

High-intensity stress elicits robust cortisol increases, and impairs working memory and visuo-spatial declarative memory in Special Forces candidates: A field experiment

JOHN TAVERNIERS^{1,2}, JORIS VAN RUYSEVELDT¹, TOM SMEETS³, & JASPER VON GRUMBKOW¹

¹Faculty of Psychology, Open University of The Netherlands, Heerlen, The Netherlands, ²Department of Behavioural Sciences, Royal Military Academy, Brussels, Belgium, and ³Faculty of Psychology and Neuroscience, Maastricht University, Maastricht, The Netherlands

(Received 16 July 2009; revised 18 January 2010; accepted 21 January 2010)

Abstract

While running a selection procedure, 27 male Belgian Special Forces candidates, with a mean age of 27.4 years (SD = 5.1), were randomly assigned to a no-stress control ($n = 14$) or a high-intensity stress group ($n = 13$). Participants in the latter group were exposed to an extremely strenuous mock prisoner of war (POW) exercise. Immediately after stress or control treatment, working memory and visuo-spatial declarative memory performances were measured by the digit span (DS) test and the Rey–Osterrieth complex figure (ROCF), respectively. Concurrently, stress levels were assessed by obtaining salivary cortisol measurements and subjectively by the NASA Task Load Index (TLX). As expected, exposure to high-intensity stress led to both robust cortisol increases and significant differences in TLX scores. Stress induction also significantly impaired DS and ROCF performances. Moreover, delta cortisol increases and ROCF performance in the POW stress group showed a significant negative correlation, while DS performances followed the same tendency. Summarizing, the current findings complement and extend previous work on hormonal stress effects, and the subsequent performance deterioration on two memory tests in a unique high-intensity stress environment.

Keywords: *Digit span test, glucocorticoids, high-intensity stress, NASA TLX, Rey–Osterrieth complex figure*

Introduction

Trial Registration: Dutch Trial Register NTR 1518

Empirical research on intense, real-life, and uncontrollable stress is scarce (Dickerson and Kemeny 2004). While practical and/or ethical issues commonly restrict stress research, ongoing special forces (SF) selection programs provide ideal venues for this type of research in healthy humans. Past research on military populations has provided consistent evidence that psychological stress elicits robust endocrinological alterations, and leads to reduced performance in diverse outcome measures (Morgan et al. 2000, 2001, 2002, 2004, 2006). However, to the authors' knowledge, stress hormones after intense, real-life, and uncontrollable stress have never been linked to memory deficits assessed with

multiple, highly standardized neuropsychological measures.

The phenomenon “stress” has been referred to as a reaction to tangible or mentally evoked threats to the bodily homeostasis (Levine 2005), and is known to trigger a number of psychophysiological reactions. Many of these are related to the stress-responsive hypothalamic–pituitary–adrenal (HPA) axis. Although its exact mechanisms are not yet entirely clear, the HPA axis seems to respond to stress according to a dose–response relationship (Joëls et al. 2006; Kudielka et al. 2009). Glucocorticoid (GC) regulation (primary cortisol in humans) is a protective HPA reaction that mobilizes energy for coping with the stressor. Increased GC secretion also shuts down the initial fight or flight response of the sympathetic nervous and immune systems to prevent them from overshooting and damaging vital bodily functions (Munck 2000).

Ample research has demonstrated that, besides providing protection, the secretion of GCs modulates learning and memory. Although the precise direction is still somewhat ill-understood, both animal studies (McGauch and Roozendaal 2002; Roozendaal et al. 2006; Touma et al. 2008) and human research (Het et al. 2005; Smeets et al. 2008; Wolf 2009) indicate that GC secretion during acute stressful events may have facilitating as well as disruptive effects on cognitive performance. These dual effects have led to suggestions of an inverted U-curve to characterize the relationship between learning and GC secretion (Abercrombie et al. 2003; Andreano and Cahill 2006). Animal research, primarily in rodents, attributes this relationship essentially to divergent corticosteroid affinities of mineralocorticoid receptors (MRs; with a high affinity for cortisol) and glucocorticoid receptors (GRs; with a significantly lower affinity for cortisol) in the brain (McEwen and Sapolsky 1995). While memory facilitation seems to occur when MRs are fully and GRs only partially occupied, it is only when GRs become fully saturated that a decline in learning is observed (Abercrombie et al. 2003).

GC effects have been frequently linked to declarative memory, involving the amygdala, the hippocampus, and associated brain regions (Joëls et al. 2006; Shin et al. 2006; Wolf 2009). Recent findings, however, demonstrate that stress not only affects declarative memory, but also influences working memory (WM; Oei et al. 2006; Schoofs et al. 2008, 2009). Brain areas involved in WM functioning are considered to be, although not completely without controversy (Andres 2003), in general prefrontal and parietal brain structures (Lezak et al. 2004; Müller and Knight 2006). A meta-analysis on brain imaging studies by Cabeza and Nyberg (2000) revealed that WM is almost always associated with increased activity in the prefrontal cortex (PFC). More specifically, the left dorsolateral PFC is activated for verbal WM tests and the right for spatial WM. Recent functional magnetic resonance imaging (fMRI) research confirmed these findings, as acute stress, elicited by viewing extremely violent movie clips, reduced the WM-related dorsolateral PFC activity and reallocated neural resources away from executive function networks (Qin et al. 2009).

The influences of moderate stress levels on cognitive functioning have been extensively researched under laboratory conditions (Dickerson and Kemeny 2004; Chida and Hamer 2008). However, a meta-analysis by Dickerson and Kemeny (2004, p. 359) found that the endogenously evoked right-hand side of the inverted U-curve, including the effects of realistic high-intensity stress on memory performance, “[...] has been remarkably scarcely investigated in empirical research.” Prior research mostly applied exogenously administered cortisol or sought indirect clinical evidence from cortisol-related disorders (such as Cushing’s

disease) to investigate the right-hand side of the inverted U-curve (Mauri et al. 1993; Starkman et al. 2001).

An important exception is the work by Morgan and colleagues. In a decade of research examining the impact of intense stress in special military populations, Morgan et al. (2000) found strong (serum and salivary) cortisol concentration increases in extremely stressful military exercises. Follow-up studies, under similar circumstances, extended these findings and found potent stimulation of secretion of several hormones (plasma cortisol, catecholamines, and neuropeptide Y), as well as evidence of dissociation and relationships with diminished (military) performance (Morgan et al. 2001, 2002, 2004). In more recent work, Morgan et al. (2006) examined the effects of acute stress during US Army “Survival, Evasion, Resistance, and Escape” (SERE 1985) training. Interestingly, the authors included the Rey–Osterrieth complex figure (ROCF; Knight and Kaplan 2003), a renowned test for WM and visuo-spatial abilities, and showed that SERE training yielded dissociative symptoms as well as significant deficits in cognitive performance.

