The Fatigued Brain: Fatigue and Cognitive Functions in Young Adults

Plukaard, S.C.

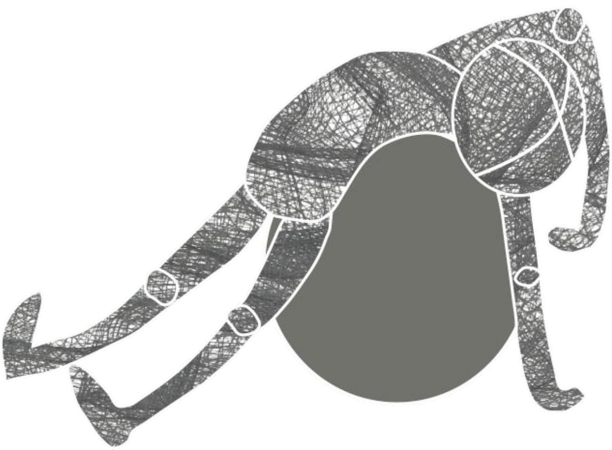
2015

CHAPTER SIX

Medical Students with Chronic Fatigue Show Declines in Performance and Neural Activation at Higher Working Memory Load

Sarah Plukaard, Dick J. Veltman, Lydia Krabbendam & Jelle Jolles

(submitted)



ABSTRACT

Fatigue is common in the general population and is associated with cognitive decline. This study investigated whether fatigued individuals are characterized by altered working memory capacity. To this end, we compared 26 chronically fatigued medical students to 26 peers without fatigue (mean age: 22 years). The participants performed a working memory task that consisted of a low, intermediate, and high load condition. Both groups were additionally subjected to a fatigue manipulation and a non-fatiguing control manipulation to evaluate effects of an acute fatigue induction. As expected, the pattern of load-dependent activation was different between students with and without chronic fatigue. At the highest load condition, fatigued students showed reduced activation compared to non-fatigued students in the prefrontal cortex and superior parietal cortex. Correspondingly, fatigued students responded slower at intermediate and high load. Only the group of fatigued students showed a correlation between accuracy and activation of the prefrontal cortex and anterior insula/ frontal operculum. Lastly, the fatigue manipulation resulted in overall reduced accuracy in both groups and did not result in altered brain activation. Altogether, the results indicate that fatigued students differ from their non-fatigued peers in working memory processing.

INTRODUCTION

Fatigue in healthy individuals as well as in individuals with a medical condition has been associated with compromised cognitive control (e.g., Capuron et al., 2006; DeLuca, 2005; Lorist, Boksem, & Ridderinkhof, 2005). An important system underlying cognitive control is working memory (Baddeley, 2003). The present study therefore examined fatigue in the context of working memory processing. We anticipated that functional magnetic resonance imaging (fMRI) would be sensitive to fatigue-related differences in a healthy young intelligent population, such as medical students. This population is considerably affected by fatigue in terms of high fatigue ratings (Plukaard, Van Batenburg-Eddes, Vos, Croiset, & Jolles, submitted; Tanaka, Fukuda, Mizuno, Kuratsune, & Watanabe, 2009; Tanaka, Mizuno, Fukuda, Shigihara, & Watanabe, 2008) and prevalence of burnout (Galán, Sanmartín, Polo, & Giner, 2011; Guthrie et al., 1998). An advantage of studying working memory is that one can differentiate between load conditions. This allowed us to investigate whether effects of fatigue would become manifest at particular levels of working memory load.

Typically, fatigued individuals report problems with attention and concentration, are more prone to making mistakes, require more effort for everyday accomplishments, or fail to continue with desired activities. As such, individuals who experience high fatigue over a prolonged period of time may encounter more difficulties with everyday functioning at school, university, at work or in their social life. In the long run, these difficulties may lead to reduced work productivity or, in case of students, lower grades. This is for instance illustrated by a study that associated fatigue with reduced academic performance (Nagane, 2004). We anticipate a cognitive basis for these problems in terms of less efficient or less effective information processing.

Many decades ago, fatigue was labeled as a 'central control problem' (Bartlett, 1943). Correspondingly, fatigue is recurrently associated with reductions in cognitive control, whereas automatic processes typically remain unaffected (Schellekens, Sijtsma, Vegter, & Meijman, 2000; van der Linden & Eling, 2006), emphasizing the distinction made by

Schneider and Shiffrin (1977) between so-called 'automatic processes' and 'controlled processes'. Automatic processes occur reflexively whereas controlled processes are effortful, occur sequentially and depend on limited resource capacity (Schneider & Shiffrin, 1977; Schneider, 2003). Accordingly, fatigue may correspond to deficits in resource capacity. Working memory is commonly characterized by its limited capacity and is considered to be at the core of cognitive control (Baddeley, 2003). Less effective or efficient working memory function may thus underlie a wide range of difficulties that are encountered by fatigued individuals.

Working memory refers to the ability to temporarily maintain and manipulate information in memory (Baddeley, 2003; Mark D'Esposito, 2001). The capacity of working memory can be demonstrated with functional brain imaging as it engages areas in the fronto-parietal circuit in which the level of activation depends on task load (Callicott et al., 1999; Curtis & D'Esposito, 2003). That is, activation in working memory areas increases in correspondence with increasing working memory load. A less prominent increase or a decrease in one or more of these areas in response to higher working memory load is considered an indication of working memory constraints, in particular when this is accompanied by performance declines (i.e., lower accuracy or longer reaction times; Goldberg et al., 1998).

