Brain Mechanisms of Social Threat Effects on Working Memory

V. A. van Ast[1,†], J. Spicer[2,†], E. E. Smith[2,‡], S. Schmer-Galunder[2], I. Liberzon[3], J. L. Abelson[3] and T. D. Wager[4]

[1]Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands,
[2]Department of Psychology, Columbia University, New York, NY 10027, USA, [3]Department of Psychiatry, University of Michigan,
Ann Arbor, MI 48109, USA and [4]Department of Psychology and Neuroscience, University of Colorado, Boulder, 345 UCB, Boulder,
CO 80305, USA

Address correspondence to Tor D. Wager, Department of Psychology and Neuroscience, University of Colorado, Boulder, 345 UCB, Boulder,
CO 80305, USA. Email: tor.wager@colorado.edu
[†]V. A. van Ast and J. Spicer contributed equally to this work.
[‡]Deceased.

Social threat can have adverse effects on cognitive performance, but the brain mechanisms underlying its effects are poorly understood. We investigated the effects of social evaluative threat on working memory (WM), a core component of many important cognitive capabilities. Social threat impaired WM performance during an N-back task and produced widespread reductions in activation in lateral prefrontal cortex and intraparietal sulcus (IPS), among other regions. In addition, activity in frontal and parietal regions predicted WM performance, and mediation analyses identified regions in the bilateral IPS that mediated the performance-impairing effects of social threat. Social threat also decreased connectivity between the IPS and dorsolateral prefrontal cortex, while increasing connectivity between the IPS and the ventromedial prefrontal cortex, a region strongly implicated in the generation of autonomic and emotional responses. Finally, cortisol response to the stressor did not mediate WM impairment but was rather associated with protective effects. These results provide a basis for understanding interactions between social and cognitive processes at a neural systems level.

Keywords: executive function, fMRI, intraparietal sulcus, mediation, stress

Introduction

The performance of complex mental tasks is essential to modern life. With sufficient attentional focus, humans can maintain and manipulate information in ways that enable symbolic reasoning and problem solving. But, all too often, the desire to perform optimally impairs performance just when it is most critical—particularly under conditions of social evaluative threat (SET), a form of psychosocial stress involving evaluation and judgment by others. For instance, even minimal priming of social and racial stereotypes can induce "stereotype threat" and impair cognitive performance (Steele and Aronson 1995). Likewise, manipulations of negative social feedback can lead to suboptimal decision-making (Kassam et al. 2009). Conversely, resilience to stress leads to enhanced academic performance (Bardi et al. 2011), and interventions that bolster self-worth can help reduce the effects of stereotype threat (Cohen et al. 2009; Miyake et al. 2010). However, in spite of great interest in how SET and related social processes affect cognitive performance, the specific cognitive and brain mechanisms that mediate impairments are just beginning to be explored.

Social evaluative threat may impair cognitive performance by interfering with working memory (WM), a class of cognitive control processes central to many higher-level capabilities, including learning, problem-solving, and decision-making (Baddeley 2003). Working memory performance correlates strongly with general fluid intelligence (Gray et al. 2003) and is impaired in several psychiatric and neurological disorders (e.g., Barch and Smith 2008). Behaviorally, SET has been shown to impair WM (Elzinga and Roelofs 2005; Oei et al. 2006; Luethi et al. 2008; Schoofs et al. 2008) providing support for the idea that SET may impair cognitive performance through WM. However, findings have been somewhat inconsistent (Porcelli et al. 2008; Weerda et al. 2010; Cornelisse et al. 2011), possibly reflecting the fact that SET can engage both "challenge" and "threat" responses (Kassam et al. 2009) and shift cognitive processing from executive processing-heavy strategies to faster, more heuristic ones (Beilock et al. 2004; Schwabe and Wolf 2009). In addition, recent neuroimaging studies using WM tasks have reported both stress-induced increases (Porcelli et al. 2008; Weerda et al. 2010) and decreases (Qin et al. 2009) in lateral prefrontal cortex (lPFC) activity without finding any stress-induced alterations in WM performance. Thus, the brain mechanisms by which social threat influences cognitive performance remain unclear.

Here, we investigated whether changes in fronto–parietal regions related to WM "mediate" (Fig. 1A) stress effects on WM, including 1) whether SET influenced brain responses to an WM task (Path *a*); whether WM-related brain activity predicts WM performance (Path *b*); and whether these effects are jointly strong enough to mediate the SET–WM performance relationship (mediation effect *a*b*) (Baron and Kenny 1986; MacKinnon et al. 2007). We aimed at creating an ecologically valid SET challenge. We therefore used a social challenge modeled closely after the Trier Social Stress Test (TSST) (Kirschbaum et al. 1993) in combination with additional social evaluative elements during WM. Working memory was manipulated using a verbal N-back task (Jonides et al. 1998; Gray et al. 2003) (Fig. 1B) that reliably activates frontal and parietal cortical regions (Owen et al. 2005). We used Mediation Effect Parametric Mapping (MEPM)(Wager, Waugh, et al. 2009; Atlas et al. 2010) to obtain maps of *P*-values for all effects (using bootstrap resampling), focusing on frontal and intraparietal sulcus (IPS) regions consistently linked to WM and attentional control more generally (Corbetta and Shulman 2002; Wager et al. 2004; Lepsien et al. 2005; Owen et al. 2005; Mayer et al. 2007; Bledowski et al. 2009; Liston et al. 2009). We also focused on ventromedial prefrontal cortex (vmPFC), an area engaged by stress and threat (Pruessner et al. 2008; Wager, Van Ast, et al. 2009; Wager, Waugh, et al. 2009; Roy et al. 2012) that is commonly deactivated by cognitive tasks (Gusnard et al. 2001; Fox et al. 2005).
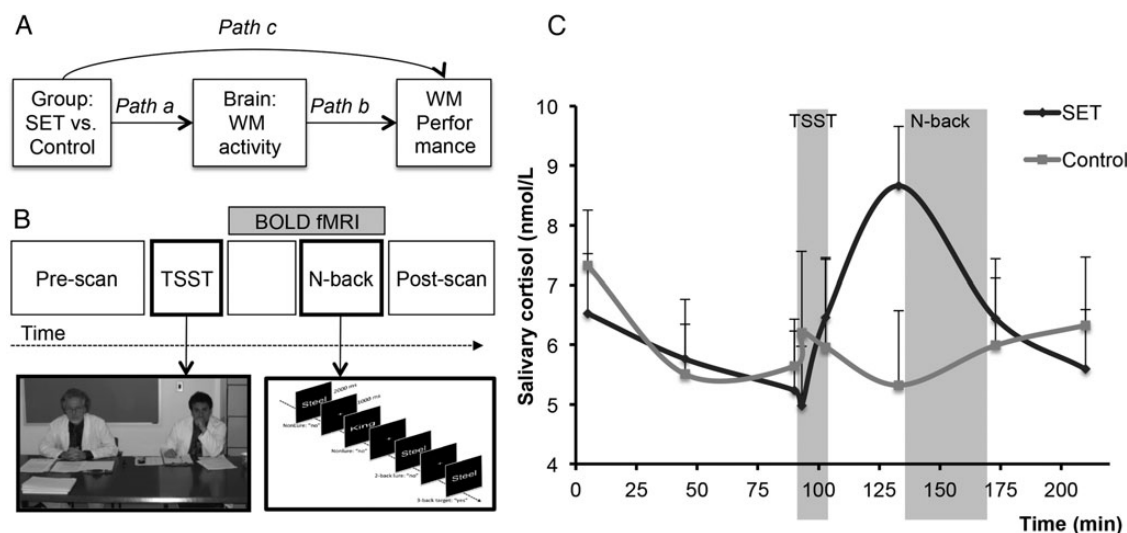
Figure 1. Mediation model, study design, and cortisol results. (A) The mediation model used to search for brain activity on a voxel-wise basis that formally explains the relationship between Group and working memory (WM) Performance. For convincing evidence that changes in fronto–parietal regions "mediate" stress effects on WM, 4 conditions should hold. First, stress should alter WM performance (Path c, the total effect). Second, stress should alter activity during WM in fronto–parietal regions (Path a). Third, fronto–parietal activity should predict WM performance, controlling for stress (Path b). And fourth, the stress effects on the brain and brain–WM performance relationships should be nonadditive, that is, the effect of brain activity should formally mediate the stress–WM relationship (Path a*b) (Baron and Kenny 1986; MacKinnon et al. 2007). (B) Schematic outline of the experiment. The scanning session commenced with the acquisition of structural and baseline functional scans. Next, the social evaluative threat (SET) group was administered the trier social stress test (TSST), and the Control group was administered the control task. Next, all participants returned to the scanner at which point they completed a divided attention task lasting 28 min followed by the N-back task. In the figure, a 3-back trial schematic is depicted. In the 3-back task, participants indicated whether each word matched the word seen 3 trials ago with a "yes" or "no" response. At the end of scanning, participants completed a final set of questionnaires. (C) Cortisol results. Mean free salivary cortisol concentration in nanomoles per liter for the SET and Control groups.

Methods

Participants

Twenty-one participants ($n = 21$; mean age = 22.24 [4.24] years) were recruited and randomized into 1 of 2 groups: SET ($n = 10$, 5 females) and control ($n = 11$, 5 females). Participants were healthy, right-handed, native English speakers with normal or corrected vision. Participants were screened for Axis I disorders from the DSM-IV and immune and endocrine system disorders. All female participants were tested during the luteal phase of the menstrual cycle to control for variations in reactivity of the hypothalamic–pituitary–adrenal (HPA) axis throughout the cycle (Kirschbaum et al. 1999), and all sessions began at 3 p.m. Participants were asked to refrain from behaviors known to impact the HPA axis, including: alcohol, smoking, heavy exercise within 48 h of the imaging session, and heavy meals, caffeine, and exercise within 2 h of the imaging session. Also, participants were asked to get a normal night's sleep the night before. Compliance with these restrictions was confirmed upon arrival at the imaging center.

All study procedures were approved by the IRB of Columbia University and were in accordance with the Declaration of Helsinki, and participants provided written informed consent. Owing to technical failure, N-back behavioral data from 8 full and 2 half runs were not recorded, comprising about 10% of the total data (which we assessed to be missing at random). Salivary cortisol samples from 1 SET participant were not usable and therefore excluded.

