Psychoneuroendocrinology (2005) **30**, 771–784

PNEC

www.elsevier.com/locate/psyneuen

A meta-analytic review of the effects of acute cortisol administration on human memory

S. Het, G. Ramlow, O.T. Wolf*

Department of Experimental Psychology, University of Duesseldorf, Universitaetsstrasse 1, D-40225 Duesseldorf, Germany

Received 10 November 2004; received in revised form 23 February 2005; accepted 8 March 2005

KEYWORDS
Cortisol;
Glucocorticoids;
HPA;
Memory;
Meta-analysis

Summary Adrenal glucocorticoids (GC) secreted during stress modulate memory. Animal and human studies investigating the effects of acute GC treatment on memory have reported conflicting (enhancing as well as impairing) results. Several theories have been proposed to integrate these contradictory findings. Among the variables discussed are the timing of the GC treatment (before learning or before retrieval) and the time of day (morning versus afternoon). Here we review meta-analytically the results of 16 studies, which experimentally investigated the acute impact of cortisol treatment on human memory. The results revealed that the timing of GC application in the course of a study is a relevant variable which explains a substantial amount of the significant heterogeneity within the effect sizes. The studies which administered cortisol before retrieval ($n=4$) reported a significant decrease (average effect size of $d=-.49$) in memory performance. Studies which administered cortisol before learning ($n=12$) found on average no effect ($d=.08$), but there is heterogeneity within these effect sizes. Further analysis on these experiments indicated that studies, which administered cortisol in the morning found a significant memory impairment ($d=-.40$), while studies conducted in the afternoon observed a small but significant memory enhancement ($d=.22$). This meta-analysis supports the idea that the timing of GC administration (before learning or before retrieval) is a major determinant of the effects of GCs on human memory. We discuss methodological limitations of the current analysis and suggest several areas for future research.

1. Introduction

Stress leads to activation of the hypothalamus pituitary adrenal axis resulting in the increased release of glucocorticoids (GCs). These hormones (cortisol in humans; corticosterone in rodents) influence multiple target tissues including the brain. A longstanding history of studies in laboratory animals as well as in humans has demonstrated the ability of GCs to influence memory. Special interest has been placed on hippocampal dependent declarative memory in humans and hippocampal dependent spatial memory in rodents (McEwen and Sapolsky, 1995; Lupien and McEwen, 1997; Belanoff et al., 2001; Gold et al., 2002; Wolf, 2003). Studies on this topic can be separated into two areas: On the one hand there are

* Corresponding author. Tel.: +49 211 81 11799; fax: +49 211 81 12019.
E-mail address: oliver.wolf@uni-duesseldorf.de (O.T. Wolf).

0306-4530/$ - see front matter © 2005 Elsevier Ltd. All rights reserved.
doi:10.1016/j.psyneuen.2005.03.005

studies, which investigated memory performance under the impact of chronic, long lasting stress- or GC-treatment. On the other hand, there are studies reporting on the impact of acute stress- or GC-treatment. Studies investigating chronic GC-effects predominantly report impairing effects on memory (McEwen and Sapolsky, 1995; Lupien and McEwen, 1997; Belanoff et al., 2001; Gold et al., 2002; Wolf, 2003). The situation with studies investigating the acute effects of stress- or GC-treatment is another. Both animal and human studies found enhancing as well as impairing effects following acute stress- or GC-treatment (Lupien and McEwen, 1997; De Kloet et al., 1999; Roozendaal, 2002; Wolf, 2003). The challenging conceptual question is how can these contradictory results be integrated?

Recently Roozendaal (2002) described one integrative model. The author reviewed findings on the acute effects of GCs in rats on distinct memory phases, i.e. on memory consolidation and memory retrieval. He concluded that memory consolidation is enhanced by acute stress or acute GC-treatment, while delayed retrieval is impaired (Roozendaal, 2002). These GC-effects depend on glucocorticoid receptor (GR) activation in the hippocampus (see also Oitzl et al. (2001)) and concurrent noradrenergic activation in the basolateral amygdala (Roozendaal, 2002). The basolateral amygdala (BLA) seems to be a key structure in a memory-modulatory system that regulates stress and GC effects on memory consolidation, memory retrieval and working memory (Roozendaal, 2002; Roozendaal et al., 2003, 2004). Roozendaal (2002) suggests that stress may activate the BLA, which in turn switches the brain into a 'memory-consolidation state' allowing for strong consolidation of the current event, but simultaneously compromising memory retrieval. These animal studies have not addressed the scenario of GCs given before initial learning with recall being tested at a time when GCs are still elevated. However, this is most often the case in human studies reviewed here. It has been suggested, that if GCs are still elevated at retrieval testing, then a negative effect would prevail (Roozendaal, 2002). In contrast, beneficial effects on consolidation might only be observable if GCs return to baseline levels at the time of memory retrieval testing (de Quervain et al., 2000; Okuda et al., 2004).

There are of course alternative theories available, which do not necessarily have to be in discrepancy to the one outlined above. One prominent idea is that the relationship between memory performance and plasma GC-concentration is an *inverted U-shape* dose response curve. The review by Lupien and McEwen (1997) is one example of how this theory can integrate the opposing results in this research area. The inversed U-pattern often observed in neuropsychopharmacology was first observed for GCs in the 1970s (Kovacs et al., 1976; Flood et al., 1978). Both groups demonstrated that very high or very low GC-concentrations cause memory impairment, while moderate concentrations cause memory enhancement. The idea of an inverted dose response relationship is supported by behavioural and electrophysiological studies (Diamond et al., 1992; Pavlides et al., 1994; Vaher et al., 1994a,b). The underlying mechanism is most likely the different affinity of the two receptors for GCs (Reul and De Kloet, 1985; De Kloet et al., 1987, 1998; Lupien and McEwen, 1997). It appears that memory enhancing effects not only depend on saturated mineralocorticoidreceptor (MR) occupancy, but on a parallel low to moderate glucocorticoidreceptor (GR) occupancy (De Kloet et al., 1998; Sapolsky, 2003). As a result of the inverted U-shaped function it has been suggested that administered GCs might have negative effects at a time of high basal cortisol concentrations, as one can see in the morning in humans. At a time of low basal cortisol concentrations (e.g. in the evening in humans) GCs might have positive effects (Lupien and Lepage, 2001; Lupien et al., 2002a).

Meta-analysis is an alternative way of integrating multiple and sometimes opposing results (Glass, 1976). This methodological approach might be more objective in integrating results of multiple studies than qualitative reviews (Hedges and Olkin, 1985). In a meta-analysis the considered studies are weighted to objective aspects, like effect-sizes or *p*-values. Recently, Sauro and colleagues (2003) reviewed meta-analytically animal and human studies on the effects of acute or chronic stress on declarative memory. They found that stress reduces memory performance in animals as well as in humans. However, this research synthesis did not cover studies which investigated the impact of GCs on memory performance pharmacologically. It also did not address the issue of modulatory factors like memory phase or time of day. Pharmacological studies, however, are ideally suited to show the 'pure' acute effect of increased cortisol concentrations on human memory.

For the meta-analysis at hand the following hypotheses were suggested: A significant heterogeneity of the effect sizes was expected due to the discrepant results of pharmacological studies investigating the acute effect of cortisol on memory. According to Roozendaal (2002), we assumed that this heterogeneity can partially be explained by the time of GC application in the course of a study (treatment before learning versus treatment before retrieval). According to Lupien et al. (2002a), a further hypothesis was that the time of day for the investigation (morning versus afternoon) could explain a part of the remaining variance.