Differences in the pattern of load-dependent brain activation may thus provide insight into working memory capacity. Evidence from aging studies illustrates such differences with increased neural activity at lower working memory load in older adults compared to young adults (Reuter-Lorenz & Cappell, 2008). In addition, load-dependent increases in activation appear to level off in older adults at lower levels of working memory (Mattay, et al., 2006; Nagel et al., 2009; Nyberg et al., 2009). Over-activation is not unique to aging. Instead, it is also observed in low performers versus high performers (Jaeggi et al., 2007), in the early stages of skill acquisition (Church, Coalson, Lugar, Petersen, & Schlaggar, 2008), or in relation to any other extrinsic or intrinsic cognitive challenge, such as task demands or structural and metabolic differences (Park & Reuter-Lorenz, 2009). If fatigue coincides with differences in working memory capacity, one possible expectation thus relates to stronger or more widespread brain activation, particularly at lower levels of working memory load. This would be consistent with neuroimaging research in patient populations demonstrating that fatigue corresponds to elevated levels of neural activation (Chaudhuri & Behan, 2004; Cook, O'Connor, Lange, & Steffener, 2007; DeLuca, Genova, Capili, & Wylie, 2009a; DeLuca, Genova, Hillary, & Wylie, 2008). Higher brain activation may reflect increased effort expenditure to maintain adequate levels of performance, which could explain the experience of fatigue (see DeLuca, 2005).

There are several indications for compromised working memory in relation to fatigue. First, prevalence of fatigue in elementary and high school students has been associated with decreases in behavioral measures of working memory (Mizuno, Tanaka, Fukuda, Imai-Matsumura, & Watanabe, 2011). Moreover, sleep deprivation was found to result in reduced working memory span (Engle, 2010), which in another study was accompanied by reduced inferior parietal and inferior frontal activation (Choo, Lee, Venkatraman, Sheu, & Chee, 2005). Increased activation in working memory areas as well as recruitment of additional prefrontal cortex (PFC) regions has been observed in chronic fatigue syndrome (CFS) patients in the absence of performance differences (Lange et al., 2005). Such increases appear to depend on load, as Caseras et al. (2006) observed diminished activation in frontal and parietal areas at higher working memory load in CFS patients (Caseras et al., 2006). It is unclear whether fatigue in the healthy population would correspond to similar patterns of working memory processing. Such knowledge could generate possible explanations as to the perception of fatigue or the cognitive difficulties in fatigued individuals and associated long-term problems such as reduced academic performance.

The present study investigated whether fatigued students are characterized by differences in working memory capacity based on the pattern of activity in load-dependent brain regions. There is little evidence for a relation between fatigue and working memory in individuals from the healthy population, nor is there a clear understanding of whether a period of cognitively demanding activities, which typically results in higher fatigue (e.g., Boksem, Meijman, & Lorist, 2006; Lorist et al., 2000; van der Linden, Frese, & Meijman, 2003), would affect working memory in a similar fashion. The current sample of participants included medical students who are generally highly homogeneous with respect to motivation and intelligence. Furthermore, all participants were females. The choice for using only females was based upon recent studies that show differences between males and females in brain structure as well as function (Lenroot & Giedd, 2010), which has also been observed in medical students (Veroude, Jolles, Croiset, & Krabbendam, 2014). The participants performed a working memory task that distinguished between a low (1 item), intermediate (3 items), and high (4 items) load condition. We compared medical students with prolonged fatigue to students without fatigue. 'Fatigue' was operationalized in terms of severity, i.e., fatigue ratings exceeding a certain threshold, as well as duration, i.e., high fatigue of longer than two months duration, and is hereafter termed 'chronic fatigue'. Second, we studied the effect of an acute cognitive fatigue manipulation in both groups. Prior research has indicated that group differences in working memory can depend on such a manipulation (Klaassen et al., 2014). Moreover, both operationalizations of fatigue in the same group coincided with increased neural activation during task switching (Plukaard, Veltman, Krabbendam & Jolles, submitted). Correspondingly, we predicted that fatigued medical students would show a stronger increase in brain activity at lower working memory load in load-dependent areas.

METHOD

Participants

This study included 26 chronically fatigued females (age range 19.0 – 30.6 years, M = 21.8, SD = 2.3) and 26 non-fatigued females (age range 18.8 – 28.2 years, M = 21.6, SD = 1.9). All participants were right-handed and reported no history of medical, neurological or psychiatric disorder, and no use of medication (apart from contraceptive drugs).

Participants were recruited based on data form a survey that was distributed among all first, second and third year medical students at the VU University Medical Center Amsterdam. Out of 1050 eligible students, 701 students returned the completed questionnaire (67% response rate). Students who reported fatigue complaints for longer than 2 months and who scored above 76 on the Checklist Individual Strength (CIS; Vercoulen et al., 1994) were selected for the *fatigue group*. The cutoff point of >76 was determined by Bültmann and colleagues (Bultmann et al., 2000) and indicates a fatigue level at which individuals are at risk of subsequent sickness absence. Of the 701 survey respondents, 31% scored above this threshold. Students who reported no fatigue complaints and who scored below 65 on the CIS (i.e., the mean score of all respondents) were selected for the *control group*.

All participants gave written informed consent prior participation and received financial compensation. The Medical Ethics Committee of the VU University Medical Center Amsterdam approved this study.

Procedure

All participants completed a 1.5 h. practice session and two 3 h. test sessions on three separate days. The practice session was scheduled within a week before the first test session. During the practice session, participants completed a self-report depression scale (CES-D; Radloff, 1977), the CIS (once more) and several neuropsychological tests that served to compare the groups on basic cognitive abilities. These tests included the 20-minute version of the Raven Advanced Progressive Matrices Test (RAPM; Hamel & Schmittmann, 2006; Raven, 2000), the fourth version of the Peabody Picture Vocabulary Test (PPVT-IV; Dunn & Dunn, 2005) as measures of non-verbal and verbal intelligence, the Letter Digit Substitution Task (LDST; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006[a]) as a measure of information processing speed, the Concept Shifting Task (CST; Van der Elst, Van Boxtel, Van Breukelen, & Jolles, 2006[b]) and the Stroop Color-Word Test (Stroop, 1935) as measures of cognitive flexibility. Lastly, the participants practiced three fMRI tasks.

