In all cases females showed a higher activation than males within

Table 3. Continued

ROI results based on the averaged estimates per cluster							
Filtering steps							
ME sex neg		ME arousal rating neg		Arousal rating *x*-valence category	ME valence rating neg		Valence rating *x*-valence category
p	*d*	*p*	*r*	Neg, neu: *p*	*p*	*r*	Neg, neu: *p*
6.2×10^{-8}*	0.42	0.0034	0.11	6.7×10^{-12}*	**5×10^{-4}***	−0.13	0.0075
2×10^{-5}*	0.33	0.76	0.01	0.00033*	0.67	−0.02	0.021
5.3×10^{-12}*	0.54	**0.00055***	0.13	**4.5×10^{-11}***	**0.00026***	−0.14	0.0059
4.6×10^{-7}*	0.39	**6.1×10^{-6}***	0.17	**3.4×10^{-14}***	0.0043	−0.11	8.1×10^{-5}*
6.8×10^{-6}*	0.35	0.069	0.07	1×10^{-6}*	**0.00073***	−0.13	**0.00013***
3.5×10^{-7}*	0.4	0.079	0.07	3.1×10^{-7}*	0.62	−0.02	0.046
0.0025	0.24	0.28	0.04	6.9×10^{-6}*	0.22	−0.05	0.0066
4.6×10^{-6}*	0.36	0.14	0.06	3×10^{-5}*	0.033	−0.08	0.0034
0.15	0.11	0.41	−0.03	0.53	0.21	−0.05	0.013
0.023	0.18	0.58	0.02	5.6×10^{-5}*	0.11	−0.06	0.0065
0.01	0.2	0.063	0.07	1.4×10^{-13}*	0.076	−0.07	0.004
3.8×10^{-6}*	0.36	0.03	0.08	2×10^{-10}*	0.035	−0.08	0.24
4.1×10^{-7}*	0.4	**0.0017***	0.12	**4.7×10^{-16}***	**0.00081***	−0.13	0.086
0.00023*	0.29	0.89	0.01	0.0016*	0.027	−0.09	0.05
4×10^{-6}*	0.36	**8.6×10^{-6}***	0.17	**$<1 \times 10^{-16}$***	0.2	−0.05	0.064
3.4×10^{-10}*	0.49	**1.1×10^{-14}***	0.29	**2.9×10^{-12}***	**0.00046***	−0.13	**4.4×10^{-12}***
2.9×10^{-06}*	0.37	6.4×10^{-12}*	0.26	1.2×10^{-10}*	0.0015*	−0.12	1.1×10^{-11}*
0.33	0.07	6.1×10^{-7}*	0.19	8.3×10^{-13}*	0.0055	−0.11	1.3×10^{-6}*
6.7×10^{-9}*	0.45	0.55	0.02	0.24	0.48	−0.03	0.0099
1.5×10^{-16}*	0.65	0.024	0.08	1.3×10^{-12}*	0.087	−0.06	0.13
8.9×10^{-8}*	0.43	0.082	0.07	0.012	0.44	−0.03	0.00057*
3.5×10^{-5}*	0.32	0.0098	0.1	0.00061*	**0.00074***	−0.13	**4.4×10^{-6}***
0.074	0.14	0.19	0.05	0.3	0.45	−0.03	0.056
1.4×10^{-6}*	0.38	0.15	0.06	0.00071*	0.053	−0.07	0.00055*

the negative valence category (Table 3). Next, we identified all regions that showed: (1) a significant correlation with the averaged subjects' valence or arousal rating of the negative pictures only, and eventually (2) an additional significant interaction between the averaged valence or arousal rating and the neutral and negative valence category. The overall picture indicated that by applying these additional filters, we identified motor-relevant regions (Table 3, see regions marked with an asterisk), which were specifically associated with the valence and arousal ratings of negative pictures and were more active in females compared with males. Figure 4 shows exemplarily the results for the filtering steps within two ROIs, which survived all steps for valence (*A–C*, right cerebellum cortex 2) or arousal (*D–F*, left precentral gyrus 3) ratings. When applying the same filter steps for the short-delay memory performances none of the regions survived the filtering.

To visually confirm these results we investigated, separately for each valence category, the linear relationship between fMRI signal intensity and ratings using parametric modulation in the first-level analyses. We superimposed the ROIs showing a significant interaction between sex and valence category on the activation maps of valence and arousal ratings for the negative, neutral, and positive valence category separately. By combining these two activation maps, it was possible to visualize that ROIs, showing a significant interaction between sex and valence category, were preferentially located in brain regions, in which activity was associated primarily with the ratings of the negative valence category (Fig. 2).

To summarize, the behavioral results showed that women rated especially negative pictures as more arousing and more negative than men. The fMRI interaction analysis for sex and valence category comparing meaningful versus scrambled pictures during the picture-rating task identified regions that were specifically more activated in females compared with males when viewing negative pictures. These regions can be grouped as mainly motor-relevant regions, as well as the posterior cingulate. Additionally, differences in activity (meaningful vs scrambled) in several of these regions were especially associated with the ratings of the negative pictures.

Task 2: picture-memory task, delayed free recall

Overview

Emotionally arousing information is generally better remembered than neutral information. Therefore, the question arises, whether the stronger ratings of females for emotional stimuli are associated with differences in memory performance, favoring females in case of emotional information.

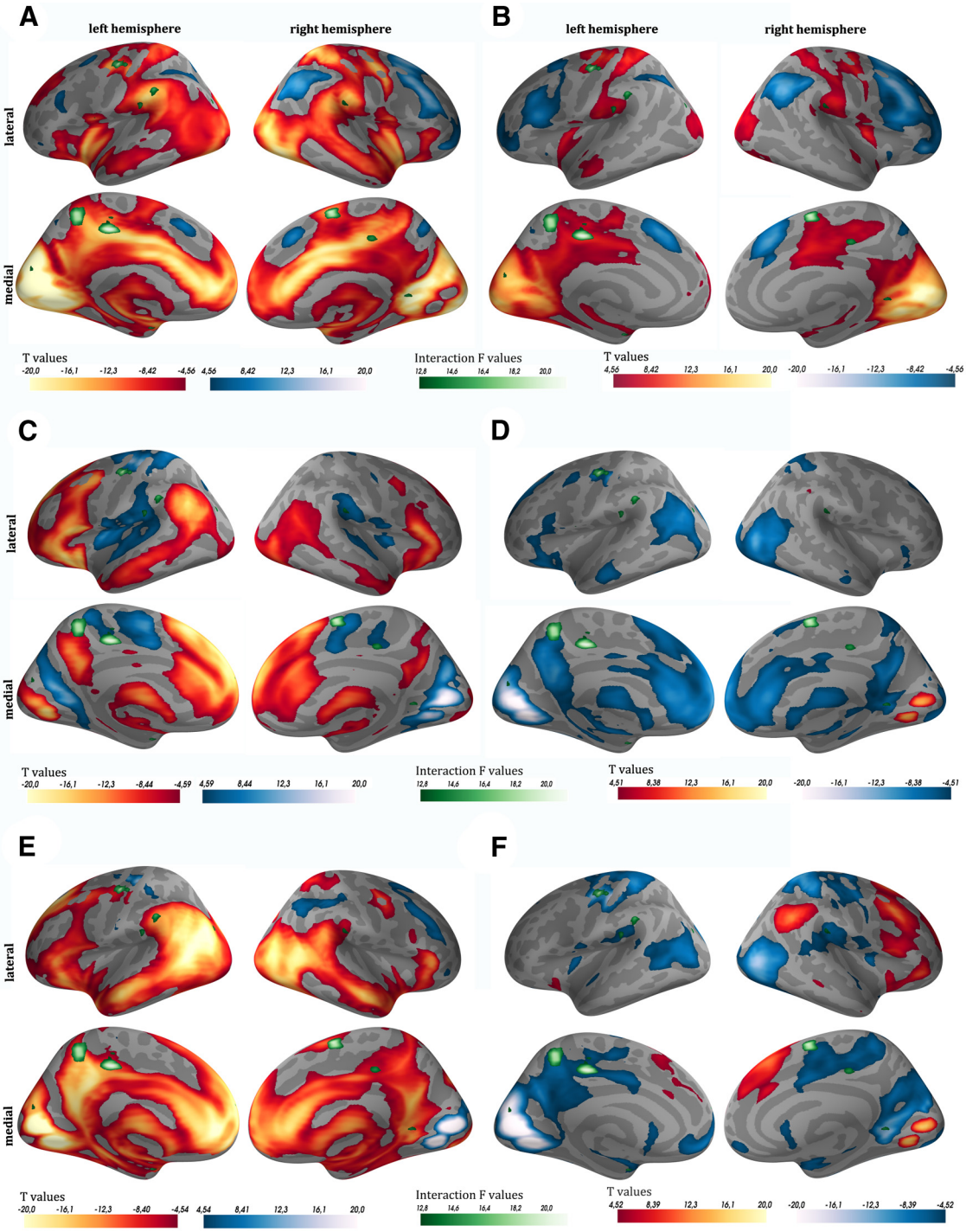


Figure 2. Picture-rating task during encoding. fMRI results of the parametric modulation for arousal (***A***, ***C***, ***E***) and valence (***B***, ***D***, ***F***) ratings, separately for the three valence categories (negative ***A***, ***B***, neutral ***C***, ***D***, positive ***E***, ***F***). Red colors indicate that higher arousal ratings and more negative valence ratings are associated with an increase in fMRI signal. Blue colors indicate that lower arousal ratings and more positive valence ratings are associated with an increase in fMRI signal. Superimposed in green are the clusters that showed a significant interaction between sex and valence category in the meaningful versus scrambled contrasts of the picture-rating task during encoding.

Behavioral data

Across both sexes, subjects' memory performances showed substantial differences between valence category (main effect of valence SD: $F_{(2,6460)} = 3742.64$, $p < 1 \times 10^{-16}$, $\eta^2 = 28.62\%$; LD: $F_{(2,3952)} = 1289.04$, $p < 1 \times 10^{-16}$, $\eta^2 = 18.58\%$). *Post hoc* tests showed that pictures from the positive valence category (SD: $t_{(3231)} = -79.71$, $p < 1 \times 10^{-16}$; LD: $t_{(1977)} = -47.91$, $p < 1 \times 10^{-16}$), as well as from the negative valence category (SD: $t_{(3231)} = 68.34$, $p < 1 \times 10^{-16}$; LD: $t_{(1977)} = 39.06$, $p < 1 \times 10^{-16}$; Fig. 1*C*,*D*) were significantly better remembered than neutral pictures.

There was a significant interaction effect between sex and valence category on the short-delay (10 min delayed) free recall of the pictures ($F_{(2,6460)} = 35.47$, $p = 4.4 \times 10^{-16}$, $\eta^2 = 0.38\%$). *Post hoc* tests showed that although females generally performed better than males, this advantage was most pronounced for positive material (positive: $t_{(3229)} = 12.15$, $p < 1 \times 10^{-16}$, $d = 0.45$;

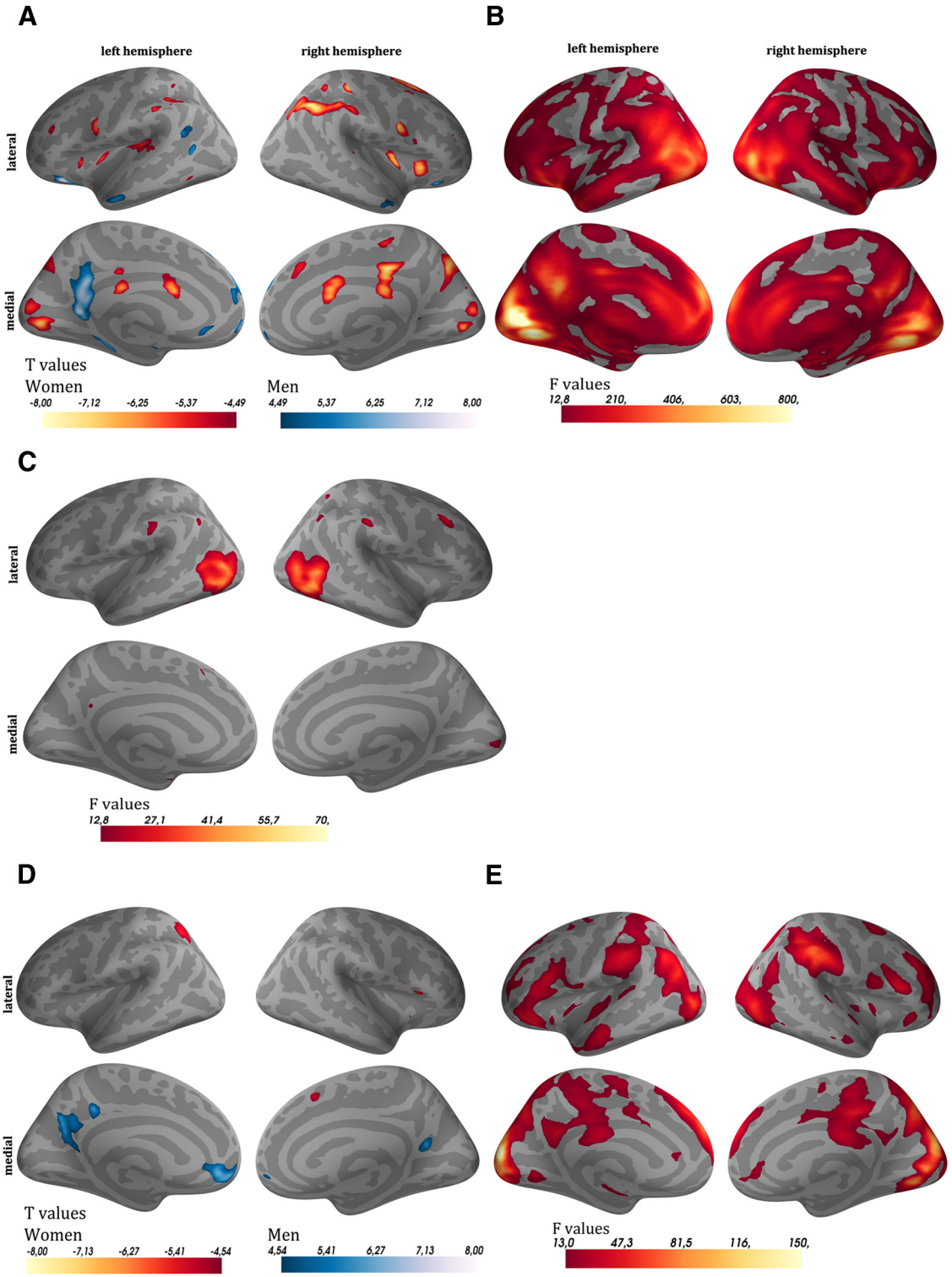


Figure 3. Main effects of sex and valence for the picture-rating task during encoding and for recognition. ***A***, ***B***, The contrast meaningful versus scrambled pictures during encoding (***A***, main effect of sex; ***B***, main effect of valence). ***C***, The contrast remembered versus nonremembered pictures during encoding (main effect of valence only). ***D***, ***E***, The contrast old versus new pictures of the recognition task (***D***, main effect of sex; ***E***, main effect of valence). For the main effect of sex (***A***, ***D***) red indicates that this contrast was more pronounced in females than in males, whereas blue indicates the opposite. For the main effect of valence (***B***, ***C***, ***E***) the brighter the regions are, the higher the differences for the contrasts were between the three valence categories.

neutral: $t_{(3229)} = 6.16$, $p = 8.3 \times 10^{-10}$, $d = 0.23$; negative: $t_{(3229)} = 4.06$, $p = 5 \times 10^{-5}$, $d = 0.15$). The specific advantage of remembering positive material for females could also be seen in the long delay (20–24 h delayed) free-recall task (interaction between sex and valence category: $F_{(2,3952)} = 21.66$, $p = 4.4 \times 10^{-10}$, $\eta^2 =$ 0.38%; main effect of sex positive: $t_{(1975)} = 10.42$, $p < 1 \times 10^{-16}$, $d = 0.5$; neutral: $t_{(1975)} = 5.54$, $p = 3.4 \times 10^{-8}$, $d = 0.27$; negative: $t_{(1975)} = 5.09$, $p = 3.8 \times 10^{-7}$, $d = 0.25$). The effect size for the females' advantage of positive material was medium. There was no significant three-way interaction ($F_{(2,9880)} = 0.38$, $p = 0.68$)

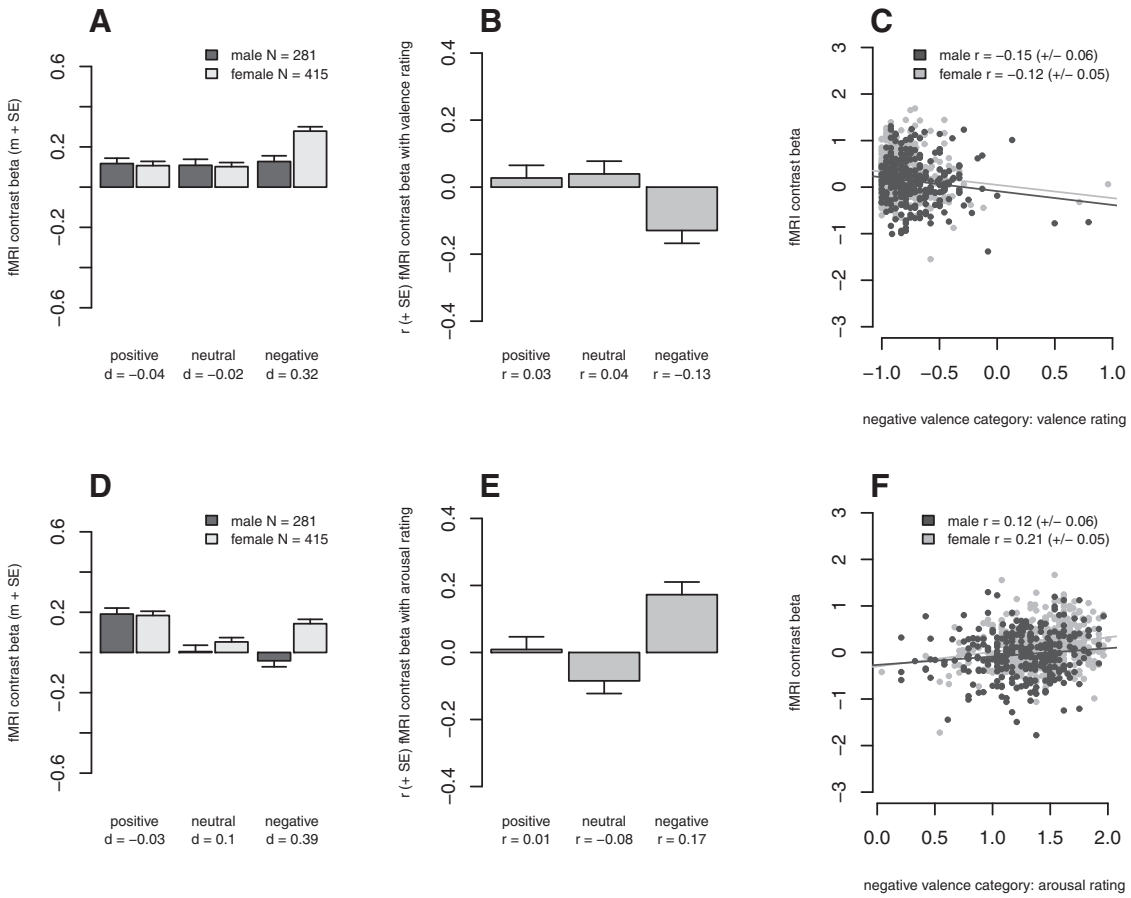


Figure 4. Picture-rating task during encoding, contrast meaningful versus scrambled. Depicted are the steps of the ROI analyses exemplary for the right cerebellum cortex 2 (***A***–***C***) and left precentral gyrus 3 (***D***–***F***). ***A***, ***D***, The significant interaction between sex and valence category. Positive values indicate that meaningful pictures compared with scrambled pictures were associated with a higher brain activation of the subjects. ***B***, ***E***, The association between the fMRI contrast parameter estimates and the averaged ratings of the subjects. For the valence rating (***B***), a negative correlation means that a larger difference in activation between meaningful and scrambled pictures leads to more negative ratings. For the arousal rating (***E***), a positive correlation implies that a larger difference in activation between meaningful and scrambled pictures leads to higher arousal ratings. ***C***, ***F***, The averaged ratings of negative pictures (*x*-axis) for all subjects against the fMRI contrast parameter estimate of negative versus scrambled pictures (*y*-axis) and the regression slopes for both sexes separately. m + SE, mean and standard error of the mean; d,r, effect sizes.

between valence, sex, and time-point (short- vs long-delay recall). Therefore, the women's special advantage for positive pictures did not change over the two time points. These results showed a different profile as compared with the analyses of the ratings. Women showed a more extreme appraisal especially of the negative pictures but a better memory performance especially for the positive pictures. Therefore, these two effects are most likely not connected to each other. Furthermore, females showed a better memory performance for neutral pictures, although there was no difference in emotional appraisal for this category.