2. Methods

2.1. Sample of studies

All primary studies were obtained by using computer-based literature search machines (Medline and PsychINFO/PSYNDEX), crossing the keywords *cortisol, hydrocortisone, HPA, memory* and *glucocorticoids*. This search generated 775 articles. The abstracts were reviewed and all articles that could not be excluded based on the abstract were retrieved. Additionally, reference lists were searched for relevant papers. Only studies which were available by July 2004 were included.

2.2. Selection criteria for the sample

All the studies which were included in this meta-analytic investigation had to meet the following criteria: (a) published in a peer reviewed journal, (b) use of an experimental design, (c) placebo control, (d) sample of healthy human subjects, (e) written in English, (f) psychometric measure of short- and/or long-term memory, (g) oral or intravenous application of cortisol (HC, hydrocortisone) or cortisone, (h) investigation of acute effects of cortisol or cortisone application, (i) conducted during the day with awake subjects.

According to this, studies were excluded when they met the following criteria: (a) use of a quasi-experimental design, (b) no placebo control, (c) animal study, (d) sample of patients or elderlies, (e) application of synthetic glucocorticoids with different pharmacological profiles (e.g. dexamethasone, prednisone), (f) investigation of long-term effects of HC application exclusively, (g) laboratory stress exposure instead of pharmacological manipulation, (h) use of an indirect memory measure (e.g. reaction time, EEG changes) exclusively.

2.3. Coding

For descriptive purpose and error correction each considered study was coded by two independent referees (S.H. and G.R.) for relevant variables. For this, we constructed a coding scheme, which recorded the following information about a study: (a) year of publication (ICC = 1.0),[1] (b) total number of subjects (ICC = 1.0), (c) total number of groups (ICC = 1.0), (d) number of subjects in each group (ICC = .925), (e) number of drop-outs (ICC = .928), (f) sex (*6* = 1.0), (g) age (ICC = .998), (h) route of HC administration (*6* = 1.0), (i) experimental design (*6* = .940), (j) number of dependent variables (ICC = .647) (k) type of dependent variables (*6* = .732), (l) description of memory measures (e.g. free recall of words, cued paragraph recall, recognition memory performance) (*6* = .962), (m) time of recall (immediate and/or delayed recall) (*6* = .884), (n) time of the investigation (a.m. versus p.m.) (*6* = 1.0), (o) time of treatment with HC (treatment before learning or before recall) (*6* = 1.0), (p) retention interval (time between learning and recall) (ICC = .977), (q) relevant results for this meta-analytic investigation (ICC = 1.0). Discrepancies were resolved by discussion.

2.4. Calculation of effect sizes

Effect sizes were calculated for the memory results and then coded as effect sizes for immediate recall, delayed recall or delayed recognition. Following Hedges and Olkin (1985), the effect sizes (g_{Hedges}) were defined as the difference between the mean of the experimental group ($\bar{X}_{EG}$) and the placebo control group ($\bar{X}_{CG}$) standardized by the pooled standard deviation (S_{pooled}).[2]

The effect sizes were calculated by using the free meta-analytic software program META (Schwarzer, 1989; http://userpage.fu-berlin.de/~health/meta_e.htm). Positive *gs* indicate better memory in the cortisol group, while negative *gs* indicate the opposite. According to Cohen (1988) an effect size of .20 was classified as small, .50 as moderate and .80 as large. To avoid a sample size dependent overestimation of the population's effect size, each g_{Hedges} was converted to an adjusted *d* value by using Hedge's formula (Hedges and Olkin, 1985; Rosenthal, 1994).

When possible, effect sizes were calculated from the *means, standard deviations* or *standard error of means* reported in the article. When this information was not provided, the effect size was computed from inferential statistics, like *t* or *F* values (Rosenthal, 1994). In cases where this was also not possible, we attempted to contact the author(s) to obtain the required information, which was most often successful (see Acknowledgement).

Most included studies had multiple *ds* because they investigated more than one cortisol dose and/or used multiple memory measures. When a study included more than one measure of the same

[1] Values in the brackets indicate interrater reliabilities, which were calculated using the IntraClass Correlation ($ICC_{two\text{-}way\text{-}random\text{-}effects\text{-}model}$) for continuous variables and Cohens' Kappa (*6*) for categorical variables, as suggested in meta-analytical literature (e.g. Orwin, 1994).

[2] $g_{Hedges} = \frac{\bar{X}_{EG} - \bar{X}_{CG}}{S_{Pooled}}$, $S_{Pooled} = \sqrt{\frac{(n_{EG}-1)S_{EG}^2 + (n_{CG}-1)S_{CG}^2}{n_{EG}+n_{CG}-2}}$.

dependent variable (e.g. free recall of neutral words and free recall of pictures), *ds* were calculated for each variable and summarised by calculating an average *d*. Following Haberlandt (1999), we regarded delayed free and cued recall as well as delayed recognition as measures of declarative memory. Immediate free recall, cued recall or immediate recognition performances were regarded as measures of short-term or working memory. These measures were not summarized because they reflect different memory constructs. Free and cued recall was summarized to a variable called 'recall'. Recognition, in contrast, was always separated from 'recall' because of different degrees of difficulty (Haberlandt, 1999) and probably a different localization in neural structures (Brown and Aggleton, 2001; Buckner and Wheeler, 2001; Rugg and Yonelinas, 2003). We expected that effects were likelier to be seen for 'effortful' measures like free recall. Therefore we used recall and not recognition for the general analysis if a particular study contained both measures (e.g. Abercrombie et al., 2003). But we exploratively also looked at the recognition data. We only calculated effect sizes for memory of neutral stimuli. Separate analyses of emotional memories were not calculated because of the small number of studies performed on this topic (Buchanan and Lovallo, 2001; Abercrombie et al., 2003; Tops et al., 2003; Rimmele et al., 2003). Studies administering multiple doses (Beckwith et al., 1986; Newcomer et al., 1999; Lupien et al., 1999a; Abercrombie et al., 2003) contributed only one effect size to this research synthesis. We decided to choose the effect size of the dose which was closest to the median (i.e. 25 mg) of cortisol doses used in studies with only one cortisol dose. This approach was preferred over an average across multiple doses because a nonlinear dose response function has been reported in some multiple dose studies in humans (Lupien et al., 1999a; Abercrombie et al., 2003) and is supported by electrophysiological work in animals (see Lupien and McEwen (1997) and Sapolsky (2003) for review).

2.5. Analysis of effect sizes

For each study we calculated the average effect size for memory of neutral stimuli and its standard deviation. We defined its 95% confidence interval ($^{95\%}$CI) as a test of significance. Then we integrated all effect sizes to examine whether *ds* were consistent across the studies and could be considered as coming from the same population. For this each *d* was weighted by the reciprocal of its variance, in order to give greater weight to studies with a large sample size (Hedges and Olkin, 1985). Then we calculated the weighted average effectsize (*d*), its standard deviation and $^{95\%}$CI. To test whether the *ds* share a common effect size, we calculated a test of homogeneity (Q_T). In the case of a significant result, we assumed that the *ds* are heterogeneous, i.e. they are from different study populations. If so, we differentiated the *ds* into independent categories according to the categorical model of research synthesis, which is recommended in the case of a small number of integrated studies (Hedges and Olkin, 1985). As stated in the hypotheses, the time of treatment in the course of a study (before learning versus before retrieval) was hypothesized as a relevant determinant of the heterogenous results. Therefore we defined two categories (cortisol application before learning versus cortisol application before retrieval). Subsequently, we looked at the homogeneity of *ds* within each class (Q_T), then tested the overall homogeneity by the within-class-goodness-of-fit statistic (Q_W) and finally tested the homogeneity between the classes (Q_B). In the case of an undesired significant Q_T-statistic, we differentiated the according *ds* in categories of time of investigation (morning or afternoon). Analyses of contrasts and post hoc Scheffé-test were calculated, respectively, if the Q_B-statistic was significant.