The test sessions took place on weekend days with one or two weeks apart, and started at 14:00 or at 15:30 h. The test sessions commenced outside the scanner with a fatigue manipulation in one session (i.e., fatigue session) and a control manipulation in the other session (i.e., control session; the order of the sessions was randomized). Apart from the manipulation, both test sessions were identical. During the fatigue manipulation, participants performed 20 minutes of mental arithmetic, followed by 20 minutes of brainteaser puzzles (such as arithmetic sequences and syllogisms), 20 minutes of a computerized Stroop task with extra auditory interference, adopted from Evers and colleagues (Evers, van der Veen, Jolles, Deutz, & Schmitt, 2009) and an N-back computer task (2- and 3-back) for 30 minutes (see also Klaassen et al., 2013, 2014 for details). During the control manipulation, participants spent 1.5 h. reading magazines or watching documentaries (a collection of magazines and documentary style DVD's was provided). Afterwards, the participants were transferred as quickly as possible to the scanner in which they performed the working memory task and two other tasks that are described elsewhere.

Questionnaires

The CIS consists of 20 statements that measure different aspects of fatigue: fatigue severity (8 items, e.g., "I feel tired"); concentration (5 items, e.g., "my thoughts easily wander off"); motivation (4 items, e.g., "I'm looking forward to many fun things to do"); and physical activity (3 items, e.g., "I don't do much during the day"). Respondents were instructed to indicate how they felt during the previous two weeks. Responses were scored on a 7-point Likert scale, ranging from 1 "Yes, that is true", to 7 "No, that is not true". The sum of all scales yields a fatigue score ranging from 20 to 140. In addition, participants were asked to indicate whether they suffered from fatigue for longer than two weeks. If they answered this question with yes, they were asked to specify how long they experienced fatigue by choosing one of the following options: "from two weeks to one month", "from one to two months", "from two to three months", "from three to four months" or "longer than four months".

The Rating Scale Mental Effort (RSME; Zijlstra, 1993) was administered to assess the effect of the manipulation on subjective levels of fatigue and mental effort. The scale was completed before (T0) and after the manipulation (T1), as well as after scanning (T2). The RSME contains seven visual analog scales (range: 0 – 150) that measure different aspects of fatigue, including required effort for focusing of attention, tiredness and boredom (e.g., "How much effort does it take to suppress feelings of boredom?"). The RSME scores were based on the mean of the seven scales and ranged from 0 to 150.

Working memory task

The working memory task was adopted from Smeets, Vuurman, Van Boxtel, Evers, & Jolles (2010) and programmed in Eprime version 1.2, running under windows XP on a HP Compaq Desktop PC (Intel Core 2 processor, 17 inch 60 Hz monitor). Stimuli were back-projected onto a screen located behind the scanner. Participants viewed the screen through an angled mirror attached to the head coil.

Each trial of the task started with a white fixation cross, which remained on the screen for 1000-1500 ms (pseudorandomly varied in steps of 10 ms), followed by a colored dot (1000 ms) that varied in color (red, green, blue, yellow, pink, brown and turquoise. All stimuli were presented centrally on a gray background. During the entire task, one response button was used. In the baseline condition (*load 0*), participants viewed a sequence of dots that could have any of the above-mentioned colors. They were instructed to respond to each dot, regardless of its color, by pressing the button as soon as the dot appeared on the screen. In the load 1, load 3 and load 4 conditions, responses were based on the color of the dots. In the *load 1* condition, dots of only two different colors were presented. Participants were instructed to press the button for only one of the two colors. In the *load 3* condition, three colors were added resulting in five differently colored dots. Participants were instructed to respond to three of five colors. The *load 4* condition consisted of all seven colors and participants were instructed to respond to four of the seven colors. Two, five and seven items for each load condition were based on a pilot study by Smeets et al., (2010) in which performance differences between load conditions were observed using these numbers. All three load-conditions consisted of 50% target trials to average out no-go effects (i.e., inhibitory processes) between the conditions.

Conditions were presented in blocks of 12 trials each. The duration of each block was approximately 25 seconds and started with an instruction screen indicating which colored dots to respond to. The blocks were presented in sequences with fixed order of increasing difficulty (load 0, load 1, load 3, load 4). The colors used for targets and non-targets in load 1 or load 3 were always used in the same way in the consecutive load condition (i.e., load 3 or load 4). Six sequences were presented in random order, resulting in a total of 72 trials per condition. Total duration of the task was approximately 12 minutes.

All participants responded with their right hand and were instructed to respond as fast and accurately as possible.

fMRI Data Acquisition

Imaging data were collected on a GE Signa HDxt 3.0-Tesla MRI-scanner with an 8-channel head coil (General Electric, Milwaukee, Wisconsin) at the VU University Medical Center. Functional images were acquired with 40 slices in ascending order using a T2*-weighted echo planar imaging (EPI) sequence (repetition time (TR) = 2100 ms, echo time (TE) = 30 ms, flip angle (FA) = 80°, field of view (FOV) = 22.4 x 22.4 cm, voxel size = 3.5 x 3.5 x 3 mm). A T1-weigthed anatomical scan with 172 slices was acquired for co-registration and normalization (TR = 8.2 ms, TE = 3.2 ms, FA = 12°, FOV = 25.6 x 25.6 cm, voxel size = 1 x 1 x 1 mm).

Questionnaire and Behavioral Data Analysis

Independent *t*-tests were carried out to compare the groups on descriptive variables, questionnaire scores and neuropsychological tests. Only for the CES-D and PPVT non-parametric Mann-Whitney *U* tests were used, because these variables were non-normally distributed. RSME scores were analyzed with a group (fatigue, control) x manipulation (fatigue, control) x time of assessment (T0, T1 and T2) ANOVA.