The current study

Cognitive capacity is generally assessed through tasks that require participants to store and process increasing amounts of information until the point at which recall errors are made (Lezak et al. 2004; Alloway 2006). For the current study, standardization and scientific validity were, evidently, the most important test prerequisites. Next, the ease of administration had to promote uniformity and experimental control in field conditions. Finally, the sensitive SF environment obliged the research team to introduce tests with a high face validity, of a short duration, and that could be visually administered—given the Belgian bilingualism. The digit span (DS) test (representing map coordinates or radio frequencies) for WM (Wechsler 1987; Lezak et al. 2004) and the ROCF (simulating memorization and reporting of an unknown target), as a test for visuo-spatial capacity within the declarative memory (Knight and Kaplan 2003; Shin et al. 2006), were chosen because they addressed all constraints mentioned above.

Although an extensive elaboration on the method of stress induction exceeds the scope of this article, a brief description is warranted to appreciate the data. The experiment was embedded within the assessment center phase of the Belgian SF selection procedure that consisted of a series of pre-designed physically, mentally, and cognitively demanding SF job samples (Schmidt et al. 2006). One of the job samples was an individually executed combined exercise (CombEx) that contained the main ingredients of a relatively brief, albeit intense SF mission. More specifically, CombEx simulated a mock prisoner of war (POW)

situation. The exercise started with a brusque capture that unexpectedly interrupted a bogus façade exercise. Conceptually, CombEx has similarities with the SERE (1985) program and incorporates related British experiences. Irrespective of the similarities, CombEx was developed to exploit complete unexpectedness and uncontrollability to assess candidates' shock resistance, resilience, and subsequent capacity for reengagement (e.g. the physical, mental, and cognitive ability to restart a mission under extremely stressful circumstances). Since the HPA axis is known to be particularly sensitive to unpredictability and uncontrollability (Dickerson and Kemeny 2004), the set-up provided an excellent opportunity to examine how high-intensity stress affects regulation of cortisol secretion and cognitive performance in healthy humans.

Given that (1) the effects of stress, particularly those of high-intensity stress on memory and learning, have received little empirical attention (Dickerson and Kemeny 2004), (2) traditional laboratory stressors are not known to elicit the neurobiological and cognitive alterations that stress in extreme military training situations has provoked (e.g. Morgan et al. 2000, 2001, 2006), and (3) memory capacity under high naturalistic stress loads has been largely neglected in the literature, the current field experiment was framed within an extremely strenuous phase of a SF selection program. Consistent with the dose-response theorem (Het et al. 2005; Joëls et al. 2006), robust cortisol increases were expected for the high-intensity stress (vs. no-stress control) group. Given that, under the conditions used, endogenously instigated stimulation of a high level of cortisol secretion was assumed to have the potential to fully saturate GRs in the PFC (Abercrombie et al. 2003), a significant decline in both WM and visuo-spatial declarative memory performance under high-intensity stress was hypothesized.

Methods

Participants

Participants were 27 healthy, physically fit males with normal body mass index. Ages ranged from 21 to 37 years ($M = 27.4$ yrs, $SD = 5.1$). All were active duty Belgian Armed Forces who volunteered as SF candidates. Participants were preliminarily medically tested (e.g. for cardiovascular deficiencies), and on location they were assessed for endocrine disorders and the use of medication. Given reported use of medication, data from one participant were excluded from further analyses.

Measures and materials

Saliva sampling and cortisol analyses. Salivary cortisol samples were collected by cotton roll devices (Salivette[®]; Sarstedt, Etten-Leur, The Netherlands)

and stored at -20°C immediately after collection. Subsequently, the samples were thawed, numbered, and centrifuged at 3000 rpm, 4°C for 5 min at the Dresdner Technical University LabServices. Salivary free cortisol was analyzed using a commercial chemiluminescence immunoassay (IBL Hamburg, Germany). Samples from same subjects were analyzed in a single run to reduce error variance. Inter- and intra-assay coefficients of variation were below 10%.

Task load index. Subjective allostatic task (over)load was measured as a proxy of stress by the NASA Task Load Index (TLX) (Hart and Staveland 1988). The TLX is a multidimensional rating scale, designed to obtain estimates from participants while performing a task or shortly afterward. TLX combines information about the magnitude of six task load-related subscales; mental demands, physical demands, time demands, own performance, effort, and frustration. The scale has been extensively applied in research and is considered a highly sensitive assessment technique (Rubio et al. 2004). Total scores were obtained by summing raw scores of the six subscales, presented as Likert scales ranging from 0 to 20. High scores represented high stress levels.

Digital span test. The DS paradigm has been linked to performance on intelligence tests and problem solving (Sternberg 2003), and is often used as a free-standing performance measure in WM research (Axelrod et al. 2006). A computerized version of the standard digital span (DS) test was used as a measure of WM (Wechsler 1987). Testing procedures were consistent with the digits forward (DF) and digits backward (DB) paradigm. Participants were presented with 4–9 digits for immediate recall (one by one, at a rate of one per second). As prescribed for healthy participants (Lezak et al. 2004), total scores were obtained by combining forward and backward component scores.

Rey–Osterrieth complex figure. The ROCF is one of the most widely used neuropsychological tests for the evaluation of non-verbal learning, planning, and WM. In the immediate recall paradigm, the figure permits the assessment of visuo-spatial abilities within declarative memory (Knight and Kaplan 2003; Shin et al. 2006). A computerized version of the ROCF was presented in black-on-white for 45 s and with a size-on-screen of 12×8 cm, according to the intentional, single trial, immediate recall protocol (Knight 2003a). Considering that low variability of the ROCF copy scores has been reported in healthy subjects (Shin et al. 2006), the copy phase was excluded to avoid ceiling effects. Participants subsequently had 3 min for immediate recall. The ROCF was scored

double-blind, applying the Denman scoring system (DnSS). DnSS was chosen for its explicit focus on learning and memory capacities, elaborate narrative scoring definitions, increased individual point ceiling (maximal score = 72 points), and high inter-scorer reliability (Knight 2003a).

Control measures. The degree to which the HPA axis is activated during stressful events can show considerable individual variation depending upon character issues and life history (Kudielka et al. 2009). In the CombEx situation, two control measures for individual differences were thought to be of interest. First, the impact of event scale-revised (IES-R) (Weiss and Marmar 1997) measured the impact of formerly experienced traumatic stress. The IES-R is probably the most widely used self-report measure in the field of traumatic stress impact and potential risk of post-traumatic stress disorder (PTSD). The scale was chosen for its psychometric qualities as well as for its sensitivity to detect lower symptom levels (Creamer et al. 2003). The 22 items of the IES-R encompass three subscales (intrusions, hyperarousal, and avoidance) that parallel the DSM-IV criteria for PTSD. Participants were instructed to complete the scale while recalling the event in their lives that had most impact. High scores represented high impact of formerly experienced traumas. Answering possibilities ranged from 0 (not at all) to 4 (extremely). The scale's internal consistency (Cronbach's α) was 0.96 (Creamer et al. 2003).