If the number of studies integrated is small, as in our meta-analysis, a publication bias should be evaluated descriptively by a normal quantile plot according to Wang and Bushman (1998). This is a scatter plot with effect sizes on the abscissa and their estimated quantile of the standardized normal distribution on the ordinate. In the case of a large number of studies to be integrated a funnel plot should be used. A significance test for a publication bias was done according to Begg (1994). It is based on Spearman's correlation (r_s) between the standardized effect sizes and their variance. A significant positive correlation would suggest the presence of a publication bias.

3. Results

3.1. Description of study features

Sixteen experimental studies on the acute effects of cortisol administration on memory published in 15 papers (Abercrombie et al., 2003; Beckwith et al., 1986; Buchanan and Lovallo, 2001; de Quervain et al., 2000, 2003; Fehm-Wolfsdorf et al., 1993; Hsu et al., 2003; Kirschbaum et al.,

1996; Lupien et al., 1999a, 2002a; Monk and Nelson, 2002; Newcomer et al., 1999; Rimmele et al., 2003; Tops et al., 2003; Wolf et al., 2001a) met the inclusion criteria and were thus used in this meta-analysis. Average effect sizes and the most important characteristics for these studies are shown in Table 1.

A total of 563 participants were studied. Ten studies only investigated male subjects, while the remaining investigated mixed sex samples. The age of participants ranged from 18 to 40 years, with an overall mean of 24.23 years (SD = ±2.15).

All considered studies were randomized double-blind studies, except one, which was a randomized single-blind study (Kirschbaum et al., 1996). Four studies investigated the effect of different cortisol doses on memory performance. The used doses of cortisol ranged from 5 to 100 mg (Median = 25 mg; DM[3] = ±7.5). Three studies administrated cortisol intravenously, while the remaining used oral administration. With the exception of Beckwith et al. (1986), all studies controlled the cortisol levels in saliva or plasma. In 12 studies cortisol treatment took place before learning and four studies implemented treatment before recall. Mean delay between treatment and learning and treatment and recall was 1.34 h ($SD = \pm 2.11$). Retention interval ranged from 0 (immediate recall) up to 168 h (delayed recall). Eight studies measured memory performance by using simple word lists. The number of lists in these studies ranged from 1 to 8 lists. Two studies investigated memory with word pairs. Immediate recall was tested in seven studies, delayed recall in 11 studies and both were tested in four studies. Three studies investigated only recognition memory performance and six studies measured recognition and recall performances. Ten studies took place on more than one day. Five studies took place in the morning. If a particular study was conducted in the morning as well as in the afternoon (Fehm-Wolfsdorf et al., 1993) we decided by randomization which data should be considered. Similarly, we decided by randomization which recall test to include in the meta-analysis if multiple delayed recall tests were conducted (Abercrombie et al., 2003) within a study.

We regarded the paper of de Quervain et al. (2000) as a publication of two independent studies, because the effect of treatment with cortisol before learning and its administration before recall was studied in independent samples. Thus we considered two average effect sizes for this study. The independent sample which was treated with cortisone immediately after training was not considered, since it was the only study using this design. In total we calculated 86 effect sizes, with each study contributing from 1 to 24 effect sizes.

[3] Deviation of median.

3.2. Reduction of effect sizes

Most studies investigated more than one treatment group and/or more than one measure of memory. In addition there were studies, which investigated the effect of stimuli valence. As described, we concentrated on the effect of cortisol on memory of neutral stimuli. This resulted in 64 effect sizes. In studies, which measured different, not summable memory constructs (e.g. free recall versus recognition) we selected *ds* with regard to our hypotheses that difficult memory measures are likelier to show effects. Hence, 43 *ds* remained. For studies with multiple *ds* due to different measures we averaged these in order to create a single effect size for each study. For multiple dose studies we integrated those effect sizes which were nearest to the median of doses used (25 mg). The analysed effect sizes are shown with their $^{95\%}$CI in Fig. 1. A large variation of effect sizes is observable ranging from −1.10 (Tops et al., 2003) to .51 (Abercrombie et al., 2003).

3.3. Primary analysis

The integration of all effect sizes resulted in an average weighted effect size of $d = -.01$ ($-.17 \leq d \leq .14$). The value of Q_T was larger than the critical value of the χ^2 distribution ($\chi^2_{15} = 29.65$; $p < .05$), indicating significant heterogeneity. Thus the *ds* seem to originate from different populations and do not share a common underlying effect size.

Next we categorized the effect sizes according to the factor time of cortisol administration in the course of each study and created two groups of effect sizes (cortisol application before learning versus cortisol application before retrieval). The results of the corresponding categorical integration are shown in Fig. 1. As one can see, we found a negative average effect size of $d = -.49$ ($-.86 \leq d \leq -.10$) for studies which administered cortisol before recall. This effect size is significantly different to zero ($p < .01$). The test for homogeneity of effect sizes within this group of studies was not significant ($\chi^2_3 = .42$; $p > .05$). For studies which administered cortisol before learning we found an average effect size of $d = .08$ ($-.09 \leq d \leq .26$). This effect size is not significantly

Table 1 Summary of descriptive features and average effect sizes of memory performances for all included studies.

Study	n_{Total}	n_{IG}	Age			Sex	Design	Dose(s) of cortisol (mg)	Considered measure of memory	Time of treatment	Time of study	Retention interval (h)	n_d	d
			Range	Mean	SD									
Abercrombie et al. (2003)	90	3	18–33			Male	Between subject	20* and 40 (PO)	→Delayed free recall of words in session I →Delayed free recall of pictures in session I	Before learning	PM	.5	24	$d^{recall} = .51$ $d^{recognitionn} = .40$
Beckwith et al. (1986)	80	5				Male	Between subject	5,10, 20*, and 40 (PO)	→Immediate free recall of words	Before learning	NN	0	10	$d^{recall} = .18$
Buchanan and Lovallo (2001)	48	2	20–40	26.70		Mixed	Between subject	20 (PO)	→Delayed free recall of pictures →Delayed cued recall of pictures	Before learning	PM	168	6	$d^{recall} = .46$ $d^{recognition} = -.19$
de Quervain et al. (2000)	12	2	20–40	28.80	5.50	Mixed	Within subject	25 (PO)	→Delayed free recall of words	Before recall	PM	24	6	$d^{recall} = -.49$ $d^{recognitionn} = .01$
de Quervain et al. (2000)	12	2	20–40	28.80	5.50	Mixed	Within subject	25 (PO)	→Delayed free recall of words	Before learning	PM	24	6	$d^{recall} = -.02$ $d^{recognition} = .05$
de Quervain et al. (2003)	14	1	21–27	22.70	1.80	Male	Within subject	25 (PO)	→Delayed cued recall of word pairs	Before recall	PM	24	2	$d^{recall} = -.56$ $d^{recognitionn} = -.70$
Fehm-Wolfsdorf et al. (1993)	18	1		24		Male	Within subject	50 (PO)	→Immediate free recall of words	Before learning	AM	0	2	$d^{recall} = -.03$
Hsu et al. (2003)	20	1	18–32	22	.70	Male	Within subject	100 (PO)	→Delayed recognition of words	Before learning	PM	.05	2	$d^{recognition} = .19$
Kirschbaum et al. (1996)	40	2		24.70	2.70	Male	Between subject	10 (PO)	→Delayed cued recall of words	Before learning	AM	.5	1	$d^{recall} = -.67$
Lupien et al. (1999a)	40	4		24.35	3.13	Male	Between subject	5.79, 43.75* and 88.90 (IV)	→Delayed cued recall of word pairs	Before learning	AM	.25	6	$d^{recall} = .08$
Lupien et al. (2002a)	18	2	20–30	24.20	4.30	Male	Between subject	35 (IV)	→Delayed recognition of word stem	*Before* recall	PM	.5	1	$d^{recognition} = -.22$