To confirm that the above described sex- and valence-specific memory effects were independent of the influence of available confounding variables, we expanded our (reduced) linear model. We included the averaged valence and arousal ratings, the ratings reaction speed and words short-delay recall performance, as well as their interaction terms with valence category in our linear model (full model). For the effects of sex, valence category, and their interaction on reaction speed and words short-delay recall performance see Table 4. We performed an overall test (log-likelihood test) to determine whether these additional variables explained a significant amount of variance of the subjects' memory performance (for each variable separately and conjointly). Next, we investigated whether in the full model the significant sex- and valence-category interaction effect is still detectable. Finally, we determined whether the effect-sizes of the females' advantage in memory performance for the three valence categories separately changed when taking the additional variables into account (Table 5). In all models including the ratings or the words short-delay recall these covariates explained a significant amount of variance ($p < 0.007$). Including the reaction speed of the ratings as the only covariates could not explain a significant amount of variance ($p > 0.1$). Regardless of the covariates included, the interaction between sex and valence category was significant ($p < 0.0002$), and the interaction term F and p values of the corresponding full and reduced models were in a comparable range. When comparing the effect sizes of the females' memory advantage between the reduced and full model, there was a considerable decrease in d_{sex} for all three valence categories when including words short-delay performance as a covariate in the model (positive pictures: maximum $d_{reduced\text{-}full} = 0.07$; neutral pictures: maximum $d_{reduced\text{-}full} = 0.05$; negative pictures: maximum $d_{reduced\text{-}full} = 0.13$).

Together, compared with males, females rated especially negative pictures as more arousing and more negative during the picture presentation. Females also displayed stronger brain activation in mainly motor-relevant regions when viewing negative compared with scrambled pictures. However, in the free-recall test females outperformed males not only in negative pictures, but also in neutral pictures and especially in positive pictures.

Table 4. Analyses of possible confounding variables (covariates) regarding their effects of sex, valence category, and the interaction between sex and valence category

Variable	Interaction sex *x*-valence category	Main effect of valence category	Main effect of sex	Positive pictures only: main effect of sex	Neutral pictures only: main effect of sex	Negative pictures only: main effect of sex
Picture valence rating reaction speed	$F_{(2,3442)} = 19.97$ $p = 2.4 \times 10^{-9}$	$F_{(2,3442)} = 202.82$ $p < 1 \times 10^{-16}$	$F_{(1,1720)} = 5.71$ $p = 0.017$	$t_{(1720)} = -0.23$ $p = 0.82$ $d = -0.01$	$t_{(1720)} = -1.65$ $p = 0.099$ d = −0.08	$t_{(1720)} = -4.5$ $p = 7.4 \times 10^{-6}$ $d = -0.23$
Picture arousal rating reaction speed	$F_{(2,3366)} = 59.99$ $p < 1 \times 10^{-16}$	$F_{(2,3366)} = 50.61$ $p < 1 \times 10^{-16}$	$F_{(1,1682)} = 5.89$ $p = 0.015$	$t_{(1682)} = -0.71$ $p = 0.48$ $d = -0.04$	$t_{(1682)} = 0.39$ $p = 0.7$ $d = 0.02$	$t_{(1682)} = -6.46$ $p = 1.4 \times 10^{-10}$ $d = -0.33$
Words SD	$F_{(2,3878)} = 1.23$ $p = 0.29$	$F_{(2,3878)} = 83.7$ $p < 1 \times 10^{-16}$	$F_{(1,1938)} = 63.55$ $p = 2.7 \times 10^{-15}$	$t_{(1938)} = 6.51$ $p = 9.5 \times 10^{-11}$ $d = 0.32$	$t_{(1938)} = 5.78$ $p = 8.7 \times 10^{-9}$ $d = 0.28$	$t_{(1938)} = 4.76$ $p = 2.1 \times 10^{-6}$ $d = 0.23$

For reaction speed, there was a significant interaction between sex and valence category: males showed the slowest reaction times when viewing negative pictures (see Table 2). For the words short-delay (SD) memory there was a main effect of sex, with females in general outperforming males.

Table 5. Influence of possible confounding variables (covariates) on the interaction effect of sex and valence category regarding free-recall memory performance

			Full vs reduced		Reduced model sex x-valence category		Full model sex x-valence category		Positive d_{sex}		Neutral d_{sex}		Negative d_{sex}	
Task	Covariates	*N*	*LR*	*p*	*F*	*p*	*F*	*p*	Reduced	Full	Reduced	Full	Reduced	Full
Picture SD	Picture arousal rating (1)	3212	56.58	3.2×10^{-12}	34.75	1×10^{-15}	35.38	4.4×10^{-16}	0.45	0.44	0.23	0.23	0.15	0.14
	Picture valence rating (2)	3212	79.47	$<1 \times 10^{-16}$	34.75	1×10^{-15}	34.18	1.8×10^{-15}	0.45	0.44	0.23	0.24	0.15	0.12
	(1 + 2)	3212	118.62	$<1 \times 10^{-16}$	34.75	1×10^{-15}	34.54	1.2×10^{-15}	0.45	0.44	0.23	0.24	0.15	0.12
	Picture arousal rating reaction speed (3)	1683	2.99	0.39	31.47	2.9×10^{-14}	31.81	2.1×10^{-14}	0.47	0.47	0.16	0.16	0.13	0.12
	Picture valence rating reaction speed (4)	1721	6.16	0.1	28.96	3.4×10^{-13}	29.62	1.8×10^{-13}	0.47	0.47	0.17	0.17	0.15	0.13
	(1–4)	1679	84.27	6.3×10^{-13}	31.35	3.2×10^{-14}	32.45	1.1×10^{-14}	0.47	0.46	0.16	0.18	0.14	0.08
	Words SD (5)	1869	43.3	2.1×10^{-9}	16.92	4.8×10^{-8}	15.73	1.6×10^{-7}	0.52	0.47	0.29	0.24	0.24	0.21
	(1–2, 5)	1858	100.81	$<1 \times 10^{-16}$	16.64	6.4×10^{-8}	14.63	4.7×10^{-7}	0.53	0.46	0.3	0.25	0.25	0.18
	(1–5)	798	81.95	3.1×10^{-11}	13.62	1.4×10^{-6}	12.01	6.6×10^{-6}	0.6	0.53	0.26	0.22	0.3	0.17
Picture LD	Picture arousal rating (1)	1971	17.1	0.00067	21.56	4.9×10^{-10}	20.51	1.4×10^{-9}	0.51	0.51	0.28	0.28	0.25	0.25
	Picture valence rating (2)	1971	12.18	0.0068	21.56	4.9×10^{-10}	20.74	1.1×10^{-9}	0.51	0.5	0.28	0.28	0.25	0.23
	(1 + 2)	1971	26.64	0.00017	21.56	4.9×10^{-10}	20.18	1.9×10^{-9}	0.51	0.5	0.28	0.28	0.25	0.23
	Picture arousal rating reaction speed (3)	832	3.45	0.33	12.63	3.6×10^{-6}	12.42	4.40×10^{-6}	0.52	0.52	0.24	0.24	0.31	0.29
	Picture valence rating reaction speed (4)	851	4.61	0.2	12.49	4.1×10^{-6}	12.02	6.6×10^{-6}	0.53	0.52	0.25	0.25	0.32	0.31
	(1–4)	831	32.84	0.001	12.46	4.3×10^{-6}	11.25	1.4×10^{-5}	0.52	0.5	0.24	0.24	0.31	0.24
	Words SD (5)	1856	42.44	3.2×10^{-9}	19.24	4.9×10^{-9}	17.87	1.9×10^{-8}	0.5	0.45	0.27	0.22	0.25	0.21
	(1–2, 5)	1849	69.23	2.2×10^{-11}	19.14	5.4×10^{-9}	16.83	5.3×10^{-8}	0.51	0.45	0.28	0.23	0.25	0.2
	(1–5)	798	63.74	5.7×10^{-8}	11.55	1×10^{-5}	8.58	2×10^{-4}	0.51	0.44	0.23	0.18	0.31	0.21

Covariates were the valence and arousal ratings, as well as the reaction speeds of valence and arousal ratings during the picture-rating task. We additionally included the memory performance of a words short-delay task in the model. We tested the influence of each covariate separately and combinations of variables. Aim of the analyses was to determine, whether the sex and valence category interaction effect of the free recall memory performance was still detectable, when correcting for possible confounding variables. Full models included the covariates and their interaction term with valence category, whereas the reduced model did not include the covariates. LR, log-likelihood ratio; SD, short delay; LD, long delay.

When correcting for the ratings, reaction speed of ratings and words short-delay recall, the significant interaction between sex and valence category on memory performance was still significant. These data suggest that the sex- and valence-category-dependent differences in free recall were independent from sex- and valence-category-dependent differences in emotional appraisal, and could not be explained by confounding factors like reaction speed or memory performance of words.

fMRI data

From the previous fMRI analysis during the picture-rating task, contrasting meaningful versus scrambled pictures, we did not find an involvement of medial temporal lobe (MTL) regions regarding the interaction between sex and valence category. Thus, there was no hint for a special recruitment of MTL regions for emotional pictures that could explain the women's advantage in memory performance later on. To further investigate this issue, we added another fMRI analysis during the picture-rating task contrasting remembered versus not remembered pictures (first-level analysis: positive, negative and neutral remembered versus not remembered; subsequent memory effect). We calculated an ANOVA (second-level analysis) with sex as between-factor (male, female), valence category as within-factor (positive, neutral, negative), and the interaction term between sex and valence category. In the behavioral data, we observed a sex *x*-valence category interaction effect regarding memory performance, with females showing a better memory performance especially for positive pictures. Therefore, our main interest was also on the sex *x*-valence category interaction effects in the fMRI analyses, which showed no significant results. In addition, the SVC, which restricted the analysis to the MTL, did not show any significant clusters for the interaction term. For the main effect of sex, no suprathreshold cluster was found. Results of the main effect of valence are presented in Figure 3*C*.

To summarize, females showed a memory performance advantage particularly for positive pictures, which was independent of their more extreme ratings in the encoding phase of the experiment. The fMRI interaction analysis for sex and valence category

comparing remembered versus nonremembered pictures (subsequent memory) showed no significant cluster at the whole-brain level. Even at lower threshold (SVC) we did not identify regions in the MTL, which were recruited by females in particular when viewing positive pictures during the picture-rating task.

Task 3: picture memory task, recognition

Overview

In the fMRI analyses of the picture-rating task during picture encoding we did not find evidence for memory-relevant valence category-specific sex differences. The question arises, whether the valence category-specific sex effects carried over to a second memory task, the picture recognition task. The main analysis was based on the correctly recognized old pictures; as control conditions, we also analyzed the incorrectly remembered new pictures (false alarm) and analyzed a combined model including correctly recognized old pictures and false alarms. The pictures that had to be recognized were the same pictures as in the picture-rating task, which already had to be freely recalled.

Behavioral data

Across both sexes, subjects' memory performances (correctly recognized old pictures) differed substantially between the three valence categories ($F_{(2,2436)} = 159.56$, $p < 1 \times 10^{-16}$, $\eta^2 = 2.16\%$; Fig. 1*E*). *Post hoc* tests showed that pictures from the positive ($t_{(1219)} = -4.36$, $p = 1.4 \times 10^{-5}$), as well as negative ($t_{(1219)} = 16.04$, $p < 1 \times 10^{-16}$) valence category were significantly better remembered than neutral pictures.

There was a significant interaction effect between sex and valence category ($F_{(2,2436)} = 8.87$, $p = 0.00015$, $\eta^2 = 0.38\%$). *Post hoc* test showed a significant advantage of males in recognizing negative pictures ($t_{(1217)} = -4.29$, $p = 1.9 \times 10^{-5}$, $d = -0.25$; but see additional analysis in the following paragraph). There was neither a significant sex difference for positive pictures ($t_{(1217)} = -0.65$, $p = 0.51$, $d = -0.04$), nor for neutral pictures ($t_{(1217)} = -1.88$, $p = 0.06$, $d = -0.11$). There was also no Bonferroni-corrected ($p < 0.01$) significant main effect of sex ($F_{(1,1217)} = 5.56$, $p_{\text{nominal}} = 0.019$). Therefore, it was not possible to show that the sex and valence category interaction effect of the free recall, favoring females especially for positive pictures, carried over to the subsequent recognition task. The significant interaction effect between sex and valence category for correctly recognizing old pictures could not be shown for the false alarms in the same recognition task ($F_{(2,2436)} = 0.21$, $p = 0.81$; Fig. 1*F*).

We additionally analyzed correctly recognized old pictures and false alarms in one model to account for a possible response bias in the recognition task (Windmann and Kutas, 2001). The three-way interaction analyzing sex, valence category and task (correctly recognizing old pictures and false alarms), was not significant ($F_{(2,6085)} = 1.24$, $p = 0.29$). There was a significant two-way interaction between valence category and task ($F_{(2,6090)} = 52.67$, $p = 4.79 \times 10^{-13}$), and a significant main effect of sex ($F_{(2,6090)} = 6.7$, $p = 0.0098$). All other two-way interactions were not significant (sex *x*-task: $F_{(1,6090)} = 2.11$, $p = 0.15$; sex *x*-valence category: $F_{(1,6090)} = 1.92$, $p = 0.15$). Given the observed pattern in the data after having taken into account the false alarms (Fig. 1 *E*,*F*), the recognition performance for negative pictures cannot be considered as especially superior in males than in females.

fMRI data

In the first-level analysis, we assessed activity during the recognition of pictures by contrasting activity during the presentation of old pictures against activity during the presentation of new pictures. In the second-level analysis, we calculated an ANOVA with sex as between-factor (male, female), valence as within-factor (positive, neutral, negative), and the interaction term between sex and valence. In the behavioral analyses, we found a significant interaction between sex and valence category regarding recognition performance when analyzing correctly recognized old pictures only, with males showing a better memory performance particularly for negative pictures. Our main interest was also in the sex *x*-valence category interaction effects in the fMRI analyses, which showed no significant results. In addition, the SVC did not show any significant clusters for the interaction term. Figure 3 *D*,*E* shows the results of the main effects of sex and valence.

Together, the females' memory advantage in the free recall setting particularly for positive pictures was not found in the recognition setting. This suggests that the sex- and valence-dependent differences in memory performances were: (1) task-specific and (2) not due to sex- and valence-category-dependent differences in appraisal during encoding. Furthermore, the fMRI interaction analysis for sex and valence comparing old versus new pictures showed no significant cluster on WB level no more than when applying a small volume correction for the MTL regions only.

Discussion

By analyzing behavioral data of four different samples comprising >3300 subjects we were able to show that the women's stronger appraisal of emotional material, especially for negative pictures, is accompanied by a stronger activation of motor-relevant brain regions and the posterior cingulate when viewing negative pictures. However, this stronger reactivity in the encoding phase to negative material was not linked to a corresponding sex and valence category dependent difference in memory performance later on, although we could show that across sexes emotional stimuli were remembered better than neutral stimuli. By comparing the memory data of two subsequent tasks, a free-recall task and a recognition task, we were able to show that sex differences regarding memory performance were dependent on valence category and task. Specifically, women showed a special advantage for remembering positive pictures in a free-recall task, which was absent in a recognition task. We could further show that the females' advantage for positive pictures in the free-recall tasks lasted for at least 24 h.