Accuracy and reaction times (RT) were recorded during the working memory task. Responses faster than 120 ms, and extreme outliers (responses slower than 2.5 standard deviations above the mean) were excluded from the analyses. For the RT analyses, we excluded all error trials. Mean accuracy scores (i.e., fractions correct) and RTs of each load condition were compared in separate group x manipulation x load (load 0, load 1, load 3, load 4) mixed repeated measures ANOVAs.

fMRI Data Analysis

Imaging data were analyzed using Statistical Parametric Mapping software (SPM8; www.fil.ion.ucl.ac.uk/spm), implemented in Matlab (The Mathworks, Inc.). The following preprocessing steps were applied: reorienting, realignment and unwarping, slice time correction, coregistration, normalization into MNI standard space, reslicing to voxels of 3 x 3 x 3 mm, and spatial smoothing with an 8 mm full width at half maximum (FWHM) isotropic Gaussian kernel.

At first level, a block-design was used to model trials for each load condition, with boxcar regressors convolved with a hemodynamic response function. A high-pass filter of 128 s cutoff was applied to remove low frequency noise. In case of motion artifacts (i.e., spikes of > 0.5 mm movement between scans), additional regressors were computed (i.e., one regressor per spike containing all zero's and a '1' at the location of the spike) and added to the model as regressors of no interest. For each participant in each session, contrasts were computed to evaluate neural activation that correlated with increasing load (load 4 > load 3 > load 1 > load 0). In addition, contrasts for each load condition versus baseline were computed (i.e., load 1 > load 0; load 3 > load 0; load 4 > load 0).

The first level contrasts were entered into second level mixed full factorial ANOVAs with group (two levels) as independent factor and manipulation (two levels) as dependent factor. To include performance differences in the model, accuracy was added as covariate, which we allowed to interact with group. For the contrast evaluating increased activation over all load conditions we added overall accuracy and for the contrasts per load condition we added accuracy on the corresponding condition. All second level contrasts were masked by the positive effect of condition. For the effect of load averaged over group and manipulation, significance level was thresholded at $p < .05$ following a Family Wise Error rate (FWE) correction for multiple comparisons. For group and manipulation differences, the FWE correction was too conservative and we therefore used a Monte Carlo simulation of brain volume (Slotnick, Moo, Segal, & Hart, 2003) to obtain appropriately corrected results. Assuming an individual voxel type I error of $p < .005$, a cluster extent of 30 contiguous resampled voxels (3 x 3 x 3 mm) was required to correct for multiple comparison at $p < .05$.

To evaluate the relationship between brain activation and working memory performance, we investigated areas in which load-dependent activation increases correlated with accuracy. The mean activity within the region showing a group interaction was extracted using the data extraction tool available in MarsBaR (http://marsbar.sourcefourge.net) and correlated with performance accuracy using Spearman's rank correlations.

RESULTS

Sample Characteristics

Table 1 displays the sample characteristics. The fatigue group scored significantly higher than the control group on all subscales of the CIS and on the CES-D. The two groups did not differ on age and neuropsychological test scores ($p < .0073032$, Bonferroni corrected for 13 tests with a mean correlation coefficient of $\rho = .25$; http://www.quantitativeskills.com/sisa/calculations/bonfer.htm).

Table 1

Sample Characteristics

		Fatigue group	Control group	*p*
N		26	26	
Age		21.8 (2.3)	21.6 (1.9)	.686
CIS score	total	85.1 (15.5)	39.1 (9.7)	<.001
	fatigue severity	38.4 (7.2)	16.6 (5.2)	<.001
	concentration	20.2 (6.5)	10.1 (3.9)	<.001
	motivation	12.8 (4.1)	6.3 (1.7)	<.001
	physical activity	13.8 (4.4)	6.2 (3.6)	<.001
CES-D		14.0 (7.5) [a]	4.2 (3.5)	<.001
Hb (nmol/l)		8.4 (1.0) [b]	8.5 (0.7) [a]	.699
RAVEN		23.3 (3.1)	22.9 (3.8)	.633
PPVT		108.9 (7.4)	109.8 (6.8)	.898
LDST		45.4 (4.4)	46.5 (4.9)	.497
CST		6.4 (4.8) [a]	4.1 (3.5)	.057
STROOP interference		29.1 (9.2) [a]	22.7 (7.6)	.010

Note. * Indicates significant group difference with $p < .0038$ (Bonferroni correction).

[a] Data of 1 person is missing (N=25); [b] Data of 2 persons is missing (N=24).

Manipulation Check

Subjective fatigue ratings are illustrated in Figure 1. A significant manipulation x time of assessment interaction ($F(2,100)=21.35$, $p < .001$) indicated that fatigue ratings increased more during the fatigue session compared to the control session. Post hoc *t*-tests indicated that subjective feelings of fatigue and mental effort differed significantly between sessions at T1 ($t(51) = 6.67$, $p < .001$, $r = .34$) and not at T0 and T2 ($t < 0.95$, $p > .346$, $r < .14$).

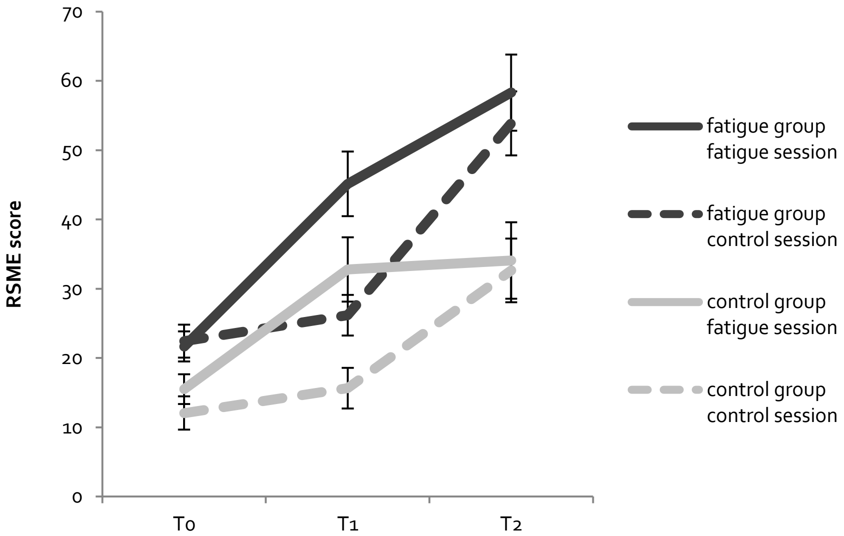


Figure 1. Fatigue scores measured by the Rating Scale of Mental Effort (RSME) at baseline (T0), after the manipulation (T1) and after scanning (T2). Dark gray represents the fatigue group and light gray represents the control group with solid lines for the fatigue manipulation and dashed lines for the control manipulation. The scores differed significantly between sessions only at T1. Error bars show standard error of the mean (SEM).