Monk and Nelson (2002)	26	1	18-34			Mixed	Within subject	30 (PO)	→ Delayed recall of faces → Delayed recall of objects	Before learning	PM	1	4	$d^{recall} = -.28$
Newcomer et al. (1999)	51	3	18-30	22.20	2.20	Mixed	Between subject	25* and 100 (PO)	→ Delayed free paragraph recall	Before learning	PM	.5	6	$d^{recall} = .18$
Rimmele et al. (2003)	63	4		25.90	5.80	Mixed	Between subject	25 (PO)	→ Delayed free recall of slides → Delayed free recall of details	Before learning	PM	168	6	$d^{recall} = .48$ $d^{recognition} = .42$
Tops et al. (2003)	22	2	18-27	21.18		Male	Between subject	10 (PO)	→ Immediate recognition of words	Before learning	AM	0	6	$d^{recognition} = -1.10$
Wolf et al. (2001a)	9	1	19-30	24	1.20	Male	Within subject	35-40 (IV)	→ Delayed free recall of words	*Before* recall	AM	2	3	$d^{recall} = -.63$

*, Considered dose; n_{IG}, number of independent groups; n_d, total number of calculated *ds*; *d*, (average) effect size of a particular study, which was considered in the following analyses; NN, missing information.

different to zero ($p = .05$). The test for homogeneity revealed a significant heterogeneity for the effect sizes of this category ($\chi^2_{11} = 22.15$; $p < .05$).

We differentiated the effect sizes of studies administering cortisol before learning according to the factor time of study (morning versus afternoon) because of the significant heterogeneity of *ds* of these studies. Statistical analyses of homogeneity for each of the three categories are described in Table 2.

As apparent from Fig. 1, negative effect sizes were calculated for studies which were performed in the morning. Most studies conducted in the afternoon show positive or no effects. The study by Beckwith et al. (1986) was excluded from this analysis, because its study time was not described in the paper and we could not contact the author. Stratification was finished and heterogeneity between the classes was analysed because there was no category with significant heterogeneity of *ds*.

As expected, we found a significant Q_B value, which indicates a significant heterogeneity between the three classes ($\chi^2_2 = 15.37$; $p < .01$). As shown in Table 2, the $^{95\%}$CI of the averaged weighted effect sizes of each category did not include zero, which indicates significant average effect sizes for each category ($p < .05$). A post hoc Scheffé-procedure revealed that studies which administered cortisol before learning and which were conducted in the afternoon differed significantly from the other two categories ($\chi^2_1 = 8.30$; $p < .01$ and $\chi^2_1 = 10.33$; $p < .01$). As shown in Table 2, Q_T was not significant in all categories, indicating homogeneity of effect sizes within each category. Thus each of the three *ds* can be interpreted. In accordance with this, Q_W was not significant either ($\chi^2_{13} = 14.28$; $p > .05$).

Additionally we also analyzed our data set with Revman (Version 4.2), a free software (http://www.cc-ims.net/RevMan) which includes a statistical program for meta-analyses. This software is among others used for the Cochrane Reviews. The algorithms underlying this software differ in some aspects slightly from the approach originally used by us (e.g. the way the variance of d is estimated). Using this software we recalculated the effect sizes for our two main hypothesis (1. learning versus retrieval and 2. morning versus afternoon) firstly using the fixed effect model and secondly the random effect model. Almost identical *d* and *p* values were obtained (data not shown), thereby supporting the conclusions derived from our original approach described above.

Exploratory integration of effect sizes of recognition memory performances (also displayed in Table 1) revealed an average weighted effect size

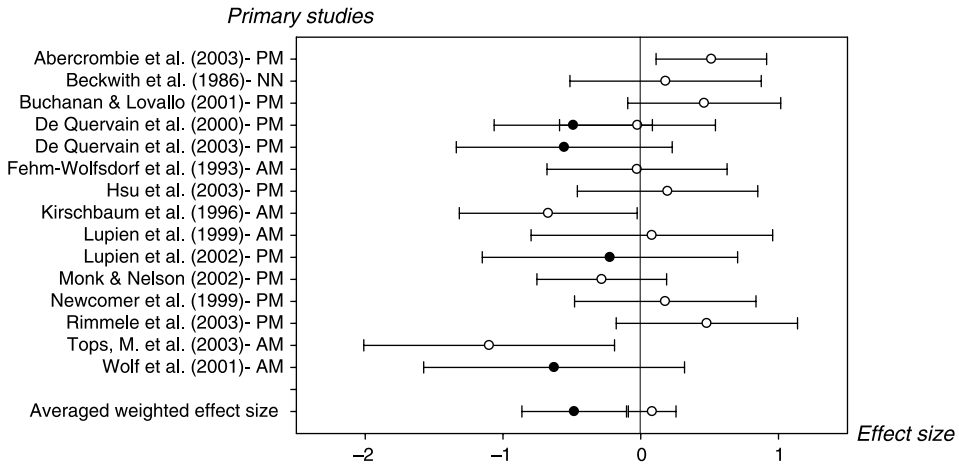

Figure 1 Scatter plot of average effect sizes and their $^{95\%}$CI of cortisol-memory studies. Bright dots (○) indicate treatment before learning, while dark dots (●) indicate treatment before recall. Studies, which were conducted in the morning, are signed with 'AM'. Studies conducted in the afternoon are marked as 'PM'. 'NN' indicates missing information. The average weighted effect sizes of both categories are shown beneath.

of $d = .01$. The according $^{95\%}$CI was ($-.20 \leq d \leq .21$). The test of homogeneity revealed a trend ($\chi^2_8 = 14.43$; $p = .07$). A comparison of the effect sizes for recall and recognition was not performed because of violated assumption of independency of the integrated studies.

Publication bias can conceivably distort the validity and therefore the conclusions of a meta-analysis. We used the normal quantile-plot to evaluate the presence of a publication bias. The effect sizes form a linear curve (data not shown), which indicates that they are normally distributed. A significant positive correlation between the standardized effect sizes and their variances would indicate a publication bias (Begg, 1994). For the present analysis the test of significance revealed a negative, non-significant correlation between the standardized effect sizes and their variances of $r_s = -34$ ($p = .20$), which indicates the absence of a publication bias.

4. Discussion

We reviewed meta-analytically the impact of acute cortisol treatment on human memory using the results from 16 independent placebo controlled studies with a total of 563 subjects. The goal was to test whether or not influential theories about key modulatory variables, like memory phase and time of day are supported by the currently available empirical evidence.

As expected, we found a significant heterogeneity within the effect sizes of all included studies. This indicates significant differences between the studies, and therefore does not allow a general conclusion about the effects of acute cortisol treatment on memory. Therefore, we differentiated the effect sizes into two independent categories which accord with the model of Roozendaal (2002). This model suggests that cortisol impairs memory retrieval whilst enhancing memory consolidation.

Table 2 Results of categorical integration of average effect sizes.

Category	Significance test for the average effect size		Analysis of effect size heterogeneity		
	d	$^{95\%}$CI	χ^2	df	p
Cortisol before recall ($n=4$)	−.49*	{−.86 ≤ d ≤ −.10}	.42	3	.94
Cortisol before learning—a.m. ($n=4$)	−.40*	{−.77 ≤ d ≤ −.03}	5.36	3	.14
Cortisol before learning—p.m. ($n=8$)	.22*	{.02 ≤ d ≤ .41}	8.49	7	.29

Integration was done with regard to time of cortisol application (before learning versus before recall) and time of investigation (morning versus afternoon). *Indicates that the average effect size is significantly different from 0. χ^2, df and p-levels are from the test for homogeneity of the effect sizes within each category (Q_T).