The finding of a more extreme appraisal of emotional material in females compared with males, in particular for the negative valence category, is interesting in the context of vulnerability to neuropsychiatric disorders (Earls, 1987; Culbertson, 1997; Weinstock, 1999; Holden, 2005). Emotional dysregulation is a common component of many neuropsychiatric disorders (Cole et al., 1994; Kring and Sloan, 2009) and women are more likely to develop major depression, anxiety disorder, and post-traumatic stress disorder (Eysenck et al., 1991; Donaldson et al., 2007; Mohlman et al., 2007; Liu et al., 2012). In our data, the stronger reactivity of females especially to negative material, measured by judgments of the perceived valence and arousal, was related to higher brain activations in motor-relevant regions and the posterior cingulate. This pattern might suggest that females might be better prepared to physically react to negative events than males. Other studies using ERPs, EMG, startle response, and facial expression (Grossman and Wood, 1993; Kring and Gordon, 1998; Bradley et al., 2001; Gard and Kring, 2007; Lithari et al., 2010) also indicated increased facial and motor reactions especially upon negative emotional stimuli presentation in females compared with males. For the interpretation of these findings it is important to note that subjective judgments of valence and

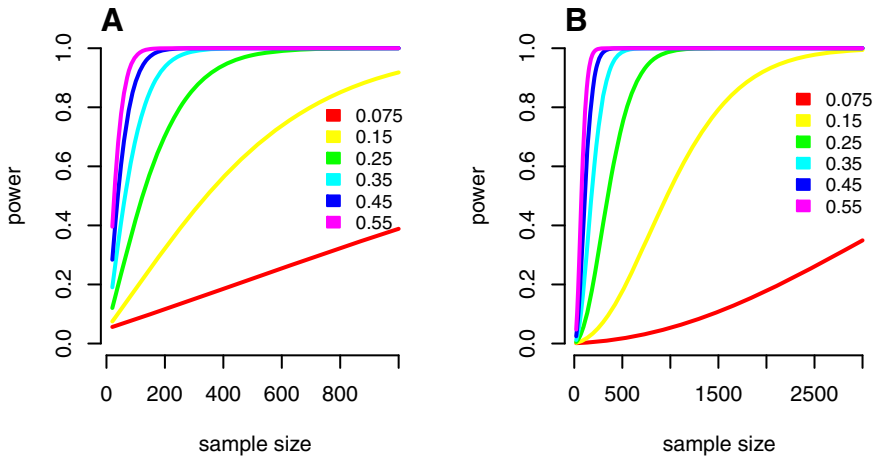


Figure 5. Power-analyses for the sex effects of the behavioral data. The graphs illustrate the necessary sample sizes to be adequately powered (80%) to replicate the reported ranges of effect-sizes *d* in an independent sample, assuming a false-positive rate $\alpha = 0.05$ (***A***) or $\alpha = 0.001$ (***B***). The analyses were done with the pwr package (Champely, 2009) in R (R Development Core Team, 2011).

arousal are differentially related to the actual physiological responses and do not exclusively reflect physiological arousal. Valence ratings have been linked to heart rate and facial EMG, whereas arousal ratings are more closely associated with skin conductance (Lang et al., 1993). Another explanation for the more extreme ratings in females are normative expectations with females being expected to be more emotional, pointing to more social aspects of the sex-differences in emotional appraisal ratings (Fischer, 1993; Grossman and Wood, 1993; Barrett et al., 1998).

Regarding the females' advantage in memory tasks, it has been discussed that the memory advantage might be confounded with a females' advantage in verbal tasks, and that it is hardly possible to disentangle these two mechanisms (Andreano and Cahill, 2009). In our study as well, better verbal abilities may have contributed to females' general advantage in the free-recall task. An indirect hint can be seen in our data by including the word short-delay recall performance as covariate in the analyses. Correcting for word short-delay recall led to a valence-category-independent decrease in differences in memory performance between males and females, whereas the specific females' advantage for positive pictures was still present.

Regarding the differences in the interaction effect between sex and valence category in free recall versus recognition, several explanations are possible. For example, processes taking place shortly before or during encoding may vary in their impact on different tasks, on different valence categories and also on males and females (Zoladz et al., 2013). There are hints that free recall and recognition are based not only on shared, but also on task-specific encoding mechanisms (Staresina and Davachi, 2006). It is also possible that the free-recall task interfered with the memory formation and influenced the later recognition task, albeit in an unexpected manner, because the females' special advantages in free recall could not be replicated in recognition. Additionally, interaction effects between sex and valence category might depend on task difficulty. The overall performance in the recognition task was higher than in the free-recall task, indicating differences in task difficulty. Furthermore, it has been argued, that differences in remember rates can indicate differences in response bias, rather than reflecting successful recollection (Windmann and Kutas, 2001; Dougal and Rotello, 2007). In our data, we found evidence suggesting a general sex-dependent difference in response rate, with higher response rates in males.

It is known that the more similar the processes during encoding and retrieval are, the more likely the material will be remembered later, but that these effects depend on task difficulty, context, and retrieval mode (Morris et al., 1977; Barak et al., 2013; Parks, 2013). Therefore, it is possible that the transfer from free recall to recognition is also influenced by the subjects' sex, and by the encoded material. Especially because we could show with our data that the appraisal of the material was dependent on sex and valence category during encoding.

We could identify corresponding patterns in fMRI during encoding regarding the interaction between sex and valence category on picture ratings. However, it was not possible to show corresponding patterns between behavior and fMRI for the subsequent memory effect during encoding. We cannot rule out the possibility that the lack of valence-category-specific sex differences in brain activity might have been influenced by the heterogeneity of the females group concerning their use of birth control methods, as well as admixture of women in different stages of their cycle as reported in literature for several cognitive domains (Rumberg et al., 2010; Bonenberger et al., 2013; Marecková et al., 2014). It would be interesting in future studies to investigate the detailed role of hormonal contraceptives and menstrual cycle in the context of the here observed valence-specific sex differences (Ertman et al., 2011).

Small sample size has been identified as an issue undermining the reliability of findings in neuroscience (Ioannidis, 2008; Button et al., 2013). Importantly, our study was well powered for effect sizes typically observed in neuroscience (Kühberger et al., 2014). Whereas the observed effects of valence category in our study are in a medium to large effect size range, the sex effects are, as expected (Hyde and Linn, 1988; Hyde, 2005; Lindberg et al., 2010), in a small to medium effect-size range. For the sex-and valence-category-interaction effect we see small effects only, which can at least partially be explained by the observed interaction pattern: Most times we see a consistent main effect, e.g., females outperforming males in memory performance, which is modulated by the valence category, e.g., females showing a special advantage for positive pictures. The effect size of an interaction effect not only depends on the pattern of interaction, but also on the effect size of the main effects (Whisman and McClelland, 2005), and in a mixed model design on the correlation between the repeated measurements. Therefore, the interpretation of an effect size in the context of a mixed model interaction term is difficult. Given the nature of complex cognitive traits and complex diseases, which emerge due to the combination of genetic and environmental background and also gene-environment interactions, one would not expect a single factor to explain a large portion of the observed variation. In the case of sex effects, obvious differences in genetic background additionally affect hormone levels and most likely interact with environmental factors. All these factors conjointly result in a given complex phenotype. The interaction analyses allowed us to study an additional modulatory factor, the three valence categories of the stimulus material, which influenced the observed association between sex and the investigated phenotypes. These observations can serve as a

starting point, to further disentangle possible influential factors related to valence category on the sex and phenotype associations. Considering the small to medium effect sizes detected in this study, it is critical to design a priori well powered studies. Figure 5 provides information about the sample sizes necessary for replication of the here reported main effects of sex only.

Together, the present findings suggest that the valence category-specific sex differences in emotional appraisal and in free recall of pictures are likely two independent phenomena. The females' stronger reaction to negative stimuli is paralleled by a stronger activation of motor-relevant brain regions during the encoding and rating of the material, but is not paralleled by a better recall or recognition particularly of negative material later on. By comparing two different memory tasks, a free recall and a recognition task, which were based on the same encoded material, we were able to show that the sex and valence category-specific differences in memory performance were highly task-dependent. In a free-recall setting, females outperformed males especially for positive material, although in the recognition setting this effect was absent. fMRI during encoding did not reveal activation differences that reflected the females' advantage of positive pictures in free recall.

References

Andreano JM, Cahill L (2009) Sex influences on the neurobiology of learning and memory. Learn Mem 16:248–266. CrossRef Medline

Bakeman R (2005) Recommended effect size statistics for repeated measures designs. Behav Res Methods 37:379–384. CrossRef Medline

Balliet D, Li NP, Macfarlan SJ, Van Vugt M (2011) Sex differences in cooperation: a meta-analytic review of social dilemmas. Psychol Bull 137:881–909. CrossRef Medline

Bao AM, Swaab DF (2011) Sexual differentiation of the human brain: relation to gender identity, sexual orientation and neuropsychiatric disorders. Front Neuroendocrinol 32:214–226. CrossRef Medline

Barak O, Vakil E, Levy DA (2013) Environmental context effects on episodic memory are dependent on retrieval mode and modulated by neuropsychological status. Q J Exp Psychol (Hove) 66:2008–2022. CrossRef Medline

Barrett LF, Robin L, Pietromonaco PR, Eyssell KM (1998) Are women the "more emotional" sex? Evidence from emotional experiences in social context. Cogn Emot 12:555–578. CrossRef

Blalock HM Jr (1966) The identification problem and theory building: the case of status inconsistency. Am Sociol Rev 31:52–61. CrossRef

Bloise SM, Johnson MK (2007) Memory for emotional and neutral information: gender and individual differences in emotional sensitivity. Memory 15:192–204. CrossRef Medline

Bonenberger M, Groschwitz RC, Kumpfmueller D, Groen G, Plener PL, Abler B (2013) It's all about money: oral contraception alters neural reward processing. Neuroreport 24:951–955. CrossRef Medline

Bradley MM, Codispoti M, Sabatinelli D, Lang PJ (2001) Emotion and motivation II: sex differences in picture processing. Emotion 1:300–319. CrossRef Medline

Büchel C, Holmes AP, Rees G, Friston KJ (1998) Characterizing stimulus-response functions using nonlinear regressors in parametric fMRI experiments. Neuroimage 8:140–148. CrossRef Medline

Button KS, Ioannidis JP, Mokrysz C, Nosek BA, Flint J, Robinson ES, Munafò MR (2013) Power failure: why small sample size undermines the reliability of neuroscience. Nat Rev Neurosci 14:365–376. CrossRef Medline

Cabeza R, Nyberg L (2000) Neural bases of learning and memory: functional neuroimaging evidence. Curr Opin Neurol 13:415–421. CrossRef Medline

Cahill L (2006) Why sex matters for neuroscience. Nat Rev Neurosci 7:477–484. CrossRef Medline

Cahill L (2014) Fundamental sex difference in human brain architecture. Proc Natl Acad Sci U S A 111:577–578. CrossRef Medline

Cahill L, Haier RJ, Fallon J, Alkire MT, Tang C, Keator D, Wu J, McGaugh JL (1996) Amygdala activity at encoding correlated with long-term, free recall of emotional information. Proc Natl Acad Sci U S A 93:8016–8021. CrossRef Medline

Champely S (2009) pwr: Basic functions for power analysis. R package version 1.1.1.

Cohen J (1992) A power primer. Psychol Bull 112:155–159. CrossRef Medline

Cole PM, Michel MK, Teti LO (1994) The development of emotion regulation and dysregulation: a clinical perspective. Monogr Soc Res Child Dev 59:73–100. CrossRef Medline

Cross CP, Copping LT, Campbell A (2011) Sex differences in impulsivity: a meta-analysis. Psychol Bull 137:97–130. CrossRef Medline

Culbertson FM (1997) Depression and gender: an international review. Am Psychol 52:25–31. CrossRef Medline

Davis MC, Matthews KA, Twamley EW (1999) Is life more difficult on Mars or Venus? A meta-analytic review of sex differences in major and minor life events. Ann Behav Med 21:83–97. CrossRef Medline

de Frias CM, Nilsson L-G, Herlitz A (2006) Sex differences in cognition are stable over a 10-year period in adulthood and old age. Neuropsychol Dev Cogn B Aging Neuropsychol Cogn 13:574–587. CrossRef Medline

de Quervain DJ, Henke K, Aerni A, Treyer V, McGaugh JL, Berthold T, Nitsch RM, Buck A, Roozendaal B, Hock C (2003) Glucocorticoid-induced impairment of declarative memory retrieval is associated with reduced blood flow in the medial temporal lobe. Eur J Neurosci 17:1296–1302. CrossRef Medline

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980. CrossRef Medline

Donaldson C, Lam D, Mathews A (2007) Rumination and attention in major depression. Behav Res Ther 45:2664–2678. CrossRef Medline

Dougal S, Rotello CM (2007) "Remembering" emotional words is based on response bias, not recollection. Psychon Bull Rev 14:423–429. CrossRef Medline

Earls F (1987) Sex differences in psychiatric disorders: origins and developmental influences. Psychiatr Dev 5:1–23. Medline

Egan SK, Perry DG (2001) Gender identity: a multidimensional analysis with implications for psychosocial adjustment. Dev Psychol 37:451–463. CrossRef Medline

Ertman N, Andreano JM, Cahill L (2011) Progesterone at encoding predicts subsequent emotional memory. Learn Mem 18:759–763. CrossRef Medline

Eysenck MW, Mogg K, May J, Richards A, Mathews A (1991) Bias in interpretation of ambiguous sentences related to threat in anxiety. J Abnorm Psychol 100:144–150. CrossRef Medline

Fischer AH (1993) Sex differences in emotionality: fact or stereotype? Fem Psychol 3:303–318. CrossRef

Fischl B, Salat DH, Busa E, Albert M, Dieterich M, Haselgrove C, van der Kouwe A, Killiany R, Kennedy D, Klaveness S, Montillo A, Makris N, Rosen B, Dale AM (2002) Whole brain segmentation. Neuron 33:341–355. CrossRef Medline

Gard MG, Kring AM (2007) Sex differences in the time course of emotion. Emotion 7:429–437. CrossRef Medline

Grossman M, Wood W (1993) Sex differences in intensity of emotional experience: a social role interpretation. J Pers Soc Psychol 65:1010–1022. CrossRef Medline

Hager W, Hasselhorn M (1994) Handbuch deutschsprachiger Wortnormen, XIII. Göttingen, Germany: Hogrefe.

Henke K, Kroll NE, Behniea H, Amaral DG, Miller MB, Rafal R, Gazzaniga MS (1999) Memory lost and regained following bilateral hippocampal damage. J Cogn Neurosci 11:682–697. CrossRef Medline

Herlitz A, Nilsson LG, Bäckman L (1997) Gender differences in episodic memory. Mem Cognit 25:801–811. CrossRef Medline

Herlitz A, Reuterskiöld L, Lovén J, Thilers PP, Rehnman J (2013) Cognitive sex differences are not magnified as a function of age, sex hormones, or puberty development during early adolescence. Dev Neuropsychol 38:167–179. CrossRef Medline

Holden C (2005) Sex and the suffering brain. Science 308:1574–1577. CrossRef Medline

Hyde JS (2005) The gender similarities hypothesis. Am Psychol 60:581–592. CrossRef Medline

Hyde JS, Linn MC (1988) Gender differences in verbal ability: a meta-analysis. Psychol Bull 104:53–69. CrossRef

Ingalhalikar M, Smith A, Parker D, Satterthwaite TD, Elliott MA, Ruparel K,

Hakonarson H, Gur RE, Gur RC, Verma R (2014) Sex differences in the structural connectome of the human brain. Proc Natl Acad Sci U S A 111:823–828. CrossRef Medline

Ioannidis JP (2008) Why most discovered true associations are inflated. Epidemiology 19:640–648. CrossRef Medline

Jazin E, Cahill L (2010) Sex differences in molecular neuroscience: from fruit flies to humans. Nat Rev Neurosci 11:9–17. CrossRef Medline

Kramer JH, Delis DC, Kaplan E, O'Donnell L, Prifitera A (1997) Developmental sex differences in verbal learning. Neuropsychology 11:577–584. CrossRef Medline

Kreft IGG, Kreft I, de Leeuw J (1998) Introducing multilevel modeling. London: SAGE Publications.

Kring AM, Gordon AH (1998) Sex differences in emotion: expression, experience, and physiology. J Pers Soc Psychol 74:686–703. CrossRef Medline

Kring AM, Sloan DM (2009) Emotion regulation and psychopathology: a transdiagnostic approach to etiology and treatment. New York: Guilford.

Kudielka BM, Kirschbaum C (2005) Sex differences in HPA axis responses to stress: a review. Biol Psychol 69:113–132. CrossRef Medline

Kühberger A, Fritz A, Scherndl T (2014) Publication bias in psychology: a diagnosis based on the correlation between effect size and sample size. PLoS One 9:e105825. CrossRef Medline

LaBar KS, Cabeza R (2006) Cognitive neuroscience of emotional memory. Nat Rev Neurosci 7:54–64. CrossRef Medline

Lang PJ, Öhmann A, Vaitl D (1988) The international affective picture system (slides). Gainesville, FL: Center for Research in Psycholophysiology, University of Florida.

Lang PJ, Greenwald MK, Bradley MM, Hamm AO (1993) Looking at pictures: affective, facial, visceral, and behavioral reactions. Psychophysiology 30:261–273. CrossRef Medline

Lawrence MA (2012) ez: Easy analysis and visualization of factorial experiments. R package version 3.0–0.

Lindberg SM, Hyde JS, Petersen JL, Linn MC (2010) New trends in gender and mathematics performance: a meta-analysis. Psychol Bull 136:1123–1135. CrossRef Medline

Lithari C, Frantzidis CA, Papadelis C, Vivas AB, Klados MA, Kourtidou-Papadeli C, Pappas C, Ioannides AA, Bamidis PD (2010) Are females more responsive to emotional stimuli? A neurophysiological study across arousal and valence dimensions. Brain Topogr 23:27–40. CrossRef Medline

Liu WH, Wang LZ, Zhao SH, Ning YP, Chan RC (2012) Anhedonia and emotional word memory in patients with depression. Psychiatry Res 200:361–367. CrossRef Medline

Marecková K, Perrin JS, Nawaz Khan I, Lawrence C, Dickie E, McQuiggan DA, Paus T (2014) Hormonal contraceptives, menstrual cycle and brain response to faces. Soc Cogn Affect Neurosci 9:191–200. CrossRef Medline

McCarthy MM, Konkle AT (2005) When is a sex difference not a sex difference? Front Neuroendocrinol 26:85–102. CrossRef Medline

McGaugh JL (2004) The amygdala modulates the consolidation of memories of emotionally arousing experiences. Annu Rev Neurosci 27:1–28. CrossRef Medline

McGaugh JL, Roozendaal B (2002) Role of adrenal stress hormones in forming lasting memories in the brain. Curr Opin Neurobiol 12:205–210. CrossRef Medline

McLean CP, Anderson ER (2009) Brave men and timid women? A review of the gender differences in fear and anxiety. Clin Psychol Rev 29:496–505. CrossRef Medline

Meyer-Bahlburg HF (2010) From mental disorder to iatrogenic hypogonadism: dilemmas in conceptualizing gender identity variants as psychiatric conditions. Arch Sex Behav 39:461–476. CrossRef Medline

Miettunen J, Jääskeläinen E (2010) Sex differences in Wisconsin Schizotypy scale: a meta-analysis. Schizophr Bull 36:347–358. CrossRef Medline

Milner B (1972) Disorders of learning and memory after temporal lobe lesions in man. Clin Neurosurg 19:421–446. Medline

Mohlman J, Carmin CN, Price RB (2007) Jumping to interpretations: social anxiety disorder and the identification of emotional facial expressions. Behav Res Ther 45:591–599. CrossRef Medline

Morris CD, Bransford JD, Franks JJ (1977) Levels of processing versus transfer appropriate processing. J Verbal Learn Verbal Behav 16:519–533. CrossRef

Papassotiropoulos A, de Quervain DJ (2011) Genetics of human episodic memory: dealing with complexity. Trends Cogn Sci 15:381–387. CrossRef Medline

Parks CM (2013) Transfer-appropriate processing in recognition memory: perceptual and conceptual effects on recognition memory depend on task demands. J Exp Psychol Learn Mem Cogn 39:1280–1286. CrossRef Medline

Phelps EA (2004) Human emotion and memory: interactions of the amygdala and hippocampal complex. Curr Opin Neurobiol 14:198–202. CrossRef Medline

Pinheiro J, Bates D, DebRoy S, Sarkar D, R Core Team (2011) nlme: linear and nonlinear mixed effects models. In: R package, 3.1–102 Edition.

