# Supplementary Methods

## Power Analysis

To calculate the minimum sample size needed for results with sufficient power, an a priori power analysis was conducted in G\*Power 3.1 (Faul, Erdfelder, Buchner, & Lang, 2009; Faul, Erdfelder, Lang, & Buchner, 2007) for a repeated-measures ANOVA with given significance threshold α = .05 and power (1 – β) = .95. Even though our study employed a within-subjects design, the expected effect size was based on reports of a between-subjects study by Schoofs et al. (2013) for the effect of stress on *n*-back performance in male subjects (medium effect size, *d* = .59). Hence, the assumed effect size might represent an over-estimation of the underlying population parameter. The power analysis indicated a sample size of at least *n* = 26 participants needed to reject the null hypothesis of no differences between the stress and control condition on the differences between WM loads. Therefore, the total sample of 34 participants allowed for a dropout of up to 8 participants from analyses while maintaining enough statistical power. Note that this sample size is also comparable to the number of participants measured by Qin et al. (2009) for the effect of stress on the neural mediation of WM.

## Procedure details

Within 7 days prior to the first intervention sessions, participants came to the institute for a verbal intelligence assessment (Wortschatztest, Schmidt & Metzler, 1992) and a high-resolution structural scan, which was used for coregistration purposes in the current study.

Upon arrival, participants were taken to a preparation room, rested for 10 minutes, then filled in baseline mood questionnaires and practiced the task on a computer outside the MRI scanning room before undergoing either intervention. After the intervention, participants were led to the MRI scanner and performed the task.

### TSST and control condition

During the TSST, participants were instructed to imagine applying for their dream job. They anticipated a job interview for 5 minutes while having the possibility to prepare for it by making notes. Afterwards, participants gave a free speech (5 minutes, without notes) explaining why they would be suitable for the job. This was done in front of a committee wearing white laboratory coats and acting in a neutral non-emotional manner, which was designed to alleviate stress responses due to the absence of non-verbal (social) feedback from the committee. The TSST committee consisted of one actor and one experienced psychology student, both of whom were introduced as psychologists specializing in the analysis of non-verbal behavior. After the free speech, participants were asked to perform an evaluated mental arithmetic task (5 minutes). During this task they had to serially subtract the number 17 starting at 2043 as fast and accurately as possible out loud. Upon failure, participants had to restart at 2043. During the interview and mental arithmetic task, participants stood in front of a microphone and a video camera and were supposedly video- and audiorecorded. However, neither device was turned on. After the stress induction procedure, participants were taken to the MRI scanner. At the end of the experimental session containing the stress condition, participants were debriefed by a psychologist about the aims of the stress intervention. They were told that the stress committee consisted of specially trained actors that were instructed to act in a reserved and neutral manner during the interview. Participants were told that they had not been video- or audio-recorded and most importantly, that the interview would not be relevant to the rest of the study. They were encouraged to talk to the experimenter about their experiences and feelings during the stress intervention. This debriefing aimed to relieve participants from the experienced stress and prevent stress effects caused by anticipation of another stress intervention, if the non-stress control condition was the second test session. During the control condition participants were instructed that they would read a piece of text and could relax. They anticipated reading for 5 minutes, then read a neutral, non-fiction text about the Mesozoic era for 10 minutes. After 10 minutes of reading, the experimenter led the participants to the MRI scanning area.

## Past subjective stress assessment

The PSS-10 comprises 10 items and assesses answers on a 5-point Likert-like scale (0 = never, 1 = almost never, 2 = sometimes, 3 = fairly often, 4 = very often). Items 4, 5, 7, and 8 are inversely coded. After reversing these items, all items are summed up for a total score between 0 and 40, with higher total scores indicating more past subjective stress. Cronbach’s α of the PSS-10 ranges from α = .84 to .86 (Cohen et al., 1983). The PSS-10 displays medium correlations (range: *r* = .52 – 76) to similar constructs, like depression (Cohen et al., 1983).

Two participants in the sample used for WM analyses (*n* = 32) failed to fill out the trait questionnaires. Therefore, all analyses on past subjective stress measurements were performed in a sample of 30 participants.

