



A multifaceted investigation of the link between mental fatigue and task disengagement

JESPER F. HOPSTAKEN,^a DIMITRI VAN DER LINDEN,^a ARNOLD B. BAKKER,^{a,b} AND MICHIEL A. J. KOMPIER^c

^aDepartment of Work and Organizational Psychology, Institute of Psychology, Erasmus University Rotterdam, Rotterdam, The Netherlands

^bDepartment of Applied Psychology, Lingnan University, Hong Kong, China

^cDepartment of Work and Organizational Psychology, Behavioural Science Institute, Radboud University Nijmegen, Nijmegen, The Netherlands

Abstract

Mental fatigue is often characterized by reduced motivation for effortful activity and impaired task performance. We used subjective, behavioral (performance), and psychophysiological (P3, pupil diameter) measures during an *n*-back task to investigate the link between mental fatigue and task disengagement. After 2 h, we manipulated the rewards to examine a possible reengagement effect. Analyses showed that, with increasing fatigue and time-on-task, performance, P3 amplitude, and pupil diameter decreased. After increasing the rewards, all measures reverted to higher levels. Multilevel analysis revealed positive correlations between the used measures with time-on-task. We interpret these results as support for a strong link between task disengagement and mental fatigue.

Descriptors: Fatigue, Motivation, P3, Pupil diameter, Task engagement

Sustained performance on cognitively demanding tasks often leads to mental fatigue, which is a complex state characterized by a reluctance for further effort and changes in mood, motivation, and information processing (Meijman, 1997; van der Linden, Frese, & Meijman, 2003). Levels of mental fatigue may fluctuate due to normal everyday activities (e.g., as a consequence of daily job demands), but may also be chronic and comorbid to diseases or disorders such as Parkinson's disease (Chaudhuri & Behan, 2004), depression (Demyttenaere, De Fruyt, & Stahl, 2005), and burnout (Maslach, Schaufeli, & Leiter, 2001). In the workplace, mental fatigue has been found to be one of the most frequent causes for work accidents (Baker, Olson, & Morisseau, 1994; McCormick et al., 2012). Despite the mundane nature of mental fatigue, the exact psychological mechanisms that are involved remain relatively unknown. For example, although it appears that cognitive control is sensitive to fatigue, there is an ongoing investigation to determine the exact cognitive processes that are compromised. The most difficult part of fatigue to grasp scientifically may be the interplay between cognition and motivation. Some researchers have referred to fatigue as a stop-emotion that serves to prevent exhaustion by exerting too many resources into a task (cf. van der Linden, 2011). As a consequence, people often tend to disengage from the task at hand when they are getting fatigued. This disen-

gement is characterized by impairments in motivation (e.g., lower effort invested in the task), cognition (e.g., diminished attention and task focus), and effective behavior (e.g., decreased task performance; Boksem & Tops, 2008; Hockey, 1997).

Based on the multifaceted nature of fatigue, the central aim of the present study is to examine mental fatigue using multiple indicators of task (dis)engagement within a well-established fatigue paradigm. Specifically, we will investigate fatigue-related effects on motivation, cognition, and performance by using a set of subjective, behavioral, and psychophysiological measures of task engagement. This way, we assess on what levels task (dis)engagement and fatigue are related. By directly linking these three types of measures (i.e., subjective, behavioral, and psychophysiological), we make an innovative contribution to the mental fatigue research field.

Fatigue and Engagement

As mentioned earlier, fatigue is often observed as a reluctance for further effort. On the behavioral level, fatigue is related to a general disengagement and low vigor in contrast to the possibility of exploiting the benefits of a certain task, or exploring the environment for rewarding activities (Boksem, Meijman, & Lorist, 2006; van der Linden, Frese, & Sonnentag, 2003). Several studies have shown that fatigue mainly impacts top-down cognitive control processes (e.g., Lorist et al., 2000; Lorist, Boksem, & Ridderinkhof, 2005; van der Linden & Eling, 2006; van der Linden, Frese, & Meijman, 2003), as they generally require the subjective experience of investing high effort (Dehaene, Kerszberg, & Changeux, 1998). Bottom-up cognitive processes, on the other hand, requiring less effort, are relatively unaffected by fatigue. Boksem and Tops

This work was supported by the Netherlands Organization for Scientific Research (NWO; Project 400-08-015). We thank Henk van Steenbergen for his help with the export of the Tobii eye-tracking data to Brain Vision Analyzer.