Study Procedure

The experiment included a practice session and a scanning session the next day. During practice, participants performed the N-back task until they reached a performance criterion of 70% correct, in order to reduce preexisting inter-individual differences in WM. Specifically, participants performed a run of 2-back trials, and then a run of 3-back trials with feedback. On a per run basis, if at least 20 trials had been completed and accuracy exceeded 70%, the criterion was considered met and the run was terminated. The scanning session consisted of: 1) structural and baseline functional scans, 2) an fMRI-compatible adaptation of the TSST (Kirschbaum et al. 1993) or control task, 3) a divided attention task (DAT) (~28 min, involving male/female decisions or indoor/outdoor decisions of faces superimposed on indoor/outdoor scenes, data to be presented elsewhere), 4) the N-back task, 5) behavioral and emotional assessments, and 6) debriefing.

SET and Control Manipulations

During the TSST (Kirschbaum et al. 1993) (for SET participants), 2 confederates wearing white lab coats entered the fMRI scanner room and were introduced as professors. They instructed participants to prepare a speech that would later be delivered to them. Social evaluative threat participants were given 3 min to prepare a 5-min speech on their strengths and weaknesses as candidates for "the job of their dreams" while functional fMRI measures were acquired. The job was specific to each participant, based on a questionnaire administered on the first testing day. Following the preparation period, SET participants left the scanner and entered a conference room, in which they delivered the speech (5 min) and performed an oral serial subtraction task (5 min). For SET participants, we aimed at extending the SET environment beyond the TSST itself. Therefore, the professors accompanied the participants back to the scanner and told the participants they would continue monitoring their performance on all further tasks through the video and by computer analysis. Also, the following elements were added to the experimental procedure: 1) the N-back task was introduced as a measure that was predictive of IQ and professional achievements, 2) participants were told that persons with an average IQ should be able to score easily in at least the 60th percentile on the N-back task, 3) participants were told that the "professors" would continue evaluating them on their performance, 4) after each N-back run, sham feedback indicating inferior performance was presented, and 5) participants were told that their score was below average and that if they would not improve their performance, they would be asked to give another speech in front of the professors.

Control participants were not introduced to the confederates but instead were instructed by the experimenter to imagine spending a day in New York City with a friend who was visiting. They were told that they would be asked to write down their story in a 5-min period following the scan. At the end of the preparation period, control participants left the scanner room and recorded their stories on their own (5 min).

To further promote the nonthreatening nature of this task, they were told that they could take longer than 5 min if they would like and that their answers would not be used as data in the experiment.

Verbal N-Back
The N-back task was administered over 4 scanning runs, each with intermixed blocks of 2-back and 3-back trials. A run consisted of 9 blocks, and blocks were presented in ABBA or BAAB order, interspersed by 25 s of resting fixation. In total, 18 blocks of 2-back and 18 blocks of 3-back were administered. Each block started with an N-back instruction (either 2-back or 3-back) followed by 10 N-back trials. There were 3 types of trials: 1) target, 2) lure (nontarget), and 3) nonlure (nontarget). Lure trials were included as WM involves both maintenance and "control" processes (Baddeley 2003), and lures specifically challenge attentional control, since the tendency to make a target response based on familiarity or recency is prepotent (Jonides et al. 1998; Gray et al. 2003). In a 2-back block, lure trials consisted of a 3- or 4-back match. In a 3-back block, lure trials consisted of a 2- or 4-back match.

On each of 360 trials, a word was presented for 2 s followed by an inter-stimulus-interval fixation cross for 1 s. Participants were instructed to respond as quickly and accurately as possible to each word with a button press indicating, "yes, this word matches the word presented n-trials ago" (a target response) or "no, this word does not match the one presented n-trials ago" (a nontarget response). The first 3 trials of each block were not included in the analysis; these trials were always nontargets in the 3-back block and therefore did not carry the same task demand as the remaining trials. The remaining trials consisted of 42% targets and 29% for each of lures and nonlures. Words were selected from the MRC Psycholinguistic Database (http://www.psych.rl.ac.uk/MRC_Psych_Db.html), contained maximally 3 syllables, and were all defined as relatively familiar, easy to imagine, concrete, and acquired at a young age. The task was administered with E-Prime software (Psychology Software Tools, Inc.) through LCD goggles. Responses were recorded by a MR-compatible button box.

Salivary Cortisol and Subjective Affect
At 8 time points throughout the scanning session (∼5, 45, 90, 93, 103, 133, 173, and 210 min after the session start, see Fig. 1C), measures of salivary cortisol were collected using Salivettes. Along with each sample, participants were asked to fill out a subset of the Positive and Negative Affect Scale (PANAS) (anxious, excited, and alert) (Watson et al. 1988). Participants reported levels of affect on a scale of 1 to 4, with 1 equaling "not at all" and 4 equaling "very much."

Salivary Cortisol and Subjective Affect Analysis
Alterations in salivary cortisol levels and subjective mood were analyzed with a repeated-measures ANOVA with Time (pre-TSST, during TSST, post-TSST, pre-N-back, and post-N-back) as a within-subjects factor and Group (SET, Control) as a between-subjects factor. Percentage increase of cortisol was calculated according to the following formula: (Δcort/baseline) × 100, where Δcort is the change in absolute cortisol from the baseline. The Pre-TSST time point was selected as a baseline measure since it occurred just before the administration of the speech preparation (SET) or writing (control) instructions. The pre-N-back time point was selected since it was hypothesized to serve as the peak time point for the SET group following the SET manipulation given the lag of the cortisol response (Dickerson and Kemeny 2004).

N-Back Behavioral Analysis
Reaction times that were less than 200 ms or 3 standard deviations above or below the mean were excluded. Time out trials were excluded from the analysis for both accuracy and reaction time (RT) variables. Accuracy and RT data were first submitted to a 2 × 2 × 3 mixed model ANOVA with the factors Workload (2-back, 3-back), Group (SET, Control), and Trial Type (Target, Lure, Nonlure). Accuracy data were converted into a unit of sensitivity, A-sensitivity (Zhang and Mueller

2005). Since we expected impairments in WM due to the effects of SET a priori, we used planned contrasts for follow-up analyses on significant interactions (i.e., these were not corrected for multiple comparisons). To create a single, composite measure of WM performance, we combined A-sensitivity and RT by employing principal component analysis (PCA) across measures on the 2 workload types (2-back and 3-back), using z-scores for each measure across the sample. We interpreted the first 3 components yielded by the analysis and found that the first loaded positively on A-sensitivity and negatively on RT. Thus, the higher the value of this component, the better the performance as it indicates higher accuracy and shorter RTs. This component, labeled "WM Performance," was utilized in all subsequent analyses assessing composite WM behavior.

Imaging Acquisition
Measures of brain activity were acquired with a 3T Philips Achieva scanner using BOLD gradient echo planar imaging. The following parameters were specified: voxel size = 2 × 2 × 3 mm, slice number = 37, TR = 2000 ms, TE = 20 ms, flip angle = 72°, and acquisition = interleaved. A T1-weighted structural image (MPRAGE) was also acquired.

Imaging Analysis
BOLD images were converted to ANALYZE format and preprocessed using SPM5 software (Wellcome Department of Cognitive Neurology). The preprocessing steps were as follows: 1) slice-timing correction, 2) realignment yielding 6 standard head-movement parameters (i.e., x, y, z-translation, pitch, yaw, and roll), 3) coregistration to the structural image, 4) normalization to the template MNI space, and 5) smoothing with an 8-mm kernel.

Voxel-Wise Maps
At the first level, we estimated models as a function of workload and trial type (2-back targets, 2-back lures, 2-back nonlures, 3-back targets, 3-back lures, 3-back nonlures, and start trials) convolved with the standard canonical hemodynamic response function, using a similar approach as a previous report (Gray et al. 2003). Also, the 6 movement parameters from realignment were included as nuisance regressors. A high-pass filter using a cutoff of 1/160 Hz was specified. Given our interest in SET effects on overall WM performance, the contrast of primary interest compared targets, lures, and nonlures together with rest (i.e., the 25-s blocks of rest in between N-back blocks and fixation between words) [N-back vs. Rest]. Also, we included both accurate and error trials in the imaging analysis, since A-sensitivity, 1 part of the behavioral composite measure, WM performance, was based on both kinds of trials.

At the second level, we utilized the mediation analysis toolbox in order to implement MEPM (Wager, Waugh, et al. 2009; Atlas et al. 2010). Matlab code implementing mediation analyses is freely available at: http://wagerlab.colorado.edu/. By following standard mediation logic (Baron and Kenny 1986), MEPM enables assessment of the relationship between an experimental manipulation (X) and a dependent variable (Y), mediated by brain activity (M). Relationships between X and M and M and Y are expressed by path coefficients a and b, respectively. If M mediates the relationship between X and Y, then the product $a*b$ should become significantly different from zero. The mediation tests for each of these 3 effects (a, b and $a*b$) in every single voxel within a multi-level model. Mathematical details of the MEPM model can be found elsewhere (Wager, Waugh, et al. 2009; Atlas et al. 2010). To answer our main question, we entered Group into the model as the predictor × (SET = 1, control = −1), WM performance as Y, and the [N-back vs. Rest] contrast as M. With this multi-level path model, the following voxel-wise effects were tested: 1) WM-related brain activity in response to the SET manipulation (Path a), 2) brain activity related to overall WM performance, controlling for group (Path b), and 3) the mediation effect, indicative of brain regions mediating the relationship between SET and WM performance (Path $a*b$). As we expected SET-related WM impairment, it followed that mediators should be negative: either SET reduces activity (negative Path a) in regions that generally aid performance (positive Path b), or the other way around.

Unless otherwise noted, brain maps are shown using a threshold of $P < 0.005$, two-tailed and uncorrected, with an extent of $k = 10$ voxels, to preserve the balance between sensitivity and false positive rate (Lieberman and Cunningham 2009). The $P < 0.005$ two-tailed threshold is the equivalent of $P < 0.0025$ in SPM, as results presented in SPM are one-tailed. We also thresholded analyses in a priori regions of interest at $P < 0.05$ family-wise error rate (FWER) corrected based on cluster extent and a primary threshold of $P < 0.01$, two-tailed ($P < 0.005$ in SPM), using AlphaSim (Cox 1996) and the Neurosynth WM mask as described later. For mediation tests, the primary threshold required both the Path a and Path b links to be significant at $P < 0.005$, as well as the $a*b$ mediation test. Because the Path a and Path b effects are conditionally independent, the probability of both being significant by chance is $P = 0.000025$ or lower.