R Development Core Team (2011) R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.

Read S, Pedersen NL, Gatz M, Berg S, Vuoksimaa E, Malmberg B, Johansson B, McClearn GE (2006) Sex differences after all those years? Heritability of cognitive abilities in old age. J Gerontol B Psychol Sci Soc Sci 61:P137–P143. Medline

Roozendaal B, McGaugh JL (2011) Memory modulation. Behav Neurosci 125:797–824. CrossRef Medline

Rumberg B, Baars A, Fiebach J, Ladd ME, Forsting M, Senf W, Gizewski ER (2010) Cycle and gender-specific cerebral activation during a verb generation task using fMRI: comparison of women in different cycle phases, under oral contraception, and men. Neurosci Res 66:366–371. CrossRef Medline

Schacter DL, Wagner AD (1999) Medial temporal lobe activations in fMRI and PET studies of episodic encoding and retrieval. Hippocampus 9:7–24. CrossRef Medline

Staresina BP, Davachi L (2006) Differential encoding mechanisms for subsequent associative recognition and free recall. J Neurosci 26:9162–9172. CrossRef Medline

Su R, Rounds J, Armstrong PI (2009) Men and things, women and people: a meta-analysis of sex differences in interests. Psychol Bull 135:859–884. CrossRef Medline

Tolin DF, Foa EB (2006) Sex differences in trauma and posttraumatic stress disorder: a quantitative review of 25 years of research. Psychol Bull 132:959–992. CrossRef Medline

Trent S, Davies W (2012) The influence of sex-linked genetic mechanisms on attention and impulsivity. Biol Psychol 89:1–13. CrossRef Medline

Volk HE, McDermott KB, Roediger HL 3rd, Todd RD (2006) Genetic influences on free and cued recall in long-term memory tasks. Twin Res Hum Genet 9:623–631. CrossRef Medline

Weinstock LS (1999) Gender differences in the presentation and management of social anxiety disorder. J Clin Psychiatry 60:9–13. CrossRef Medline

Whisman MA, McClelland GH (2005) Designing, testing, and interpreting interactions and moderator effects in family research. J Fam Psychol 19:111–120. CrossRef Medline

Windmann S, Kutas M (2001) Electrophysiological correlates of emotion-induced recognition bias. J Cogn Neurosci 13:577–592. CrossRef Medline

Zoladz PR, Warnecke AJ, Woelke SA, Burke HM, Frigo RM, Pisansky JM, Lyle SM, Talbot JN (2013) Prelearning stress that is temporally removed from acquisition exerts sex-specific effects on long-term memory. Neurobiol Learn Mem 100:77–87. CrossRef Medline

4.6. Hippocampal activation, memory performance in young and old, and the risk for sporadic Alzheimer's disease converge genetically to calcium signaling

Heck, A., Fastenrath, M., Coynel, D., Auschra, B., Bickel, H., Freytag, V., Gschwind, L., Hartmann, F., Jessen, F., Kaduszkiewicz, H., Maier, W., Milnik, A., Pentzek, M., Riedel-Heller, S.G., Spalek, K., Vogler, C., Wagner, M., Weyerer, S., Wolfsgruber, S., de Quervain, D.F.-J., & Papassotiropoulos, A. *Submitted.*

Title: Hippocampal activation, memory performance in young and old, and the risk for sporadic Alzheimer's disease converge genetically to calcium signaling

Authors: Angela Heck, PhD[1,2*], Matthias Fastenrath, PhD[1,3], David Coynel, PhD[1,3], Bianca Auschra, MS[1], Horst Bickel, PhD [4], Virginie Freytag, MS[1], Leo Gschwind, MS[1,3], Francina Hartmann, MS[1], Frank Jessen, MD[5,6], Hanna Kaduszkiewicz, MD[7], Wolfgang Maier, MD[5,6], Annette Milnik, MD[1,8], Michael Pentzek, PhD[9], Steffi G. Riedel-Heller, MD[10], Klara Spalek, PhD , Christian Vogler PhD[1,2], Michael Wagner PhD[5,6], Siegfried Weyerer PhD[11], Steffen Wolfsgruber, MS[5,6], Dominique F. de Quervain, MD[2,3+], Andreas Papassotiropoulos, MD[1,2,12*+]

Affiliations:

[1] Division of Molecular Neuroscience, Department of Psychology, University of Basel, CH-4055 Basel, Switzerland

[2] Psychiatric University Clinics, University of Basel, CH-4055 Basel, Switzerland

[3] Division of Cognitive Neuroscience, Department of Psychology, University of Basel, CH-4055 Basel, Switzerland

[4] Department of Psychiatry, Technical University of Munich, 85748 Munich, Germany

[5] Department of Psychiatry, University of Bonn, 53105 Bonn, Germany

[6] DZNE, German Center for Neurodegenerative Diseases, 53175 Bonn, Germany

[7] Kiel University, Medical Faculty, Institute of General Practice, 24105 Kiel, Germany

[8] Department of Neurology, University of Magdeburg, 39120 Magdeburg, Germany

[9] Institute of General Practice, Medical Faculty, Heinrich-Heine-University Düsseldorf, 40225 Düsseldorf, Germany

[10] Institute of Social Medicine, Occupational Health and Public Health, Medical Faculty, University of Leipzig, 04109 Leipzig, Germany

[11] Central Institute of Mental Health, Medical Faculty Mannheim / Heidelberg University, 68159 Mannheim, Germany

[12] Department Biozentrum, Life Sciences Training Facility, University of Basel, CH-4056 Basel, Switzerland

***Correspondence to**: andreas.papas@unibas.ch, angela.heck@unibas.ch

\+ Co-senior author

Author contributions: A.H., M.F., H.B., F.J., H.K., W.M., S.R-H, M.W., S.We., D. J.-F. d. Q. & A.P. designed and conceptualized the study. A.H., M.F., B.A., D.C., L.G., V.F., F.J., H.K., A.M., M.P., S.R-H, C.V., M.W., S. We., S. Wo. & A.P. did the data analysis. F.H., H.B., W.M., M.P., K.S., M.W., S.We. & S.Wo. were responsible for data collection and preparation. D. J.-F. d. Q. & A.P. supervised the project. All authors contributed to the writing of the manuscript.

Conflicts of interest: There are no conflicts of interest to declare.

Word count: 3’143

Abstract:

Importance Human episodic memory performance is linked to the function of specific brain regions, including the hippocampus, declines as a result of increasing age, and is markedly disturbed in Alzheimer's disease (AD), an age-associated neurodegenerative disorder affecting primarily the hippocampus. Exploring the molecular underpinnings of human episodic memory is key to the understanding of hippocampus-dependent cognitive physiology and pathophysiology.

Objective To determine whether biologically defined groups of genes are enriched in episodic memory performance across ages, in memory encoding-related brain activity and in AD.

Design, Setting, and Participants In this multicenter collaborative study, gene set enrichment analysis was done by using primary and meta-analysis data from 57968 participants. The Swiss cohorts consisted of 3043 healthy young adults assessed for episodic memory performance. In a subgroup (1119 participants) of one of these cohorts, functional magnetic resonance imaging (fMRI) was used to identify gene set-dependent differences in brain activity related to episodic memory. The German AgeCoDe cohort consisted of 763 non-demented elderly participants assessed for episodic memory performance. The International Genomics of Alzheimer's Project (IGAP) case-control sample consisted of 54162 participants (17008 patients with sporadic AD, 37154 controls).

Main Outcomes and Measures Gene set enrichment analysis in all samples was done by using genome-wide single nucleotide polymorphism (SNP) data. Episodic memory performance in the Swiss and AgeCoDe cohorts was quantified by picture and verbal delayed free recall tasks. In the fMRI experiment, activation of the hippocampus during encoding of pictures served as the phenotype of interest. In the IGAP sample, diagnosis of sporadic AD served as the phenotype of interest.

Results We detected significant and consistent enrichment for genes constituting the Calcium Signaling Pathway (KEGG entry: hsa04020), especially those related to the elevation of cytosolic calcium. This enrichment was observed in episodic memory performance in young and old, in hippocampal activation, and in the risk for sporadic AD.

Conclusion and Relevance By detecting consistent significant enrichment in independent cohorts of young and elderly participants, this study identifies calcium signaling as a central player of hippocampus-dependent human memory processes, both in cognitive health and disease and contributes to the understanding -and hopefully treatment- of

Running title: Calcium signaling genes in episodic memory and AD

Introduction

Episodic memory (EM), i.e. the ability to encode and retrieve a particular event along with its contextual information,[1] is a polygenic cognitive trait, characterized by large interindividual variability and substantial heritability.[2-5] (for review see[5]) As a consequence of physiological ageing processes in such brain regions as the hippocampus and the medial temporal lobe, performance in EM tasks declines with age.[6-9] Pathological EM impairment is a behavioral hallmark of age-related neurodegenerative conditions, such as Alzheimer's disease (AD).[10,11]

Genome-wide studies utilizing single-marker statistics have been successful in identifying single loci linked to intact and impaired EM.[5,12-14] However, despite the obvious conception that the analysis of genetically complex traits should account for the underlying biological complexity, the vast majority of large-scale genetic association studies to date are restricted to the use of single-marker statistics. Clearly, this approach does not fully account for the polygenic nature of the phenotype under study. Triggered by statistical approaches for the analysis of gene expression, gene-set enrichment analysis (GSEA) has recently become available. By taking into account prior biological knowledge, GSEA examines whether test statistics for a group of related genes have consistent deviation from chance.[15,16] Thus, GSEA methods aim at identifying biologically meaningful sets of genes associated with a certain trait, rather than focusing on a single locus. As shown recently in studies on working memory,[17] autism,[18] bipolar disorder,[19-21], ADHD[22] and schizophrenia.[21,23,24] GSEA can identify convergent molecular pathways relevant to neuropsychiatry.

Here we studied the enrichment of biologically defined gene sets in EM across ages, in EM-related brain activity, and in an EM-related neurodegenerative disorder (Fig. 1). Genome-wide GSEA of EM performance was performed in multiple independent data sets

of young and aged cognitively healthy subjects (n=3806). In a large case-control sample (n=54162) we also performed GSEA for the risk of sporadic AD. We show that genes constituting the calcium signaling pathway, which comprises molecules regulating signal transduction and neuronal synaptic transmission, are consistently linked to EM performance across ages, and to sporadic AD. In a subsample of 1119 healthy young subjects who participated in a functional magnetic resonance imaging (fMRI) study, calcium signaling pathway-related allelic load was also associated with EM-related activity in the hippocampus, a brain region typically involved in EM and in AD pathology.[25-27]

Methods

Samples:

Discovery sample: This sample is part of an ongoing, continuously recruiting behavioral genetics study in the city of Basel, Switzerland. For the purposes of this study (data lock August 2013), data from 1458 healthy young Swiss adults (66.6% female; mean age: 22.5 ± 3.5 years) were available. Subjects were free of any neurological or psychiatric condition and did not take medication at the time of the experiment. All participants gave written informed consent before participation and completed a picture delayed free recall task, which reflects EM performance. For a detailed description of the procedure please refer to the online-only material.

Replication sample: This sample is part of an ongoing, continuously recruiting imaging genetics study in the city of Basel, Switzerland. For the purposes of this study (data lock August 2013), data from 1176 healthy young Swiss adults (60% female; mean age: 22.5 ± 3.3 years) were available. Subjects were free of any neurological or psychiatric condition and did not take medication at the time of the experiment. All participants gave written informed consent before participation and, while undergoing fMRI acquisition, completed a similar picture delayed free recall task as in the discovery sample. For a detailed description of the procedure please refer to the online-only material.

Zurich sample: We recruited 409 healthy young Swiss adults (72.4% female; mean age: 21.2 ± 1.9 years) for a behavioral genetics study in the city of Zurich, Switzerland. Subjects were free of any neurological or psychiatric condition and did not take medication at the time of the experiment. All participants gave written informed consent before participation and completed a picture delayed free recall task similar to the one used in the

discovery sample and in the replication sample 1. For a detailed description of the procedure please refer to the online-only material.

Healthy elderly sample: This sample consisted of elderly participants of the German Study on Ageing, Cognition and Dementia in primary care patients (AgeCoDe). The AgeCoDe study is an ongoing primary care-based prospective longitudinal study on early detection of mild cognitive impairment and dementia established by the German Competence Network Dementia. The sampling frame and sample selection process of the AgeCoDe study have been described in detail previously.[28] Please also refer to the online-only material.

Genetic heterogeneity: For each of the four cognitively healthy samples, the genomic control inflation factor lambda (λ_{GC}) was calculated to assess admixture. Lambda is defined as the median X^2 association test statistic divided by the theoretical distribution under the null distribution.[29] λ_{GC} showed a range between 1.0046 and 1.0449, indicating the absence of noteworthy admixture in these samples.

Array-based SNP genotyping: Samples were processed as described in the Genome-Wide Human SNP Nsp/Sty 6.0 User Guide (Affymetrix). For a detailed description of the procedure please refer to the online-only material.

Brain imaging

fMRI preprocessing and first level analyses: Preprocessing and data analysis was performed using SPM8 (Statistical Parametric Mapping, Wellcome Trust Centre for Neuroimaging; http://www. fil.ion.ucl.ac.uk/spm) implemented in MATLAB R2011b (MathWorks). Volumes were slice-time corrected to the first slice and realigned using the

‘‘register to mean’’ option. A mean image was generated from the realigned series and coregistered to the structural image. This ensured that functional and structural images were spatially aligned. The functional images and the structural images were spatially normalized by applying DARTEL, which leads to an improved registration between subjects.[30,31] For a detailed description please refer to the online-only material.

fMRI group statistics

The first-level contrast parameters were used for behavioral analyses in a random effects model (second-level analysis). We used a regression model to analyze associations between brain activation differences (meaningful vs scrambled pictures) and the multi-allelic score. Age and sex were included as covariates. A one-sample t-test was computed to assess the significance of task-related activation (meaningful vs scrambled pictures) at the group level. The analysis was focused on the left and right hippocampi, defined using the template-based hippocampal ROIs (see below: Construction of a Population-Average Anatomical Probabilistic Atlas). For a detailed description please refer to the online-only material.

Statistical genetic analysis

Genome-wide association analyses: Genome-wide association analyses: For each genome-wide analysis, *P* values were obtained using linear regression analyses as implemented in PLINK.[32] Sex and age were included as covariates. We applied the following quality control criteria: Non-significant deviation from Hardy-Weinberg equilibrium (HWE; *P*(HWE>0.0001)) and a minor allele frequency (MAF)>0.01. Mean per SNP call rate was >99%.

Pathway analysis: GSEA was performed using MAGENTA.[33] Briefly, the method first maps SNPs onto genes and then assigns each gene a SNP association score (i.e. the maximum SNP *P* value within ±0kb of the annotated gene). By applying a step-wise multiple linear regression analysis, the analysis is corrected for the following confounders: gene size, number of SNPs, number of independent SNPs, number of recombination hotspots, linkage disequilibrium and genetic distance. Lastly, a gene-set enrichment-like statistical test is applied to determine if a gene set is enriched for highly ranked *P* values compared to a gene-set of identical size, randomly drawn from the genome. False-discovery rate (FDR) based on the 75th percentile of association *P* values from all genes was used for multiple testing correction. As recommended, we used the 75th percentile cutoff because it yields optimal power for weak genetic effects that are expected for highly polygenic traits (e.g., EM performance).[30] The utilized gene sets are extracted and curated from the MSigDB v3.1 database (http://www.broadinstitute.org/gsea/msigdb), including gene-sets from different online databases (KEGG, Gene Ontology GO, BioCarta and Reactome).[34,35] We used a gene set size ranging between 20 and 200 genes to avoid both overly narrow and broad functional gene-set categories, resulting in 1'411 to be analyzed gene-sets.

We also applied INRICH (http://atgu.mgh.harvard.edu/inrich/), a software tool that examines enrichment of association signals for genetic gene-sets.[36] For a detailed description please refer to the online-only material.

Multilocus genetic score calculation: To capture the multi-allelic effect of the Calcium Signaling Pathway gene set, we generated an individual multilocus genetic score using the scoring procedure implemented in PLINK.[32] The score comprises all variants of the discovery sample, which proved significant ($P < 0.05$) after correction for number of independent SNPs per gene (Table e2). The PLINK algorithm calculates the score by summing up the individual number of reference alleles over all SNPs, weighted by the

direction of effect on EM performance with "1" (the reference allele enhances WM performance) or "-1" (the reference allele decreases EM performance), and finally averages the score by the number of non-missing SNPs.

Results

GSEA of EM in young healthy adults

Discovery sample (*n*=1458). After calculation of *P* values for association with EM performance (picture free recall task, see Online Only Methods) under the additive genetic model, we ran GSEA using MAGENTA.[33] Among the 1411 database-derived gene-sets, MAGENTA identified significant enrichment (FDR < 0.05; multiple testing-corrected) for one gene set, the Calcium Signaling Pathway gene set (KEGG entry: hsa04020) (Table 1). No additional gene set withstood correction for multiple testing. The Calcium Signaling Pathway gene set was also significant when applying INRICH,[36] an alternative GSEA method. Of 864 independent intervals that contained the best genome-wide association signals, INRICH identified 26 intervals overlapping with the target genes of the Calcium Signaling Pathway gene set. Subsequent permutation analysis showed significant enrichment for this gene set ($P_{empirical}$=0.021).

Replication sample (*n*=1176). Next, GSEA of the identical task (picture free recall task, see Methods) was performed in an independently recruited replication sample. The Calcium Signaling Pathway gene set was enriched significantly (P=0.015).