Second, personality hardiness (PH) was measured as a cognitive style that is of importance to inherently stressful environments since it conceptualizes the individual's capacity to cope with stress (Maddi 2007). PH is characterized by perceptions on three sub-dimensions: commitment, control, and challenge (Maddi et al. 2006). Higher levels of PH have been shown to provide a natural advantage and are associated with increased outcome performance in both civilian and military stress research, notably with SF candidates (Eid and Morgan 2006; Bartone et al. 2008). PH was assessed by the revised 15-item dispositional resilience scale (DRS15-R), a cross-culturally improved version of the DRS15. High scores represented a hardy PH style. Answering possibilities ranged from 0 (not at all true) to 4 (completely true) with an internal consistency (Cronbach's α) of 0.82 (Bartone 2007).

Procedure

Registration and testing protocols were submitted to the Dutch Trial Register (NTR 1518) and approved by the standing ethics committee of the Open University, The Netherlands.

Screening and preparation. Participants voluntarily engaged in the official SF selection process. Approximately 1 month prior to the selection phase, all received a thorough medical screening that required the most stringent medical profile within the Belgian Armed Forces, including tests for hormonal and cardiovascular deficiencies. Next, candidates had to pass a psychological test battery that included the Neuroticism, Extraversion, and Openness Personality Inventory-Revised (NEO PI-R; Costa and McCrae 1992) and the generalized cognitive test battery, a standard Belgian Armed Forces cognitive ability test (Irvine 2006). The remaining candidates were invited into the SF selection week and signed a written informed consent at entry. Next, candidates were instructed to remove all external identification marks (grades, brevets, and identity tags) and received a chest number to maximize selection objectivity. Candidates then enrolled in the assessment center phase. In the course of the first day, they collectively completed the IES-R and DRS15-R scales, and orally reconfirmed their informed consent during a semi-structured intake interview.

Experimental procedure. Participants were randomly assigned to the no-stress control group ($n = 14$) or the stress group ($n = 13$) for a between-subjects field experiment. Participants were tested individually and, to control for circadian cortisol rhythms, all tests fell between 19.00 and 22.00 h. Participants were deprived of food, drinks, smoking, and heavy physical exercise at least 90 min prior to the cortisol measurements. They were not deprived of sleep the night before and all performed exactly the same daytime activities.

Participants were intentionally kept unaware of the selection program. Shortly after the cortisol baseline (C_0) measurement, those assigned to the no-stress control condition ran a 60 min non-intensive filler task (completing administration and a non-stressful weapon handling task), while the stress group was exposed to the 60 min CombEx. Behind the façade of a bogus non-stressful activity, the latter were brusquely captured and physically constrained (10 min, transportation included), confined (30 min), and subjected to a concise mock POW interrogation (15 min). Transportation toward the cognitive test battery and test instructions took about 5 min. Individual timings were meticulously registered and all fell within the 60 ± 5 min interval.

Computer-administered testing procedures were identical for both conditions. Two cortisol samples were collected during the memory tests, one at onset (C_1 at T_0) and one 15 min later (C_2 at $T + 15$). The time point for the start of the memory tests and, concurrently, the first cortisol measurement, 60 min after the start and about 15 min after termination

of CombEx, was based on previous findings that indicated that these periods coincide with robust increases of stress hormones (Dickerson and Kemeny 2004; Joëls et al. 2006; Morgan et al. 2006). After CombEx, participants were instructed to reflect on their respective stress phases and to complete the TLX. Finally, all passed an obligatory clinical debriefing. Figure 1 shows the experimental set-up for both test conditions.

Statistical analyses

The effect of high-intensity stress (TLX) was analyzed by an independent sample *t*-test (two tailed) and cortisol concentrations were analyzed with a mixed model analysis of variance (ANOVA) with Time (C_0 , C_1 , C_2) as repeated measure factor and Group (no-stress control vs. stress) as between-group factor. Greenhouse–Geisser corrected *p*-values are reported when appropriate. Next, for each participant individual peak cortisol responses (ΔC) were computed. ΔC was defined as: $\Delta C = C_{\text{Peak}} - C_0$, and analyzed by two-tailed independent samples *t*-tests. For two participants in the no-stress control group, ΔC could not be calculated due to missing data. An additional area under curve (AUC) computation, with respect to cortisol concentration increase ($\text{AUC}_{\text{I-Cort}}$; Preussner et al. 2003), was performed as a single measure of the total cortisol response. Potential differences in $\text{AUC}_{\text{I-Cort}}$ were investigated using a two-tailed independent *t*-test. Memory performance on both the DS and the ROCF tests were analyzed by separate two-tailed independent samples *t*-tests (no-stress control vs. stress). As an indication of the DnSS's inter-rater reliability, a *post hoc* Pearson correlation of 0.97 was found between the immediate (for selection purposes) and the experimental double-blind ROCF scores. IES-R and PH were analyzed by separate two-tailed independent samples *t*-tests. For all tests, alpha was set at 0.05.

Results

There was no difference between IES-R [$(M_{\text{no-stress}} = 9.79, \text{SE} = 2.52)$; $(M_{\text{stress}} = 12.38, \text{SE} = 1.72)$; $(t(25) = -0.84; p = 0.41)$ and PH [$(M_{\text{no-stress}} = 36.21, \text{SE} = 0.80)$; $(M_{\text{stress}} = 35.30, \text{SE} = 0.73)$; $(t(25) = 0.83; p = 0.41)$] scores. Although there were no maximum scores on the PH measure, all participants scored above the 85th percentile. Cortisol data were log transformed (Ln) due to excessive skewness. Unless explicitly indicated, all further cortisol-related analyses were conducted with log-transformed data. Salivary cortisol pre-stress concentrations (C_0) did not differ between no-stress control and stress groups [$t(25) = 0.36; p = 0.72$] (Figure 2).