However, we found only one study (de Quervain et al., 2000) which investigated the acute effects of cortisol on consolidation by treating the subjects immediately after learning with cortisol. We therefore defined one category for studies in which cortisol was administered before retrieval and the other category for studies administering cortisol before learning. It is noteworthy that studies in which cortisol was administered before learning with recall being tested shortly afterwards are difficult to interpret because it remains unclear which memory phase is affected by cortisol (initial learning, consolidation or retrieval). Therefore, results of the studies in this category can only partially test the model of Roozendaal (2002).

For the primary categorical integration we found a negative average effect size of $d = -.49$ for studies which administered cortisol before retrieval. Subjects treated before retrieval showed a decrease in performance in all four studies. The effect sizes were very similar except for the study by Lupien et al. (2002a), which may be due to the used delay (30 min) as well as the used recall paradigm (recognition) used in this study.

Despite this, our analysis strongly supports the idea that treatment with GCs given before memory retrieval causes impairment in humans as previously described in rats (Roozendaal, 2002). According to Cohen (1988), the size of this deficit can be interpreted as moderate. Subjects treated with cortisol performed on average nearly half a standard deviation below subjects treated with placebo. Such an effect could conceivably be of relevance for everyday life (e.g. for performance in exams, medical treatment with GCs or in testimony situations). This effect and its size can of course not be generalized to short-term memory, retrieval of procedural memory or retrieval of declarative emotional stimuli. All these domains await additional future experimental investigation.

Recently Buss et al. (2004) examined the effects of cortisol on the specificity of autobiographical memories. Since this test differs substantially from other declarative memory tasks (e.g. there is no learning phase) it was not included into the current meta-analysis. However, the observed effect size in this study ($d = -.52$) is very similar to those observed in the other cortisol retrieval studies. In our view this suggests that the effects of cortisol on memory retrieval are similar for episodic and autobiographic memory.

In the face of an obvious variation of effect sizes of studies which administered cortisol before learning and a significant p-value for the test for homogeneity, additional stratification was performed. We investigated the influence of time of day (morning versus afternoon). This analysis is based on the observations from Lupien et al. (2002a). The researchers reported that GCs administered at the time of the circadiane trough or after a pharmacological cortisol depletion have a positive effect on memory. This relationship might be mediated by differential activation of MRs and GRs (Lupien et al., 2002a). At the time of the circadiane trough, i.e. in the afternoon, MRs are occupied by cortisol, while GRs are hardly occupied (De Kloet et al., 1998). A moderate pharmacological elevation of cortisol might lead to saturated MRs and a partial activation of GRs. MR-saturation plus moderate GR-activation might in turn lead to enhanced cognitive performance (De Kloet et al., 1998; Sapolsky, 2003).

The average effect size for studies taking place in the morning was significantly different from zero and had a magnitude of $d = -.40$. However, one has to note that the effect sizes of studies which were conducted in the morning varied substantially, and that the effect sizes of the study of Lupien et al. (1999a) and Fehm-Wolfsdorf et al. (1993) are smaller compared to the other two studies of this category (Kirschbaum et al., 1996; Tops et al., 2003). Different memory tests as well as different cortisol doses used might partially explain these discrepancies.

The other category contained studies which administered cortisol before learning and which were conducted in the afternoon. These studies found either no effect (de Quervain et al., 2000) or a slight enhancing effect of cortisol (Buchanan and Lovallo, 2001; Abercrombie et al., 2003; Hsu et al., 2003; Rimmele et al., 2003), except the study of Monk and Nelson (2002). For this study we observed a negative effect size, which mainly resulted from the delayed recognition of faces ($d = -.57$). For the other dependent variable in this study, i.e. delayed recognition of objects we found an effect size of $d = .01$, which accords more to the other effect sizes of this category. The average effect size for this category is smaller ($d = .22$) but positive and therefore opposed to the effects observed in the other two categories. Moreover this small effect size was also still significantly different from zero.

The results of this analysis on the influence of the time of day on the effect sizes of studies which administered cortisol before learning suggest that this factor is also a determinant of the contradictory results in this area. As suggested by Lupien et al. (2002a), cortisol elevations in the morning appear to lead to impairing effects, while cortisol elevations in the afternoon might lead to absent or enhancing effects. This finding cannot be generalized to the category of studies which administered cortisol before retrieval because it contained both types of

studies, i.e. three of these studies were conducted in afternoon while only one study was conducted in the morning. So the negative effect of cortisol on retrieval might be relatively independent of the time of day. Future studies should consider the time of day when designing their experiments and should explicitly report.

The current meta-analysis has of course several methodological limitations. We used very stringent inclusion and exclusion criteria (see Section 2) in order to avoid the main problems of most meta-analytical studies, namely the trash-in-trash-out problem and the well-known apples-and-oranges problem (Matt and Cook, 1994; Hall et al., 1994).

The inclusion of experimental studies exclusively and the reliance on studies published in peer reviewed journals substantially reduces the trash-in-trash-out problem, which originates if studies with poor qualities are integrated and which threats the validity of the results of the research synthesis. However, several of the analyzed studies can be criticized for their small samples sizes, even though sample size was integrated in the weighted average effect sizes created.

The reduction of the apples-and-oranges problem has always been problematic in research synthesis. We only included studies in which cortisol or cortisone was administered in a placebo controlled design in order to reach homogeneity in independent variables. If a primary study used different doses of cortisol we only considered the effect size of the group which received a dosage nearest to the median (i.e. 25 mg) of cortisol doses used in studies with only one dosage. On the other hand, such stringent inclusion criteria for treatment do not allow conclusions on closely related pharmacological treatments, i.e. prednisone or dexamethasone. However, studies investigating cortisol and dexamethasone in identical experimental designs observed similar results for both hormones (e.g. Newcomer et al., 1994, 1999). Additionally, the negative effect of GCs on retrieval was recently shown with prednisone as well (de Quervain et al., 2004). Another issue is the homogeneity in the dependent variables. Actually, we had little homogeneity in measures of memory since studies differed profoundly in the used memory tests, retention intervals and retrieval testing methods. This heterogeneity of dependent measures was not avoidable and might in part be responsible for the remaining variance within the categories. To reduce this heterogeneity we did not summarize results of memory tasks of different severity (e.g. recognition and free recall) but rather considered only the results of the free recall task.

For recognition memory, we found that the effect sizes were on average descriptively smaller—almost zero—than the effect sizes for free or cued recall performances. This finding may indicate that recognition memory performance is less suitable to uncover effects of cortisol on memory. Whether this observation reflects differences in task difficulty/sensitivity of differences in involved brain regions (prefrontal versus medial temporal; Brown and Aggleton, 2001; Buckner and Wheeler, 2001; Rugg and Yonelinas, 2003) awaits to be investigated in future neuroimaging studies (de Quervain et al., 2003).

More studies are needed which assess different memory constructs. Currently the number of studies measuring other memory constructs was too small. We found only two studies (Lupien et al., 1999a; Wolf et al., 2001a), which investigated working memory and met the inclusion criteria. Only one study (Kirschbaum et al., 1996) tested effects on procedural memory. Especially working memory is of interest in this research area due to the high number of GRs in the prefrontal cortex of primates (see Lupien and Lepage (2001)). Only after the publishing of more experiments, will we be able to meta-analytically test if cortisol differentially influences short-term/working, declarative and procedural memory.