## Working Memory Assessment: *n*-back Task

The task was programmed using MATLAB 2008b (T**he MathWorks Inc., Natick, MA, USA) using the** Psychophysics Toolbox (Brainard, 1997). A stream of white single-digit numbers (0 – 9) on a black background was projected onto a screen behind the scanner, which was visible with a mirror placed above the MRI head coil. Participants had to indicate with their dominant hand via button presses on an MRI compatible button box whether the presented stimulus matched the stimulus presented in the *n* steps from earlier. The task consisted of two WM loads, *0-Back* and *2-Back*. During 0-Back*,* participants had to respond whenever the digit currently shown on the screen was the number zero. During 2-Back, participants had to indicate if the number presented on the screen was equal to the number shown two digits before. Numbers in both loads were presented for 500 milliseconds (ms) with an inter-trial interval of 900 ms. Both WM loads consisted of six blocks of stimulation. In each block 22 stimuli were presented in random order, each containing three target stimuli and a rest period at the end, during which participants were instructed to fixate on a centered white fixation cross. Each block was introduced by an instruction cue (0-back or 2-back) which was visible for 2 seconds before every block. The 0-Back and 2-Back blocks alternated, and the starting block was counterbalanced between test sessions across subjects. Total task duration was 10 minutes and 40 seconds. Before each test session, participants received instructions and training on the task (one block per condition), including a post-training teach-back.

## Physiological Stress Response

Salivary cortisol was collected at six time points using Salivette saliva sampling tubes (SalivetteCortisol®, Sarstedt, Nuembrecht, Germany). Time points relative to the start of intervention (stress or control) were: t1: -30 minutes; t2: -2 minutes; t3: +10 minutes; t4: +15 minutes; t5: +30 minutes; t6: +45 minutes. Participants chewed each Salivette sample for approximately 2 minutes to ensure maximum collection of saliva. After collection, saliva samples were stored at −80 °C before being sent for analysis (Biochemical Laboratory, University of Trier, FB 1, Clinical and Physiological Psychology). Analysis was performed with a competitive solid phase time-resolved fluorescence immunoassay with fluorometric end point detection (DELFIA), as described in more detail by Dressendörfer and colleagues (1992). The intra-assay variation coefficient ranged from 4.0% to 6.7%, and the inter-assay variation coefficients ranged from 7.1% to 9.0% (Dressendörfer et al., 1992).

## fMRI Acquisition

fMRI data was acquired using a 3 Tesla high-resolution Siemens PRISMA MR-system (Siemens, Erlangen, Germany), equipped with a 20-channel head coil. Blood oxygenation level dependent (BOLD) signals were acquired using a T2\*-weighted echo-planar imaging (EPI) sequence (TR=2000 ms, TE=30 ms, flip angle=90°, matrix=192x192 mm, voxel size=3x3x5 mm). We collected 30 slices covering the whole brain, approximately parallel to the anterior–posterior commissure plane. In total 319 volumes were collected per test session. After the task, field maps were acquired for estimation of voxel-displacement maps during preprocessing (TR=488 ms, TE=4.92 ms, flip angle=60°, matrix=192x192 mm). Additionally, on a separate day within seven days before the first test session, three-dimensional T1-weighted images (TR=5000 ms, TE=2.03 ms, FoV=256x240x176, voxel size=1x1x1mm) were obtained using a magnetization-prepared two rapid acquisition gradient echoes (MP2RAGE) sequence with a 32-channel head coil. This scan was used as anatomical reference to the EPI data during the coregistration procedure.

fMRI data were preprocessed and analyzed using SPM12 (http://www.fil.ion.ucl.ac.uk/spm/). The first 5 volumes of each functional time series were discarded. Before preprocessing, the origins of the functional imaging series were reoriented to the anterior–posterior commissure plane in native space. Preprocessing included slicetiming, realignment, coregistration, and warping to Montreal Neurological Institute (MNI) space. The obtained normalization parameters were applied to the realigned images, which were resliced with a voxel size of 3x3x4 mm. All images were smoothed with a Gaussian kernel of 6 mm full width at half-maximum (FWHM).

fMRI data of three participants from the n = 32 behavioral sample had to be excluded from further analyses due to corrupted data leaving a sample of 29 subjects for fMRI analyses (see Table 1 for sub-sample characteristics). Including the behavioral outliers in the analysis did not change the reported results.