After stress induction, the *t*-test on TLX scores yielded a significant difference between the no-stress control ($M = 9.86, \text{SE} = 1.28$) and stress ($M = 58.62, \text{SE} = 2.59$) groups [$t(17.6) = -16.9; p < 0.001$]. Noteworthy differences between both groups for the physical and time demands TLX-subscales were non-significant. This signaled the exclusive mental-psychological nature of the stress exposure. The 2 (Group; no-stress, stress) \times 3 (Time; C_0, C_1, C_2) mixed model ANOVA also yielded a significant Group \times Time interaction effect [Wilks' $\lambda = 0.28, F(2,24) = 30.99; p < 0.001$, partial $\eta^2 = 0.72$], a significant effect of Time [Wilks' $\lambda = 0.43, F(2,24) = 15.82; p < 0.001$, partial $\eta^2 = 0.57$], and a significant main between subjects effect of Group [$F(2,24) = 50.32; p < 0.001$, partial $\eta^2 = 0.69$]. *Post hoc* tests revealed no differences between the cortisol concentrations for the no-stress control group over time. In the stress group, however, salivary cortisol concentrations showed significant differences between C_0 and both C_1 ($M_{\text{Difference}} = -21.16, \text{SE} = 2.70, p < 0.001$) and C_2 ($M_{\text{Difference}} = -24.74, \text{SE} = 3.16, p < 0.001$), while C_1 and C_2 did not differ significantly from each other ($p = 0.07$). The results of these analyses are shown in Figure 2. For information purposes only,

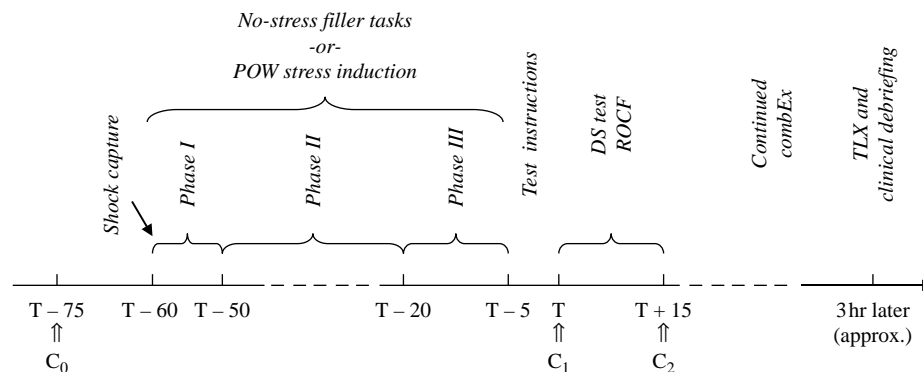


Figure 1. Experimental time frame (minutes) for both groups (no-stress vs. stress), including: baseline saliva sampling (C_0 at T - 75), filler tasks or POW stress induction (Phases I–III), test instructions, performance tests (DS test and ROCF), cortisol saliva sampling (C_1 at T and C_2 at T + 15), completion of NASA TLX, and clinical debriefing. Timings are in minutes, unless stated otherwise.

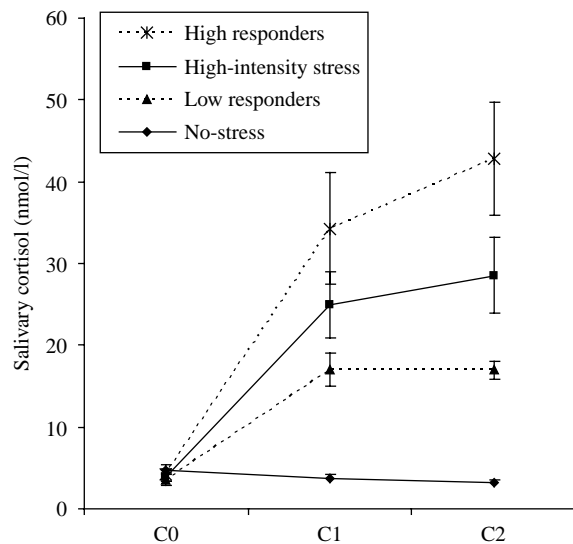


Figure 2. Salivary cortisol concentrations (means \pm SEM) for the no-stress control ($n = 14$) and the stress group ($n = 13$) (C_0 at T-75 min, C_1 at T, and C_2 at T + 15 min). Exclusively for information purposes, the figure includes differences in mean salivary cortisol increment for high ($n = 6$) vs. low cortisol responders ($n = 7$) in the stress group (cut-off by median split).

Figure 2 represents differences in mean salivary cortisol increment in the stress group for high vs. low cortisol responders, cut-off by median split (after Smeets et al. 2006a). Due to the restricted number of participants in the stress group, differences between high vs. low cortisol responders were not included in further analyses.

The Pearson correlation between subjective (TLX) and objective (increases in salivary cortisol concentrations, ΔC) stress indicators was statistically significant, $r(25) = 0.89$, $p < 0.01$. ΔC also differed significantly between groups in the independent samples t -test [$t(25) = -10.17$; $p < 0.001$], with untransformed means of -0.81 nmol/l (SE = 0.59) for the no-stress and 27.01 nmol/l (SE = 4.39) for the stress group. For the stress group, mean salivary cortisol increases of over 600% were observed. As expected, individual differences in cortisol responses were high, with a maximum increase of over 11-fold. Similarly, the no-stress control and the stress group differed significantly in terms of AUC_{LCort} [$t(25) = -9.96$; $p < 0.001$].

Mean DS scores of no-stress control and stress groups differed significantly [$t(25) = 2.20$; $p = 0.038$] (Figure 3(a)). Specifically, mean performance scores for these groups in the DB paradigm ($M = 5.64$, SE = 0.51 vs. $M = 4.08$, SE = 0.45) also differed significantly [$t(25) = 2.30$; $p = 0.03$], while mean scores for the DF paradigm ($M = 5.71$, SE = 0.64 vs. $M = 4.07$, SE = 0.61) were (marginally) not significantly different [$t(25) = 1.84$; $p = 0.078$]. Similarly, mean ROCF scores of the no-stress control and stress [$t(25) = 2.83$; $p = 0.009$] groups differed

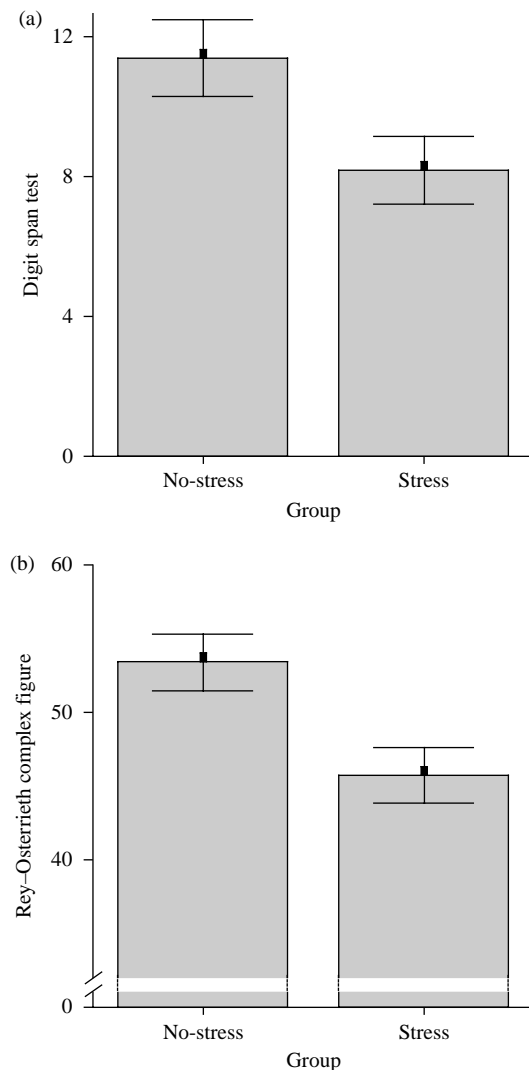


Figure 3. Representation of the mean performance scores (\pm SEM) for the no-stress control ($n = 14$) and the stress group ($n = 13$) on the memory performance tasks; the DS test for WM (a) and the ROCF for visuo-spatial declarative memory (b). In (a) and (b) the No-stress and Stress group means are significantly different from each other ($p < 0.05$, $p < 0.01$).