Address correspondence to: Jesper F. Hopstaken, Institute of Psychology, Erasmus University Rotterdam, Burgemeester Oudlaan 50, 3062 PA Rotterdam, The Netherlands. E-mail: hopstaken@fsw.eur.nl

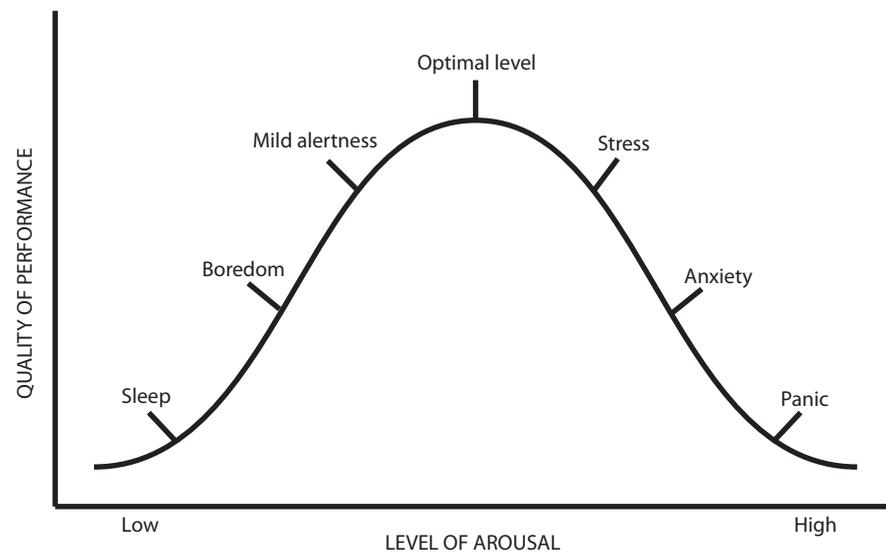


Figure 1. The classical Yerkes-Dodson law describes low levels performance for low and high levels of arousal and high levels of performance for intermediate levels of arousal.

(2008) propose that feelings of fatigue lead to abandoning behavior when the energy cost exceeds the perceived benefits of continued performance. Therefore, it is widely accepted that, generally speaking, fatigue negatively impacts task engagement when mental effort is required.

To date, only few studies have been conducted on the psychophysiology of mental fatigue (e.g., Boksem et al., 2006; Faber, Maurits, & Lorist, 2012; Lorist, 2008). In the present study, we will use two measures that have often been considered as indicators of task engagement. However, these measures—the parietal P3 event-related potential (ERP) and pupil diameter—have not yet been thoroughly tested within a typical fatigue design.¹ The P3 event-related potential is one of the most heavily investigated ERPs and consists of a frontal P3a component that has been related to focal/stimulus-driven attention and novelty detection, and a parietal P3b component that has been related to focused attention related to subsequent working memory activation and salience detection (Polich, 2007). Begleiter, Porjesz, Chou, and Aunon (1983) proposed that the parietal P3 component may index motivational properties of a stimulus. Later, it was found that the effect of motivational significance on the P3 amplitude is also modulated by the amount of attention that is paid to the stimulus (Johnson, 1993). Combining the sensitivity to both motivational and attentional aspects of a task has led researchers to link the P3 to task engagement (e.g., Murphy, Robertson, Balsters, & O’Connell, 2011).

The other psychophysiological measure we use—the diameter of the pupil—has for many years been considered as an index of psychophysiological arousal or neural gain. Classical work of Beatty and Kahneman has already showed that the pupil is sensitive to momentary load and effort during mental tasks (Beatty, 1982; Kahneman, 1973; Kahneman & Beatty, 1966). In recent years,

however, these statements have been nuanced by specifically relating the pupil diameter to control states of engagement (i.e., exploration vs. exploitation). For example, Gilzenrat, Nieuwenhuis, Jepma, and Cohen (2010) and Jepma and Nieuwenhuis (2011) conducted multiple experiments to relate pupil diameter to task engagement. They observed that task engagement and exploitation behavior were related to an intermediate pupil diameter. On the other hand, disengagement from the task in the form of distraction and explorative behavior was related to increased pupil diameter. This relation was described as the classical Yerkes-Dodson curve (Yerkes & Dodson, 1908), in which intermediate arousal leads to optimal engagement and performance (see Figure 1). Low and high arousal, on the contrary, lead to disengagement and impaired task performance. With regard to our study, the low arousal (and corresponding small pupil) state is especially interesting because the behavioral consequence (i.e., general disengagement and impaired performance) seems to strongly overlap with the behavioral consequence of fatigue.