Behavioral Results

Accuracy. Behavioral results are illustrated in Figure 2. Accuracy decreased as a function of load ($F(3,150) = 20.90$, $p < .001$, $r = .35$) with post hoc *t*-tests confirming that apart from 'load 0 vs load 1' and 'load 3 vs load 4' ($p > .999$) all load conditions differed significantly from each other ($p < .005$ for all other comparisons, Bonferroni corrected). A main effect of manipulation indicated that across groups and load, accuracy decreased after the fatigue manipulation ($F(1,50) = 5.39$, $p < .05$, $r = .31$). No other effects of fatigue on accuracy were observed ($p > .178$, $r < .12$ for all other comparisons).

Reaction times (RT). The assumption of sphericity was violated for the different load conditions with an epsilon of < .70. Therefore, RTs were analyzed with multivariate tests (i.e., MANOVA), which do not assume sphericity. Moreover, the sample size was reasonably large ($n > 10 + k$; $n = 52$; $k = 4$), which renders MANOVA more powerful than univariate tests (Stevens, 2012). Higher working memory load was associated with longer RTs ($F(3,48) = 393.30$, $p < .001$, Pillai's Trace = .96, $r = .24$). Post hoc t-tests indicated that all load conditions differed significantly from each other ($p < .001$ for all comparisons). A main effect of group on RT data indicated that the fatigue group responded slower than the control group ($F(4,47) = 3.00$, $p < .05$, Pillai's Trace = .20, $r = .25$). This group effect depended on load as indicated by an interaction between group and load ($F(3,48) = 2.82$, $p < .05$, Pillai's Trace = .15, $r = .24$). Figure 2B shows that the groups differed most in the higher load conditions, which was confirmed by post-hoc *t*-tests ($p = .135$ for load 0, $p < .05$ for load 1 and load 4, $p < .005$ for load 3, all *p*-values are Bonferroni corrected). No other effects of fatigue were found on the RT data ($p > .77$ for all other comparisons).

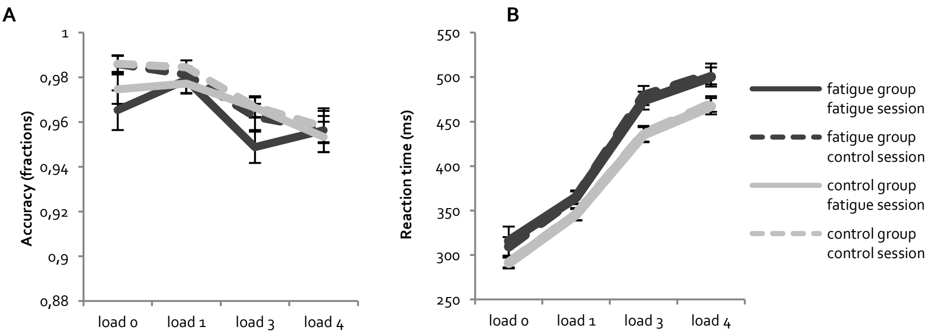


Figure 2. Behavioral results. This figure displays behavioral performance for each load condition. Dark gray represents the fatigue group and light gray represents the control group with solid lines for the fatigue manipulation and dashed lines for the control manipulation. Error bars show SEM. (**A**) Accuracy scores; overall, participants were less accurate after the fatigue manipulation compared to the control manipulation. (**B**) RT scores; the fatigue group responded slower than the control group, particularly in the higher load conditions.

fMRI Results

Effects of working memory load. Across group and manipulation, a robust effect of working memory load (load 4 > load 3 > load 1 > load 0) was found (summarized in Table 2). A bilateral fronto-parietal network showed increased activation with increasing working memory load. Activated areas included the dorsolateral prefrontal cortex (DLPFC), anterior cingulate cortex (ACC), premotor cortex, precuneus, and superior parietal lobule (SPL).

Table 2

Brain regions showing increased activation as a function of load (i.e., load 4 > load 3 > load 1 > load 0)

Peak activation area	BA	x	y	z	*z*	k
Angular Gyrus	39	-30	-55	40	> 8	585
Precuneus	7	9	-64	43	6.74	
Middle Frontal Gyrus	9	-42	20	25	> 8	848
Precentral Gyrus	6	-39	2	31	> 8	
Insula		-30	23	1	> 8	
Medial Frontal Gyrus	6	-6	11	52	> 8	314
Precuneus	19	33	-58	43	> 8	166
Insula		33	26	-2	6.93	64
Middle Frontal Gyrus	9	42	29	22	6.36	70
Inferior Occipital Gyrus	17	-18	-97	-5	5.49	67
Lingual Gyrus	17	-12	-100	1	5.10	
Cuneus	17	15	-97	10	5.27	26
Lingual Gyrus	17	21	-97	1	4.83	

Note. Activations are thresholded at $p < .05$, FWE corrected. Coordinates are in MNI space. BA = Brodmann Area.