Calcium Signaling Pathway allelic load correlates with hippocampal activation

In an additional experiment, conducted in a subgroup (n=1119) of the replication sample, functional magnetic resonance imaging (fMRI) was used to identify gene set-dependent differences in brain activity related to EM (see Methods). We focused our search on the hippocampus because i) the Calcium Signaling Pathway gene set was associated with EM, which depends on the hippocampus,[37-40] and ii) components of this gene set are part of the signaling cascade involved in the formation of hippocampus-dependent memory in vertebrates.[41-45] Thus, the left and right hippocampi served as regions of interest (ROI). Independently of allelic load, we detected highly robust picture encoding-related activation (contrast: meaningful vs scrambled pictures) in the hippocampus (Fig. e1). To capture the multi-allelic effect of the Calcium Signaling Pathway gene set on hippocampal activity, PLINK[32] was used to generate an individual multilocus genetic score. The score was weighted by the direction of effect, whereby larger scores indicated better EM performance (see Methods). Of note, the multiallelic score for this imaging sample, which was a subset of the replication sample, was calculated by using only those SNPs (including the respective directions of effect), which proved significant in the independent discovery sample (see Methods). This procedure prevented model overfitting (i.e. inflation of test statistics), which would have occurred if also significant SNPs of the replication sample would have been used for score calculation.

Genetic score-dependent analysis revealed a significant positive correlation between genetic score values and activation in the right hippocampus (peak at [33 -16.5 -16]; t=3.35; $P_{\text{uncorrected}=}$0.0004, $P_{\text{small volume correction (SVC)}}$<0.05, Fig. 2; genotype-independent task-related activation at this coordinate t=42.51). The peak association in the left hippocampus [-24.75 -13.75 -28] did not survive small volume correction (t=2.36; $P_{\text{uncorrected}=}$0.009, P_{SVC}> 0.05). In order to ensure that the EM fMRI results were not driven by structural

changes related to the gene set, we performed a voxel-based morphometry (VBM) structural analysis. No significant effect of the multi-allelic score on gray matter volume was found within the hippocampal ROI.

EM core gene set

GSEA tests for statistical enrichment at gene set level. Thus, a certain gene set might prove significant in two different samples without any overlap of the gene set components, which gave rise to the significant enrichment (i.e. significantly associated genes), between samples. We tested this possibility by comparing the significant components of the Calcium Signaling Pathway gene set between the discovery and the replication sample. The overlap was significant (P=0.007; exact hypergeometric probability). Of the 144 Calcium Signaling Pathway genes, 26 genes contributed to gene set significance in both samples, whereas 66 genes did not contribute to gene set significance in either sample. The remaining 52 genes contributed to gene set significance in one of the two samples. Thus, the former group of 26 genes was defined as the replicated EM core gene set (Table 2, Fig. 3, Table e1). Further exploratory analysis (see Methods) revealed that the EM core gene set was highly significantly enriched with genes involved in the elevation of cytosolic calcium (42.3% of the genes, P=8.9 x 10^{-18}). In comparison, the enrichment of the group of 66 non-contributing genes with molecules involved in the elevation of cytosolic calcium (10.6% of the genes, P=2.7 x 10^{-7}) was 10 orders of magnitude weaker.

GSEA in additional samples

Zurich sample (*n*=409). Participants performed a picture free recall task similar to the task used in the discovery and replication samples (see Methods). The EM core gene set was significantly enriched (P=0.038, Fig. 3). No significant enrichment (P=0.704, Fig. 3) was found for the set of 66 genes, which did not contribute to the significance of the Calcium Signaling Pathway gene set in any of the discovery and the replication samples.

GSEA in non-demented elderly subjects, AgeCoDe sample (*n*=763). This sample of cognitively healthy elderly individuals was included to investigate whether the observed association of the EM core gene set with EM performance can be also observed in older adults. In analogy to the discovery, replication, and the Zurich sample, genome-wide *P* values for association with EM performance (delayed verbal free recall, see Methods) under the additive genetic model were used for GSEA. MAGENTA revealed significant enrichment (P=0.004, Fig. 3) of the EM core gene set. Also in this sample, no significant enrichment (P=0.284, Fig. 3) was found for the set of 66 genes, which did not contribute to the significance of the Calcium Signaling Pathway gene set.

GSEA in Alzheimer's disease

EM deficits represent a behavioral hallmark of AD[11] and are observed early in the course of the disease. We investigated the enrichment of the EM core gene set in a large AD case-control sample (for a detailed study description, please refer to the online-only material).

International Genomics of Alzheimer's Project (IGAP) case-control sample[46] (*n*=54162; n_{cases}=17008, $n_{controls}$=37154). 7036050 autosomal SNP *P* values of association with sporadic AD served as input for MAGENTA. MAGENTA was run with the identical parameters as in the studies of cognitively healthy subjects. The EM core gene set was significantly enriched (*P*=0.013, Fig. 3). Also in this sample, no significant enrichment ($P_{=}$0.384, Fig. 3) was found for the set of 66 genes, which did not contribute to the significance of the Calcium Signaling Pathway gene set in the EM samples.

Discussion

We detected consistent and robust associations between Calcium Signaling Pathway genes and human EM performance. In particular, a core gene set comprising 26 genes was significantly enriched in 4 independent cohorts of young and elderly cognitively healthy individuals (n=3806). This finding is compatible with the critical role of calcium signaling in molecular processes underlying memory, as shown in model organisms and *in vitro* studies:[47] For example, increases in intracellular calcium are causally related to the induction of long-term potentiation (LTP) and long-term depression (LTD),[48-50] two cellular correlates of learning and memory.[51,52] The results of the present study suggest that genes involved in calcium signaling are also related to human episodic memory throughout adulthood.

Moreover, fMRI data revealed that the individual Calcium Signaling-related allelic load correlated with hippocampal activity measured during memory encoding. Animal studies have amply demonstrated that calcium signaling genes are crucial for the formation of hippocampus-dependent memory.[41-45] Our findings suggest that calcium signaling genes are related to the formation of hippocampus-dependent memory also in humans.

Interestingly, the EM core gene set was also significantly enriched in a large case-control study of sporadic AD. Calcium signaling dysregulation has been repeatedly observed in cell culture and animal models of AD.[53,54] In addition, treatment with dantrolene, a drug decreasing free intracellular calcium concentration, diminished beta amyloid (Aβ) load, one of the histopathological hallmarks of AD, and reduced learning and memory deficits in Tg2576 mice, which overexpress a mutant form of the amyloid precursor protein.[55] Our results are in support of a role for calcium signaling genes in AD.

Interestingly, the EM core gene set, which proved significant in all studied cohorts, including the AD case-control sample, was significantly enriched with genes involved in the elevation of cytosolic calcium. A local increase in calcium concentrations results in a number of short-term and long-term synapse-specific alterations that are essential for dendritic development, neuronal survival, synaptic plasticity, learning, and memory.[56,57] A genetic profile favoring the local increase in calcium concentrations might therefore be related to better cognitive capacities and consequently to a delayed clinical manifestation of cognitive decline in AD patients. On the other hand, a sustained increase in intracellular calcium can, in the long run, lead to neurodegeneration and cell death.[58-60] In this case, a genetic profile favoring the local increase in calcium concentrations could also increase AD risk. Thus, given the complexity of the phenotypic, biological, and temporal relationship between cognition, physiological brain aging, and neurodegeneration,[61,62] it is not possible to make definite inferences regarding the direction of effect of the results presented herein.

While we did identify a core set of EM-related calcium signaling pathway genes, we stress that genes not contributing significantly to pathway enrichment cannot be excluded from being related to human EM. This is due to a number of possible reasons. For example, some genes might be associated with other forms or stages of EM as those investigated in the present study. Some genes may not be related to variability of memory performance because their expression in the brain might be too tightly regulated and independent of common genetic variability. However, a lack of physiologically meaningful genetic variability by no means implies a minor importance of these genes for the phenotype under study.

Further, the results of the present study must not lead to the erroneous assumption of an exclusive role of calcium signaling genes in human EM and AD. Given the ubiquity and versatility of calcium signaling,[63] it is apparent that it must play a role in a variety of

neurocognitive traits. For example, we have shown that calcium-related genes are also enriched in human working memory (WM) and WM-linked brain activation.[17] Nonetheless, by showing robust and consistently significant enrichment in independent cohorts of young and elderly participants, our study identifies calcium signaling as a central player of hippocampus-dependent human memory processes, both in cognitive health and disease and thereby contributes to the understanding -and hopefully treatment- of hippocampus-dependent cognitive pathology.

Acknowledgments

We thank Elmar Merkle, MD, Christoph Stippich, MD, and Oliver Bieri, MD for granting access to the fMRI facilities of the University Hospital Basel.

This work was funded by the Swiss National Science Foundation (Sinergia grant CRSII1_136227 to D.Q. and A.P.), by the 7th framework programme of the European Union (ADAMS project, HEALTH-F4-2009-242257), and by the German Research Network on Dementia (KND) and the German Research Network on Degenerative Dementia (KNDD), German Federal Ministry of Education and Research grants 01GI0420 and 01GI0711.

We also thank the International Genomics of Alzheimer's Project (IGAP) for providing summary results data for these analyses. The investigators within IGAP contributed to the design and implementation of IGAP and/or provided data but did not participate in analysis or writing of this report. IGAP was made possible by the generous participation of the control subjects, the patients, and their families. The i–Select chips was funded by the French National Foundation on Alzheimer's disease and related disorders. EADI was supported by the LABEX (laboratory of excellence program investment for the future) DISTALZ grant, Inserm, Institut Pasteur de Lille, Université de Lille 2 and the Lille University Hospital. GERAD was supported by the Medical Research Council (Grant n° 503480), Alzheimer's Research UK (Grant n° 503176), the Wellcome Trust (Grant n° 082604/2/07/Z) and German Federal Ministry of Education and Research (BMBF): Competence Network Dementia (CND) grant n° 01GI0102, 01GI0711, 01GI0420. CHARGE was partly supported by the NIH/NIA grant R01 AG033193 and the NIA AG081220 and AGES contract N01–AG–12100, the NHLBI grant R01 HL105756, the Icelandic Heart Association, and the Erasmus Medical Center and Erasmus University. ADGC was supported by the NIH/NIA grants: U01 AG032984, U24 AG021886, U01 AG016976, and the Alzheimer's Association grant ADGC–10–196728.

Angela Heck had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Tulving E. Episodic memory: from mind to brain. *Annu Rev Psychol.* 2002;53:1-25.

2. Alarcon M, Plomin R, Fulker DW, Corley R, DeFries JC. Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado Adoption Project. *Behav Genet.* Jul 1998;28(4):255-264.

3. Bouchard TJ, Jr., Lykken DT, McGue M, Segal NL, Tellegen A. Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science.* Oct 12 1990;250(4978):223-228.

4. McClearn GE, Johansson B, Berg S, et al. Substantial genetic influence on cognitive abilities in twins 80 or more years old. *Science.* Jun 6 1997;276(5318):1560-1563.

5. Papassotiropoulos A, de Quervain DJ. Genetics of human episodic memory: dealing with complexity. *Trends Cogn Sci.* Sep 2011;15(9):381-387.

6. Balota DA, Dolan PO, Duchek JM. Memory changes in healthy young and older adults. In: Tulving E, Craik FIM, eds. *Handbook of memory*. Oxford: Oxford University Press; 2000:395-409.

7. Bernard FA, Desgranges B, Eustache F, Baron JC. Neural correlates of age-related verbal episodic memory decline: a PET study with combined subtraction/correlation analysis. *Neurobiol Aging.* Oct 2007;28(10):1568-1576.

8. Grady CL, Craik FI. Changes in memory processing with age. *Curr Opin Neurobiol.* Apr 2000;10(2):224-231.

9. Light LL. Memory and aging: four hypotheses in search of data. *Annu Rev Psychol.* 1991;42:333-376.

10. Langbaum JB, Fleisher AS, Chen K, et al. Ushering in the study and treatment of preclinical Alzheimer disease. *Nature reviews. Neurology.* Jul 2013;9(7):371-381.

11. McKhann GM, Knopman DS, Chertkow H, et al. The diagnosis of dementia due to Alzheimer's disease: recommendations from the National Institute on Aging-Alzheimer's Association workgroups on diagnostic guidelines for Alzheimer's disease. *Alzheimer's & dementia : the journal of the Alzheimer's Association.* May 2011;7(3):263-269.

12. Bettens K, Sleegers K, Van Broeckhoven C. Genetic insights in Alzheimer's disease. *Lancet neurology.* Jan 2013;12(1):92-104.

13. Debette S, Ibrahim Verbaas CA, Bressler J, et al. Genome-wide studies of verbal declarative memory in non-demented older people: the CHARGE consortium. *Biological Psychiatry.* 2014.

14. Ramirez A, van der Flier WM, Herold C, et al. SUCLG2 identified as both a determinator of CSF Abeta1-42 levels and an attenuator of cognitive decline in Alzheimer's disease. *Hum Mol Genet.* Dec 15 2014;23(24):6644-6658.

15. Ramanan VK, Shen L, Moore JH, Saykin AJ. Pathway analysis of genomic data: concepts, methods, and prospects for future development. *Trends Genet.* Jul 2012;28(7):323-332.

16. Wang L, Jia P, Wolfinger RD, Chen X, Zhao Z. Gene set analysis of genome-wide association studies: methodological issues and perspectives. *Genomics.* Jul 2011;98(1):1-8.

17. Heck A, Fastenrath M, Ackermann S, et al. Converging genetic and functional brain imaging evidence links neuronal excitability to working memory, psychiatric disease, and brain activity. *Neuron.* Mar 5 2014;81(5):1203-1213.

18. Voineagu I, Wang X, Johnston P, et al. Transcriptomic analysis of autistic brain reveals convergent molecular pathology. *Nature.* Jun 16 2011;474(7351):380-384.

19. Holmans P, Green EK, Pahwa JS, et al. Gene ontology analysis of GWA study data sets provides insights into the biology of bipolar disorder. *Am J Hum Genet.* Jul 2009;85(1):13-24.
20. Sklar P, Ripke S, Scott LJ, et al. Large-scale genome-wide association analysis of bipolar disorder identifies a new susceptibility locus near ODZ4. *Nature genetics.* Oct 2011;43(10):977-983.
21. The Pathway Analysis Subgroup of the Psychiatric Genomics C. Psychiatric genome-wide association study analyses implicate neuronal, immune and histone pathways. *Nat Neurosci.* 2015;18(2):199-209.
22. Stergiakouli E, Hamshere M, Holmans P, et al. Investigating the contribution of common genetic variants to the risk and pathogenesis of ADHD. *Am J Psychiatry.* Feb 2012;169(2):186-194.
23. O'Dushlaine C, Kenny E, Heron E, et al. Molecular pathways involved in neuronal cell adhesion and membrane scaffolding contribute to schizophrenia and bipolar disorder susceptibility. *Mol Psychiatry.* Mar 2011;16(3):286-292.
24. Schizophrenia Working Group of the Psychiatric Genomics C. Biological insights from 108 schizophrenia-associated genetic loci. *Nature.* Jul 24 2014;511(7510):421-427.
25. Braak H, Braak E. Staging of Alzheimer's disease-related neurofibrillary changes. *Neurobiol Aging.* May-Jun 1995;16(3):271-278; discussion 278-284.
26. Manns JR, Eichenbaum H. Evolution of declarative memory. *Hippocampus.* 2006;16(9):795-808.
27. Rasch B, Papassotiropoulos A, de Quervain DF. Imaging genetics of cognitive functions: Focus on episodic memory. *Neuroimage.* Nov 15 2010;53(3):870-877.
28. Luck T, Riedel-Heller SG, Kaduszkiewicz H, et al. Mild cognitive impairment in general practice: age-specific prevalence and correlate results from the German study on ageing, cognition and dementia in primary care patients (AgeCoDe). *Dement Geriatr Cogn Disord.* 2007;24(4):307-316.
29. Devlin B, Roeder K. Genomic control for association studies. *Biometrics.* Dec 1999;55(4):997-1004.
30. Ashburner J. A fast diffeomorphic image registration algorithm. *Neuroimage.* Oct 15 2007;38(1):95-113.
31. Klein A, Andersson J, Ardekani BA, et al. Evaluation of 14 nonlinear deformation algorithms applied to human brain MRI registration. *Neuroimage.* Jul 1 2009;46(3):786-802.
32. Purcell S, Neale B, Todd-Brown K, et al. PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet.* Sep 2007;81(3):559-575.
33. Segre AV, Consortium D, investigators M, et al. Common inherited variation in mitochondrial genes is not enriched for associations with type 2 diabetes or related glycemic traits. *PLoS Genet.* Aug 2010;6(8).
34. Ashburner M, Ball CA, Blake JA, et al. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature genetics.* May 2000;25(1):25-29.
35. Ogata H, Goto S, Sato K, Fujibuchi W, Bono H, Kanehisa M. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Res.* Jan 1 1999;27(1):29-34.
36. Lee PH, O'Dushlaine C, Thomas B, Purcell SM. INRICH: interval-based enrichment analysis for genome-wide association studies. *Bioinformatics.* Jul 1 2012;28(13):1797-1799.
37. Habib R, Nyberg L. Neural correlates of availability and accessibility in memory. *Cereb Cortex.* Jul 2008;18(7):1720-1726.