significantly (Figure 3(b)). Within the stress group, correlation analyses addressed the relationship between ΔC , DS, and ROCF scores, respectively. ΔC and ROCF scores were found to be significantly correlated, $r(13) = -0.78$, $p \leq 0.01$; thus, confirming the expected relationships, while ΔC and DS data showed only a tendency [$r(13) = -0.48$, $p \leq 0.09$]. Follow-up analyses using distribution-free rank correlations confirmed these findings between ΔC and ROCF scores [Kendall's τ -b: $r(13) = -0.59$, $p \leq 0.01$; Spearman's ρ : $r(13) = -0.79$, $p \leq 0.01$], and between ΔC and DS scores [Kendall's τ -b: $r(13) = -0.47$, $p \leq 0.05$; Spearman's ρ : $r(13) = -0.61$, $p \leq 0.05$]. Comparably, ROCF but not DS performance correlated negatively with AUC_{LCort} within the stress group [$r(13) = -0.67$,

$p \leq 0.01$ and $r(13) = -0.42$, $p \leq 0.15$, respectively]. Follow-up analyses using distribution-free rank correlations showed significance for the relation between AUC_{I_Cort} and DS scores [Kendall's τ -b: $r(13) = -0.39$, $p \leq 0.07$; Spearman's ρ : $r(13) = -0.51$, $p \leq 0.08$], and between AUC_{I_Cort} and ROCF scores [Kendall's τ -b: $r(13) = -0.56$, $p \leq 0.01$; Spearman's ρ : $r(13) = -0.75$, $p \leq 0.01$].

To illustrate the differences in visuo-spatial declarative memory between the groups, Figures 4(a) (no stress control group) and (b) (stress group) show representative examples of ROCF performance. An original version of the ROCF can be consulted in Knight and Kaplan (2003). While all participants performed within the expected, standardized ranges for healthy persons (Knight 2003b), the differences in performance under high-intensity stress seem clear.

Discussion

The current field experiment aimed to improve insight into the dynamics of endogenously instigated regulation of high levels of cortisol, and to measure its possible effects on WM and visuo-spatial declarative memory performance. As expected, robust increases

in salivary cortisol concentrations were observed in the stress group. The results further indicated that both groups performed within acceptable ranges for both performances tests (Sternberg 2003; Knight 2003b). In identical test conditions, however, exposure to high-intensity stress led to a significant decrease in performance in both the DS test and ROCF. Whereas laboratory studies have generally shown low correlations between self-reported stress and cortisol measures (Dickerson and Kemeny 2004), exposure to high-intensity stress yielded significant correlations, first, between self-reported and salivary cortisol stress measures, and, second, between cortisol increase and performance outcomes.

Cortisol regulation

As expected, high-intensity stress provoked robust increases in salivary cortisol concentration, which largely exceeded those evoked by traditional laboratory stressors (Het et al. 2005). Mean cortisol rises of over 600%, with a maximum of over 11-fold were observed. These findings contribute to the HPA axis dose-response theorem in a realistic and unconventional process of stress induction. The current study has obvious contextual similarities with the work of Morgan and colleagues (2000, 2001, 2002, 2004, 2006). Conversion from *Système International* to conventional units revealed that the salivary cortisol concentrations were relatively well in line with hormonal data of Morgan et al. (2000, 2001, 2002). The latter measured cortisol, and other psychophysiological markers, in comparable circumstances, and registered similar or slightly more elevated mean cortisol increases, dependent upon the specific stressor and the investigated population. Apart from the work of Morgan et al., there are, to the authors' knowledge, no other reports of comparable methods of stress induction.

Robust increases of salivary cortisol are also consistent with the notion that novelty, unexpectedness, and genuine uncontrollability create larger effects (Dickerson and Kemeny 2004). Moreover, the selection context and the omnipresent threat of negative evaluation (Dickerson et al. 2008) most likely added to the overall stress effect; failure in the CombEx would substantially impede participants' progress to the personal goals they had singled out. On the other hand, a similar evaluation threat was present for both groups given that there was no attitudinal difference in approach by the evaluators. Hence, it can be reasonably expected that deviations in salivary cortisol concentrations due to a different evaluative threat perception can be excluded.

Finally, objective (salivary cortisol increase) and subjective (TLX self-reports) stress measures correlated significantly. This finding is contradictory to results from laboratory studies that have generally

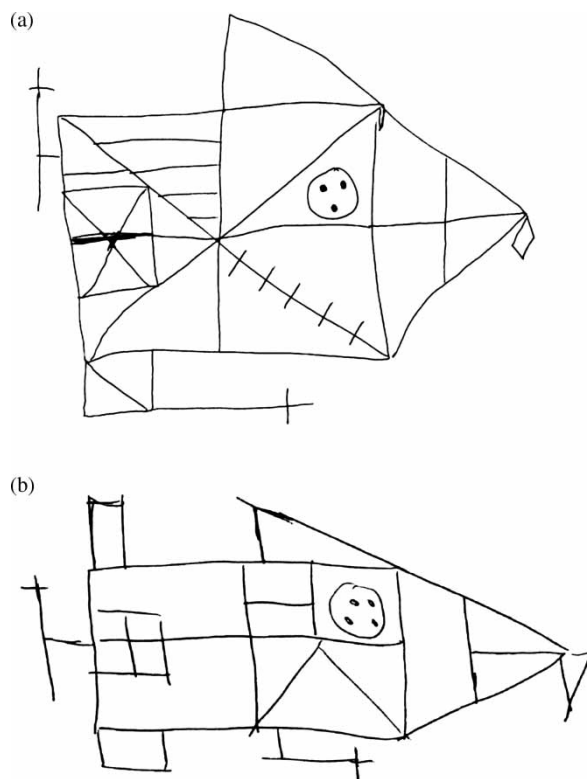


Figure 4. Two representative performance examples of the ROCF for the no-stress control (a) vs. the stress (b) group. While the original ROCF strongly resembles Figure 4(a), a full view of the original ROCF can be consulted in Knight and Kaplan (2003) (*The handbook of ROCF usage: Clinical and research applications*. Psychological Assessment Resources, Inc., Lutz, FL, USA).

shown low correlations between self-reports of experienced stress and cortisol measures (Dickerson and Kemeny 2004). It is, however, unsure whether this correlation can be explained by the shock effect, the realism, the uncontrollability (except under condition of jeopardizing the entire selection phase), or the salience of the situation for the participants. Yet another explanation might be that, in contrast to previous studies that used subjective stress measures that were more general in nature (e.g. overall mood questionnaires), the TLX is a more sensitive subjective measure for this particular high-intensity stress situation. Future research could unravel the respective roles of the different stressors and address the sensitivity of the TLX.