In an experimental time-on-task design, we tested the hypothesis that mental fatigue is related to task disengagement and reduced P3 amplitude and pupil diameter. Participants continuously worked on an *n*-back task for an extended time (2 h). The *n*-back task is known to require high levels of task engagement and sufficient levels of voluntary attentional control (Chen, Mitra, & Schlaghecken, 2008; Cohen et al., 1997; Watter, Geffen, & Geffen, 2001). We expected that, during a 2-h continuous performance task, subjective fatigue increases and task engagement decreases with increasing time-on-task. In such a demanding task, a decrease in engagement should be accompanied by compromised task performance. Moreover, we predicted that, parallel to the increase of fatigue and the decrease of reported task engagement, there will be decreases in both the pupil diameter and the P3 amplitude. The combination and direct comparison of subjective measures of fatigue, measures of task performance, and physiological measures of the P3 and pupil diameter is an informative and innovative way to examine the relation between mental fatigue and task disengagement.

1. We know of two studies: Boksem et al. (2006) briefly mention the P3 but mainly focus on response-locked potentials (Ne/ERN), and Massar et al. (2010) measure the more frontal P3a component, which is functionally different from the parietal P3b that we measure.

Reversibility of Mental Fatigue Effects

Previous studies have shown that task engagement mainly depends on the expected value of engaging in a task. This expected value is based on a tradeoff between the expected costs and rewards of a task (Aston-Jones & Cohen, 2005; Cohen, McClure, & Yu, 2007). When the cost/reward tradeoff is favorable, it stimulates exploitation of the task rewards by engaging in the task. That is, the motivation and corresponding attentional focus that is typical for task engagement occurs when executing the task provides enough intrinsic (e.g., pleasure, excitement) and/or extrinsic (e.g., monetary benefits) rewards to make it worth the effort. However, when the tradeoff becomes unfavorable, one tends to disengage from the task. The exploitation behavior then makes place for exploration behavior that is manifested in the tendency to explore the environment, in pursuit of more rewarding tasks. This subsequently increases the probability of failures in task-related behavior (e.g., failing to respond to important task-related cues).

The role of cost/reward tradeoffs of engagement have also been emphasized in several previous fatigue studies (e.g. Boksem & Tops, 2008) indicating that mental fatigue is likely to occur when the costs of engaging in a task exceed the predicted rewards. In line with this notion, increasing the rewards of task engagement under fatigue may cause shifts in motivation that may drive attention back to task-related cues. This effect has been confirmed in a previous study by Boksem and colleagues (2006) who showed that, after 2 h of cognitive performance, fatigued participants could partly restore their performance when they received a monetary reward. This effect was accompanied by an increase in the error-related negativity ERP. Boksem and coworkers considered such findings supportive of the notion that dopaminergic reward systems play a role in fatigue-related decline of performance.

In the present study, we focus on the responsiveness of the P3 amplitude and pupil diameter as measures of task engagement in response to increasing rewards during mental fatigue. We expect that a change in the task rewards may positively influence the cost/reward tradeoff and lead to task reengagement. To test this, we included a manipulation in which we increased the rewards for engaging in the task after participants had already worked on the demanding task for 2 h. This manipulation, to some extent, resembles the one used by Boksem et al. (2006). Specifically, we told participants that the remaining time that was left on the task would depend on the quality of their performance during the last series of trials. After working on the task for 2 h, this manipulation provided a strong motivation to reengage in the task and optimize performance to be permitted to stop. We then examine in what way such a reward manipulation affects task engagement, the P3 amplitude, and pupil diameter. We hypothesize that an increase in task rewards concurs with an increase in P3 amplitude and pupil diameter, suggesting reengagement in the task.

Method

Participants

Twenty undergraduate students (3 males, 17 females), between the ages of 17 and 24 ($M = 19.9$ years, $SD = 2.0$) participated in the study and received study credits. All participants were well rested and in good health as measured by self-report. The participants reported to have slept 7 or more hours and were asked to withhold the intake of caffeine and alcohol during the 24 h before the experi-

ment. All participants had normal or corrected-to-normal vision. Written informed consent was obtained prior to the study.