Studies using psychosocial stress to activate the HPA axis have observed different effects of stress on memory for high versus low cortisol stress-responders (e.g. Takahashi et al., 2004). The interesting and important issue of individual differences in the responsiveness to pharmacologically administered glucocorticoids has received little attention as of to date. Such differences might be able to account for some of the variance observed within as well as between studies. Individual differences could reflect genetic factors (e.g. Wust et al., 2004), differences in tissue glucocorticoid sensitivity (e.g. Rohleder et al., 2003) or could be caused by differences in lifetime cortisol exposure (Lupien et al., 2002b). Attempts to characterize 'cortisol responder' and 'non-responder' appears to be a fruitful venue for future pharmacological studies.

Most studies in the present meta-analysis used a delay between initial treatment and cognitive testing of at least 30 min, but see Lupien et al. (1999a) and Hsu et al. (2003) for exceptions. Most study authors suggest that genomic effects (mediated via the two GC receptors) underlie the behavioural changes observed in humans. However, in rodents behavioural studies, for instance on exploratory reactivity (Sandi et al., 1996) as well as electrophysiological studies (e.g. Joels, 2001) also

reported rapid GC effects (within minutes) which most likely reflect non-genomic GC actions. More studies on rapid GC effects in the human would be desirable.

We only analyzed studies with young healthy subjects. Some previous experiments observed changes in the response to cortisol treatment with aging or psychiatric disease (Lupien et al., 1994, 1999b, 2002b; Wolf et al., 2001a, 2002; Porter et al., 2002; Bremner et al., 2004). In the future with more studies at hand it would be interesting to investigate the effects of aging or disease meta-analytically. At present this appeared unfeasible.

There are (too) few studies investigating female subjects, most likely reflecting an attempt to avoid the fluctuation in sex hormones which by themselves might modulate cognition or GC sensitivity (Hampson, 1990; Rohleder et al., 2001, 2003). Only six of 16 studies investigated samples of both sexes and no study investigated only females. Furthermore, only one of the studies with mixed sex samples controlled menstrual cycle status and only two controlled for the use of oral contraceptives (Newcomer et al., 1999; Monk and Nelson, 2002). So, our findings cannot easily be generalized to women. However, none of the studies which used mixed sex sample observed and/or reported sex differences. The situation might be different for studies using stressors, as here sex differences have been reported in rodents (e.g. Luine, 2002; Beck and Servatius, 2003; Shors et al., 2004) as well as in humans (Wolf et al., 2001b).

Furthermore, there is a lack of studies comparing effects of different doses of cortisol on memory. Currently there are only four studies available which used multiple cortisol doses (Beckwith et al., 1986; Newcomer et al., 1999; Lupien et al., 1999a; Abercrombie et al., 2003). Therefore, we were unable to test for the presence of an inversed U-shaped dose response curve as observed in behavioural and electrophysiological animal studies (see McEwen and Sapolsky (1995), Lupien and McEwen (1997) and Sapolsky (2003) for review).

Another limitation is that there is only one human study (de Quervain et al., 2000) investigating the effect of cortisol on memory consolidation by administering the hormone immediately after learning, which is in contrast to the large number of animal studies using this approach (see Roozendaal (2002) for review). To evaluate whether an enhancing effect of cortisol on consolidation also occurs in humans, as suggested by Buchanan and Lovallo (2001), more studies are needed which attempt to directly replicate the findings obtained in rodents.

Furthermore, we did not analyze the influence of valence or arousal on the effect of cortisol on memory since only four published studies systematically investigated this issue. Those studies differed substantially in their used methods and designs. In contrast to Abercrombie et al. (2003), three studies reported that the effects of GCs are modulated by arousal or valence (Buchanan and Lovallo, 2001; Tops et al., 2003; Rimmele et al., 2003). It is conceivable that cortisol and emotional arousal might interact differently depending on the memory phase tested. Indeed while Buchanan and Lovallo (2001) observed that the beneficial effects of cortisol on memory consolidation is only detectable for arousing pictures, we just recently observed that the impairing effect of cortisol on retrieval is stronger for emotionally arousing words (Kuhlmann et al., 2005). Again more studies are needed in order to evaluate the effects of stimuli valence using meta-analysis. These studies are especially important since they might be of more relevance for psychiatric disorders like depression or PTSD than studies using mostly neutral learning material.

The most severe limitation of this meta-analysis is that our findings are based on a relatively small number of studies—especially in the created subcategories. However, 16 studies or even less is not uncommon for meta-analyses to present moderator analyses—especially if the inclusion criteria are strict (Hedges and Olkin, 1985; Kaylor et al., 1987; Benschop et al., 1998; Hogervorst et al., 2000; Sack et al., 2001; van Emmerik et al., 2002; Sauro et al., 2003; Deacon and Abramowitz, 2004). With the small number of studies available we had to use a categorical approach (Hedges, 1994) which implicates a loss of information. With a larger number of studies the use of a general linear model would have been possible. This would have enabled us to estimate the continuous influence of several variables on the effect sizes simultaneously and also allowed the detection of interactions between two or more predictors (e.g. time of day and retention interval). Such a linear model should be based on at least 20 studies and additionally requires a ratio between the number of predictors and the number of included studies of at least 1:15 in order to estimate regression coefficients appropriately (Hedges and Olkin, 1985; Rustenbach, 2003).

Publication bias is an important issue in meta-analysis because it threatens the internal and the external validity of the meta-analytical findings and raises the option that obtained results reflect publishing practice. Since the number of studies in

our meta-analysis is small we investigated the publication bias using a method recommended by Wang and Bushman (1998). The obtained results suggest that a publication bias is unlikely and support the idea that our effect sizes are interpretable and valid.

In sum, our research synthesis documents that the acute effects of cortisol on memory are not always beneficial or detrimental but depend on several modulating variables. Quantitative evidence is provided that the effects of cortisol substantially depend on the time of cortisol application in the course of a study. Cortisol given before recall impairs declarative memory retrieval. In contrast, we could not uncover a clear effect in studies which administered cortisol before learning. Further analysis of these studies indicates that studies conducted in the morning appear to find impairing effects, while studies in the afternoon might find enhancing or no effects. Although a differentiated picture of the acute effects of cortisol on memory has evolved over the last decade much work is ahead of the scientific community.

Acknowledgements

This study was supported by grants from the German Research Foundation (DFG; WO 733/3-1; WO 733/6-1). The authors wish to thank Prof. Krauth and Dr. Rustenbach for their most helpful statistical comments. We thank Stefanie Holdmann for language editing. The following authors provided additional data from their studies and thereby contributed to the success of this meta-analysis: H.C. Abercrombie; T.W. Buchanan; G. Domes; D.J. de Quervain; S.J. Lupien; R.H. McAllister-Williams; C. Monk; J.W. Newcomer and M. Tops.

References

Abercrombie, H.C., Kalin, N.H., Thurow, M.E., Rosenkranz, M.A., Davidson, R.J., 2003. Cortisol variation in humans affects memory for emotionally laden and neutral information. Behav. Neurosci. 117, 505–516.

Beck, K.D., Servatius, R.J., 2003. Stress and cytokine effects on learning: what does sex have to do with it? Integr. Physiol Behav. Sci. 38, 179–188.

Beckwith, B.E., Petros, T.V., Scaglione, C., Nelson, J., 1986. Dose-dependent effects of hydrocortisone on memory in human males. Physiol. Behav. 36, 283–286.

Begg, C.B., 1994. Publication bias. In: Cooper, H., Hedges, L.V. (Eds.), The Handbook of Research Synthesis. Russell Sage, New York, NY, pp. 399–409.