Learning and memory

Exposure to high-intensity stress led to a significant decrease in both memory outcomes. These results are in line with other studies that applied the DS test as a performance measure for WM. There are, however, some controversies concerning the DS test, more specifically on the inclusion of the DF paradigm. Some researchers formulated *a priori* reservations in regard to the DF's sensitivity as a measure, particularly for healthy young adults (Banken 1985; Smeets et al. 2006b; Unsworth and Engle 2007; Schoofs et al. 2008, 2009). Here, the standardized DS test (Wechsler 1987) proved to be sensitive enough to produce significant results for combined DF–DB scores. This result accords well with Lezak et al.'s (2004) recommendation to treat both paradigms as if they measure the same or highly correlated behaviors in normal control subjects. However, it must be highlighted that, even under the current study's stress conditions, there are indications of a reduced sensitivity of the DF paradigm in healthy young adults (Schoofs et al. 2009).

The stress-induced decrease in ROCF recall scores indicates a significant impairment of visuo-spatial declarative memory capacities (Shin et al. 2006). In relation to the overall ROCF immediate recall scores, this result is well in line with the quantitative results of Morgan et al. (2006) who measured ROCF performance in a SERE (1985) context[†]. Interestingly, the differences in ROCF results found in the current study were essentially quantitative, as all participants managed to incorporate the broad *gestalt* overview of the figure but failed to accurately process specific ROCF details (note that the results are not completely comparable, given that Morgan et al. included a copy phase in their ROCF application). Morgan et al. (2006) reported organizational deficits and piecemeal drawing strategies during the copy phase of their stress group, and found significant negative correlations between symptoms of dissociation and memory performance. A possible explanation for these

deviating results could be that the population of the current study reported no or very little previously experienced trauma, while an important subset of the earlier study had encountered “war zone experiences” or “life-threatening events”. Previously experienced trauma is known to have a pervasive influence during subsequent stressful events on GC regulation, and cognitive performance (Yehuda et al. 2001; Kudielka et al. 2009). A second explanation could be that, given CombEx's much shorter time frame (vs. SERE), no dissociations were evoked (three previous CombEx applications never signaled any sign of dissociations). Interestingly, observations during task completion and inquiries after the clinical debriefing revealed that nearly all participants performed some form of mental “repetition” or “seeing” of the running digits or ROCF; thus, giving support to Baddeley's (2002) suggested WM model, incorporating a *phonological loop* and *visuo-spatial sketch pad*.

The decline of memory performance on the right-hand side of the inverted U-curve, between learning and GC secretion, has been attributed to divergent corticosteroid affinities of MRs (with a high affinity for cortisol) and GRs (with a significantly lower affinity for cortisol) in the brain (McEwen and Sapolsky 1995; De Kloet et al. 1999). Accordingly, it is assumed that the high-intensity stress fully saturated GRs, which led to the decrease in WM performance. This assumption is in line with observations that the PFC plays an important role in WM functioning (e.g. Cabeza and Nyberg 2000), and that acute stress reallocates neural resources away from executive brain networks (Qin et al. 2009). Evidence from animal and human research confirmed the presence of high concentrations of GRs in the PFC (Perlman et al. 2007; Patel et al. 2008), which could influence higher order cognitive functions via elevated levels of cortisol (possibly combined with supra-optimal levels of catecholamines; Qin et al. 2009). In relation to the visuo-spatial declarative memory performance, it is assumed that the amygdala, the hippocampus, and associated brain regions (Joëls et al. 2006; Shin et al. 2006; Wolf 2009) were affected by excessive cortisol concentrations. The finding that participants scored well within the expected ranges can be explained by their overall hardy cognitive style—all scored above the 85th percentile on the PH control measure, combined with the absence of past traumatic experiences. High scores on PH are known to improve individual's capacity to cope with stressful events and enable them to perform better under stressful circumstances (Maddi 2007).

The current findings may have *practical implications* that exceed the narrow scope of operational SF applications. First, appropriate cognitive functioning in extreme, non-routine circumstances is vital in the many so-called high reliability professions that are characterized by high demands and fatal outcomes

in case of failure (Flin 2001). Fire fighters, emergency medics, police officers, pilots, and civilian disaster relief workers may encounter extreme, non-routine circumstances that require the best knowledge available to warrant their appropriate cognitive functioning (Bartone et al. 2008). On the other hand, it must be noted that because of the stringent inclusion criteria for Belgian SF operators, the current results, although indicative, cannot be automatically generalized to situations in which randomly chosen lay people experience a sudden traumatic experience. Second, the findings can have direct practical implications for selection procedures that aim at the early identification of individuals with increased risk for cognitive decline or who lack inherent coping abilities to reengage under stress; for designers of interfaces that enable operators to split attention when extreme levels of allostatic load can be reasonably anticipated (Wickens et al. 2004; Yuan et al. 2006); for inoculation training or *in vivo* habituation to repeated gradual stress exposures (Meichenbaum 2007); and for cognitive-behavioral interventions to improve coping abilities, and regulate cortisol modulation (Hammerfald et al. 2006).

This study, evidently, has limitations. First, its conclusions are restricted to the right-hand side of the inverted U-curve between stress and memory, which makes comparisons with moderate stress levels difficult. Next, it would be both desirable and informative to add performance tasks with repeated exposures of learning materials. In line with Lezak et al. (2004, p. 465), it is thought that this would upgrade the study's findings to a more sophisticated level of cognition, "[...] permit emergence of a learning curve.", and transcend the level of memory functioning to effective learning. Also, replications on larger and more heterogeneous (e.g. gender diversification) samples are necessary. The restricted sample size, however, is thought to be inherent to the strenuousness of the application, which is mirrored by the limited influx of Belgian SF candidates (with, so far, no female candidates). Follow-up research should address these constraints, for example, through application of counterbalanced within-subjects designs, improved learning materials, and more balanced male-female samples. Also, the issue concerning the reduced sensitivity of the DF paradigm in healthy young adults needs further attention. Despite these limitations, the current field experiment's major strengths lay in the assessment of two standardized neuropsychological outcome measures in an inimitable, though well-controlled real-world setting. It is thought that these strengths support the study's ecological validity and the conclusions that can be drawn from it.

Acknowledgements

The authors would like to express their sincere gratitude to the research team (Wouter Huybens, Ilse Provoost,

Olivier Serniclaes, and Maarten Wendelen) for their valued assistance; to Jef Syroit, Susan van Hooren, and three anonymous experts for their appreciated comments on an earlier version of this article; and to the Belgian military authorities for their permission to conduct the research.