# Supplementary Results

## Stress Response During Acute Stress

Average time points of saliva sampling and visual analog scale assessments are depicted in Table S1. Of the 34 participants, 6 (18%) showed a cortisol increase below the set threshold of 1.5 nmol/l and were therefore considered non-responders. In the remaining 28 responders, average cortisol levels during the stress session were significantly higher than in the control session (*t*27=5.69, *p*<.001, *d*=1.28). Pre-stress cortisol (t1) was not significantly different during the stress and the control condition (*t*27=.78, *p*=.44, *d*=.19). The rmANOVA on cortisol response showed significant main effects of condition and time. Also, a significant interaction effect of condition and time was observed (see Table S2 and Figure 1). Mauchly’s test indicated a violation of the sphericity assumption for the factor time (χ²14=108.41, *p*<.001), and the interaction of condition and time (χ²14= 86.22, *p*<.001). Therefore, the degrees of freedom of the effects were corrected with Greenhouse-Geisser estimates of sphericity (ε=.45/.48, respectively). Due to an error, VAS data of t5 during the control condition were not assessed for one participant, leaving a sample of 33 participants, for analysis of the visual analog scales. In this sample, subjective stress levels at the start of the experiment (t1) were not significantly different during the stress and the control condition (*t*32=.024, *p*=.98, *d*=.004, see Figure 1), ruling out possible confounding effects of baseline differences between conditions. An rmANOVA on subjective stress responses showed significant main effects of condition and time. Also, a significant interaction effect of condition and time was observed (see Table S2 and Figure 1). Mauchly’s test indicated a violation of the sphericity assumption for the VAS (arousal and valence) on the factor time (arousal: χ²14=39.74, *p*<.001, valence: χ²14=52.92, *p*<.001). The degrees of freedom of the analyses were thus corrected with Greenhouse-Geisser estimates of sphericity (ε=.61/.55, respectively). Mauchly’s test for the interaction of condition and time also indicated a violation of sphericity (arousal: χ²14=28.82, *p*=.011, valence: χ²14=40.36, *p*<.001). The degrees of freedom of the analyses were therefore corrected with Greenhouse-Geisser estimates of sphericity (ε=.70/.63, respectively). Arousal and valence showed similar results to subjective stress: Average arousal levels in the stress condition were significantly higher than in the control condition, (*t*32=4.58, *p<*.001, *d*=.80), valence levels in the stress condition were significantly lower than in the control condition (*t*32=3.60, *p<*.001, *d*=.63).

Post-stress cortisol levels at the average peak (t4) and the post-stress subjective stress level at the average peak (t3) in the stress condition were significantly higher (each Bonferroni corrected for 6 paired-sample t-tests) than at the same time point in the control condition (cortisol: *t*21=7.65, *p*<.001, *d*=1.87; subjective stress: *t*32=8.37, *p*<.001, *d*=1.47). Importantly, also shortly before the *n*-back task was performed (t5), cortisol levels and subjective stress levels in the stress condition were still significantly elevated (Bonferroni corrected) compared with t5 in the control condition (cortisol: *t*21=7.89, *p*<.001, *d*=1.88; subjective stress: *t*32=2.39, *p*=.023, *d*=.42). This effect remained stable until t6 for cortisol (*t*21=7.71, *p*<.001, *d*=1.97).

## Association of subjective and physiological stress responses with

## working memory performance

The rmANCOVAs, including the control-normalized cortisol AUC and the control-normalized subjective stress AUC as covariates, delivered similar results to the rmANOVA without the covariates. Additionally, there were no significant main effects of the two covariates (cortisol AUC: d’: *F*1, 28=.48, *p*=.49, ηp2=.02; RT: *F*1, 28=.39, *p*=.54, ηp2=.014, subjective stress AUC: d’: *F*1, 28=.85, *p*=.37, ηp2=.03; RT: *F*1, 28=.61, *p*=.44, ηp2=.021).