Regions of Interest

We selected relevant WM-related brain areas using the Neurosynth meta-analytic database (Yarkoni et al. 2011), by searching for the term "working memory." Positive results (probability of activation above base rate across the brain) in the Neurosynth "working memory" reverse-inference map were thresholded at $q < 0.05$ FDR-corrected. We thresholded contiguous clusters at $k = 50$ to yield 17 distinct regions (Fig. 3B). The mask was mirrored across the left and right hemispheres, taking the union of voxels in both hemispheres, smoothed with a 4-mm FWHM kernel, and thresholded at 0.05 to limit the size of the smoothed borders. An additional Neurosynth-based mask was created with regions of interest associated with self-related and default-mode-related processing. Reverse-inference maps related to "self" and "default" were thresholded at $q < 0.05$ FDR-corrected. Reverse-inference maps and the intersection were chosen here because they limited results to a priori areas of interest in the ventromedial prefrontal cortex, posterior cingulate, and a limited set of additional regions (including portions of the inferior frontal and parietal cortices). Then, the intersection of the WM, self, and default masks was calculated. By combining these masks, we defined an a priori mask of co-occurring WM-, and SET-related brain regions (WM + SET) (Fig. 3D). To further examine the robustness of our results, we extracted 8 distinct ROIs from the Neurosynth WM mask and conducted small volume cluster extent-based correction (SVC) in AlphaSim (Cox 1996) thresholded at $P < 0.05$ FWER-corrected. The primary threshold was $P < 0.01$ (10 000 Monte Carlo iterations), with FWER correction based on the number of contiguous significant voxels within the ROI.

Working Memory-Related Pattern of Interest

In addition to regions of interest, we conducted a "pattern of interest" analysis (Wager et al. 2013) based on an a priori template for distributed, WM-related activity defined by prior literature. The basic idea is to specify a distributed pattern related to WM "a priori" and then to test WM effects and SET effects on the "expression," or integrated activity, in the template pattern. The "pattern of interest" generalizes the "region of interest" approach to testing a distributed pattern, allowing us to test activity in a distributed WM-related network without testing multiple comparisons. The pattern was based on "reverse inference" Z-score maps from Neurosynth using the term "working memory" (Fig. 3B). Z-scores from the meta-analytic map were used as weights for each voxel in the distributed map. The "expression" of the pattern for a given data image (e.g., participants' contrast maps) is calculated as a weighted average of activity, with weights defined by the Z-scores in the pattern map. In linear algebraic terms, this is equivalent to projecting the contrast images from each subject onto the multivariate vector defined by the pattern (see below for additional details).

To define the pattern of interest, the Neurosynth "working memory" map was thresholded at $q < 0.05$ FDR-corrected. Positive (yellow in Fig. 3B) and negative (blue in Fig. 3B) values in the map reflect activation probability above and below the base rate across the brain, respectively. More specifically, Z-score values in the map were derived from a one-way chi-square test comparing the counts of activating versus nonactivating "working memory" studies in each voxel to the base rate of activation across the brain. To improve estimation properties with sparse data, counts were regularized using an m-estimator to smooth activation probability estimates toward 0.5, as described in

Yarkoni et al. (2011), and reflected in this equation: $P(A_j = 1 | T_k = 1) = (\Sigma_i A_{ij} T_{ik} + mP)/(\Sigma_i T_{ik} + m)$ where A reflects activation, T reflects the term (e.g., "working memory"), j indexes voxel, k indexes task, i indexes study, m is a virtual equivalent sample size, and P is a prior probability. The parameters m and P are set equal to 2 and 0.5 respectively; this smoothing is equivalent to adding 2 virtual studies that have term k present—1 having an activation one, 1 without. P-values from the likelihood ratio test were transformed to Z-values, with voxels with probabilities above the base rate for activation receiving positive Z-values, and those below the base rate for activation receiving negative Z-values. Activation frequencies below the whole-brain base rate are assumed here to reflect likely WM-related deactivations; this assumption is borne out in individual studies in default mode network (DMN) regions and others with negative Neurosynth Z-scores. Thus, the Z-values are useful as estimates of the degree to which voxels are activated (positive Z-scores) and de-activated (negative Z-scores) in WM.

To apply the pattern to the data, the Neurosynth WM Z-score map was resampled to the space of the functional data using trilinear interpolation. The Z-scores in this map were treated as a pattern of weights; the map was vectorized, so that the vector $\vec{w}_{map}$ contained the Z-scores. We then calculated the dot product of this vector with each subject's [N-back vs. Rest] contrast map by vectorizing contrast images ($\vec{\beta}_{map}$) and taking the dot product of each individual's contrast map and the Neurosynth template mask: ($pr = \vec{\beta}_{map}^T \vec{w}_{map}$). This yielded a single set of pattern response scores (pr), 1 per subject, which reflects the expression of (match to) the overall a priori WM pattern. If the pattern weights $\vec{w}_{map}$ all had values of 1, this would be equivalent to taking the sum (the average up to scaling) of the voxels in the mask. By using Z-scores as weights, the voxels with the most reliable activation in Neurosynth contribute more to the pr score. In addition, activation of regions with negative Z-scores reduces the pr score, and de-activation of regions with negative Z-scores increases the pr score. This similarity metric is related to others, including the pattern correlation (the correlation coefficient of the 2 maps) and the cosine similarity; among these, we used the dot product because it is closest to the original data in the sense that it preserves the scaling of the response across participants, and individual differences in activation magnitude across the pattern contribute to between-subjects error. Once we calculated the pattern response score for each participant, these scores were then tested for SET effects. Thus, a significant effect of SET on pattern expression would indicate that the amplitude of anticipated activation and deactivation based on the Neurosynth WM map differed for SET and control participants.

Connectivity Analyses

We conducted 3 types of connectivity analyses. In Analysis 1, we identified regions that were functionally correlated with key mediators of threat effects (used as "seed" regions) on average across WM and Rest. In Analysis 2, we identified regions that showed task-dependent (WM—Rest) functional connectivity with key mediators. In Analysis 3, we tested whether task-dependent connectivity estimates mediated SET effects on WM performance. Results were visualized with the WM + SET mask.

Analysis 1: average connectivity across WM and Rest. We carried out voxel-wise, time series connectivity analyses implemented in custom software in Matlab (Mathworks, Inc.) using the WM + SET mask. In the first analysis, negative mediators (i.e., those that explained the negative effects of SET on WM performance) from the mediation analysis were entered as seed regions, and maps of cross-correlation coefficients were computed by measuring the time series correlation between each "seed" brain region and all other voxels falling into the WM + SET mask. Voxel time courses were high-pass-filtered (cutoff period of 246 s, 0.0041 Hz) before analysis. Next, a second-level analysis was performed using robust regression (Wager et al. 2005). In this analysis, connectivity maps for each participant were entered as the outcome data, and predictors included an intercept (capturing mean connectivity across all participants) and the effect of group (capturing SET vs. Control effects on connectivity). Connectivity in this analysis was not intended to isolate "intrinsic" or "task-independent" connectivity, as it could reflect a combination of task-related and task-independent effects. We test whether connectivity is task-related in subsequent

analyses mentioned later; therefore, we did not regress out task effects in this first analysis.

Analysis 2: task-dependent connectivity (WM—Rest). To test whether connectivity estimates varied as a function of the task, we performed a psychophysiological interaction (PPI) (O'Reilly et al. 2012) analysis on N-back-dependent connectivity. The model used brain time series data for each subject as the outcome and included predictors for 1) Task condition (Rest = 0, N-back = 1), 2) centered seed region time series data (from mediators in the IPS), and 3) the Task × Seed time series interaction. The interaction was the main regressor of interest, as it reflects task-dependent connectivity with the seed of interest. Effects of Group (SET vs. Control) were assessed at the second level as in the previous connectivity analysis.

Analysis 3: task-dependent connectivity mediates performance reductions. Finally, to test whether task-dependent connectivity mediated SET effects on WM performance, we used voxel-wise Task × connectivity estimates from Analysis 2 in a mediation analysis, with X = Group (SET vs. Control), M = Task × connectivity, and Y = WM performance. In this model, Path a tested whether WM-dependent connectivity was affected by the SET manipulation. Path b tested whether connectivity predicted WM performance, controlling for Group, and the Path $a*b$ mediation tested whether the relative strengths of these effects satisfied the criteria for mediation.

Results

Physiological and Subjective Measurement of SET

Social evaluative threat resulted in salivary cortisol increases as evidenced by a marginally significant Group × Time ANOVA ($F_{4,68} = 2.25$, $P = 0.073$, $\eta_p = 0.12$). Indeed, percent change in cortisol (i.e., change between Pre-TSST and Pre-N-back time points) was greater in the SET group compared with the Control group ($t_{18} = 2.28$, $P = 0.04$) (Fig. 1C). None of the Group × Time ANOVAs for subjective affect reached significance (all $F < 1.4$).
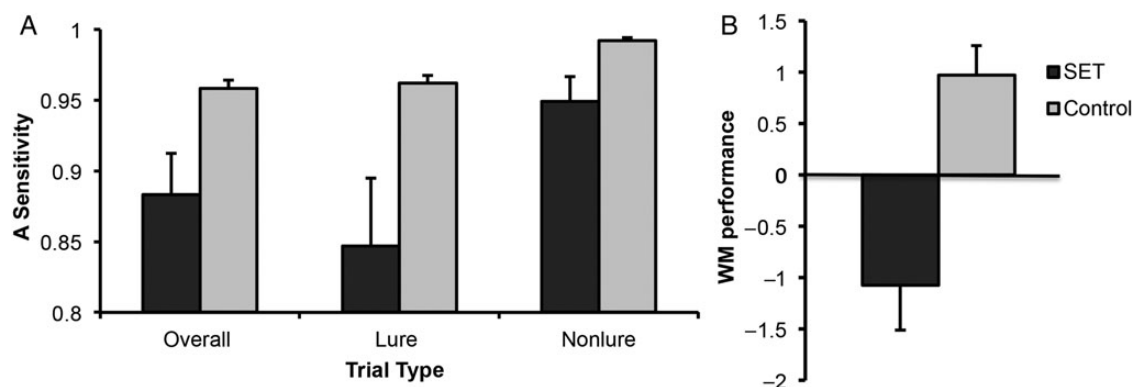
Working Memory Performance