Declaration of interest: The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Note

[†]Morgan et al. (2006) applied the Taylor scoring system (TSS). For comparison purposes, as a (rough) guide, the TSS and the DnSS relate to each other as: $TSS = 2/3 \times DnSS$. This relationship, however, is only indicative, given the deviating design approaches of both scoring systems (see Knight 2003a). Interested researchers are cordially invited to contact the first author to request copies of the current actual ROCF outcomes.

References

- Abercrombie HC, Kalin NH, Thurow ME, Rosenkranz MA, Davidson RJ. 2003. Cortisol variation in humans affects memory for emotionally laden and neutral information. *Behav Neurosci* 117:505–516.
- Alloway TP. 2006. How does working memory work in the classroom? *Educ Res Rev* 1:134–139.
- Andreano JM, Cahill L. 2006. Glucocorticoid release and memory consolidation in men and women. *Psychol Sci* 17:466–470.
- Andres P. 2003. Frontal cortex as the central executive of working memory: Time to revise our view. *Cortex* 39:871–895.
- Axelrod B, Fichtenberg N, Millis S, Wertheimer J. 2006. Detecting incomplete effort with digit span from the Wechsler Adult Intelligence Scale – Third Edition. *Clin Neuropsychol* 20: 513–523.
- Baddeley AD. 2002. Is working memory still working? *Eur Psychol* 7:85–97.
- Banken JA. 1985. Clinical utility of considering digits forward and digits backward as separate components of the Wechsler Adult Intelligence Scale-Revised. *J Clin Psychol* 41:686–691.
- Bartone P. 2007. Test-retest reliability of the dispositional resilience scale-15, a brief hardiness scale. *Psychol Rep* 101:943–944.
- Bartone P, Roland R, Picano J, Williams T. 2008. Psychological hardiness predicts success in US Army Special Forces candidates. *Int J Sel Assess* 16:78–81.
- Cabeza R, Nyberg L. 2000. Imaging cognition II: An empirical review of 275 PET and fMRI studies. *J Cogn Neurosci* 12:1–47.
- Chida Y, Hamer M. 2008. Chronic psychosocial factors and acute physiological responses to laboratory-induced stress in healthy populations: A quantitative review of 30 years of investigations. *Psychol Bull* 134:829–885.
- Costa PT, McCrae RR. 1992. Revised NEO Personality Inventory (NEO-PI-R) and the Five Factor Inventory (NEO-FFI): Professional manual. Odessa, FL: Psychological Assessment Resources Inc.
- Creamer M, Bell R, Failla S. 2003. Psychometric properties of the impact of event scale-revised. *Behav Res Ther* 41:1489–1496.
- De Kloet R, Oitzl M, Joëls M. 1999. Stress and cognition: Are corticosteroids good or bad guys? *Trends Neurosci* 22:422–426.
- Dickerson SS, Kemeny ME. 2004. Acute stressors and cortisol responses: A theoretical integration and synthesis of laboratory research. *Psychol Bull* 130:355–391.
- Dickerson SS, Mycek PJ, Zaldivar F. 2008. Negative social evaluation, but not mere social presence, elicits cortisol responses to a laboratory stressor task. *Health Psychol* 27: 116–121.