The interaction effects of cortisol and condition (d’: *F*1, 28=.24, *p*=.63, ηp2=.01; RT: *F*1, 28=3.76, *p*=.06, ηp2=.12), cortisol and WM load (d’: *F*1, 28=.43, *p*=.52, ηp2=.02; RT: *F*1, 28=1.75, *p*=.20, ηp2=.06), and the three-way interaction of cortisol, condition and WM load (d’: *F*1, 28=.34, *p*=.56, ηp2=.01; RT: *F*1, 28=.04, *p* = .85, ηp2= .001) were not significant.

Also, the interaction effects of subjective stress and condition (d’: *F*1, 28=.1.29, *p*=.27, ηp2=.044; RT: *F*1, 28=.40, *p*=.54, ηp2=.01), subjective stress and WM load (d’: *F*1, 28=.13, *p*=.72, ηp2=.005; RT: *F*1, 28=.87, *p*=.36, ηp2=.03), and the three-way interaction of subjective stress, condition and WM load (d’: *F*1, 28=.72, *p*=.40, ηp2=.03; RT: *F*1, 28=.09, *p*=.77, ηp2=.003) did not show significant effects.

Potentially, our approach of calculating a control-normalized AUCg could have resulted in different, confounded individual cortisol stress-responsivity than if measured with the AUCg from the stress condition only. If the control-normalized AUCg and the stress-only AUCg represented overlapping measurements, we would expect a positive correlation between the two variables. Indeed, we observed a highly significant correlation between the control-normalized AUC and the stress-only AUC, *r*30=.79, p<.001, indicating that both variables overlap. Consistent with this, when re-running the reported ANCOVAs including the stress-only cortisol AUC, the previous (non-significant) results did not change: The main effect of the stress-only AUC (d’: *F*1, 29=.29, *p*=.59, ηp2=.01; RT: *F*1, 29=.27, *p*=.61, ηp2=.01) and all two- and three-way interaction effects involving the stress-only AUC (d’: all *F*s<.1, *p*s>.75, ηp2s<.01; RT: all *F*s<1.67, *p*s>.21, ηp2s<.06) were not significant. Also, the association between the change scores for d’, RT and the stress-only AUC was not significant (d’: *r*30=-.097, *p*=.60; RT: *r*30=-.11, *p*=.55).

## Working Memory fMRI Results

The interaction contrast of condition and WM load indicated no clusters that were correctable for the whole brain. Subthreshold neural activation (at *puncorr.*<.005) was seen close to the left putamen (*F*1, 83=23.17, *p*uncorr.<.001), the left anterior insula (*F*1, 83=23.17, *p*uncorr.<.001), the brain stem (*F*1, 83=21.02, *p*uncorr. < .001), left cerebellum (*F*1, 83=19.37, *p*uncorr.=.001), left middle temporal gyrus (*F*1, 83=12.74, *p*uncorr.=.001), and in two clusters in the right dlPFC (*F*1, 83=19.65, *p*uncorr.<.001 and *F*1, 83=12.62, *p*uncorr.=.001).

# Supplementary Figures

Figure S1. Multi-slice axial view (MNI coordinates: z = -52 to 58) of whole-brain corrected (pFWE < .05) activation patterns from the t-contrast [2-back > 0-back] across both conditions (stress and control). Patterns represent higher BOLD signal in the 2-Back working memory load compared to the 0-Back working memory load in n = 29 participants. Results are overlaid on the SPM template for a single subject in MNI space. The color bar represents t-values.

# Supplementary Tables

Table S1. Average time points (minutes, seconds) at which saliva samples and visual analog scale of subjective stress reactivity was assessed during control and stress condition

|  |  |  |
| --- | --- | --- |
| **Time point** | **Control Condition**  ***M* (*SD*)** | **Stress Condition**  ***M* (*SD*)** |
| **t1** | -32.27 (1.11) | -38.39 (2.25) |
| **t2** | -4.35 (.01) | -4.37 (.02) |
| **t3** | 10.51 (.05) | 11.01 (.11) |
| **t4** | 14.56 (.08) | 14.49 (.21) |
| **t5** | 40.00 (.29) | 38.37 (.48) |
| **Start *n*-back** | 46.46 (1.10) | 46.02 (.41) |
| **t6** | 59.58 (.20) | 58.25 (.11) |

*Note.* Means and standard deviations (in parentheses) for the time points (minutes, seconds) at which saliva samples and visual analog scales were assessed are illustrated above. All values indicate minutes and seconds relative to the start of the respective intervention (control/stress). Abbreviations: *M* = Average, mean, *SD* = standard deviation.