Belanoff, J.K., Gross, K., Yager, A., Schatzberg, A.F., 2001. Corticosteroids and cognition. J. Psychiatr. Res. 35, 127–145.

Benschop, R.J., Geenen, R., Mills, P.J., Naliboff, B.D., Kiecolt-Glaser, J.K., Herbert, T.B., van der, P.G., Miller, G.E., Matthews, K.A., Godaert, G.L., Gilmore, S.L., Glaser, R., Heijnen, C.J., Dopp, J.M., Bijlsma, J.W., Solomon, G.F., Cacioppo, J.T., 1998. Cardiovascular and immune responses to acute psychological stress in young and old women: a meta-analysis. Psychosom. Med. 60, 290–296.

Bremner, J.D., Vythilingam, M., Vermetten, E., Anderson, G., Newcomer, J.W., Charney, D.S., 2004. Effects of glucocorticoids on declarative memory function in major depression. Biol. Psychiatry 55, 811–815.

Brown, M.W., Aggleton, J.P., 2001. Recognition memory: what are the roles of the perirhinal cortex and hippocampus?. Nat. Rev. Neurosci. 2, 51–61.

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

Buckner, R.L., Wheeler, M.E., 2001. The cognitive neuroscience of remembering. Nat. Rev. Neurosci. 2, 624–634.

Buss, C., Wolf, O.T., Witt, J., Hellhammer, D.H., 2004. Autobiographic memory impairment following acute cortisol administration. Psychoneuroendocrinology 29, 1093–1096.

Cohen, J., 1988. Statistical Power Analysis for the Behavioral Sciences. LEA, New York, NY.

Deacon, B.J., Abramowitz, J.S., 2004. Cognitive and behavioral treatments for anxiety disorders: a review of meta-analytic findings. J. Clin. Psychol. 60, 429–441.

De Kloet, E.R., Ratka, A., Reul, J.M., Sutanto, W., Van Eekelen, J.A., 1987. Corticosteroid receptor types in brain: regulation and putative function. Ann. N.Y. Acad. Sci. 512, 351–361.

De Kloet, E.R., Vreugdenhil, E., Oitzl, M.S., Joels, M., 1998. Brain corticosteroid receptor balance in health and disease. Endocr. Rev. 19, 269–301.

De Kloet, E.R., Oitzl, M.S., Joels, M., 1999. Stress and cognition: are corticosteroids good or bad guys?. Trends Neurosci. 22, 422–426.

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.

de Quervain, D.J., Henke, K., Aerni, A., Treyer, V., McGaugh, J.L., Berthold, T., Nitsch, R.M., 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.

de Quervain, D.J., Wolf, O.T., Forster, A., Kollias, S., Hock, C., Nitsch, R.M., Roozendaal, B., Conrad, C.D., 2004. Memory impairment under chronic glucocorticoid treatment is reversible after discontinuation of the treatment. Soc. Neurosci. (abstract).

Diamond, D.M., Bennett, M.C., Fleshner, M., Rose, G.M., 1992. Inverted-U relationship between the level of peripheral corticosterone and the magnitude of hippocampal primed burst potentiation. Hippocampus 2, 421–430.

Fehm-Wolfsdorf, G., Reutter, K., Zenz, H., Born, J., Lorenz, H., 1993. Are circadian variations in taste thresholds cortisol-dependent?. J. Psychophysiol. 7, 65–72.

Flood, J.F., Vidal, D., Bennet, E.L., Orme, A.E., Vasquez, S., Jarvik, M.E., 1978. Memory facilitating and anti-amnestic effects of corticosteroids. Pharmacol. Biochem. Behavior 8, 81–87.

Glass, G.V., 1976. Primary secondary, and meta-analysis of research. Educ. Res. 5, 3–8.

Gold, P.W., Drevets, W.C., Charney, D.S., 2002. New insights into the role of cortisol and the glucocorticoid receptor in severe depression. Biol. Psychiatry 52, 381–385.

Haberlandt, K., 1999. Human Memory. Exploration and Application. Allyn and Bacon, Boston, MA.

Hall, J.A., Rosenthal, R., Tickle-Degnen, L., Mosteller, F., 1994. Hypotheses and problems in research synthesis. In: Cooper, H., Hedges, L.V. (Eds.), The Handbook of Research Synthesis. Roussel Sage Foundation, New York, NY, pp. 17–28.

Hampson, E., 1990. Variations in sex-related cognitive abilities across the menstrual cycle. Brain Cogn. 14, 26–43.

Hedges, L.V., 1994. Fixed effect models. In: Cooper, H. (Ed.), The Handbook of Research Synthesis. Roussel Sage Foundation, New York, NY, pp. 285–299.

Hedges, L.V., Olkin, I., 1985. Statistical Methods for Meta-analysis. Academic Press, Orlando, FL.

Hogervorst, E., Williams, J., Budge, M., Riedel, W., Jolles, J., 2000. The nature of the effect of female gonadal hormone replacement therapy on cognitive function in post-menopausal women: a meta-analysis. Neuroscience 101, 485–512.

Hsu, F.C., Garside, M.J., Massey, A.E., McAllister-Williams, R.H., 2003. Effects of a single dose of cortisol on the neural correlates of episodic memory and error processing in healthy volunteers. Psychopharmacology (Berl) 167, 431–442.

Joels, M., 2001. Corticosteroid actions in the hippocampus. J. Neuroendocrinol. 13, 657–669.

Kaylor, J.A., King, D.W., King, L.A., 1987. Psychological effects of military service in Vietnam: a meta-analysis. Psychol. Bull. 102, 257–271.

Kirschbaum, C., Wolf, O.T., May, M., Wippich, W., Hellhammer, D.H., 1996. Stress- and treatment-induced elevations of cortisol levels associated with impaired declarative memory in healthy adults. Life Sci. 58, 1475–1483.

Kovacs, G.L., Teledgdy, G., Lissak, K., 1976. 5-Hydroxytryptamine and the mediation of pituitary-adrenocortical hormones in the extinction of active avoidance behaviour. Psychoneuroendocrinology 1, 219–230.

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. Neurobiol. Learn. Mem. 83, 158–162.

Luine, V., 2002. Sex differences in chronic stress effects on memory in rats. Stress 5, 205–216.

Lupien, S.J., Lepage, M., 2001. Stress, memory, and the hippocampus: can't live with it, can't live without it. Behav. Brain Res. 127, 137–158.

Lupien, S.J., McEwen, B.S., 1997. The acute effects of corticosteroids on cognition: integration of animal and human model studies. Brain Res. Brain Res. Rev. 24, 1–27.

Lupien, S., Lecours, A.R., Lussier, I., Schwartz, G., Nair, N.P., Meaney, M.J., 1994. Basal cortisol levels and cognitive deficits in human aging. J. Neurosci. 14, 2893–2903.

Lupien, S.J., Gillin, C.J., Hauger, R.L., 1999a. Working memory is more sensitive than declarative memory to the acute effects of corticosteroids: a dose–response study in humans. Behav. Neurosci. 113, 420–430.

Lupien, S.J., Nair, N.P., Briere, S., Maheu, F., Tu, M.T., Lemay, M., McEwen, B.S., Meaney, M.J., 1999b. Increased cortisol levels and impaired cognition in human aging: implication for depression and dementia in later life. Rev. Neurosci. 10, 117–139.

Lupien, S.J., Wilkinson, C.W., Briere, S., Menard, C., Ng Ying Kin, N.M., Nair, N.P., 2002a. The modulatory effects of corticosteroids on cognition: studies in young human populations. Psychoneuroendocrinology 27, 401–416.