38. Milner B. Disorders of learning and memory after temporal lobe lesions in man. *Clinical neurosurgery.* 1972;19:421-446.
39. Schacter DL, Tulving E. *Memory Systems*. Cambridge, MA: MIT Press; 1994.
40. Squire LR, Alvarez P. Retrograde amnesia and memory consolidation: a neurobiological perspective. *Curr Opin Neurobiol.* Apr 1995;5(2):169-177.
41. Cummings JA, Mulkey RM, Nicoll RA, Malenka RC. Ca2+ signaling requirements for long-term depression in the hippocampus. *Neuron.* Apr 1996;16(4):825-833.
42. Kandel ER. The molecular biology of memory storage: a dialogue between genes and synapses. *Science.* Nov 2 2001;294(5544):1030-1038.
43. Nicoll RA, Malenka RC. Contrasting properties of two forms of long-term potentiation in the hippocampus. *Nature.* Sep 14 1995;377(6545):115-118.
44. Shobe J. The role of PKA, CaMKII, and PKC in avoidance conditioning: permissive or instructive? *Neurobiol Learn Mem.* May 2002;77(3):291-312.
45. Tonegawa S, Nakazawa K, Wilson MA. Genetic neuroscience of mammalian learning and memory. *Philos Trans R Soc Lond B Biol Sci.* Apr 29 2003;358(1432):787-795.
46. Lambert JC, Ibrahim-Verbaas CA, Harold D, et al. Meta-analysis of 74,046 individuals identifies 11 new susceptibility loci for Alzheimer's disease. *Nat Genet.* Dec 2013;45(12):1452-1458.
47. Baker KD, Edwards TM, Rickard NS. The role of intracellular calcium stores in synaptic plasticity and memory consolidation. *Neurosci Biobehav Rev.* Aug 2013;37(7):1211-1239.
48. Malenka RC, Bear MF. LTP and LTD: an embarrassment of riches. *Neuron.* Sep 30 2004;44(1):5-21.
49. Malenka RC, Lancaster B, Zucker RS. Temporal limits on the rise in postsynaptic calcium required for the induction of long-term potentiation. *Neuron.* Jul 1992;9(1):121-128.
50. Mulkey RM, Malenka RC. Mechanisms underlying induction of homosynaptic long-term depression in area CA1 of the hippocampus. *Neuron.* Nov 1992;9(5):967-975.
51. Artola A, Singer W. Long-term depression of excitatory synaptic transmission and its relationship to long-term potentiation. *Trends Neurosci.* Nov 1993;16(11):480-487.
52. Bliss TV, Collingridge GL. A synaptic model of memory: long-term potentiation in the hippocampus. *Nature.* Jan 7 1993;361(6407):31-39.
53. Missiaen L, Robberecht W, van den Bosch L, et al. Abnormal intracellular ca(2+)homeostasis and disease. *Cell calcium.* Jul 2000;28(1):1-21.
54. Thibault O, Porter NM, Chen KC, et al. Calcium dysregulation in neuronal aging and Alzheimer's disease: history and new directions. *Cell calcium.* Nov-Dec 1998;24(5-6):417-433.
55. Oules B, Del Prete D, Greco B, et al. Ryanodine receptor blockade reduces amyloid-beta load and memory impairments in Tg2576 mouse model of Alzheimer disease. *J Neurosci.* Aug 22 2012;32(34):11820-11834.
56. Wilson RS, Barral S, Lee JH, et al. Heritability of different forms of memory in the Late Onset Alzheimer's Disease Family Study. *J Alzheimers Dis.* 2011;23(2):249-255.
57. Zeng Y, Chang W, Shu C, et al. Decreased cognitive function in extended family members from the single late-onset-Alzheimer's-disease pedigree. *Journal of the International Neuropsychological Society : JINS.* Aug 2013;19(7):809-819.

58. Amieva H, Mokri H, Le Goff M, et al. Compensatory mechanisms in higher-educated subjects with Alzheimer's disease: a study of 20 years of cognitive decline. *Brain.* Apr 2014;137(Pt 4):1167-1175.
59. Barral S, Bird T, Goate A, et al. Genotype patterns at PICALM, CR1, BIN1, CLU, and APOE genes are associated with episodic memory. *Neurology.* May 8 2012;78(19):1464-1471.
60. Stern Y. Cognitive reserve. *Neuropsychologia.* Aug 2009;47(10):2015-2028.
61. Redmond L, Ghosh A. Regulation of dendritic development by calcium signaling. *Cell calcium.* May 2005;37(5):411-416.
62. Rogaev EI. Genomics of behavioral diseases. *Frontiers in genetics.* 2012;3:45.
63. Marambaud P, Dreses-Werringloer U, Vingtdeux V. Calcium signaling in neurodegeneration. *Mol Neurodegener.* 2009;4:20.

Figure Legends

Fig. 1

Study workflow and included samples.

Fig. 2

Allelic load-dependent increases in EM-related brain activity (n=1119). The blue circles show the activation in the right hippocampus. The local maximum is located at [33 -16.5 -16]; t=3.35, $P_{\text{uncorrected}=}$0.0004, $P_{\text{small volume correction (SVC)}} < 0.05$. Activations are overlaid on coronal (upper left), sagittal (upper right) and axial (lower left) sections of the study specific group template (see Methods); displayed at $P_{uncorrected}$=0.001, using color-coded t values. L, left side of the brain; R, right side of the brain.

Fig. 3

GSEA results of two different gene sets: EM core gene set (brown color), group of 66 non-significant genes (blue color). Left panel: enrichment P values of the respective gene sets in three samples (Zurich sample, AgeCoDe sample, IGAP sample). Right panel: Box-plot of gene size distribution stratified by gene set. Filled circles indicate outliers, stars indicate extreme values.

Table 1 GSEA results (FDR q < 0.25) in the discovery sample.

Database	**Gene set name**	**FDR q value**
KEGG	calcium signaling pathway	0.024
Gene Ontology	amine transport	0.110
Gene Ontology	amine transmembrane transporter activity	0.122
Gene Ontology	structure specific DNA binding	0.140
Gene Ontology	active transmembrane transporter activity	0.142
Gene Ontology	positive regulation of nucleobase-nucleoside-nucleotide and nucleic acid metabolic process	0.148
Gene Ontology	positive regulation of transcription-DNA dependent	0.175
Gene Ontology	positive regulation of transcription from RNA polymerase II promoter	0.184
Gene Ontology	positive regulation of RNA metabolic process	0.187
Gene Ontology	G protein coupled receptor activity	0.193
Gene Ontology	serine hydrolase activity	0.195
Gene Ontology	serine type peptidase activity	0.195
Gene Ontology	peptidase activity	0.197
Gene Ontology	di_tri_valent inorganic cation transmembrane transporter activity	0.199
Gene Ontology	phosphoric diester hydrolase activity	0.207
Gene Ontology	microsome	0.209
Gene Ontology	vesicular fraction	0.248

Table 2 GSEA results in the discovery and replication samples. Genes shown in this table were identified by GSEA as significant constituents of the Calcium Signaling Pathway in the respective sample. Genes highlighted bold are members of the EM core gene set (i.e. significant in the discovery and the replication sample). All gene symbols according to HGNC nomenclature.

Discovery sample (Basel 1, n=1458)								
ADCY2	*ADCY4*	***ADCY8***	*ADCY9*	***ADRA1A***	*ADRA1B*	*ADRA1D*	*ATP2A2*	*ATP2B2*
ATP2B4	***AVPR1A***	*BST1*	*CACNA1A*	*CACNA1B*	***CACNA1E***	***CACNA1G***	***CACNA1S***	***CAMK2G***
CCKBR	*CHRM1*	*CHRM3*	***CHRM5***	*EGFR*	***GNA15***	*GNAQ*	*GRIN2A*	*GRM1*
HTR2A	***ITPKB***	***ITPR1***	*ITPR2*	*ITPR3*	*LHCGR*	*MYLK*	*MYLK2*	*NOS3*
NTSR1	*OXTR*	*P2RX5*	***P2RX7***	*PDE1B*	***PDGFRA***	***PDGFRB***	***PLCB2***	***PLCD4***
PLCG2	***PPP3CA***	***PPP3R1***	***PRKCB***	*PTAFR*	***PTGER3***	***RYR3***	*SLC8A1*	*SLC8A3*
TACR1	*TRPC1*							
Replication sample (Basel 2, n=1176)								
ADCY3	***ADCY8***	***ADRA1A***	***ATP2B4***	***AVPR1A***	***CACNA1E***	***CACNA1G***	*CACNA1I*	***CACNA1S***
CAMK2A	*CAMK2B*	***CAMK2G***	*CCKAR*	***CCKBR***	*CHP2*	***CHRM5***	*CHRNA7*	*EDNRA*
EDNRB	*GNA14*	***GNA15***	*GNAL*	*HRH2*	***HTR2A***	*HTR5A*	***ITPKB***	***ITPR1***
NOS1	*NOS2*	*P2RX4*	***P2RX7***	***PDGFRA***	***PDGFRB***	***PLCB2***	***PLCD4***	*PLCE1*
PLCG2	*PLCZ1*	*PLN*	***PPP3CA***	***PPP3R1***	***PRKCB***	***PTGER3***	*PTK2B*	***RYR3***
TACR1	*TNNC2*	*VDAC2*						

5. Discussion

The studies presented in the scope of this thesis deal with hormonal and genetic influences on the modulation of memory processes, especially on episodic memory content, in healthy human subjects. The aim of the first part of this thesis was to explore the influence of the stress hormone cortisol on memory consolidation and retrieval. The second part directs the focus to genetic mechanisms of memory processes by examining the function of the gene *HDAC5*.

In the study of Bentz et al. (2013), we induced stress using the CPT and were interested in the effects of cortisol enhancement on fear and extinction memory formation in an aversive differential conditioning paradigm. Subjective US-expectancy ratings were measured and the difference scores between expectancy ratings for CS+ (neutral stimulus paired with US) and CS- (neutral stimulus not paired with US) were calculated to indicate learning. 24 hours after acquisition, we found a significant reduction in US-expectancy after CPT induction for the stress group compared to the control group in males but not in females. Interestingly, this effect persisted even after 48 hours, indicating a prolonged effect of stress on fear memory retrieval. The sex specific difference might be due to a more pronounced cortisol reaction to the CPT in males as compared to females. We additionally hypothesized that cortisol promotes fear extinction processes, as it has been shown before that the outcome of behavioral exposure therapy of anxiety disorders, which is believed to rely on fear extinction, can be improved by additional cortisol administration (Bentz, et al., 2010; Soravia, et al., 2006). Unfortunately, extinction training in our study did not lead to a significant decrease in US-expectancy.

Abundant evidence demonstrates that stress and stress hormones have the ability to reduce memory retrieval of emotional information, whereas they enhance memory consolidation of new information (for review see de Quervain, et al., 2009). Most of the studies conducted in humans have used experimentally or pharmacologically influenced cortisol levels. In the study of Ackermann et al. (2013a) though, we were interested in the influence of naturally cycling basal cortisol levels on memory consolidation as well as retrieval in a picture-based episodic memory task. Mean as well as single cortisol levels were not significantly associated with recall performance. However, we observed significant associations with changes in cortisol levels, whereat stronger decreases during recall predicted better free recall performances in the short-delay recall 10 min as well as long-delay recall 20 hours after encoding of the pictures. In contrast, we were not able to

find any association of mean cortisol levels or changes in cortisol levels during encoding with subsequent recall performance. These results point to an involvement of changes in cortisol levels in memory retrieval processes, rather than in memory acquisition.

Taken together we could demonstrate effects of cortisol on memory retrieval in two different studies, which is in line with previous findings. Firstly by analyzing stress-induced effects of cortisol on memory performance originating from a non-classical declarative memory task (Bentz, et al., 2013) and secondly by showing effects of basal cortisol levels on memory performance in a typical declarative episodic memory task (Ackermann, et al., 2013a). In contrast to the first study (Bentz, et al., 2013), we found no substantial differences between women and men in the study of Ackermann et al. (2013a). However, both studies were in line with findings, which show that women taking hormonal contraceptives exhibit different stress responses than men or freely cycling women (Kirschbaum, Kudielka, Gaab, Schommer, & Hellhammer, 1999). Women taking hormonal contraceptives showed generally less decrease in cortisol than men or freely cycling women (Ackermann, et al., 2013a), and showed a smaller cortisol response after CPT than men (Bentz, et al., 2013). In this study only women taking hormonal contraceptives participated therefore we have not the opportunity to compare the cortisol response after CPT of men and women taking hormonal contraceptives with freely cycling women.

Among the different hormonal and neurotransmitter systems that may interact with GCs to modulate memory processes, noradrenaline plays an important role. GCs and noradrenaline are reported to regulate together important processes for synaptic plasticity, especially in the hippocampus and the amygdala (Krugers, et al., 2012). In the studies integrated in this thesis we used the CPT known to induce activation of the adrenergic system (Bentz, et al., 2013) or emotional stimuli (negative and positive pictures (Ackermann, et al., 2013a) expected to induce emotional arousal. For further studies it would be interesting to take into account measurements of noradrenergic activity such as salivary alpha amylase (Bosch, Veerman, de Geus, & Proctor, 2011) to verify whether the emotional stimuli are sufficient to induce an arousing state and therefore enhance memory performance, even though the test situation is relatively stress free.

As mentioned before in the theoretical background, cortisol is supposed to express its effects on memory formation by binding to GRs (de Kloet, et al., 2005). The results summarized above are referring only to hormone measurements. However, genetic studies found additional evidence for a role of cortisol on memory formation by identifying GR

gene polymorphisms. The *Bcl*I polymorphism of the GR gene *NR3C1* for example has been involved in traumatic memory formation in a study in cardiac surgery patients (Hauer et al., 2011). Furthermore, a study conducted at our division could expand this finding to episodic memory in healthy human subjects and reported an association between the *Bcl*I polymorphism and short-delay recall of emotional but not neutral pictures (Ackermann, Heck, Rasch, Papassotiropoulos, & de Quervain, 2013b).

The second part of this thesis shifts the focus away from cortisol to the analysis of SNPs lying in the region of *HDAC5*. In the study of Hartmann et al. (unpublished manuscript) we used the same picture-based episodic memory task as has been used in the study of Ackermann et al. (2013a) and identified in a first step an association between SNP rs184478 located in the region of *HDAC5* and short-delay episodic memory performance, especially for emotional content. Next we could show rs184478 genotype-dependent differences in methylation levels of a CpG site located within a CpG island in the *HDAC5* promoter region. Finally, we extended our analysis to mRNA expression measures. Although we could not find genotype-dependent differences in mRNA expression levels of HDAC5, we identified changes in expression levels of C17orf65, an open reading frame lying in the vicinity of *HDAC5*. To examine tissue specific effects of mRNA expression, we analyzed, in addition to our data deriving from blood, 3 publicly available data sets providing data from brain tissue. Here we could show that SNP rs184478 was associated with HDAC5 cortical expression levels, pointing to tissue-specific effects. While we could find a robust behavioral finding and this leads to the assumption that the identified SNP rs184478 may play a role in episodic memory formation, our data cannot support *HDAC5* as a candidate for episodic memory. It furthermore seems that SNP rs184478 may regulate the expression of C17orf65 instead of HDAC5 to which it is assigned. With this data we could illustrate the complexity of genetic and epigenetic mechanisms underlying genetically complex traits and the possibility that SNPs can also influence the expression of genes at a distance. Due to our results we have learned the important lesson, that the joint analysis of SNP data, expression data and phenotypic data can prevent wrong conclusions concerning the involved genes. This conclusion could also be important for the interpretation of past association studies, as the inclusion of other genetic data than SNPs, e.g. gene expression, usually has been neglected or was not possible to collect. Nowadays it becomes obvious that a significant amount of information gets lost, if research focuses only on single-platform approaches (Huang, 2015) and the use of multi-platform approaches should be the standard.

Overall, the effects of cortisol as well as the SNP rs184478 reported in the frame of this thesis are supposed to affect the brain and therefore might modulate behavior. One limitation of our results in this regard may be that our biological data was not extracted from brain tissue. We measured cortisol levels in saliva and therefore cannot make conclusions about the exact fraction entering the brain as well as how cortisol affects MRs and GRs in the brain. We furthermore determined the DNA methylation and mRNA expression data from peripheral blood. Taken together we are not able to make direct conclusions about the functions in the brain. Nevertheless several studies suggest that DNA methylation signatures in genomic regions rich in cytosine-guanine dinucleotides generally show stable epigenetic signatures across brain and non-brain tissues (Tylee, Kawaguchi, & Glatt, 2013). Moreover, a recently published paper of our group (Vukojevic et al., 2014) could show that peripheral measured methylation as well as expression of the glucocorticoid receptor gene *NR3C1*, were related to recognition performance and recognition memory-related brain activity, measured with function magnetic resonance imaging in healthy male subjects. Additionally, Klengel et al. (2013) found allele-specific demethylation of CpGs close to and in glucocorticoid response elements of the FK506 binding protein 5 (*FKBP5*), measured in peripheral blood. Additionally, they used a human hippocampal progenitor cell line and found methylation patterns comparable to the results found in peripheral blood cells. Finally, they found a correlation between the volume of the right hippocampus with peripheral blood *FKBP5* methylation.

All hormonal and genetic results explain only a small fraction of the phenotypic variation, in our case of memory processes. To gain more insight into the processes and biological pathways underlying episodic memory formation, it is important to use additional approaches such as pathway analysis or the investigation of epistasis. Additionally, it should be taken into account that different memory stages (encoding, consolidation, retrieval) can have different underlying mechanisms. The results in this thesis are mainly restricted to the process of memory retrieval as for example in the study of Ackermann et al. (2013a) we could only find effects on memory performance for cortisol changes during the recall phase but not during the consolidation phase.

The combined use of hormonal, genetic, as well as epigenetic data may be a way to better understand the mechanisms underlying memory formation and furthermore neurological and neuropsychiatric disorders. Finally, the use of this data as well as further approaches might help to identify pharmacological targets (Muglia, 2011; Papassotiropoulos et al., 2013a). Basal cortisol levels for example may have clinical

implications, since reduced cortisol levels have been observed in PTSD patients (Yehuda, 2002; Yehuda et al., 2000). The administration of GCs has already been discussed to be useful to suppress the increased memory retrieval processes that occur in PTSD patients in terms of flashbacks or intrusions (Aerni et al., 2004), and the recall of fear memories in anxiety disorders (Soravia, et al., 2006).

To conclude, the present thesis examines the modulation of memory processes from a hormonal and a genetic perspective by focusing on one hand to the hormone cortisol and on the other on the genetic region around *HDAC5*. Cortisol has already been examined in several human and animal studies and we extended these results to basal cortisol levels as well as to a fear conditioning paradigm in healthy humans. As expected, cortisol indeed modulated recall performance in our studies. HDAC5 further has been implicated in memory formation and synaptic plasticity in animal studies and we aimed to transfer these findings to healthy humans. We could find some hints that the SNP rs184478, which has been assigned to *HDAC5* may play a role in episodic memory formation, especially for emotional content. Results from our expression analysis however showed that the identified SNP may regulate expression of C17orf65 instead of HDAC5. With this results we could show, how important it is to include additional expression data to the pure SNP-phenotype analyses as the identified SNPs may regulate the expression of genes other than the nearest. Taken together the inclusion of SNP data, expression analyses and the measurement of hormones may help to avoid wrong conclusions about the mechanisms underlying memory formation in healthy and disease. Furthermore it may open the possibility to find more functional variants, as up to now only a small fraction of all identified SNPs are proven functional candidates.

6. References

Abel, T., & Lattal, K. M. (2001). Molecular mechanisms of memory acquisition, consolidation and retrieval. *Current opinion in neurobiology, 11*(2), 180-187.

Abel, T., Nguyen, P. V., Barad, M., Deuel, T. A., Kandel, E. R., & Bourtchouladze, R. (1997). Genetic demonstration of a role for PKA in the late phase of LTP and in hippocampus-based long-term memory. *Cell, 88*(5), 615-626.

Ackermann, S., Hartmann, F., Papassotiropoulos, A., de Quervain, D. J., & Rasch, B. (2013a). Associations between basal cortisol levels and memory retrieval in healthy young individuals. *Journal of cognitive neuroscience, 25*(11), 1896-1907.