Table S2.Univariate statistics: cortisol and subjective stress responses

|  |  |  |
| --- | --- | --- |
| **Measure** | **Effect** | **Statistics** |
| Cortisol | Condition | *F*1, 27=32.37, *p*<.001, ηp2=.55 |
| Time | *F*2.26, 60.96=17.83, *p*<.001, ηp2=.40 |
| Condition x Time | *F*2.38, 64.28=44.78, *p*<.001, ηp2=.62 |
|  |  |  |
| Subjective Stress | Condition | *F*1, 32=52.47, *p*<.001, ηp2=.62 |
| Time | *F*5, 160=19.67, *p*<.001, ηp2=.38 |
| Condition x Time | *F*5, 160=24.96, *p*<.001, ηp2=.21 |
|  |  |  |
| Subjective Arousal | Condition | *F*1, 32=20.97, *p*<.001, ηp2=.40 |
| Time | *F*3.04, 97.15=17.35, *p*<.001, ηp2=.35 |
| Condition x Time | *F*3.49, 111.77=18.08, *p*<.001, ηp2=.36 |
|  |  |  |
|  | Condition | *F*1, 32=12.95, *p*=.001, ηp2=.29 |
| Subjective Valence | Time | *F*2.75, 88.8=13.51, *p*<.001, ηp2=.30 |
|  | Condition x Time | *F*3.16, 100.96=8.65, *p*<.001, ηp2=.21 |

*Note.* The univariate statistical effects from two separate repeated-measures analyses of variances with condition (stress/control), time (t1 to t6) as within-subject factors and their interaction effect on cortisol in *n* = 28 cortisol responders, showing a cortisol response of > 1.5 nmol/l from pre- to post-stress during the stress condition, and on ratings of subjective stress, arousal and valence in *n* = 33 participants, who filled in all visual analog scales are illustrated above. ηp2= partial squared eta.

Table S3. fMRI results

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Main Effect of Working Memory Load (2-back vs 0-back)** | | | | | | | | | | | | | | | | | | | | | |
| Region | | Side | | Cluster size | | | | MNI - peak coordinate | | | | | *F value* | *puncorrected* | | *pFWE* | | | | | |
| corrected | | | | | |
|  | |  | |  | | | | x | y | | z | |  |  | |  | | | | | |
| dlPFC  Anterior Insula | | L/R | | 4472 | | | | -27 | -1 | | 54 | | 349.52 | < .001 | | < .001 | | | | | |
| -30 | 23 | | -2 | | 226.08 | < .001 | | < .001 | | | | | |
| 33 | 23 | | -2 | | 212.99 | < .001 | | < .001 | | | | | |
| PPC | | L/R | | 2223 | | | | -36 | 43 | | 38 | | 313.41 | < .001 | | < .001 | | | | | |
| 39 | -43 | | 42 | | 310.01 | < .001 | | < .001 | | | | | |
| Cerebellum | | L | | 20 | | | | -12 | -55 | | -46 | | 74.92 | < .001 | | < .001 | | | | | |
| R | | 993 | | | | 36 | -70 | | -26 | | 223.14 | < .001 | | < .001 | | | | | |
| Brain stem | | L | | 20 | | | | -6 | -31 | | -18 | | 43.42 | < .001 | | < .001 | | | | | |
| R | | 16 | | | | 6 | -28 | | -10 | | 60.46 | < .001 | | < .001 | | | | | |
| PCC | | L/R | | 654 | | | | 3 | -46 | | 30 | | 144.20 | < .001 | | < .001 | | | | | |
| VMPFC | | L/R | | 942 | | | | -6 | 56 | | 14 | | 133.14 | < .001 | | < .001 | | | | | |
| MTG | | L | | 85 | | | | -57 | -10 | | -22 | | 88.88 | < .001 | | < .001 | | | | | |
| Angular Gyrus | | L | | 73 | | | | -51 | -67 | | 34 | | 82.53 | < .001 | | < .001 | | | | | |
| R | | 17 | | | | 57 | -64 | | 22 | | 61.37 | < .001 | | < .001 | | | | | |
| Hippocampus | | L | | 93 | | | | -27 | -34 | | -14 | | 80.98 | < .001 | | < .001 | | | | | |
| Operculum | | R | | 451 | | | | 45 | -19 | | 22 | | 74.17 | < .001 | | < .001 | | | | | |
| **Main Effect of Acute Stress (Condition: Stress – Control)** | | | | | | | | | | | | | | | | | | | | | |
| Region | | Side | | Cluster size | | | | MNI - peak coordinate | | | | | *F value* | *puncorrected* | | *pFWE* | | | | | |
|
|  |  |  |  |  |  |  | x | | y | z | |  | | | | | |  |  |  |  | | |  |
| Putamen | | R | | 2 | | | | 24 | 11 | -6 | | 29.56 | | | < .001 | | .020 | | | | | |
| **Interaction of Acute Stress (Condition: Stress – Control) and Working Memory Load** | | | | | | | | | | | | | | | | | | | | | |
| Region | | Side | | Cluster size | | | | MNI - peak coordinate | | | | | *F value* | *puncorrected* | | *pFWE* | | | | | |
|
|  |  |  |  |  |  |  | x | | y | | z |  | | | | | |  |  |  |  | | |  |
| dlPFC | | R | | 15 | | | | 27 | 59 | 26 | | 19.65 | | | < .001 | | .026  (FWE-SVC) | | | | | |