Effects of fatigue. There were no effects of manipulation or interactions between manipulation and group. A main effect of group was found in the left SPL for the load 4 > load 3 > load 1 > load 0 contrast (Table 3). When evaluating activations for each load condition versus baseline (load 0), we observed group effects in the highest load condition (i.e., load 4 > load 0). Compared to the fatigue group, the control group showed stronger activation of the bilateral SPL and precuneus, bilateral ACC, the right DLPFC, and the left premotor cortex (Table 3; Figure 3A). Figure 3B shows that after both manipulations, the control group increased activation with increasing working memory load in these areas. The fatigue group did not show this pattern of activation after the fatigue or the control manipulation.

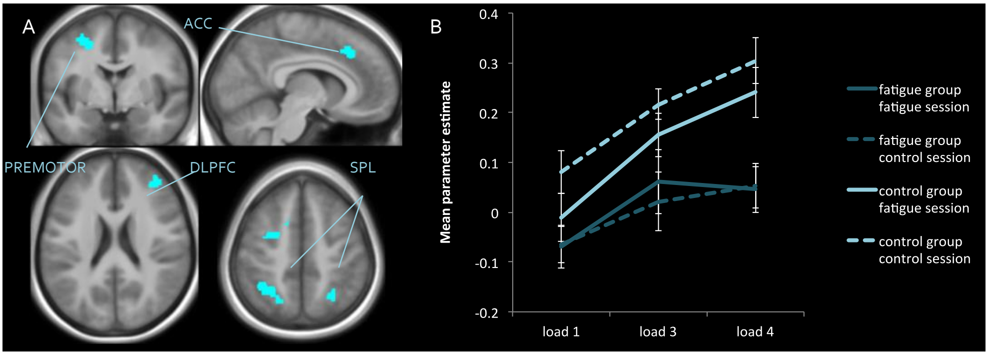


Figure 3. Group main effects based on the first level contrasts load 4 > load 0. (**A**) Main effect of group (control group > fatigue group) across sessions overlayed on the mean anatomical image of all participants. (**B**) Mean parameter estimates for each load condition > load 0 within the activated areas combined. Error bars show SEM.

Table 3

Brain regions showing greater activity increase in the control group compared to the fatigue group

Peak activation area	BA	x	y	z	*z*	k
1st level contrast: load 4 > load 3 > load 1 > load 0						
Superior Parietal Lobule	7	-30	-52	49	3.82	84
Precuneus	7	-18	-61	37	2.91	
1st level contrast: load 4 > load 0						
Superior Parietal Lobule	7	-30	-52	49	4.33	122
Precuneus	7	-18	-61	49	3.21	
Precuneus	7	-18	-61	40	3.05	
Middle Frontal Gyrus	9	39	41	28	3.85	48
Middle Frontal Gyrus	6	-27	-4	49	3.33	32
Precuneus	7	24	-55	46	3.32	55
Precuneus	7	33	-40	43	3.26	
Precuneus	7	15	-58	46	2.64	
Sub-Gyral	6	-15	8	55	3.25	35
Cingulate Gyrus	24	-12	11	43	3.05	
Cingulate Gyrus	32	-12	20	34	2.66	
Cingulate Gyrus	32	9	26	40	2.88	37
Cingulate Gyrus	32	12	17	40	2.83	

Note. Activations are thresholded at $p < .005$, corrected for magnitude of k = 30. Coordinates are in MNI space. BA = Brodmann Area.

Correlations with task performance. Areas showing a positive correlation between load-related activations and accuracy are listed in Table 4. A group x accuracy interaction revealed stronger correlations for the fatigue group compared to the control group in the right ACC, left anterior PFC and left anterior insula/ frontal operculum (Figure 4). Figure 4B illustrates the correlation between the mean activation of these areas and accuracy for each group (averaged over sessions). A significant correlation was observed in the fatigue group (Spearman's rho = .44, $p < .025$) and not in the control group (Spearman's rho = .04, $p = .84$). We found no areas showing a negative correlation with accuracy.

For completeness, we evaluated main effects of accuracy for each group separately. In the control group, we found no areas in which load-dependent activation increase correlated with accuracy. In the fatigue group, we observed a wide range of areas showing significant positive correlations, including bilateral DLPFC, ACC, anterior insula/ frontal operculum (Table 4).

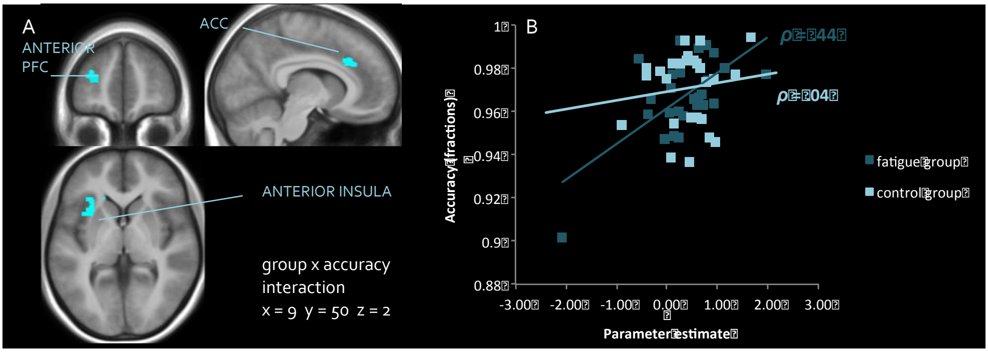


Figure 4. Group x accuracy interactions. (**A**) Regions showing a stronger correlation with accuracy in the fatigue group compared to the control group. These regions included the left anterior prefrontal cortex (anterior PFC), right anterior cingulate cortex (ACC) and the left anterior insula / frontal operculum. (**B**) Scatterplot illustrating the relation between the mean activation in these areas and overall accuracy, collapsed over sessions. Spearman's correlation showed a moderate positive relation for the fatigue group (dark turqoise) and no correlation for the control group (light turqoise).