- Eid J, Morgan CA, III. 2006. Dissociation, hardiness, and performance in military cadets participating in survival training. *Mil Med* 171(5):436–442.
- Flin R. 2001. Selecting the right stuff: Personality and high-reliability occupations. In: Roberts BW, Hogan R, editors. *Personality psychology in the workplace*. Washington, DC: American Psychology Association. p 253–276.
- Hammerfeld K, Eberle C, Grau M, Kinsperger A, Zimmermann A, Ehler U, Gaab J. 2006. Persistent effects of cognitive-behavioral stress management on cortisol responses to acute stress in healthy subjects: A randomized controlled trial. *Psychoneuroendocrinology* 31:333–339.
- Hart SG, Staveland LE. 1988. Development of a multi-dimensional workload rating scale: Results of empirical and theoretical research. In: Hancock PA, Meshkati N, editors. *Human mental workload*. Amsterdam: Elsevier Science. p 139–183.
- Het S, Ramlow G, Wolf OT. 2005. A meta-analytic review of the effects of acute cortisol administration on human memory. *Psychoneuroendocrinology* 30:771–784.
- Irvine S. 2006. The Irvine multiform series: A psychometric reference manual, Unpublished manual, Belgian Armed Forces Experimental Tests Services.
- Joëls M, Pu Z, Wiegert O, Oitzl M, Krugers H. 2006. Learning under stress: How does it work? *Trends Cogn Sci* 10:152–158.
- Knight J, Kaplan E. 2003. The handbook of Rey–Osterrieth complex figure usage: Clinical and research applications. Lutz, FL: Psychological Assessment Resources, Inc.
- Knight J. 2003a. Administration procedures and scoring systems. In: Knight JA, Kaplan E, editors. *The handbook of Rey–Osterrieth complex figure usage: Clinical and research applications*. Lutz, FL: Psychological Assessment Resources, Inc. p 57–192.
- Knight J. 2003b. ROFC psychometric characteristics and normative data. In: Knight JA, Kaplan E, editors. *The handbook of Rey–Osterrieth complex figure usage: Clinical and research applications*. Lutz, FL: Psychological Assessment Resources, Inc. p 193–254.
- Kudielka BM, Hellhammer DH, Wüst S. 2009. Why do we respond so differently? Reviewing determinants of human salivary cortisol responses to challenge. *Psychoneuroendocrinology* 34:2–18.
- Levine S. 2005. Stress: An historical perspective. In: Steckler T, Kalin NH, Reul JM, editors. *Handbook of stress and the brain*. New York: Elsevier Science. p 3–23.
- Lezak M, Howieson D, Loring D. 2004. *Neuropsychological assessment*. 4th ed. New York: Oxford University Press.
- Maddi S. 2007. Relevance of hardiness assessment and training to the military context. *Mil Psychol* 19:61–70.
- Maddi S, Harvey R, Koshaba D, Lu J, Persico M, Brow M. 2006. The personality construct of hardiness, III: Relationships with repression, innovativeness, and performance. *Pers* 74:575–597.
- Mauri M, Sinforiani E, Bono G, Vignati F, Berselli ME, Attanasio R, Nappi G. 1993. Memory impairment in Cushing's disease. *Acta Neurol Scand* 87:52–55.
- McEwen B, Sapolsky R. 1995. Stress and cognitive function. *Curr Opin Neurobiol* 5:205–216.
- McGaugh JL, Roozendaal B. 2002. Role of adrenal stress hormones in forming lasting memories in the brain. *Curr Opin Neurobiol* 12:205–210.
- Meichenbaum D. 2007. Stress inoculation training: A preventative and treatment approach. In: Lehrer RLWPM, Sime WS, editors. *Principles and practice of stress management*. London: Guilford Press. p 497–518.
- Morgan CA, III, Wang S, Mason J, Southwick SM, Fox P, Hazlett G, Charney DS, Greenfield G. 2000. Hormone profiles in humans experiencing military survival training. *Biol Psychiatry* 47: 891–901.
- Morgan CA, III, Wang S, Rasmusson A, Hazlett G, Anderson G, Charney DS. 2001. Relationship among plasma cortisol, catecholamines, neuropeptide Y, and human performance during exposure to uncontrollable stress. *Psychosom Med* 63: 412–422.
- Morgan CA, III, Rasmusson A, Wang S, Hoyt G, Hauger RL, Hazlett G. 2002. Neuropeptide-Y, cortisol, and subjective distress in humans exposed to acute stress: Replication and extension of previous report. *Biol Psychiatry* 52:136–142.
- Morgan CA, III, Southwick S, Hazlett G, Rasmusson A, Hoyt G, Zimolo Z, Charney D. 2004. Relationships among plasma dehydroepiandrosterone sulfate and cortisol levels, symptoms of dissociation, and objective performance in humans exposed to acute stress. *Arch Gen Psychiatry* 61(8):819–825.
- Morgan CA, III, Doran A, Steffian G, Hazlett G, Southwick M. 2006. Stress-induced deficits in working memory and visuo-constructive abilities in special operations soldiers. *Biol Psychiatry* 60:722–729.
- Müller NG, Knight RT. 2006. The functional neuroanatomy of working memory: Contributions of human brain lesion studies. *Neuroscience* 139:51–58.
- Munck A. 2000. Corticosteroids and stress. In: Finck G, editor. *Encyclopedia of stress*. New York: Academic Press. p 570–577.
- Oei NY, Everaerd WT, Elzinga BM, Van Well S, Bermond B. 2006. Psychosocial stress impairs working memory at high loads: An association with cortisol levels and memory retrieval. *Stress* 9: 133–141.
- Patel PD, Katz M, Karssen AM, Lyons DM. 2008. Stress-induced changes in corticosteroid receptor expression in primate hippocampus and prefrontal cortex. *Psychoneuroendocrinology* 33:360–367.
- Perlman WR, Webster MJ, Herman MM, Kleinman JE, Weickert CS. 2007. Age-related differences in glucocorticoid receptor mRNA levels in the human brain. *Neurobiol Aging* 28: 447–458.
- Preussner JC, Kirchbaum K, Meinschmid G, Hellhammer DK. 2003. Two formulas for computation of the area under the curve represent measures of total hormone concentration versus time-dependent change. *Psychoneuroendocrinology* 28:916–931.
- Qin S, Hermans EJ, Van Marle H, Luo J, Fernández G. 2009. Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biol Psychiatry* 66:25–32.
- Roozendaal B, Hui GK, Hui IR, Berlau DJ, McGaugh JL, Weinberger NM. 2006. Basolateral amygdala noradrenergic activity mediates corticosterone-induced enhancement of auditory fear conditioning. *Neurobiol Learn Mem* 86:249–255.
- Rubio S, Diaz E, Martín J, Puente JM. 2004. Evaluation of subjective mental workload: A comparison of SWAT, NASA-TLX, and workload profile methods. *J Appl Psychol* 53:61–86.
- SERE. 1985. Code of conduct, survival, evasion, resistance, and escape (SERE) training. Retrieved June 15, 2009 from <http://www.fas.org/irp/doddir/army/ar350-30.pdf>
- Schmidt FL, Greenthal AL, Hunter JE, Berner JB, Seaton FW. 2006. Job sample vs. paper-and-pencil trades and technical tests: Adverse impact and examinee attitudes. *Pers Psychol* 30:187–197.
- Schoofs D, Preuß D, Wolf OT. 2008. Psychosocial stress induces working memory impairments in an n-back paradigm. *Psychoneuroendocrinology* 33:643–653.
- Schoofs D, Wolf OT, Smeets T. 2009. Cold pressor stress impairs performance in working memory tasks requiring executive functions in healthy young men. *Behav Neurosci* 123: 1066–1075.
- Shin M, Park S, Park S, Seol S, Kwon J, Park S, Seol S, Kwon J. 2006. Clinical and empirical applications of the Rey–Osterrieth complex figure test. *Nat Protocols* 1:892–899.
- Smeets T, Jelicic M, Merkelbach H, Peters M, Fett A, Taverniers J, Henquet C, Dautzenberg J. 2006a. Enhanced memory performance on an internal-internal source monitoring test following acute psychosocial stress. *Behav Neurosci* 120: 1204–1210.

- Smeets T, Jelicic M, Merkelbach H. 2006b. The effect of acute stress on memory depends on word valence. *Int J Psychophysiol* 62:30–37.
- Smeets T, Otgaar H, Candel I, Wolf OT. 2008. True or false? Memory is differentially affected by stress-induced cortisol elevations and sympathetic activity at consolidation and retrieval. *Psychoneuroendocrinology* 33:1378–1386.
- Starkman M, Giordani B, Berent S, Schork A, Scheingart D. 2001. Elevated cortisol levels in Cushing's disease are associated with cognitive decrements. *Psychosom Med* 63:985–993.
- Sternberg R. 2003. *Cognitive psychology*. 3rd ed. Belmont: Wadsworth/Thomson Learning.
- Touma C, Bunck M, Glasl L, Nussbaumer M, Palme R, Stein H, Wolferstätter M, Zeh R, Zimbelmann M, Holsboer F, Landgraf R. 2008. Mice selected for high versus low stress reactivity: A new animal model for affective disorders. *Psychoneuroendocrinology* 33:839–862.
- Unsworth N, Engle RW. 2007. On the division of short-term and working memory: An examination of simple and complex span and their relation to higher order abilities. *Psychol Bull* 133: 1038–1066.
- Wechsler D. 1987. *WMS-R Manual: Wechsler Memory Scale-Revised*. New York: Harcourt Brace Jovanovich, Inc., The Psychological Corporation.
- Weiss D, Marmar C. 1997. The Impact of Event Scale-Revised. In: Wilson J, Keane T, editors. *Assessing psychological trauma and PTSD*. New York: Guildford. p 399–411.
- Wickens CD, Lee J, Liu Y, Becker SG. 2004. *Stress and workload. Introduction to human factors engineering*. Upper Saddle River: Pearson Education. p 324–350.
- Wolf OT. 2009. Stress and memory in humans: Twelve years of progress? *Brain Res* 1293:142–154.
- Yehuda R, Halligan S, Grossman R. 2001. Childhood trauma and risk for PTSD: Relationship to intergenerational effects of trauma, parental PTSD and cortisol excretion. *Dev Psychopathol* 13:733–753.
- Yuan K, Steedle J, Shavelson R, Alonzo A, Oppezzo M. 2006. Working memory, fluid intelligence, and science learning. *Educ Res Rev* 1:83–98.