There were no baseline differences among groups on accuracy or RT in the pre-scan practice session. A significant Trial Type (Target, Lure, Nonlure) × Group (SET, Control) interaction ($F_{2,38} = 4.84$, $P = 0.013$) demonstrated that accuracy was reduced in SET participants relative to Controls on target trials ($P = 0.011$) and lure trials ($P = 0.036$), but not nonlure trials ($P = 0.29$), using planned contrasts. This suggests that the most difficult (i.e., attention demanding) trials were sensitive to SET effects. These effects were not altered by Workload (2-back,

3-back; $F_{2,38} = 0.89$, $P = 0.419$). In addition, SET generally resulted in lower accuracy (main effect of Group; $F_{1,19} = 8.02$, $P = 0.011$). Regardless of Group, memory accuracy depended upon Workload × Trial Type (interaction; $F_{2,38} = 8.46$, $P = 0.001$), and it was generally reduced in the 3-back condition (main effect of Workload; $F_{1,19} = 12.53$, $P = 0.002$).
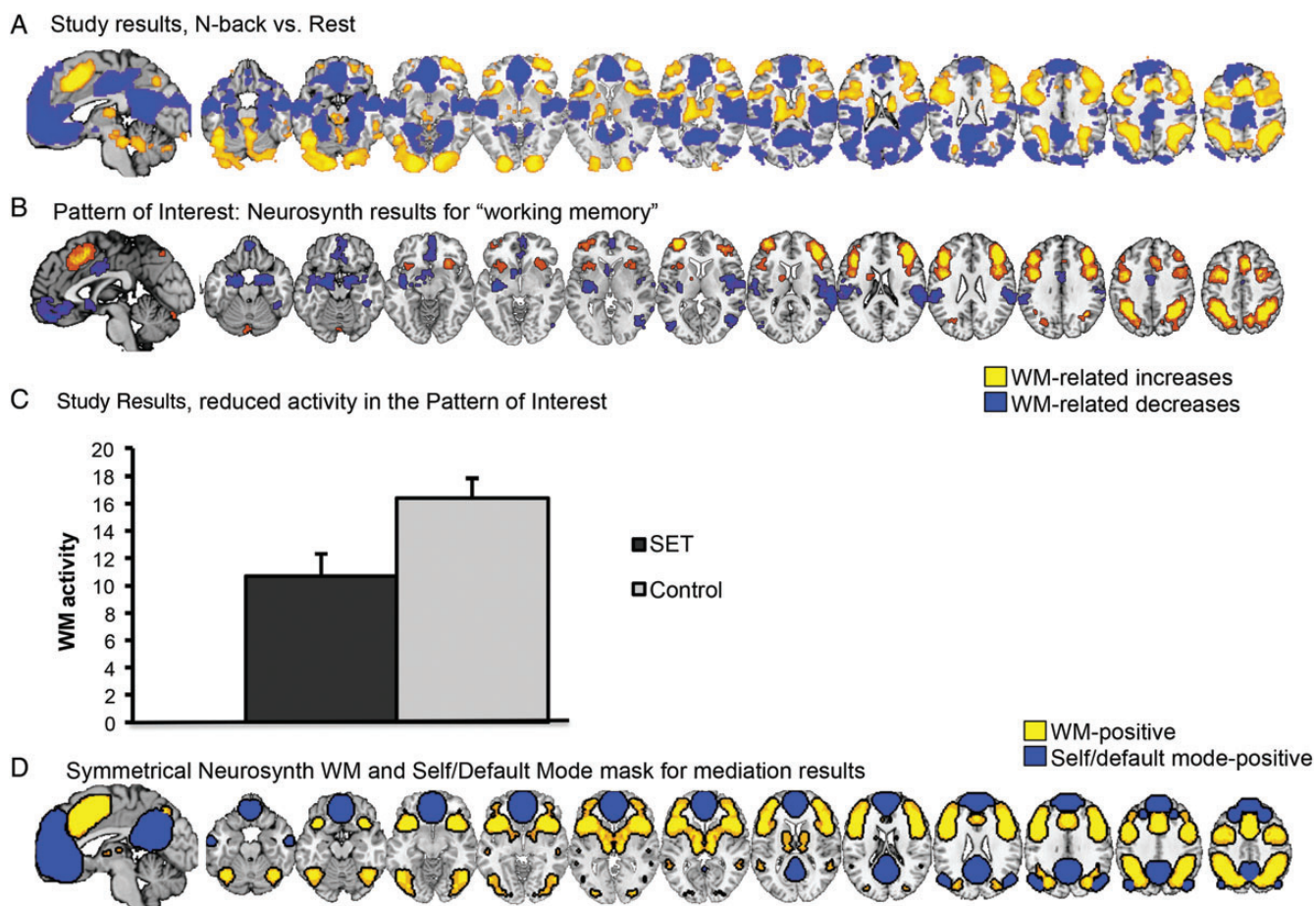
Reaction time data were submitted to the same analysis as above. A main effect for Group ($F_{1,19} = 10.52$, $P = 0.004$) indicated that SET overall significantly prolonged RT. In addition, main effects were found for Workload ($F_{1,19} = 14.13$, $P = 0.001$), with slower RTs in the 3-back condition, and a main effect of Trial Type ($F_{2,38} = 51.91$, $P < 0.001$), with longer RT for lures than for targets ($P < 0.001$), and marginally longer RT for targets than for nonlures ($P = 0.054$). This suggests that in terms of RT, lures are most difficult, followed by targets, with nonlures being easiest. A marginal significant interaction was also found for Trial Type × Workload ($F_{2,38} = 2.67$, $P = 0.082$), but all other effects remained insignificant.

Social evaluative threat-related WM impairment was also evident in A-sensitivity (Zhang and Mueller 2005), a signal-detection measure of accuracy performance, as revealed by a main effect of Group ($F_{1,19} = 6.56$, $P = 0.019$), and an interaction effect of Group × Trial Type ($F_{1,19} = 5,12$, $P = 0.036$). Indeed, SET reduced the discriminability of targets from both lures ($t_{19} = -2.50$, $P = 0.02$) and nonlures ($t_{19} = -2.54$, $P = 0.02$) (Fig. 2A), as revealed by planned contrasts. Because RT and A-sensitivity both showed performance decrements with SET, we combined them using PCA into a single, integrated measure, which we termed "WM performance" and used in brain analyses. Working memory performance was significantly impaired by SET ($t_{19} = -3.98$, $P < 0.001$; Fig. 2B).

To test whether change in cortisol predicted WM performance as a function of group, we regressed WM performance on Group (SET vs. Control), percent change in cortisol, and their interaction. As expected, the SET group showed lower WM performance ($t_{18} = -5.34$, $P < 0.001$). However, no relationship was found between Cortisol and WM performance ($t_{18} = 0.30$, $P = n.s.$). Finally, there was a trend toward a significant Group × Cortisol interaction, such that there was a stronger "positive" cortisol–performance relationship in the SET group than Control group ($t_{18} = 1.83$, $P = 0.086$). Cortisol related positively to performance in the SET group ($r = 0.714$, $P = 0.031$), but did not predict performance in the Control group ($r = -0.280$, $P = n.s.$). Taken together, these results confirmed that social

Figure 2. Effects of social evaluative threat (SET) on working memory (WM). (A) A-sensitivity. The SET Group was impaired on A-sensitivity overall and when assessed as a function of lure or nonlure conditions separately. (B) Composite WM performance. The SET Group was impaired on the composite behavioral measure, WM performance.

A Study results, N-back vs. Rest

B Pattern of Interest: Neurosynth results for "working memory"

■ WM-related increases
■ WM-related decreases

C Study Results, reduced activity in the Pattern of Interest

■ SET
□ Control

□ WM-positive
■ Self/default mode-positive

D Symmetrical Neurosynth WM and Self/Default Mode mask for mediation results

Figure 3. Main effects of N-back versus Rest. (A) Results from the [N-back vs. Rest] contrast in the current study. Red/yellow: N-back > Rest. Blue: Rest > N-back. (B) Results from a search in Neurosynth (Yarkoni et al. 2011) for the term "working memory" (WM). Notably, the results in Neurosynth match closely with those from the current study shown in A. Yellow: positively associated with WM. Blue: negatively associated with WM (see "Pattern of Interest analysis" in text for details). (C) Pattern of interest results (i.e., a weighted average across all WM regions within the Neurosynth mask in [B]) for the social evaluative threat (SET) and Control group. Social evaluative threat significantly reduced activity in WM- and SET-related brain areas. (D) A second Neurosynth-based mask was created comprising the WM mask and regions of interest associated with self-related and default-mode-related processing. Yellow: positively associated with WM. Blue: positively associated with self and default-mode processing.

evaluative threat impaired WM performance overall, but the impairment was not a consequence of increased cortisol. Instead, cortisol related to enhanced performance within the SET group.

Regions Activated by WM

The [N-back vs. Rest] contrast produced several WM and DMN regions that have been observed in previous work including increases in IPS, anterior cingulate cortex (ACC), and dorsolateral PFC (dlPFC) (Wager and Smith 2003; Champod and Petrides 2010) and decreases in medial PFC and precuneus/posterior cingulate cortex (PCN/PCC) (Qin et al. 2009; Andrews-Hanna et al. 2010) (Fig. 3A). Notably, the results observed in this study match closely with the Neurosynth WM map (Fig. 3B). Interaction effects of SET and WM load [3-back vs. 2-back] and SET and difficulty [Lure vs. Nonlure] can be found in Supplementary Material. The expression of the Neurosynth pattern of interest was significantly greater than 0 (t_{18} = 11.05, $P < 0.0001$), indicating a strong match with the a priori pattern in the expected direction. Hundred percent of the participants showed a positive pattern expression response, indicating that the WM pattern activated more for N-back than Rest in every participant.

Brain Mediators of SET Effects on WM Performance

Social evaluative threat, compared with the Control Group, induced widespread decreases in WM-related areas in the [N-back vs. Rest] contrast, including dlPFC, ACC, inferior frontal gyrus, parietal cortex including bilateral IPS, and cerebellum (Path *a*). Social evaluative threat increased N-back-related activity in vmPFC and inferior temporal gyrus. Activity in the Neurosynth pattern of interest was significantly reduced in the SET versus Control groups (t_{19} = −2.67, P = 0.015) (Fig. 3C).