Lupien, S.J., Wilkinson, C.W., Briere, S., Ng Ying Kin, N.M., Meaney, M.J., Nair, N.P., 2002b. Acute modulation of aged human memory by pharmacological manipulation of glucocorticoids. J Clin. Endocrinol. Metab. 87, 3798–3807.

Matt, G.E., Cook, T.D., 1994. Threats to the validity of research syntheses. In: Cooper, H., Hedges, L.V. (Eds.), The Handbook of Research Syntheses. Russell Sage Foundation, New York, NY, pp. 503–520.

McEwen, B.S., Sapolsky, R.M., 1995. Stress and cognitive function. Curr. Opin. Neurobiol. 5, 205–216.

Monk, C.S., Nelson, C.A., 2002. The effects of hydrocortisone on cognitive and neural function: a behavioral and event-related potential investigation. Neuropsychopharmacology 26, 505–519.

Newcomer, J.W., Craft, S., Hershey, T., Askins, K., Bardgett, M.E., 1994. Glucocorticoid-induced impairment in declarative memory performance in adult humans. J. Neurosci. 14, 2047–2053.

Newcomer, J.W., Selke, G., Melson, A.K., Hershey, T., Craft, S., Richards, K., Alderson, A.L., 1999. Decreased memory performance in healthy humans induced by stress-level cortisol treatment. Arch. Gen. Psychiatry 56, 527–533.

Oitzl, M.S., Reichardt, H.M., Joels, M., De Kloet, E.R., 2001. Point mutation in the mouse glucocorticoid receptor preventing DNA binding impairs spatial memory. Proc. Natl Acad. Sci. USA 98, 12790–12795.

Okuda, S., Roozendaal, B., McGaugh, J.L., 2004. Glucocorticoid effects on object recognition memory require training-associated emotional arousal. Proc. Natl Acad. Sci. USA 101, 853–858.

Orwin, R.G., 1994. Evaluating coding decisions. In: Cooper, H., Hedges, L.V. (Eds.), Handbook of Research Synthesis. Russell Sage, New York, NY, pp. 139–162.

Pavlides, C., Kimura, A., Magarinos, A.M., McEwen, B.S., 1994. Type I adrenal steroid receptors prolong hippocampal long-term potentiation. Neuroreport 5, 2673–2677.

Porter, R.J., Barnett, N.A., Idey, A., McGuckin, E.A., O'Brien, J.T., 2002. Effects of hydrocortisone administration on cognitive function in the elderly. J. Psychopharmacol. 16, 65–71.

Reul, J.M., De Kloet, E.R., 1985. Two receptor systems for corticosterone in rat brain: microdistribution and differential occupation. Endocrinology 117, 2505–2511.

Rimmele, U., Domes, G., Mathiak, K., Hautzinger, M., 2003. Cortisol has different effects on human memory for emotional and neutral stimuli. Neuroreport 14, 2485–2488.

Rohleder, N., Schommer, N.C., Hellhammer, D.H., Engel, R., Kirschbaum, C., 2001. Sex differences in glucocorticoid sensitivity of proinflammatory cytokine production after psychosocial stress. Psychosom. Med. 63, 966–972.

Rohleder, N., Wolf, J.M., Kirschbaum, C., 2003. Glucocorticoid sensitivity in humans-interindividual differences and acute stress effects. Stress 6, 207–222.

Roozendaal, B., 2002. Stress and memory: opposing effects of glucocorticoids on memory consolidation and memory retrieval. Neurobiol. Learn. Mem. 78, 578–595.

Roozendaal, B., Griffith, Q.K., Buranday, J., de Quervain, D.J., McGaugh, J.L., 2003. The hippocampus mediates glucocorticoid-induced impairment of spatial memory retrieval: dependence on the basolateral amygdala. Proc. Natl Acad. Sci. USA 100, 1328–1333.

Roozendaal, B., McReynolds, J.R., McGaugh, J.L., 2004. The basolateral amygdala interacts with the medial prefrontal cortex in regulating glucocorticoid effects on working memory impairment. J. Neurosci. 24, 1385–1392.

Rosenthal, R., 1994. Parametric measures of effect size. In: Cooper, H., Hedges, L.V. (Eds.), The Handbook of Research Synthesis. Russell Sage Foundation, New York, NY, pp. 231–244.

Rugg, M.D., Yonelinas, A.P., 2003. Human recognition memory: a cognitive neuroscience perspective. Trends Cogn. Sci. 7, 313–319.

Rustenbach, J.S., 2003. Metaanalyse, Eine anwendungsorientierte Einführung. Hans Huber, Bern, CH.

Sack, M., Lempa, W., Lamprecht, F., 2001. Study quality and effect-sizes—a metaanalysis of EMDR-treatment for posttraumatic stress disorder. Psychother. Psychosom. Med. Psychol. 51, 350–355.

Sandi, C., Venero, C., Guaza, C., 1996. Novelty-related rapid locomotor effects of corticosterone in rats. Eur. J. Neurosci. 8, 794–800.

Sapolsky, R.M., 2003. Stress and plasticity in the limbic system. Neurochem. Res. 28, 1735–1742.

Sauro, M.D., Jorgensen, R.S., Pedlow, C.T., 2003. Stress, glucocorticoids, and memory: a meta-analytic review. Stress 6, 235–245.

Schwarzer, R., 1989. Meta-Analysis Programs. [5.3]. Free University, Berlin.

Shors, T.J., Falduto, J., Leuner, B., 2004. The opposite effects of stress on dendritic spines in male vs. female rats are NMDA receptor-dependent. Eur. J Neurosci. 19, 145–150.

Takahashi, T., Ikeda, K., Ishikawa, M., Tsukasaki, T., Nakama, D., Tanida, S., Kameda, T., 2004. Social stress-induced cortisol elevation acutely impairs social memory in humans. Neurosci. Lett. 363, 125–130.

Tops, M., van der P, G., Baas, D., Mulder, L.J., Den Boer, J.A., Meijman, T.F., Korf, J., 2003. Acute cortisol effects on immediate free recall and recognition of nouns depend on stimulus valence. Psychophysiology 40, 167–173.

Vaher, P., Luine, V., Gould, E., McEwen, B.S., 1994a. Adrenalectomy impairs spatial memory in rats. Ann. N.Y. Acad. Sci. 746, 405–407.

Vaher, P.R., Luine, V.N., Gould, E., McEwen, B.S., 1994b. Effects of adrenalectomy on spatial memory performance and dentate gyrus morphology. Brain Res. 656, 71–78.

van Emmerik, A.A., Kamphuis, J.H., Hulsbosch, A.M., Emmelkamp, P.M., 2002. Single session debriefing after psychological trauma: a meta-analysis. Lancet 360, 766–771.

Wang, M.C., Bushman, B.J., 1998. Using the normal quantile plot to explore meta-analytic data sets. Psychol. Methods 3, 46–54.

Wolf, O.T., 2003. HPA axis and memory. Bes. Pract. Res. Clin. Endocrinol. Metab. 17, 287–299.

Wolf, O.T., Convit, A., McHugh, P.F., Kandil, E., Thorn, E.L., De Santi, S., McEwen, B.S., de Leon, M.J., 2001a. Cortisol differentially affects memory in young and elderly men. Behav. Neurosci. 115, 1002–1011.

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

Wolf, O.T., Convit, A., Thorn, E., de Leon, M.J., 2002. Salivary cortisol day profiles in elderly with mild cognitive impairment. Psychoneuroendocrinology 27, 777–789.

Wust, S., Van Rossum, E.F., Federenko, I.S., Koper, J.W., Kumsta, R., Hellhammer, D.H., 2004. Common polymorphisms in the glucocorticoid receptor gene are associated with adrenocortical responses to psychosocial stress. J. Clin. Endocrinol. Metab. 89, 565–573.