Ackermann, S., Heck, A., Rasch, B., Papassotiropoulos, A., & de Quervain, D. J. (2013b). The BclI polymorphism of the glucocorticoid receptor gene is associated with emotional memory performance in healthy individuals. *Psychoneuroendocrinology, 38*(7), 1203-1207.

Ackermann, S., Spalek, K., Rasch, B., Gschwind, L., Coynel, D., Fastenrath, M., et al. (2012). Testosterone levels in healthy men are related to amygdala reactivity and memory performance. *Psychoneuroendocrinology, 37*(9), 1417-1424.

Aerni, A., Traber, R., Hock, C., Roozendaal, B., Schelling, G., Papassotiropoulos, A., et al. (2004). Low-dose cortisol for symptoms of posttraumatic stress disorder. *The American journal of psychiatry, 161*(8), 1488-1490.

Agis-Balboa, R. C., Pavelka, Z., Kerimoglu, C., & Fischer, A. (2013). Loss of HDAC5 impairs memory function: implications for Alzheimer's disease. *Journal of Alzheimer's disease : JAD, 33*(1), 35-44.

al'Absi, M., Petersen, K. L., & Wittmers, L. E. (2002). Adrenocortical and hemodynamic predictors of pain perception in men and women. *Pain, 96*(1-2), 197-204.

Alarcon, J. M., Malleret, G., Touzani, K., Vronskaya, S., Ishii, S., Kandel, E. R., et al. (2004). Chromatin acetylation, memory, and LTP are impaired in CBP+/- mice: a model for the cognitive deficit in Rubinstein-Taybi syndrome and its amelioration. *Neuron, 42*(6), 947-959.

Alarcon, M., Plomin, R., Fulker, D. W., Corley, R., & DeFries, J. C. (1998). Multivariate path analysis of specific cognitive abilities data at 12 years of age in the Colorado Adoption Project. *Behavior genetics, 28*(4), 255-264.

Alkire, M. T., Haier, R. J., Fallon, J. H., & Cahill, L. (1998). Hippocampal, but not amygdala, activity at encoding correlates with long-term, free recall of nonemotional information. *Proceedings of the National Academy of Sciences of the United States of America, 95*(24), 14506-14510.

Artola, A., von Frijtag, J. C., Fermont, P. C., Gispen, W. H., Schrama, L. H., Kamal, A., et al. (2006). Long-lasting modulation of the induction of LTD and LTP in rat hippocampal CA1 by behavioural stress and environmental enrichment. *The European journal of neuroscience, 23*(1), 261-272.

Bailey, C. H., & Kandel, E. R. (2008). Synaptic remodeling, synaptic growth and the storage of long-term memory in Aplysia. *Progress in brain research, 169*, 179-198.

Bayley, P. J., Hopkins, R. O., & Squire, L. R. (2006). The fate of old memories after medial temporal lobe damage. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 26*(51), 13311-13317.

Bechara, A., Tranel, D., Damasio, H., Adolphs, R., Rockland, C., & Damasio, A. R. (1995). Double dissociation of conditioning and declarative knowledge relative to the amygdala and hippocampus in humans. *Science, 269*(5227), 1115-1118.

Belanoff, J. K., Gross, K., Yager, A., & Schatzberg, A. F. (2001). Corticosteroids and cognition. *Journal of psychiatric research, 35*(3), 127-145.

Bentz, D., Michael, T., de Quervain, D. J., & Wilhelm, F. H. (2010). Enhancing exposure therapy for anxiety disorders with glucocorticoids: from basic mechanisms of emotional learning to clinical applications. *Journal of anxiety disorders, 24*(2), 223-230.

Bentz, D., Michael, T., Wilhelm, F. H., Hartmann, F. R., Kunz, S., von Rohr, I. R., et al. (2013). Influence of stress on fear memory processes in an aversive differential conditioning paradigm in humans. *Psychoneuroendocrinology, 38*(7), 1186-1197.

Bliss, T. V., & Collingridge, G. L. (1993). A synaptic model of memory: long-term potentiation in the hippocampus. *Nature, 361*(6407), 31-39.

Bliss, T. V., Collingridge, G. L., & Morris, R. G. (2003). Introduction. Long-term potentiation and structure of the issue. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences, 358*(1432), 607-611.

Bliss, T. V., & Lomo, T. (1973). Long-lasting potentiation of synaptic transmission in the dentate area of the anaesthetized rabbit following stimulation of the perforant path. *The Journal of physiology, 232*(2), 331-356.

Bohacek, J., & Mansuy, I. M. (2013). Epigenetic inheritance of disease and disease risk. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 38*(1), 220-236.

Bosch, J. A., Veerman, E. C., de Geus, E. J., & Proctor, G. B. (2011). alpha-Amylase as a reliable and convenient measure of sympathetic activity: don't start salivating just yet! *Psychoneuroendocrinology, 36*(4), 449-453.

Bouchard, T. J., Jr., Lykken, D. T., McGue, M., Segal, N. L., & Tellegen, A. (1990). Sources of human psychological differences: the Minnesota Study of Twins Reared Apart. *Science, 250*(4978), 223-228.

Bouton, M. E., Mineka, S., & Barlow, D. H. (2001). A modern learning theory perspective on the etiology of panic disorder. *Psychological review, 108*(1), 4-32.

Broide, R. S., Redwine, J. M., Aftahi, N., Young, W., Bloom, F. E., & Winrow, C. J. (2007). Distribution of histone deacetylases 1-11 in the rat brain. *Journal of molecular neuroscience : MN, 31*(1), 47-58.

Buchanan, T. W., & Lovallo, W. R. (2001). Enhanced memory for emotional material following stress-level cortisol treatment in humans. *Psychoneuroendocrinology, 26*(3), 307-317.

Buchanan, T. W., Tranel, D., & Adolphs, R. (2006). Impaired memory retrieval correlates with individual differences in cortisol response but not autonomic response. *Learning & memory, 13*(3), 382-387.

Cahill, L., Gorski, L., & Le, K. (2003). Enhanced human memory consolidation with post-learning stress: interaction with the degree of arousal at encoding. *Learning & memory, 10*(4), 270-274.

Cahill, L., Haier, R. J., Fallon, J., Alkire, M. T., Tang, C., Keator, D., et al. (1996). Amygdala activity at encoding correlated with long-term, free recall of emotional information. *Proceedings of the National Academy of Sciences of the United States of America, 93*(15), 8016-8021.

Carter, R. M., O'Doherty, J. P., Seymour, B., Koch, C., & Dolan, R. J. (2006). Contingency awareness in human aversive conditioning involves the middle frontal gyrus. *NeuroImage, 29*(3), 1007-1012.

Chawla, S., Vanhoutte, P., Arnold, F. J., Huang, C. L., & Bading, H. (2003). Neuronal activity-dependent nucleocytoplasmic shuttling of HDAC4 and HDAC5. *Journal of neurochemistry, 85*(1), 151-159.

Clements, A. D. (2013). Salivary cortisol measurement in developmental research: where do we go from here? *Developmental psychobiology, 55*(3), 205-220.

Clements, A. D., & Parker, C. R. (1998). The relationship between salivary cortisol concentrations in frozen versus mailed samples. *Psychoneuroendocrinology, 23*(6), 613-616.

Cooke, S. F., & Bliss, T. V. (2006). Plasticity in the human central nervous system. *Brain : a journal of neurology, 129*(Pt 7), 1659-1673.

Cortes-Mendoza, J., Diaz de Leon-Guerrero, S., Pedraza-Alva, G., & Perez-Martinez, L. (2013). Shaping synaptic plasticity: the role of activity-mediated epigenetic regulation on gene transcription. *International journal of developmental neuroscience : the official journal of the International Society for Developmental Neuroscience, 31*(6), 359-369.

Davidson, P. S., Troyer, A. K., & Moscovitch, M. (2006). Frontal lobe contributions to recognition and recall: linking basic research with clinical evaluation and remediation. *Journal of the International Neuropsychological Society : JINS, 12*(2), 210-223.

Day, J. J., & Sweatt, J. D. (2011). Epigenetic mechanisms in cognition. *Neuron, 70*(5), 813-829.

de Kloet, C. S., Vermetten, E., Geuze, E., Kavelaars, A., Heijnen, C. J., & Westenberg, H. G. (2006). Assessment of HPA-axis function in posttraumatic stress disorder: pharmacological and non-pharmacological challenge tests, a review. *Journal of psychiatric research, 40*(6), 550-567.

de Kloet, E. R., Joels, M., & Holsboer, F. (2005). Stress and the brain: from adaptation to disease. *Nature reviews. Neuroscience, 6*(6), 463-475.

de Kloet, E. R., Oitzl, M. S., & Joels, M. (1999). Stress and cognition: are corticosteroids good or bad guys? *Trends in neurosciences, 22*(10), 422-426.

de Quervain, D. J., Aerni, A., Schelling, G., & Roozendaal, B. (2009). Glucocorticoids and the regulation of memory in health and disease. *Frontiers in neuroendocrinology, 30*(3), 358-370.

de Quervain, D. J., Bentz, D., Michael, T., Bolt, O. C., Wiederhold, B. K., Margraf, J., et al. (2011). Glucocorticoids enhance extinction-based psychotherapy. *Proceedings of the National Academy of Sciences of the United States of America, 108*(16), 6621-6625.

de Quervain, D. J., Kolassa, I. T., Ackermann, S., Aerni, A., Boesiger, P., Demougin, P., et al. (2012). PKCalpha is genetically linked to memory capacity in healthy subjects and to risk for posttraumatic stress disorder in genocide survivors. *Proceedings of the National Academy of Sciences of the United States of America, 109*(22), 8746-8751.

de Quervain, D. J., Roozendaal, B., & McGaugh, J. L. (1998). Stress and glucocorticoids impair retrieval of long-term spatial memory. *Nature, 394*(6695), 787-790.

de Quervain, D. J., Roozendaal, B., Nitsch, R. M., McGaugh, J. L., & Hock, C. (2000). Acute cortisone administration impairs retrieval of long-term declarative memory in humans. *Nature neuroscience, 3*(4), 313-314.

Delgado, M. R., Olsson, A., & Phelps, E. A. (2006). Extending animal models of fear conditioning to humans. *Biological psychology, 73*(1), 39-48.

Dere, E., Pause, B. M., & Pietrowsky, R. (2010). Emotion and episodic memory in neuropsychiatric disorders. *Behavioural brain research, 215*(2), 162-171.

Dickerson, B. C., & Eichenbaum, H. (2010). The episodic memory system: neurocircuitry and disorders. *Neuropsychopharmacology : official publication of the American College of Neuropsychopharmacology, 35*(1), 86-104.

Dickerson, S. S., & Kemeny, M. E. (2004). Acute stressors and cortisol responses: a theoretical integration and synthesis of laboratory research. *Psychological bulletin, 130*(3), 355-391.

Dolcos, F., LaBar, K. S., & Cabeza, R. (2005). Remembering one year later: role of the amygdala and the medial temporal lobe memory system in retrieving emotional memories. *Proceedings of the National Academy of Sciences of the United States of America, 102*(7), 2626-2631.

Dubois, B., Feldman, H. H., Jacova, C., Dekosky, S. T., Barberger-Gateau, P., Cummings, J., et al. (2007). Research criteria for the diagnosis of Alzheimer's disease: revising the NINCDS-ADRDA criteria. *The Lancet. Neurology, 6*(8), 734-746.

Fastenrath, M., Coynel, D., Spalek, K., Milnik, A., Gschwind, L., Roozendaal, B., et al. (2014). Dynamic modulation of amygdala-hippocampal connectivity by emotional arousal. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(42), 13935-13947.

Foa, E. B., & Kozak, M. J. (1986). Emotional processing of fear: exposure to corrective information. *Psychological bulletin, 99*(1), 20-35.

Frazer, K. A., Murray, S. S., Schork, N. J., & Topol, E. J. (2009). Human genetic variation and its contribution to complex traits. *Nature reviews. Genetics, 10*(4), 241-251.

Frey, U., Huang, Y. Y., & Kandel, E. R. (1993). Effects of cAMP simulate a late stage of LTP in hippocampal CA1 neurons. *Science, 260*(5114), 1661-1664.

Griffiths, B. B., & Hunter, R. G. (2014). Neuroepigenetics of stress. *Neuroscience, 275*, 420-435.

Guan, Z., Giustetto, M., Lomvardas, S., Kim, J. H., Miniaci, M. C., Schwartz, J. H., et al. (2002). Integration of long-term-memory-related synaptic plasticity involves bidirectional regulation of gene expression and chromatin structure. *Cell, 111*(4), 483-493.

Haettig, J., Stefanko, D. P., Multani, M. L., Figueroa, D. X., McQuown, S. C., & Wood, M. A. (2011). HDAC inhibition modulates hippocampus-dependent long-term memory for object location in a CBP-dependent manner. *Learning & memory, 18*(2), 71-79.

Hartmann, F. R., Milnik, A., Auschra, B., Freytag, V., Spalek, K., Vogler, C., Vukojevic, V., de Quervain D. J.-F., Papassotiropoulos, A., & Heck, A. (unpublished manuscript). Polymorphisms of HDAC5 are associated with episodic memory, DNA methylation and C17orf65 mRNA expression.

Hauer, D., Weis, F., Papassotiropoulos, A., Schmoeckel, M., Beiras-Fernandez, A., Lieke, J., et al. (2011). Relationship of a common polymorphism of the glucocorticoid receptor gene to traumatic memories and posttraumatic stress disorder in patients after intensive care therapy. *Critical care medicine, 39*(4), 643-650.

Henke, K. (2010). A model for memory systems based on processing modes rather than consciousness. *Nature reviews. Neuroscience, 11*(7), 523-532.

Hinkelmann, K., Moritz, S., Botzenhardt, J., Riedesel, K., Wiedemann, K., Kellner, M., et al. (2009). Cognitive impairment in major depression: association with salivary cortisol. *Biological psychiatry, 66*(9), 879-885.

Huang, Y. T. (2015). Integrative modeling of multi-platform genomic data under the framework of mediation analysis. *Statistics in medicine, 34*(1), 162-178.

Jackson, E. D., Payne, J. D., Nadel, L., & Jacobs, W. J. (2006). Stress differentially modulates fear conditioning in healthy men and women. *Biological psychiatry, 59*(6), 516-522.

Jarome, T. J., & Lubin, F. D. (2014). Epigenetic mechanisms of memory formation and reconsolidation. *Neurobiology of learning and memory, 115*, 116-127.

Joels, M., Pu, Z., Wiegert, O., Oitzl, M. S., & Krugers, H. J. (2006). Learning under stress: how does it work? *Trends in cognitive sciences, 10*(4), 152-158.

Johansson, B., Whitfield, K., Pedersen, N. L., Hofer, S. M., Ahern, F., & McClearn, G. E. (1999). Origins of individual differences in episodic memory in the oldest-old: a population-based study of identical and same-sex fraternal twins aged 80 and older. *The journals of gerontology. Series B, Psychological sciences and social sciences, 54*(3), P173-179.

Jovanovic, T., & Ressler, K. J. (2010). How the neurocircuitry and genetics of fear inhibition may inform our understanding of PTSD. *The American journal of psychiatry, 167*(6), 648-662.

Kandel, E. R. (2001). The molecular biology of memory storage: a dialogue between genes and synapses. *Science, 294*(5544), 1030-1038.

Kandel, E. R. (2012). The molecular biology of memory: cAMP, PKA, CRE, CREB-1, CREB-2, and CPEB. *Molecular brain, 5*, 14.

Kandel, E. R., Dudai, Y., & Mayford, M. R. (2014). The molecular and systems biology of memory. *Cell, 157*(1), 163-186.

Keppler, B. R., & Archer, T. K. (2008). Chromatin-modifying enzymes as therapeutic targets--Part 1. *Expert opinion on therapeutic targets, 12*(10), 1301-1312.

Kim, J. J., & Diamond, D. M. (2002). The stressed hippocampus, synaptic plasticity and lost memories. *Nature reviews. Neuroscience, 3*(6), 453-462.

Kirschbaum, C., & Hellhammer, D. H. (1994). Salivary cortisol in psychoneuroendocrine research: recent developments and applications. *Psychoneuroendocrinology, 19*(4), 313-333.

Kirschbaum, C., Kudielka, B. M., Gaab, J., Schommer, N. C., & Hellhammer, D. H. (1999). Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. *Psychosomatic medicine, 61*(2), 154-162.

Kirschbaum, C., Pirke, K. M., & Hellhammer, D. H. (1993). The 'Trier Social Stress Test'--a tool for investigating psychobiological stress responses in a laboratory setting. *Neuropsychobiology, 28*(1-2), 76-81.

Kitazawa, S. (2002). Neurobiology: ready to unlearn. *Nature, 416*(6878), 270-273.

Klengel, T., Mehta, D., Anacker, C., Rex-Haffner, M., Pruessner, J. C., Pariante, C. M., et al. (2013). Allele-specific FKBP5 DNA demethylation mediates gene-childhood trauma interactions. *Nature neuroscience, 16*(1), 33-41.

Knight, D. C., Cheng, D. T., Smith, C. N., Stein, E. A., & Helmstetter, F. J. (2004). Neural substrates mediating human delay and trace fear conditioning. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 24*(1), 218-228.

Kouzarides, T. (2007). Chromatin modifications and their function. *Cell, 128*(4), 693-705.

Kristensen, L. S., Mikeska, T., Krypuy, M., & Dobrovic, A. (2008). Sensitive Melting Analysis after Real Time- Methylation Specific PCR (SMART-MSP): high-throughput and probe-free quantitative DNA methylation detection. *Nucleic acids research, 36*(7), e42.

Krugers, H. J., Goltstein, P. M., van der Linden, S., & Joels, M. (2006). Blockade of glucocorticoid receptors rapidly restores hippocampal CA1 synaptic plasticity after exposure to chronic stress. *The European journal of neuroscience, 23*(11), 3051-3055.

Krugers, H. J., & Hoogenraad, C. C. (2009). Hormonal regulation of AMPA receptor trafficking and memory formation. *Frontiers in synaptic neuroscience, 1*, 2.

Krugers, H. J., Karst, H., & Joels, M. (2012). Interactions between noradrenaline and corticosteroids in the brain: from electrical activity to cognitive performance. *Frontiers in cellular neuroscience, 6*, 15.

Kudielka, B. M., Hellhammer, D. H., & Wust, S. (2009). Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology, 34*(1), 2-18.

Kudielka, B. M., & Kirschbaum, C. (2005). Sex differences in HPA axis responses to stress: a review. *Biological psychology, 69*(1), 113-132.

Kuhlmann, S., Kirschbaum, C., & Wolf, O. T. (2005). Effects of oral cortisol treatment in healthy young women on memory retrieval of negative and neutral words. *Neurobiology of learning and memory, 83*(2), 158-162.