Brain activity in the [N-back vs. Rest] contrast correlated positively with WM performance (Path *b*) in portions of regions activated during WM in prior studies (Wager and Smith 2003; Champod and Petrides 2010), including bilateral IPS, superior temporal gyrus, supramarginal gyrus, ACC and cerebellum. Negative correlations were found in several parts of the so-called default mode network including dorsal medial prefrontal cortex (dmPFC), frontal pole, PCN/PCC and temporal parietal junction (TPJ). However, the Neurosynth pattern of interest did not show a significant Path *b* effect (r = 0.20, P > 0.01) and therefore did not mediate SET effects on WM performance (P > 0.10).

Finally, and crucial to our research question, maps of the Path *a*b* effect showed negative mediation in bilateral IPS

(Fig. 4), resulting from social evaluative threat-induced reductions in activity (negative Path *a*) and a positive activity–performance relationship (positive Path *b*). The effect of group on performance ($R^2 = 0.454$) dropped 82% when controlling for either IPS alone and 93% when controlling for bilateral IPS (residual Group-Performance $R^2 = 0.031$). Thus, bilateral IPS explained nearly all of the effect of the SET manipulation on WM performance. Though these effect sizes are not unbiased because they resulted from a voxel-wise search (Vul et al. 2009), the results suggest that deactivation of bilateral IPS contributes substantially to the effects of SET on WM. Because the L IPS activation appeared to fall close to white matter, we conducted an additional analysis with a primary threshold of $P < 0.01$ and FWER correction based on cluster extent at $P < 0.05$ corrected. At this threshold, the activation clearly extended to gray matter bilaterally, and the R IPS was significant at FWER $P < 0.05$ (see Supplementary Results and Supplementary Fig. 1). We also note that the Path *a* effects clearly extended into gray matter bilaterally.

The mediation analysis additionally revealed 3 positive mediators: dmPFC, PCN/PCC, and TPJ, each the result of negative Path *a* effects and negative Path *b* effects (Table 1). These regions showed reduced activity during SET, but reduced activity was correlated with better performance. Such regions are

suppressor variables (MacKinnon et al. 2000) that work against the overall WM-impairing effect of SET, and such effects may provide clues about the variable effects of SET on cognition across individuals.

Effects of SET (Path *a*) were found in all 8 ROIs of the Neurosynth WM mask at SVC $P < 0.05$ FWER-corrected. However, only bilateral IPS predicted performance controlling for SET (Path *b*), and right IPS showed mediation effects (Path *a*b*) (Table 2).

Functional Connectivity

We used both left and right IPS regions from the mediation analysis as seeds for functional connectivity analysis and the WM + SET mask (Analysis 1; see Methods), in order to 1) locate functional connectivity with IPS and 2) test whether IPS connectivity was moderated by Group (SET vs. Control). Both IPS seeds were positively connected to bilateral dlPFC and negatively connected to vmPFC. Critically, as shown in Figure 5*A*, SET reduced right IPS–left dlPFC connectivity and increased connectivity between bilateral IPS and vmPFC.

As the right IPS was the strongest result in the mediation analysis (because it fell within the Neurosynth mask and survived SVC), PPI analysis (Analysis 2) was run with this ROI as
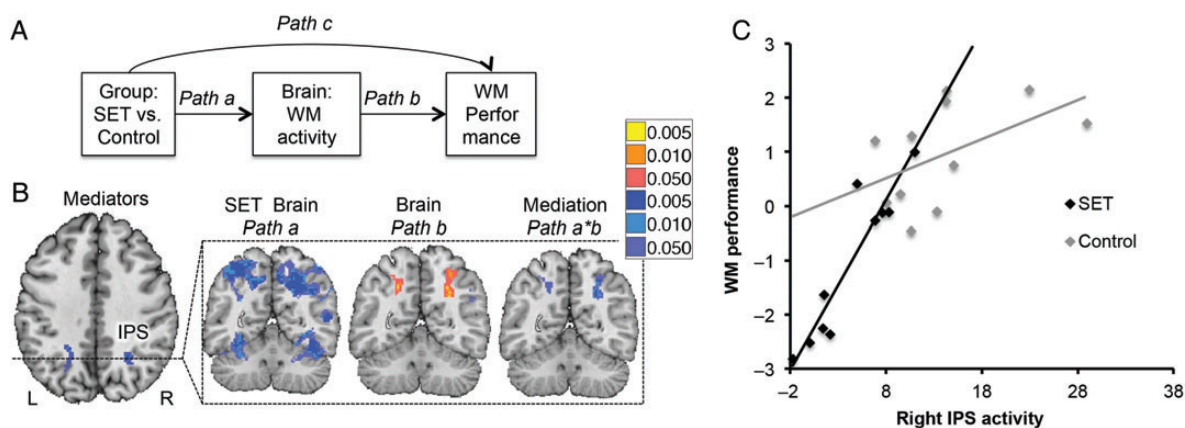
Figure 4. Brain mediators of social evaluative threat (SET) effects on working memory (WM). (*A*) Mediation path diagram. Mediation was used to search for Brain activity [N-back vs. Rest] that formally explained the relationship between Group and WM performance on a voxel-wise basis. (*B*) Negative mediators. Deactivation in bilateral intraparietal sulcus (IPS) mediates the relationship between Group and WM performance, suggesting that less activity in this region underlies SET-related WM impairment. The IPS has been implicated in studies of both WM and attention. For instance, IPS activates when attention is directed to a spatial location where the presentation of a stimulus is expected in the absence of visual information as shown in Kastner et al. 1999. Red/yellow: positive effects; blue: negative effects. (*C*) Scatterplot depicting WM performance as a function of activation in the left IPS and group. Overall, WM performance relates positively to left IPS activation; the SET group shows less activation in this region and impaired WM performance relative to the Control group. L, left; R, right.

Table 1

Brain mediators of the relationship between group and working memory (WM) performance (Path a*b)

Mediator	Side	x	y	z	Sig. Voxels	Vol. (mm³)	Path *a* (Group) Z	Path *a* (Group) P	Path *b* (WM performance) Z	Path *b* (WM performance) P	Mediation (*a*b*) Z	Mediation (*a*b*) P
Positive												
Precuneus/PCC	Right	12	−60	30	41	328	−3.72	0.000	−3.25	0.001	3.24	0.001
TPJ (angular gyrus)	Right	54	−58	30	16	128	−3.18	0.002	−3.01	0.003	2.90	0.004
dMPFC (superior frontal gyrus)	Right	2	36	48	13	104	−3.42	0.001	−3.03	0.002	2.98	0.003
Negative												
IPS (superior parietal lobule)	Right	28	−54	40	52	416	−3.56	0.000	3.20	0.001	−3.56	0.000
IPS	Left	−20	−50	40	11	88	−3.45	0.001	3.59	0.000	−3.18	0.002

Note: MNI coordinates (*x, y, z*) are shown. Sig. Voxels, the number of significant voxels at $P < 0.005$; Vol. (mm³), volume in mm³ significant at the same threshold; max Z, the peak Z-value of the region.

Table 2
Significant regions within the Neurosynth Working Memory mask

Region	#Vox in region	Vcrit	Path a, P = 0.01						Path b, P = 0.01						Path a*b, P = 0.01					
			x	y	z	Z	#Vox	Pass	x	y	z	Z	#Vox	Pass	x	y	z	Z	#Vox	Pass
R dorsolateral prefrontal cortex	3455	63	42	14	26	−6.21	289	*	—	—	—	—	—		—	—	—	—	—	
			36	44	28	−6.21	120	*	—	—	—	—	—		—	—	—	—	—	
			38	8	44	−6.21	187	*	—	—	—	—	—		—	—	—	—	—	
			32	−2	54	−6.12	166	*	—	—	—	—	—		—	—	—	—	—	
L dorsolateral prefrontal cortex	3455	63	−44	8	26	−6.21	163	*	—	—	—	—	—		—	—	—	—	—	
			−44	30	34	−6.21	86	*	—	—	—	—	—		—	—	—	—	—	
			−32	2	58	−6.21	287	*	—	—	—	—	—		—	—	—	—	—	
Bilateral paracingulate cortex	671	27	−2	20	46	−6.27	308	*	—	—	—	—	—		—	—	—	—	—	
R IPS	1927	51	26	−58	46	−7.98	1136	*	28	−54	42	6.88	140	*	28	−54	42	−6.82	122	*
L IPS	1933	50	−16	−68	48	−6.22	349	*	−18	−64	46	6.21	121	*	—	—	—	—	—	
			−40	−46	44	−6.22	589	*	—	—	—	—	—		—	—	—	—	—	
R insula	73	4	30	20	−4	−6.21	67	*	—	—	—	—	—		—	—	—	—	—	
L insula	73	5	−30	20	−4	−6.20	48	*	—	—	—	—	—		—	—	—	—	—	
Bilateral cerebellum	11	1	0	−82	−18	−6.20	11	*	—	—	—	—	—		—	—	—	—	—	

Note: The Neurosynth Working Memory (WM) mask was separated to create 8 regions of contiguous voxels. Results with MNI coordinates (x, y, z) are shown for Path a (WM-related brain activity in response to SET), Path b (brain activity related to overall WM performance, controlling for social evaluative threat), and Path a*b (the mediation effect, indicative of brain regions mediating the relationship between social evaluative threat and WM performance). #Vox in region, number of voxels within each region; Vcrit, number of critical voxels for P < 0.05 FWER-corrected; Z, the peak Z-value of the region; #Vox, number of significant voxels at primary threshold, P < 0.01, uncorrected; pass: * indicates whether the cluster survives correction (#Vox > Vcrit).
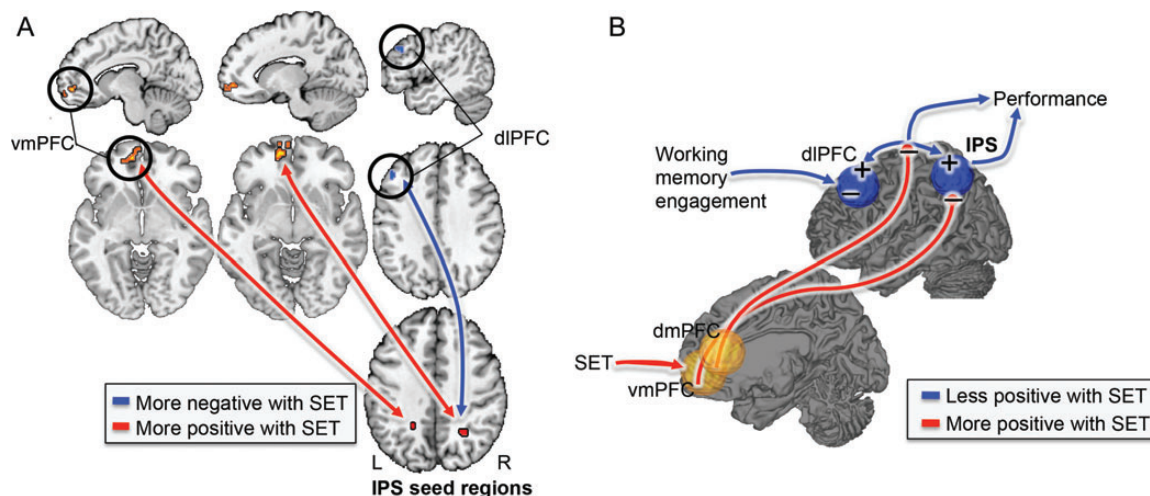
Figure 5. (A) Regions showing different connectivity for [social evaluative threat (SET) vs. Control]. Bilateral negative intraparietal sulcus (IPS) mediators (see Fig. 4) were entered as seed regions into a functional connectivity analysis, and connectivity as a function of SET was assessed. Results indicated that connectivity between the IPS and the dorsolateral prefrontal cortex decreased and connectivity between the IPS and the ventromedial prefrontal cortex increased with SET. Red/yellow: increased (more positive) associations with SET; blue: reduced (more negative) associations with SET. (B) Hypothesized interactions underlying competition between WM task performance and social threat. Bilateral IPS may aid in directing attention to and switching between various ongoing processes, depending on specific contextual demands. Under conditions of SET, increased processing of affective information constructed in part in vmPFC competes with task-focused cognitive processes maintained in fronto-parietal circuits, ultimately resulting in impaired WM performance. Red/yellow: increased (more positive) associations with SET; blue: reduced (more negative) associations with SET. + indicates enhanced activity under control conditions and − indicates reduced activity under control conditions.