*Note.* Uncorrected, whole-brain as well as small-volume corrected fMRI results (F-values, cluster sizes, side) from the main effect of working memory load, main effect of condition and their interaction effect from a repeated-measures analysis of variances in *n* = 29 participants are illustrated above. Abbreviations: dlPFC = dorsolateral prefrontal cortex, fMRI = functional magnetic resonance imaging, FWE = family-wise error correction, FWE-SVC = family-wise error correction for small-volumes, L = left, MNI = Montreal Neurological Institute, MTG = middle temporal gyrus, PCC = posterior cingulate cortex/gyrus, R = right, VMPFC = ventromedial prefrontal cortex

# Supplementary References

Brainard, D. H. (1997). The Psychophysics Toolbox. *Spatial Vision*, *10*, 433–436.

Cohen, S., Kamarck, T., & Mermelstein, R. (1983). A Global Measure of Perceived Stress. *Journal of Health and Social Behavior*, *24*(4), 385–396. http://doi.org/10.2307/2136404

Dressendörfer, R. A., Kirschbaum, C., Rohde, W., Stahl, F., & Strasburger, C. J. (1992). Synthesis of a cortisol-biotin conjugate and evaluation as a tracer in an immunoassay for salivary cortisol measurement. *Journal of Steroid Biochemistry and Molecular Biology*, *43*(7), 683–692. http://doi.org/10.1016/0960-0760(92)90294-S

Faul, F., Erdfelder, E., Buchner, A., & Lang, A.-G. (2009). Statistical power analyses using G\*Power 3.1: tests for correlation and regression analyses. *Behavior Research Methods*, *41*(4), 1149–60. http://doi.org/10.3758/BRM.41.4.1149

Faul, F., Erdfelder, E., Lang, A. G., & Buchner, A. (2007). G\*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, *39*(2), 175–191. http://doi.org/10.3758/bf03193146

Qin, S., Hermans, E. J., van Marle, H. J. F., Luo, J., & Fernández, G. (2009). Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. *Biological Psychiatry*, *66*(1), 25–32. http://doi.org/10.1016/j.biopsych.2009.03.006

Schmidt, K.-H., & Metzler, P. (1992). *Wortschatztest (WST)*. Weinheim: Beltz Test GmbH.

Schoofs, D., Pabst, S., Brand, M., & Wolf, O. T. (2013). Working memory is differentially affected by stress in men and women. *Behavioural Brain Research*, *241*, 144–153. http://doi.org/10.1016/j.bbr.2012.12.004