Kuhlmann, S., Piel, M., & Wolf, O. T. (2005). Impaired memory retrieval after psychosocial stress in healthy young men. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 25*(11), 2977-2982.

LaBar, K. S., & Cabeza, R. (2006). Cognitive neuroscience of emotional memory. *Nature reviews. Neuroscience, 7*(1), 54-64.

LaBar, K. S., Gatenby, J. C., Gore, J. C., LeDoux, J. E., & Phelps, E. A. (1998). Human amygdala activation during conditioned fear acquisition and extinction: a mixed-trial fMRI study. *Neuron, 20*(5), 937-945.

LaBar, K. S., LeDoux, J. E., Spencer, D. D., & Phelps, E. A. (1995). Impaired fear conditioning following unilateral temporal lobectomy in humans. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 15*(10), 6846-6855.

Lattal, K. M., & Wood, M. A. (2013). Epigenetics and persistent memory: implications for reconsolidation and silent extinction beyond the zero. *Nature neuroscience, 16*(2), 124-129.

LeDoux, J. E. (1996). *The emotional brain*. New York: Simon and Schuster.

LeDoux, J. E. (2000). Emotion circuits in the brain. *Annual review of neuroscience, 23*, 155-184.

LeDoux, J. E., Iwata, J., Cicchetti, P., & Reis, D. J. (1988). Different projections of the central amygdaloid nucleus mediate autonomic and behavioral correlates of conditioned fear. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 8*(7), 2517-2529.

Lee, B. K., Glass, T. A., McAtee, M. J., Wand, G. S., Bandeen-Roche, K., Bolla, K. I., et al. (2007). Associations of salivary cortisol with cognitive function in the Baltimore memory study. *Archives of general psychiatry, 64*(7), 810-818.

Levenson, J. M., O'Riordan, K. J., Brown, K. D., Trinh, M. A., Molfese, D. L., & Sweatt, J. D. (2004). Regulation of histone acetylation during memory formation in the hippocampus. *The Journal of biological chemistry, 279*(39), 40545-40559.

Levenson, J. M., & Sweatt, J. D. (2005). Epigenetic mechanisms in memory formation. *Nature reviews. Neuroscience, 6*(2), 108-118.

Levine, A., Zagoory-Sharon, O., Feldman, R., Lewis, J. G., & Weller, A. (2007). Measuring cortisol in human psychobiological studies. *Physiology & behavior, 90*(1), 43-53.

Lovallo, W. (1975). The cold pressor test and autonomic function: a review and integration. *Psychophysiology, 12*(3), 268-282.

Lovibond, P. F. (2006). Fear and avoidance: an integrated expectancy model. In M. G. Craske, Hermans, D., Vansteenwegen, D. (Ed.), *Fear and learning. From basic processes to clinical implications* (pp. 117-132). Washington DC: American Psychological Association.

Lupien, S. J., Maheu, F., Tu, M., Fiocco, A., & Schramek, T. E. (2007). The effects of stress and stress hormones on human cognition: Implications for the field of brain and cognition. *Brain and cognition, 65*(3), 209-237.

Lupien, S. J., & McEwen, B. S. (1997). The acute effects of corticosteroids on cognition: integration of animal and human model studies. *Brain research. Brain research reviews, 24*(1), 1-27.

Lynch, M. A. (2004). Long-term potentiation and memory. *Physiological reviews, 84*(1), 87-136.

Mahan, A. L., & Ressler, K. J. (2012). Fear conditioning, synaptic plasticity and the amygdala: implications for posttraumatic stress disorder. *Trends in neurosciences, 35*(1), 24-35.

Mahgoub, M., & Monteggia, L. M. (2014). A role for histone deacetylases in the cellular and behavioral mechanisms underlying learning and memory. *Learning & memory, 21*(10), 564-568.

Maren, S. (2001). Neurobiology of Pavlovian fear conditioning. *Annual review of neuroscience, 24*, 897-931.

Martin, K. C., Michael, D., Rose, J. C., Barad, M., Casadio, A., Zhu, H., et al. (1997). MAP kinase translocates into the nucleus of the presynaptic cell and is required for long-term facilitation in Aplysia. *Neuron, 18*(6), 899-912.

Mason, J. W., Giller, E. L., Kosten, T. R., Ostroff, R. B., & Podd, L. (1986). Urinary free-cortisol levels in posttraumatic stress disorder patients. *The Journal of nervous and mental disease, 174*(3), 145-149.

Mathews, H. L., & Janusek, L. W. (2011). Epigenetics and psychoneuroimmunology: mechanisms and models. *Brain, behavior, and immunity, 25*(1), 25-39.

McEwen, B. S. (2001). Plasticity of the hippocampus: adaptation to chronic stress and allostatic load. *Annals of the New York Academy of Sciences, 933*, 265-277.

McGaugh, J. L. (2004). The amygdala modulates the consolidation of memories of emotionally arousing experiences. *Annual review of neuroscience, 27*, 1-28.

Merz, C. J., Tabbert, K., Schweckendiek, J., Klucken, T., Vaitl, D., Stark, R., et al. (2010). Investigating the impact of sex and cortisol on implicit fear conditioning with fMRI. *Psychoneuroendocrinology, 35*(1), 33-46.

Miller, C. A., & Sweatt, J. D. (2007). Covalent modification of DNA regulates memory formation. *Neuron, 53*(6), 857-869.

Milner, B., Squire, L. R., & Kandel, E. R. (1998). Cognitive neuroscience and the study of memory. *Neuron, 20*(3), 445-468.

Milnik, A., Heck, A., Vogler, C., Heinze, H. J., de Quervain, D. J., & Papassotiropoulos, A. (2012). Association of KIBRA with episodic and working memory: a meta-analysis. *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 159B*(8), 958-969.

Mineka, S., & Oehlberg, K. (2008). The relevance of recent developments in classical conditioning to understanding the etiology and maintenance of anxiety disorders. *Acta psychologica, 127*(3), 567-580.

Muglia, P. (2011). From genes to therapeutic targets for psychiatric disorders - what to expect? *Current opinion in pharmacology, 11*(5), 563-571.

Myers, K. M., & Davis, M. (2002). Behavioral and neural analysis of extinction. *Neuron, 36*(4), 567-584.

Nguyen, P. V., Abel, T., & Kandel, E. R. (1994). Requirement of a critical period of transcription for induction of a late phase of LTP. *Science, 265*(5175), 1104-1107.

Nyberg, L., McIntosh, A. R., Houle, S., Nilsson, L. G., & Tulving, E. (1996). Activation of medial temporal structures during episodic memory retrieval. *Nature, 380*(6576), 715-717.

Oakley, R. H., & Cidlowski, J. A. (2013). The biology of the glucocorticoid receptor: new signaling mechanisms in health and disease. *The Journal of allergy and clinical immunology, 132*(5), 1033-1044.

Ookubo, M., Kanai, H., Aoki, H., & Yamada, N. (2013). Antidepressants and mood stabilizers effects on histone deacetylase expression in C57BL/6 mice: Brain region specific changes. *Journal of psychiatric research, 47*(9), 1204-1214.

Otsuki, K., Uchida, S., Wakabayashi, Y., Matsubara, T., Hobara, T., Funato, H., et al. (2010). Aberrant REST-mediated transcriptional regulation in major depressive disorder. *Journal of psychiatric research, 44*(6), 378-384.

Panizzon, M. S., Lyons, M. J., Jacobson, K. C., Franz, C. E., Grant, M. D., Eisen, S. A., et al. (2011). Genetic architecture of learning and delayed recall: a twin study of episodic memory. *Neuropsychology, 25*(4), 488-498.

Papassotiropoulos, A., & de Quervain, D. J. (2011). Genetics of human episodic memory: dealing with complexity. *Trends in cognitive sciences, 15*(9), 381-387.

Papassotiropoulos, A., Gerhards, C., Heck, A., Ackermann, S., Aerni, A., Schicktanz, N., et al. (2013a). Human genome-guided identification of memory-modulating drugs. *Proceedings of the National Academy of Sciences of the United States of America, 110*(46), E4369-4374.

Papassotiropoulos, A., Stefanova, E., Vogler, C., Gschwind, L., Ackermann, S., Spalek, K., et al. (2013b). A genome-wide survey and functional brain imaging study identify CTNNBL1 as a memory-related gene. *Molecular psychiatry, 18*(2), 255-263.

Papassotiropoulos, A., Stephan, D. A., Huentelman, M. J., Hoerndli, F. J., Craig, D. W., Pearson, J. V., et al. (2006). Common Kibra alleles are associated with human memory performance. *Science, 314*(5798), 475-478.

Pause, B. M., Zlomuzica, A., Kinugawa, K., Mariani, J., Pietrowsky, R., & Dere, E. (2013). Perspectives on episodic-like and episodic memory. *Frontiers in behavioral neuroscience, 7*, 33.

Penney, J., & Tsai, L. H. (2014). Histone deacetylases in memory and cognition. *Science signaling, 7*(355), re12.

Preuss, D., Schoofs, D., & Wolf, O. T. (2009). Associations between endogenous cortisol levels and emotional memory in young women: influence of encoding instructions. *Stress, 12*(5), 379-387.

Putman, P., Van Honk, J., Kessels, R. P., Mulder, M., & Koppeschaar, H. P. (2004). Salivary cortisol and short and long-term memory for emotional faces in healthy young women. *Psychoneuroendocrinology, 29*(7), 953-960.

Renthal, W., Maze, I., Krishnan, V., Covington, H. E., 3rd, Xiao, G., Kumar, A., et al. (2007). Histone deacetylase 5 epigenetically controls behavioral adaptations to chronic emotional stimuli. *Neuron, 56*(3), 517-529.

Reul, J. M., & de Kloet, E. R. (1985). Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. *Endocrinology, 117*(6), 2505-2511.

Robin, J., Hirshhorn, M., Rosenbaum, R. S., Winocur, G., Moscovitch, M., & Grady, C. L. (2015). Functional connectivity of hippocampal and prefrontal networks during episodic and spatial memory based on real-world environments. *Hippocampus, 25*(1), 81-93.

Roozendaal, B., Hernandez, A., Cabrera, S. M., Hagewoud, R., Malvaez, M., Stefanko, D. P., et al. (2010). Membrane-associated glucocorticoid activity is necessary for modulation of long-term memory via chromatin modification. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 30*(14), 5037-5046.

Roozendaal, B., McEwen, B. S., & Chattarji, S. (2009). Stress, memory and the amygdala. *Nature reviews. Neuroscience, 10*(6), 423-433.

Roozendaal, B., Okuda, S., de Quervain, D. J., & McGaugh, J. L. (2006). Glucocorticoids interact with emotion-induced noradrenergic activation in influencing different memory functions. *Neuroscience, 138*(3), 901-910.

Sapolsky, R. M. (2000). Glucocorticoids and hippocampal atrophy in neuropsychiatric disorders. *Archives of general psychiatry, 57*(10), 925-935.

Sara, S. J. (2000). Retrieval and reconsolidation: toward a neurobiology of remembering. *Learning & memory, 7*(2), 73-84.

Schwabe, L., Haddad, L., & Schachinger, H. (2008). HPA axis activation by a socially evaluated cold-pressor test. *Psychoneuroendocrinology, 33*(6), 890-895.

Schwabe, L., Joels, M., Roozendaal, B., Wolf, O. T., & Oitzl, M. S. (2012). Stress effects on memory: an update and integration. *Neuroscience and biobehavioral reviews, 36*(7), 1740-1749.

Schwabe, L., Wolf, O. T., & Oitzl, M. S. (2010). Memory formation under stress: quantity and quality. *Neuroscience and biobehavioral reviews, 34*(4), 584-591.

Scoville, W. B., & Milner, B. (1957). Loss of recent memory after bilateral hippocampal lesions. *Journal of neurology, neurosurgery, and psychiatry, 20*(1), 11-21.

Shing, Y. L., Werkle-Bergner, M., Brehmer, Y., Muller, V., Li, S. C., & Lindenberger, U. (2010). Episodic memory across the lifespan: the contributions of associative and strategic components. *Neuroscience and biobehavioral reviews, 34*(7), 1080-1091.

Smeets, T. (2011). Acute stress impairs memory retrieval independent of time of day. *Psychoneuroendocrinology, 36*(4), 495-501.

Soravia, L. M., Heinrichs, M., Aerni, A., Maroni, C., Schelling, G., Ehlert, U., et al. (2006). Glucocorticoids reduce phobic fear in humans. *Proceedings of the National Academy of Sciences of the United States of America, 103*(14), 5585-5590.

Soravia, L. M., Heinrichs, M., Winzeler, L., Fisler, M., Schmitt, W., Horn, H., et al. (2014). Glucocorticoids enhance in vivo exposure-based therapy of spider phobia. *Depression and anxiety, 31*(5), 429-435.

Squire, L. R. (1992). Declarative and nondeclarative memory: multiple brain systems supporting learning and memory. *Journal of cognitive neuroscience, 4*(3), 232-243.

Squire, L. R. (2004). Memory systems of the brain: a brief history and current perspective. *Neurobiology of learning and memory, 82*(3), 171-177.

Squire, L. R., Ojemann, J. G., Miezin, F. M., Petersen, S. E., Videen, T. O., & Raichle, M. E. (1992). Activation of the hippocampus in normal humans: a functional anatomical study of memory. *Proceedings of the National Academy of Sciences of the United States of America, 89*(5), 1837-1841.

Squire, L. R., & Zola, S. M. (1996). Structure and function of declarative and nondeclarative memory systems. *Proceedings of the National Academy of Sciences of the United States of America, 93*(24), 13515-13522.

Strachan, T., & Read, A. P. (2011) *Human molecular genetics* (4th ed.). New York.

Swan, G. E., Reed, T., Jack, L. M., Miller, B. L., Markee, T., Wolf, P. A., et al. (1999). Differential genetic influence for components of memory in aging adult twins. *Archives of neurology, 56*(9), 1127-1132.

Sweatt, J. D. (2013). The emerging field of neuroepigenetics. *Neuron, 80*(3), 624-632.

Tsankova, N. M., Berton, O., Renthal, W., Kumar, A., Neve, R. L., & Nestler, E. J. (2006). Sustained hippocampal chromatin regulation in a mouse model of depression and antidepressant action. *Nature neuroscience, 9*(4), 519-525.

Tulving, E. (2002). Episodic memory: from mind to brain. *Annual review of psychology, 53*, 1-25.

Tylee, D. S., Kawaguchi, D. M., & Glatt, S. J. (2013). On the outside, looking in: a review and evaluation of the comparability of blood and brain "-omes". *American journal of medical genetics. Part B, Neuropsychiatric genetics : the official publication of the International Society of Psychiatric Genetics, 162B*(7), 595-603.

van Ast, V. A., Cornelisse, S., Marin, M. F., Ackermann, S., Garfinkel, S. N., & Abercrombie, H. C. (2013). Modulatory mechanisms of cortisol effects on emotional learning and memory: novel perspectives. *Psychoneuroendocrinology, 38*(9), 1874-1882.

Van Honk, J., Kessels, R. P., Putman, P., Jager, G., Koppeschaar, H. P., & Postma, A. (2003). Attentionally modulated effects of cortisol and mood on memory for emotional faces in healthy young males. *Psychoneuroendocrinology, 28*(7), 941-948.

Vecsey, C. G., Hawk, J. D., Lattal, K. M., Stein, J. M., Fabian, S. A., Attner, M. A., et al. (2007). Histone deacetylase inhibitors enhance memory and synaptic plasticity via CREB:CBP-dependent transcriptional activation. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 27*(23), 6128-6140.

Vogler, C., Gschwind, L., Rothlisberger, B., Huber, A., Filges, I., Miny, P., et al. (2010). Microarray-based maps of copy-number variant regions in European and sub-Saharan populations. *PloS one, 5*(12), e15246.

Vogler, C., Spalek, K., Aerni, A., Demougin, P., Muller, A., Huynh, K. D., et al. (2009). CPEB3 is associated with human episodic memory. *Frontiers in behavioral neuroscience, 3*, 4.

Vukojevic, V., Kolassa, I. T., Fastenrath, M., Gschwind, L., Spalek, K., Milnik, A., et al. (2014). Epigenetic modification of the glucocorticoid receptor gene is linked to traumatic memory and post-traumatic stress disorder risk in genocide survivors. *The Journal of neuroscience : the official journal of the Society for Neuroscience, 34*(31), 10274-10284.

Weeber, E. J., Beffert, U., Jones, C., Christian, J. M., Forster, E., Sweatt, J. D., et al. (2002). Reelin and ApoE receptors cooperate to enhance hippocampal synaptic plasticity and learning. *The Journal of biological chemistry, 277*(42), 39944-39952.

Williams-Gray, C. H., Foltynie, T., Lewis, S. J., & Barker, R. A. (2006). Cognitive deficits and psychosis in Parkinson's disease: a review of pathophysiology and therapeutic options. *CNS drugs, 20*(6), 477-505.

Wilson, R. S., Barral, S., Lee, J. H., Leurgans, S. E., Foroud, T. M., Sweet, R. A., et al. (2011). Heritability of different forms of memory in the Late Onset Alzheimer's Disease Family Study. *Journal of Alzheimer's disease : JAD, 23*(2), 249-255.

Wolf, O. T. (2009). Stress and memory in humans: twelve years of progress? *Brain research, 1293*, 142-154.

Yehuda, R. (2002). Post-traumatic stress disorder. *The New England journal of medicine, 346*(2), 108-114.

Yehuda, R., Bierer, L. M., Schmeidler, J., Aferiat, D. H., Breslau, I., & Dolan, S. (2000). Low cortisol and risk for PTSD in adult offspring of holocaust survivors. *The American journal of psychiatry, 157*(8), 1252-1259.

Yehuda, R., & LeDoux, J. (2007). Response variation following trauma: a translational neuroscience approach to understanding PTSD. *Neuron, 56*(1), 19-32.

Zoladz, P. R., Clark, B., Warnecke, A., Smith, L., Tabar, J., & Talbot, J. N. (2011). Pre-learning stress differentially affects long-term memory for emotional words, depending on temporal proximity to the learning experience. *Physiology & behavior, 103*(5), 467-476.

Zorawski, M., Cook, C. A., Kuhn, C. M., & LaBar, K. S. (2005). Sex, stress, and fear: individual differences in conditioned learning. *Cognitive, affective & behavioral neuroscience, 5*(2), 191-201.

Declaration by candidate

I declare herewith that I have independently carried out the PhD-thesis entitled "Hormonal and genetic modulation of memory processes in healthy humans: Focus on cortisol and *HDAC5*". This thesis consists of original research articles that have been written in cooperation with the enlisted co-authors and have been published in peer-reviewed scientific journals or are in preparation/submitted for publication. Only allowed resources were used and all references used were cited accordingly.

Date: _____

Signature: _____