the seed in order to assess whether its connectivity pattern was WM-task dependent. Positive connectivity between the right IPS and right dlPFC was stronger during [N-back vs. Rest] (i.e., a positive PPI effect), and negative connectivity between the right IPS and bilateral dmPFC and vmPFC was also stronger during N-back blocks. The negative right IPS–dmPFC connectivity was less in the SET group. Thus, N-back performance creates a negative relationship between IPS and dmPFC, but this negative relationship is blocked by SET. Mediation analysis (Analysis 3) showed that in a nearby area of left MPFC, WM-dependent connectivity partially mediated the effects of Group on WM performance. That is 1) WM-dependent right IPS–mPFC connectivity was more positive (less negative) with SET (Path a) and 2) more negative IPS–mPFC connectivity

predicted better WM performance (Path b), and these effects were strong enough to pass the mediation (a*b) test.

Discussion

Theoretical understanding of the functional neurobiology underlying SET effects on WM requires a model that directly links SET-induced brain activity to WM performance. This is the first study to provide direct neural evidence linking a social challenge to impaired cognition. Social evaluative threat impaired WM performance and reduced activity in regions known to support WM—bilateral dlPFC and superior parietal cortices/IPS. Activity in several of these regions, particularly IPS, positively predicted WM performance, and IPS reductions

formally mediated the effects of the stressor on WM. In addition, SET enhanced connectivity between IPS and vmPFC, while reducing connectivity between IPS and dlPFC. Finally, N-back performance created a negative relationship between IPS and dmPFC, but this negative relationship was blocked by SET. All of these results are consistent with a model in which vmPFC encodes processes related to affective self-monitoring (Northoff et al. 2006; Mason et al. 2007; Andrews-Hanna et al. 2010; Denny et al. 2012) (and related functions), which intrude on WM maintenance, reducing fronto-parietal activity and connectivity that support effective cognitive performance.

Toward a Brain-Based Account of Emotion–Cognition Interactions Underlying WM Performance

The strength of the results in IPS suggests that this region and its connectivity with lateral and medial frontal systems plays a more important role in social modulation of cognition than is widely appreciated. In accord with our findings, Liston et al. (2009) found that chronic social stress impairs attentional control, and this effect was linked to the disruption of a frontal–parietal circuit including dorsal posterior parietal cortex. Similarly, decline in WM performance as assessed by either verbal (Chee et al. 2006), Sternberg (Mu et al. 2005), or visual (Chee and Chuah 2007) WM tasks after sleep deprivation has consistently been associated with decrements in IPS activation. The IPS is reliably activated during several core WM processes like rehearsal, updating, and monitoring (Wager and Smith 2003; Champod and Petrides 2010) and basic attentional control processes like shifting attention (Wager et al. 2004; Esterman et al. 2009), which may determine WM capacity limitations (Todd and Marois 2004; Chee and Chuah 2007; Palva et al. 2010). Contemporary models of WM propose an attentional account of WM, in which attention is a fundamental, required building block for the higher-order executive functions underlying WM (Awh and Jonides 2001; Engle 2002; Postle 2006). Thus, our finding that reduced IPS activity mediates SET effects on WM performance suggests that SET is likely to disrupt a core attentional control process related to WM and perhaps other cognitive functions.

The DMN consistently deactivates in response to externally oriented tasks (Raichle et al. 2001) with brain regions whose activation has been negatively associated with task performance (Wager et al. 2013). Moreover, the DMN is typically "anti-correlated" (negatively correlated) with several regions normally activated as a group (a task positive network; TPN) by externally oriented tasks (Fox et al. 2005), and with regions associated with task performance (Wager et al. 2013). Interestingly, it has been suggested that switching between externally and internally oriented cognition is mediated via a competitive relationship between the anti-correlated network and DMN (Fox et al. 2005; 2007). But, not only the anti-correlation between these networks is predictive of adequate cognitive performance, connectivity in either the TPN or DMN is related to cognitive performance as well (Seeley et al. 2007). Generally, our data seem to support the idea that adequate cognitive performance is obtained not only by anti-correlated connectivity between the TPN and DMN but also by connectivity within these networks alone. Importantly, SET seems to alter these relationships in various but important ways.

In a first connectivity analysis, we found that SET reduced connectivity between IPS and dlPFC. These regions are not only part of the TPN but pose a subnetwork related to executive control, which can be established independent of any task (Seeley et al. 2007). In the same analysis, we found that SET enhanced connectivity between IPS and vmPFC. This can be interpreted as a decreased anti-correlation between the DMN and TPN, in line with previous findings (Kelly et al. 2008). Such a pattern of results provides a novel link between fronto-parietal activity and vmPFC, an area broadly associated with episodic memory, emotion, value, stress, and visceromotor function (see Roy et al. 2012 for review) and hypothesized to be a central orchestrator of affective meaning (Roy et al. 2012).

VmPFC activity is increased by self-referential monitoring and reduced by WM demand (Northoff et al. 2006; Mason et al. 2007; Andrews-Hanna et al. 2010; Denny et al. 2012) and is deactivated during SET (Wager, Van Ast, et al. 2009; Wager, Waugh, et al. 2009; Roy et al. 2012). Though our results are consistent with a model in which intrusive social/affective processes interfere with WM circuits (Iordan et al. 2013), mPFC activity likely has a complex relationship with affect generation and valuation, and there is evidence for both dorsal-ventral and laterality effects. Fear- and threat-related conditions tend to reliably increase activity in dorsal vmPFC/pregenual ACC and decrease activity in more ventral, medial orbitofrontal regions (Roy et al. 2012). Since, in this analysis, we did not test for interactions with N-back itself, these results suggest that SET generally reduced intrinsic connectivity in 2 important nodes of the TPN, while enhancing connectivity between the vmPFC and IPS, 2 important nodes of the DMN and TPN, respectively.

Therefore, we tested whether connectivity estimates varied as a function of the task in the second connectivity analysis, thereby looking for evidence that SET not merely alters general connectivity patterns, but that this alteration specifically resulted from an interaction between SET and WM activity. We found that during N-back performance, a negative relationship between IPS (part of the typical TPN) and dmPFC (part of the DMN) strengthened. Interestingly, this negative relationship was blocked by SET. It has been shown previously that the strength of the negative correlation between the TPN and DMN varies across individuals and is a predictor of behavioral performance (Kelly et al. 2008). Consistent with this work, we additionally found that the reduction in the negative relationship between dmPFC and IPS by SET mediated SET effects on WM.

Taken together, under conditions of SET, increased processing of affective information constructed in part in vmPFC competes with task-focused cognitive processes maintained in fronto-parietal circuits, ultimately resulting in impaired WM performance. Among these relationships, SET seems to affect both activity within TPN and DMN networks (Hampson et al. 2006; Eichele et al. 2008) and between these networks (Kelly et al. 2008), which can affect behavior in important ways. In addition, the data seem to suggest that SET can affect these relationships overall, but also in a task-dependent way. Future studies could prosper from adding a separate resting state and -n-back scan, to dissociate the 2 processes. The findings additionally suggest that intrinsic connectivity measures are not merely a trait-like entity but can be altered in a state-dependent way. Individual differences in connectivity both under control and SET conditions likely play an important role and perhaps can explain why recent neuroimaging studies into stress effects reported mixed results (Porcelli et al. 2008; Qin et al. 2009;

Weerda et al. 2010) and why SET can engage both "challenge" and "threat" responses (Kassam et al. 2009), and shift cognitive processing from executive processing-heavy strategies to faster, more heuristic ones (Beilock et al. 2004; Schwabe and Wolf 2009). See Figure 5B for the here-described interactions that we hypothesize may underlie competition between WM task performance and social threat.

Evidence for Complexity: Opposing Mediational Pathways

In addition to the negative mediators of a priori theoretical interest, we found several positive mediators: dmPFC, PCC, and the TPJ. Positive mediation in this context implies a brain effect that works against the overall, negatively signed effect of SET on WM. Such findings have been referred to as suppressor variables in the mediation literature (MacKinnon et al. 2000) and provide evidence for the complexity of SET–performance relationships by identifying multiple brain pathways that support and antagonize the main behavioral effect. Like vmPFC, dmPFC and PCC are part of the "default mode network" (DMN) (Fox et al. 2005; Raichle and Snyder 2007; Fransson and Marrelec 2008), which is commonly deactivated during cognitive effort. The right TPJ, identified as a key region in a "ventral attention system," is likely involved in reorienting attention to the external environment when behaviorally important stimuli are encountered (Corbetta and Shulman 2002) and may be involved in enhanced filtering of irrelevant information, ensuring that attentional resources are directed to task-relevant targets (Todd et al. 2005; Shulman et al. 2007). As in prior work (Anticevic et al. 2010), deactivation in these regions predicted better WM performance, irrespective of group. However, the SET group showed greater deactivation in these regions than controls, reflecting possible compensatory mechanisms or pro-performance effects of arousal or challenge under SET (Blascovich et al. 1999; Mendes et al. 2002; Hermans et al. 2011). Ultimately, modeling neural activity in multiple, opposing pathways may help explain the complex effects that stressors can have on performance.