Table 4

Brain regions showing positive correlations with overall accuracy (i.e., regions in which load-dependent activation increase is associated with higher accuracy)

Peak activation area	BA	x	y	z	z	k
Covariate interaction: fatigue > control						
Insula		-34	11	1	3.48	53
Caudate Head		-18	26	-5	3.35	
Insula		-30	23	1	3.04	
Middle Frontal Gyrus	10	-30	47	16	3.34	44
Cingulate Gyrus	32	18	20	34	3.26	34
Cingulate Gyrus	32	6	29	28	3.10	
Covariate main effect: fatigue group						
Hippocampus		-36	-46	4	4.99	37
Middle Frontal Gyrus	6	21	8	61	4.26	38
Insula		-33	11	1	4.06	132
Caudate		-18	26	-5	3.79	
Insula	13	-45	14	-2	3.16	
Insula		33	17	4	3.83	66
Insula	13	36	26	1	3.24	
Putamen		24	23	1	3.13	
Middle Frontal Gyrus	10	-27	47	16	3.8	96
Middle Frontal Gyrus	9	-30	50	28	3.10	
Cuneus	17	15	-97	13	3.76	96
Lingual Gyrus	18	27	-85	1	3.47	
Inferior Occipital Gyrus	17	18	-97	-5	3.41	
Cingulate Gyrus	24	15	20	31	3.73	278
Cingulate Gyrus	32	6	29	28	3.68	
Cingulate Gyrus	32	-6	29	28	3.51	
Cuneus	17	-12	-100	7	3.62	36
Middle Frontal Gyrus	9	36	41	19	3.37	33
Superior Frontal Gyrus	10	33	56	19	3.02	

Note. Activations are thresholded at $p < .005$, corrected for magnitude of k = 30. Coordinates are in MNI space. BA = Brodmann Area.

DISCUSSION

The current fMRI study investigated effects of fatigue on working memory load. As expected, the results demonstrated significant differences in brain activation between students with and without chronic fatigue in the context of working memory load. Contrary to our expectations, we observed flattening of load-dependent increases, accompanied by performance declines, in fatigued students. Non-fatigued students responded faster to targets than chronically fatigued students, with the largest difference in the highest load conditions (i.e., load 3 and load 4). Compared to chronically fatigued students, non-fatigued students also showed stronger activation increases in response to higher load in several areas of the fronto-parietal network. A cognitive fatigue induction resulted in overall reduced accuracy, which was not accompanied by neural activation changes and did not differentiate between chronically fatigued and non-fatigued students.

The working memory task evoked significant effects of load in behavioral performance and neural activation. Regardless of group or session, reaction times increased and accuracy decreased with higher load, reflecting that deciding whether or not a stimulus was a target took longer and was more difficult when more items were maintained in working memory. Increasing activation coupled with increasing working memory load was found in several bilateral PFC and parietal regions, including the DLPFC, ACC and SPL. These load-dependent areas are commonly associated with working memory (Curtis & D'Esposito, 2003; Owen, McMillan, Laird, & Bullmore, 2005) and with increasing working memory load (Callicott et al., 1999; Jaeggi et al., 2007).

The characteristic load-dependent increase in fronto-parietal activation was less pronounced in fatigued students. In addition, fatigued students responded slower than their non-fatigued peers in the higher load conditions. Reduced activation of the working memory network at higher working memory load is consistent with earlier studies showing similar reductions in CFS patients compared to healthy controls (Caseras et al., 2006) and in healthy individuals following sleep deprivation (Choo et al., 2005). Our findings of performance decline coupled with attenuation of load-dependent increase in the working memory network suggests that stronger recruitment of these areas was required to preserve behavioral performance, which fatigued students failed to accomplish. These differences in behavior and neural activation thus suggest that working memory processing was less effective in chronically fatigued students. Based on the assumption that higher working memory load engages more working memory capacity and that load-dependent increases in neural activation represent the extent of working memory capacity, a decrease in brain activity accompanied by performance decline may reflect that a capacity limit is reached (Goldberg et al., 1998, see also Callicott et al., 1999; D'Esposito, 2001), causing fatigued students to disengage from the task. Alternatively, Jaeggi et al., (2007) demonstrated that the inverted U-shaped curve does not necessarily represent task-disengagement, but rather a change of strategy (Jaeggi et al., 2007). Another possibility is that fatigued individuals disengaged from the task due to reduced ability to sustain attention. Task-disengagement associated with sustained attention is a regularly reported characteristic of fatigue (Dorrian, Roach, Fletcher, & Dawson, 2007; Pattyn, Neyt, Henderickx, & Soetens, 2008; Smallwood et al., 2004; van der Linden, Keijsers, Eling, & van Schaijk, 2005). In the current study, all participants reported afterwards that the task was 'easy' or 'very easy' to perform. Performing a simple target-response task can reduce the need for active attention (Manly, Robertson, Galloway, & Hawkins, 1999; Smilek, Carriere, & Cheyne, 2010) and the present task may therefore have been sensitive to task-disengagement due to attentional lapses. However, we consider this unlikely since the present group difference occurred at the highest load levels only, whereas if attentional lapses would differentiate between load conditions at all, this should occur at task conditions that require the least amount of active attention (e.g., low working memory load). Moreover, attentional lapses are generally represented by a decrease in reaction times (e.g., Manly et al., 1999) whereas here we observed an increase in reaction times.