The Role of the Neuroendocrine System in Performance under Stress

While this study can provide only limited data on the role of cortisol and related neuroendocrine responses on performance, our results are provocative in that they are consistent with a protective role for cortisol. Although cortisol was significantly increased by the social stress challenge, higher cortisol within the SET group predicted better performance. These findings are in line with other recent evidence suggesting that cortisol can have protective or compensatory effects (Henckens et al. 2010; 2011) and provide empirical evidence that constrain theories of brain–endocrine relationships. However, the cortisol results should be interpreted with caution; though cortisol increases were significant at the group level, the study was not powered to assess individual differences in cortisol–brain–performance relationships, and only a subset of the SET group showed a strong cortisol response. Further, larger-scale studies (including direct manipulation of cortisol) are needed to more comprehensively relate hormones, brain activity, and cognitive performance.

Limitations and Future Directions

Despite a relatively small sample size, due to the complex logistics of this procedure and limitations in grant funding, SET effects on WM were consistent across a wide range of effects. In addition, the definition of a priori ROIs as well as the pattern of interest analysis using a WM mask derived from Neurosynth reduced the amount of multiple comparisons and thereby somewhat reduced the number of participants required. For these reasons, we believe that the results are valid and reliable but should be replicated on a larger scale by future studies.

Another concern may be that the additional cognitive demand by the mental arithmetic in the SET group, or more generally, mental fatigue—referring to the effects people experience following and during the course of prolonged periods of demanding cognitive activity—rather than SET, may underlie the observed effects. So, the decreased activation in the prefrontal–parietal network might simply reflect higher levels of cognitive task demand, interference, or fatigue in the SET group as compared with the control group. However, the mental arithmetic occurred much earlier than the N-back, so to be a confound, it would have to produce long-lasting fatigue effects on later N-back performance during scanning. Also, the control task was designed to control for cognitive effort in the TSST, thereby singling out the SET central to social stressors. Finally, mental fatigue typically develops while performing hours of the very same task, whereas here different tasks were used, and we implemented sufficient rest in between tasks and scans. Together, we believe it is likely that it was uniquely SET causing the observed effects during N-back.

In the present study, both men and women were included for participation. Circulating estrogens (and consequently, use of oral contraceptives) are suggested to account for some of the observed sex differences in cortisol responses to stressors between men and women (Kirschbaum et al. 1999; Kudielka and Kirschbaum 2005), and sex differences in stress effects on WM have been revealed (Cornelisse et al. 2011). Moreover, animal research indicates that these differences may be explained by sex differences in estrogen (Shansky and Morrison 2009). Given that in the present study, an effect of cortisol on WM performance was observed, it is possible that estrogen may have played a role. Although all female participants were tested during the luteal phase of the menstrual cycle in order to control for variations in reactivity of the HPA axis throughout the cycle and to optimize direct comparison with male participants (Kirschbaum et al. 1999), we cannot exclude the possibility that sex differences may have existed. Sample size was too small to further delve into these matters, but sex differences in stress effects on memory remain an important consideration for future research.

The present results cannot address whether SET influences executive function in particular or attentional performance more broadly. Although our IPS results suggest that an attentional component of WM was compromised by SET, our study was not explicitly designed to disentangle separate WM components (such as encoding and rehearsal), and we do not have a direct behavioral measure of attention processing. Notably, we found comparable effects of SET on N-back trials thought to require high ("lures") or more modest ("targets") executive demand (Jonides et al. 1998; Gray et al. 2003), suggesting that WM impairments may not be limited to the conditions with the highest demand on executive function. However, we found

some preliminary evidence for stronger SET-induced reductions in 3-back than 2-back activity (see Supplemental material). Taken together, in addition to alterations in TPN and DMN networks due to SET, additional components of executive function may exist, of which some are more susceptible to SET effects than others, and future studies may explore this issue.

This study also cannot address which elements of the SET challenge were particularly instrumental in impairing WM. Here, we aimed to create an ecologically valid stressor composed of a variety of individual elements (e.g., threat to social status, uncontrollability). In addition to adrenergic and HPA-axis activation, the SET challenge can induce psychological changes such as a loss of self-esteem and negative emotional states such as shame, anxiety, and worry. The WM impairment is related to an unknown combination of these factors and may also be caused by alterations in participants' strategies during WM performance (e.g., Beilock et al. 2004). As with many ecological interventions, it is difficult to decompose these various effects; future studies may be able to test whether it is indeed possible to manipulate specific autonomic and experiential components of SET and still influence WM. Doing so would likely require a much larger-scale study than the present experiment.

Conclusions

In sum, these data provide an empirical basis for a model of competitive interactions between self-focused and task-focused cognition and add to a growing body of evidence suggesting that mental stress can affect cognitive performance and decision-making in fundamental ways. Ultimately, this brain-based analysis may help explain at a neurophysiological level why activating social and racial stereotypes can impair academic performance (Steele and Aronson 1995), why manipulations of negative social feedback can lead to suboptimal decision-making (Kassam et al. 2009), why resilience to stress leads to enhanced academic performance (Bardi et al. 2011), and why interventions targeting the social self can help (Cohen et al. 2009; Miyake et al. 2010).

Authors' Contributions

V.A.v.A., J.S., E.E.S., and T.D.W. designed the research; V.A.v.A., J.S., and S.S.G. performed the research; V.A.v.A., J.S., and T.D.W analyzed the data. V.A.v.A., J.S., E.E.S., I.L., J.L.A., and T.D.W. wrote the paper.

Supplementary Material

Funding

Notes

References

Andrews-Hanna JR, Reidler JS, Huang C, Buckner RL. 2010. Evidence for the default network's role in spontaneous cognition. J Neurophysiol. 104:322–335.

Anticevic A, Repovs G, Shulman GL, Barch DM. 2010. When less is more: TPJ and default network deactivation during encoding predicts working memory performance. Neuroimage. 49:2638–2648.

Atlas L, Bolger N, Lindquist M, Wager T. 2010. Brain mediators of predictive cue effects on perceived pain. J Neurosci. 30:12964.

Awh E, Jonides J. 2001. Overlapping mechanisms of attention and spatial working memory. Trends Cogn Sci. 5:119–126.

Baddeley A. 2003. Working memory: looking back and looking forward. Nat Rev Neurosci. 4:829–839.

Barch DM, Smith E. 2008. The cognitive neuroscience of working memory: relevance to CNTRICS and schizophrenia. Biol Psychiatry. 64:11–17.

Bardi M, Koone T, Mewaldt S, O'Connor K. 2011. Behavioral and physiological correlates of stress related to examination performance in college chemistry students. Stress. 14:557–566.

Baron RMR, Kenny DAD. 1986. The moderator-mediator variable distinction in social psychological research: conceptual, strategic, and statistical considerations. J Pers Soc Psychol. 51:1173–1182.

Beilock S, Kulp C, Holt L, Carr T. 2004. More on the fragility of performance: choking under pressure in mathematical problem solving. J Exp Psychol Gen. 133:584–600.

Blascovich J, Mendes WB, Hunter SB, Salomon K. 1999. Social "facilitation" as challenge and threat. J Pers Soc Psychol. 77:68–77.

Bledowski C, Rahm B, Rowe JB. 2009. What "works" in working memory? Separate systems for selection and updating of critical information. J Neurosci. 29:13735–13741.

Champod AS, Petrides M. 2010. Dissociation within the frontoparietal network in verbal working memory: a parametric functional magnetic resonance imaging study. J Neurosci. 30:3849–3856.

Chee MWL, Chuah YML. 2007. Functional neuroimaging and behavioral correlates of capacity decline in visual short-term memory after sleep deprivation. Proc Natl Acad Sci USA. 104:9487–9492.

Chee MWL, Chuah LYM, Venkatraman V, Chan WY, Philip P, Dinges DF. 2006. Functional imaging of working memory following normal sleep and after 24 and 35 h of sleep deprivation: correlations of fronto-parietal activation with performance. Neuroimage. 31:419–428.

Cohen GL, Garcia J, Purdie-Vaughns V, Apfel N, Brzustoski P. 2009. Recursive processes in self-affirmation: intervening to close the minority achievement gap. Science. 324:400–403.

Corbetta M, Shulman G. 2002. Control of goal-directed and stimulus-driven attention in the brain. Nat Rev Neurosci. 3:201–215.

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

Cox RW. 1996. AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res. 29:162–173.

Denny BT, Kober H, Wager TD, Ochsner KN. 2012. A meta-analysis of functional neuroimaging studies of self- and other judgments reveals a spatial gradient for mentalizing in medial prefrontal cortex. J Cog Neurosci. 24:1742–1752.

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

Eichele T, Debener S, Calhoun VD, Specht K, Engel AK, Hugdahl K, Yves von Cramon D, Ullsperger M. 2008. Prediction of human errors by maladaptive changes in event-related brain networks. Proc Natl Acad Sci USA. 105:6173–6178.

Elzinga BM, Roelofs K. 2005. Cortisol-induced impairments of working memory require acute sympathetic activation. Behav Neurosci. 119:98–103.

Engle RW. 2002. Working memory capacity as executive attention. Curr Dir Psychol. 11:19–23.

Esterman M, Chiu Y-C, Tamber-Rosenau BJ, Yantis S. 2009. Decoding cognitive control in human parietal cortex. Proc Natl Acad Sci USA. 106:17974–17979.

Fox M, Snyder A, Vincent J, Raichle M. 2007. Intrinsic fluctuations within cortical systems account for intertrial variability in human behavior. Neuron. 56:171–184.

Fox MD, Snyder AZ, Vincent JL, Corbetta M, Van Essen DC, Raichle ME. 2005. The human brain is intrinsically organized into dynamic, anticorrelated functional networks. Proc Natl Acad Sci USA. 102:9673–9678.

Fransson P, Marrelec G. 2008. The precuneus/posterior cingulate cortex plays a pivotal role in the default mode network: evidence from a partial correlation network analysis. Neuroimage. 42:1178–1184.

Gray J, Chabris C, Braver T. 2003. Neural mechanisms of general fluid intelligence. Nat Neurosci. 6:316–322.

Gusnard DA, Raichle ME, Raichle ME. 2001. Searching for a baseline: functional imaging and the resting human brain. Nat Rev Neurosci. 2:685–694.

Hampson M, Driesen NR, Skudlarski P, Gore JC, Constable RT. 2006. Brain connectivity related to working memory performance. J Neurosci. 26:13338–13343.

Henckens MJAG, van Wingen GA, Joëls M, Fernández G. 2011. Time-dependent corticosteroid modulation of prefrontal working memory processing. Proc Natl Acad Sci USA. 108:5801–5806.

Henckens MJAG, van Wingen GA, Joëls M, Fernández G. 2010. Time-dependent effects of corticosteroids on human amygdala processing. J Neurosci. 30:12725–12732.

Hermans EJ, van Marle HJF, Ossewaarde L, Henckens MJAG, Qin S, van Kesteren MTR, Schoots VC, Cousijn H, Rijpkema M, Oostenveld R et al. 2011. Stress-related noradrenergic activity prompts large-scale neural network reconfiguration. Science. 334:1151–1153.

Iordan AD, Dolcos S, Dolcos F. 2013. Neural signatures of the response to emotional distraction: a review of evidence from brain imaging investigations. Front Hum Neurosci. 7:200.

Jonides J, Smith E, Marshuetz C, Koeppe R, Reuter-Lorenz P. 1998. Inhibition in verbal working memory revealed by brain activation. Proc Natl Acad Sci USA. 95:8410.

Kassam K, Koslov K, Mendes W. 2009. Decisions under distress. Psychol Sci. 20:1394.

Kastner S, Pinsk MA, De Weerd P, Desimone R, Ungerleider LG. 1999. Increased activity in human visual cortex during directed attention in the absence of visual stimulation. Neuron. 22:751–761.

Kelly AMC, Uddin LQ, Biswal BB, Castellanos FX, Milham MP. 2008. Competition between functional brain networks mediates behavioral variability. Neuroimage. 39:527–537.

Kirschbaum C, Kudielka BM, Gaab J, Schommer NC, Hellhammer DH. 1999. Impact of gender, menstrual cycle phase, and oral contraceptives on the activity of the hypothalamus-pituitary-adrenal axis. Psychosom Med. 61:154–162.

Kirschbaum C, Pirke KM, Hellhammer DH. 1993. The "trier social stress test" – a tool for investigating psychobiological stress responses in a laboratory setting. Neuropsychobiology. 28:76–81.

Kudielka B, Kirschbaum C. 2005. Sex differences in HPA axis responses to stress: a review. Biol Psychol. 69:113–132.

Lepsien J, Griffin I, Devlin J, Nobre A. 2005. Directing spatial attention in mental representations: interactions between attentional orienting and working-memory load. Neuroimage. 26:733–743.

Lieberman MD, Cunningham WA. 2009. Type I and Type II error concerns in fMRI research: re-balancing the scale. Soc Cogn Affect Neurosc. 4:423–428.

Liston C, Mcewen BS, Casey BJ. 2009. Psychosocial stress reversibly disrupts prefrontal processing and attentional control. Proc Natl Acad Sci USA. 106:912–917.

Luethi M, Meier B, Sandi C. 2008. Stress effects on working memory, explicit memory, and implicit memory for neutral and emotional stimuli in healthy men. Front Beh Neurosci. 2:5. doi: 10.3389/neuro.08.005.2008

MacKinnon D, Krull J, Lockwood C. 2000. Equivalence of the mediation, confounding and suppression effect. Prev Sci. 1:173–181.

MacKinnon DP, Fairchild AJ, Fritz MS. 2007. Mediation analysis. Annu Rev Psychol. 58:593–614.

Mason MF, Norton MI, Horn JDV, Wegner DM, Grafton ST, Macrae CN. 2007. Wandering minds: the default network and stimulus-independent thought. Science. 315:393–395.

Mayer J, Bittner R, Nikolic D, Bledowski C, Goebel R. 2007. Common neural substrates for visual working memory and attention. Neuroimage. 36:441–453.

Mendes W, Blascovich J, Lickel B, Hunter S. 2002. Challenge and threat during social interactions with White and Black men. Pers Soc Psychol B. 28:939–952.

Miyake A, Kost-Smith LE, Finkelstein ND, Pollock SJ, Cohen GL, Ito TA. 2010. Reducing the gender achievement gap in college science: a classroom study of values affirmation. Science. 330:1234–1237.

Mu Q, Nahas Z, Johnson K, Yamanaka K, Mishory A. 2005. Decreased cortical response to verbal working memory following sleep deprivation. Sleep. 28:55–67.

Northoff G, Heinzel A, de Greck M, Bermpohl F, Dobrowolny H, Panksepp J. 2006. Self-referential processing in our brain – a meta-analysis of imaging studies on the self. Neuroimage. 31:440–457.

Oei N, Everaerd W, Elzinga B, 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.

O'Reilly JX, Woolrich MW, Behrens TEJ, Smith SM, Johansen-Berg H. 2012. Tools of the trade: psychophysiological interactions and functional connectivity. Soc Cogn Affect Neurosci. 7:604–609.

Owen AM, McMillan KM, Laird AR, Bullmore E. 2005. N-back working memory paradigm: a meta-analysis of normative functional neuroimaging studies. Hum Brain Mapp. 25:46–59.

Palva JM, Monto S, Kulashekhar S, Palva S. 2010. Neuronal synchrony reveals working memory networks and predicts individual memory capacity. Proc Natl Acad Sci USA. 107:7580–7585.

Porcelli A, Cruz D, Wenberg K, Patterson M, Biswal B, Rypma B. 2008. The effects of acute stress on human prefrontal working memory systems. Physiol Behav. 95:282–289.

Postle BR. 2006. Working memory as an emergent property of the mind and brain. Neuroscience. 139:23–38.

Pruessner JC, Dedovic K, Khalili-Mahani N, Engert V, Pruessner M, Buss C, Renwick R, Dagher A, Meaney MJ, Lupien S. 2008. Deactivation of the limbic system during acute psychosocial stress: evidence from positron emission tomography and functional magnetic resonance imaging studies. Biol Psychiatry. 63:7.

Qin S, Hermans EJ, van Marle HJF, Luo J, Fernández G. 2009. Acute psychological stress reduces working memory-related activity in the dorsolateral prefrontal cortex. Biol Psychiatry. 66:25–32.

Raichle ME, MacLeod AM, Snyder AZ, Powers WJ, Gusnard DA, Shulman GL. 2001. A default mode of brain function. Proc Natl Acad Sci USA. 98:676–682.

Raichle ME, Snyder AZ. 2007. A default mode of brain function: a brief history of an evolving idea. Neuroimage. 37:1083–1090; discussion 1097–discussion 1099.

Roy M, Shohamy D, Wager TD. 2012. Ventromedial prefrontal-subcortical systems and the generation of affective meaning. Trends Cogn Sci. 16:147–156.

Schoofs D, Preuß D, Wolf O. 2008. Psychosocial stress induces working memory impairments in an n-back paradigm. Psychoneuroendocrinology. 33:643–653.

Schwabe L, Wolf OT. 2009. Stress prompts habit behavior in humans. J Neurosci. 29:7191–7198.

Seeley WW, Menon V, Schatzberg AF, Keller J, Glover GH, Kenna H, Reiss AL, Greicius MD. 2007. Dissociable intrinsic connectivity networks for salience processing and executive control. J Neurosci. 27:2349–2356.

Shansky RM, Morrison JH. 2009. Stress-induced dendritic remodeling in the medial prefrontal cortex: effects of circuit, hormones and rest. Brain Res. 1293:108–113.

Shulman GL, Astafiev SV, McAvoy MP, d'Avossa G, Corbetta M. 2007. Right TPJ deactivation during visual search: functional significance and support for a filter hypothesis. Cereb Cortex. 17:2625–2633.

Steele CM, Aronson J. 1995. Stereotype threat and the intellectual test performance of African Americans. J Pers Soc Psychol. 69:797–811.

Todd JJ, Fougnie D, Marois R. 2005. Visual short-term memory load suppresses temporo-parietal junction activity and induces inattentional blindness. Psychol Sci. 16:965–972.

Todd JJ, Marois R. 2004. Capacity limit of visual short-term memory in human posterior parietal cortex. Nature. 428:751–754.

Vul E, Harris C, Winkielman P, Pashler H. 2009. Puzzlingly high correlations in fMRI studies of emotion, personality, and social cognition. Persp Psych Sci. 4:274–290.

Wager T, Jonides J, Reading S. 2004. Neuroimaging studies of shifting attention: a meta-analysis. Neuroimage. 22:1679–1693.

Wager TD, Keller MC, Lacey SC, Jonides J. 2005. Increased sensitivity in neuroimaging analyses using robust regression. Neuroimage. 26:99–113.

Wager TD, Atlas LY, Lindquist MA, Roy M, Woo C-W, Kross E. 2013. An fMRI-based neurologic signature of physical pain. N Engl J Med. 368:1388–1397.

Wager TD, Smith EE. 2003. Neuroimaging studies of working memory: a meta-analysis. Cogn Affect Behav Neurosci. 3:255–274.

Wager TD, Van Ast VA, Hughes BL, Davidson ML, Lindquist MA, Ochsner KN. 2009. Brain mediators of cardiovascular responses to social threat, part II: prefrontal-subcortical pathways and relationship with anxiety. Neuroimage. 47:836–851.

Wager TD, Waugh CE, Lindquist M, Noll DC, Fredrickson BL, Taylor SF. 2009. Brain mediators of cardiovascular responses to social threat: part I: reciprocal dorsal and ventral sub-regions of the medial prefrontal cortex and heart-rate reactivity. Neuroimage. 47:821–835.

Watson D, Clark LA, Tellegen A. 1988. Development and validation of brief measures of positive and negative affect: the PANAS scales. J Pers Soc Psychol. 54:1063–1070.

Weerda R, Muehlhan M, Wolf OT, Thiel CM. 2010. Effects of acute psychosocial stress on working memory related brain activity in men. Hum Brain Mapp. 31:1418–1429.

Yarkoni T, Poldrack RA, Nichols TE, Van Essen DC, Wager TD. 2011. Large-scale automated synthesis of human functional neuroimaging data. Nat Methods. 8:1–10.

Zhang J, Mueller S. 2005. A note on ROC analysis and non-parametric estimate of sensitivity. Psychometrika. 70:203–212.