Notably, we did not observe increased neural activation in fatigued students at lower levels of working memory load. Based on individual differences with respect to aging

(Park & Reuter-Lorenz, 2009), development (Church et al., 2008), performance levels (Jaeggi et al., 2007), and clinical fatigue (Caseras et al., 2006; Chaudhuri & Behan, 2004; Cook et al., 2007; DeLuca, Genova, Capili, & Wylie, 2009b; Lange et al., 2005), we expected additional neural recruitment in terms of increased or more dispersed brain activation, particularly at lower levels of working memory load, at which we did not expect performance differences. Such additional recruitment would indicate less efficient compensatory strategies or increased effort expenditure. This would also be compatible with recent findings from a task-switching study based on the same group of students (Plukaard et al., submitted) in which fatigued students additionally recruited the ACC. Perhaps the current task did not allow for many compensatory strategies due to its simplicity, as increasing task difficulty can facilitate compensatory responses (Drummond, Brown, Salamat, & Gillin, 2004). Fatigued students maintained their level of accuracy comparable to that of their non-fatigued counterparts and were thus equally able to accurately perform the task. Future research should indicate whether a more complicated task, that for example relies more on active manipulation of items in working memory storage, would evoke compensatory or effort-related neural differences.

The fact that we only observed behavioral effects of fatigue in terms of reaction times and not in terms of accuracy endorses the notion that working memory capacity is not purely defined by storage size (i.e., the number of items that can be maintained), but rather by the quality of memory for these items (Bays & Husain, 2008; Keshvari, van den Berg, & Ma, 2013; Palmer, 1990). Congruently, reduced capacity may also influence the speed at which these items are processed. That is, fatigued students took longer to decide whether a stimulus matched one of the items stored in working memory than non-fatigued students when more items were maintained (i.e., at load 3 or 4). Performance accuracy remained the same between the groups, indicating that both groups were equally able to maintain different quantities in memory. Thus, our data appears to provide evidence for a group difference in working memory capacity in terms of quality, but not in terms of quantity. However, this inference should be made with caution, since our design did not differentiate between the quality and quantity of working memory load.

Even though we observed no difference in performance accuracy between fatigued and non-fatigued students, the association between performance accuracy and neural recruitment differed between groups, as indicated by a significant interaction between group and accuracy in the anterior PFC, ACC and anterior insula/ frontal operculum. Performance of fatigued students was related to load-dependent recruitment of this network; fatigued students who recruited these areas to a greater extent performed better than fatigued students who failed to recruit these areas or who recruited these areas to a lesser extent. In non-fatigued students, no such correlations were observed. In general, these areas are involved in central executive processes that may underpin working memory function; the anterior PFC and ACC are involved in decision-making, response conflict and error monitoring (Koechlin & Hyafil, 2007; Ridderinkhof, Ullsperger, Crone, & Nieuwenhuis, 2004). The anterior insula is considered to play a critical role in affective, social, and high-level cognitive control processes (Menon & Uddin, 2010), and is together with frontal operculum typically activated during working memory tasks (e.g., Cohen, Perlstein, & Braver, 1997; D'Esposito et al., 1998; Owen et al., 2005; Pessoa, Gutierrez, Bandettini, & Ungerleider, 2002). Activation of these areas together may represent evaluative processing in order to adequately adjust behavior in response to errors. As such, recruitment of this network could characterize a strategy upon which fatigued students, who concurrently fail to continue to increase activation of the fronto-parietal network, can rely as compensation for this attenuation. A possible explanation for the lack of a correlation between this network and performance accuracy in non-fatigued students is that these students, as a group, were able to engage more working memory capacity, allowing them to increase activation of the fronto-parietal network and respond faster to targets at higher working memory load. In turn, this might diminish the need or perhaps the time for compensatory response evaluation.

A final finding in this study involved the overall decline in performance accuracy in response to the fatigue manipulation, which was not accompanied by brain activation changes. Reduced accuracy is a common finding after a period of demanding cognitive activity (e.g., Lorist et al., 2000) and in line with Klaassen et al. (2014), who used the same fatigue manipulation and observed that after this manipulation, both middle-aged and young male teachers made more errors on a letter Sternberg task (Klaassen et al., 2014). In contrast to the group main effect, we consider it likely that attentional lapses underlie this effect, given the observation that such lapses are a typical response to prolonged task performance (Dorrian et al., 2007; Pattyn et al., 2008; Smallwood et al., 2004; van der Linden et al., 2005) and the fact that the present effect did not distinguish between load conditions. We therefore propose that this accuracy reduction in response to the fatigue manipulation results from a general attentional decline affecting visual processing or motor output, rather than working memory per se. From this perspective, it is also not surprising that we observed no manipulation-related differences in the fMRI data, as the examined neural activation was based on working memory processes (i.e., differences between load conditions).

There are several methodological considerations that need to be addressed. First, all load conditions consisted of 50% targets that required a response. These conditions were compared to a baseline condition in which participants were required to respond to each stimulus (i.e., 100% targets). Consequently, each load condition additionally involved withholding a response, which may thus involve additional processing (e.g., inhibition). It is improbable that such additional processes will have affected our results, since the proportion of targets was held constant in each load condition. Second, the present study used a block design, which generally increases statistical power compared to event-related designs. A drawback of a block design is that it provides no information about trial specific working memory processes, such as the role of encoding, retention or response. Future research could focus on these distinct components of working memory. Lastly, to exclude confounding factors such as intelligence, gender or motivation, we used a homogeneous sample that consisted of female medical students. This has increased the possibility of finding fatigue-related differences and it enables generalization to the larger population of highly educated females. However, it does not enable generalization to males. Forthcoming studies should evaluate whether the present results will apply to males and/or individuals with another educational background.

Conclusion and implications

The present study demonstrated that working memory was less effective or efficient in students with chronic fatigue. Behavioral as well as neural activation differences depended on load condition, which suggests that these differences entailed working memory capacity. As working memory underlies a wide range of cognitive processes and behaviors that form the basis for everyday performance, fatigued individuals might benefit from improvement of working memory capacity. Possible interventions aimed at increasing working memory capacity, which can be accomplished by training (Jha, Stanley, Kiyonaga, Wong, & Gelfand, 2010; Olesen, Westerberg, & Klingberg, 2004), might be a promising way forward.

Acknowledgments

The authors are grateful to Gertjan M.J.J. Jong and Arnout de Groot for all their help. This work was supported by the NWO 433-08-205 from the Netherlands Science Organization.