Stimuli and Data Acquisition

Participants were seated in a comfortable chair in a dimly lit and sound-attenuated room facing an eye-tracking screen at a distance of approximately 65 cm. During the whole experiment, pupil diameter and electroencephalogram (EEG) were measured continuously. The participants performed a visual letter n -back task in 1-back, 2-back, and 3-back variants. Participants were asked to decide whether the letter presented on the screen was a target or nontarget stimulus. In the n -back task, a stimulus is a target when the presented letter is the same as the letter presented n letters before. Accordingly, they responded on the corresponding button in the armrest of the chair. The stimuli were presented in the center of the screen and consisted of the letters B, C, D, E, G, J, P, T, V, and W in the font Palatino Linotype point size 40. In the Dutch language, these letters are phonologically similar in order to prevent sound-related retrieval strategies. The letters were presented randomly with a target rate of 25%. The n -back task has been used successfully in previous experiments to induce mental fatigue (Massar, Wester, Volkerts, & Kenemans, 2010). It is a cognitively demanding task that requires the sustained engagement of working memory and attention in order to uphold adequate levels of performance (Watter et al., 2001).

Procedure

Before the experiment, participants filled out questionnaires about their general health, current level of fatigue, and task motivation (see description of these measures below). After the calibration of the eye-tracking device, participants were instructed on the n -back task. Participants practiced on each variant of the task until they reached a minimum of 70% accuracy. The experimental task was divided in seven time-on-task blocks. Each block consisted of 63 trials of the 1-back task, followed by 63 trials of the 2-back task, followed by 63 trials of the 3-back task and lasted for about 18 min (depending on random intervals). There was no rest between tasks. The n -back stimuli were displayed for 500 ms with an interstimulus interval randomized at 5 to 5.5 s. The length of this interval was long enough to ensure that the pupil diameter returned to baseline levels (Beatty, 1982; Stern, Ray, & Quigley, 2000)

After each block, the participants had to indicate their current level of fatigue and task engagement. The participants had only limited time to do this, to make sure they would not rest. After they completed six blocks of 18 min, we introduced our reward manipulation. We told them that the remaining time of the experiment would depend on their performance relative to their performance on the previous blocks. We explained that if they performed better the remaining time could be as short as about 5 min. However, we also told them that if they performed similar or worse the remaining time could run up to about 40 min (i.e., it could range from somewhere between 5 and 40 min depending on their performance). We assumed that, after about 2 h of continuous performance, this provides a strong incentive to optimize performance. In reality, the length of this last block was the same as the first six blocks. After the experimental task, the participants were asked to fill in questionnaires about their levels of fatigue.

Measures and Data Processing

Subjective measures. Subjective fatigue was measured before, during, and after the task in order to monitor its temporal

progression. Before and after the task, participants filled in the Rating Scale Mental Effort (RSME; Zijlstra, 1993), which consists of seven vertical scales assessing different aspects of mental fatigue (e.g., difficulty keeping attention on the task, difficulty exerting further effort in the task). The scales have numerical (0 to 150) and verbal (*not at all* to *extremely*) anchors. After each time-on-task block during the experiment, the participants were asked, "How tired do you feel?" They had to respond by moving a slider from 0 to 100, with increments of five. The slider had no anchors, but the extreme ends were labeled with *very much* and *not at all*. Due to missing data at certain blocks, two participants were excluded for the analysis of the subjective fatigue and one participant was excluded for the analysis of the subjective engagement.

After each time-on-task block, we also measured task engagement by asking, "How engaged are you in the task?" The participants had to reply by moving a slider from 0 to 100, with increments of five. The slider had no anchors except for the extremes *very much* and *not at all*. Because task engagement was measured multiple times during the experiment, the temporal progression of subjective engagement in the task could also be monitored.

Behavioral measures. The most relevant behavioral measure of performance on the *n*-back task was accuracy. We operationalized accuracy by calculating the *d* prime for each time-on-task interval. As described by signal detection theory, the *d* prime was calculated as an indication of accuracy (Wickens, 2001). While accuracy was the most important focus for the participant during the task, we wanted to make sure accuracy effects were not clouded by accuracy/speed tradeoffs. Therefore, we also examined reaction times (RTs).

Physiological measures. Pupil diameter was recorded continuously during the entire length of the experimental task with a Tobii Eyetracker 2150 with a sample rate of 50 Hz. For two female participants, the eye-tracking data were not saved due to a technical problem. We excluded these participants from the analysis of the eye-tracking data but included their data in the other analyses. The recordings were exported to Brain Vision Analyzer (Brain Products, Gilching, Germany). Artifacts and blinks were detected by the eye tracker and removed by using a linear interpolation algorithm. To measure baseline pupil diameter, we averaged the pupil diameter in the 500 ms before stimulus onset. During this period, participants saw a black screen so there was no interference from pupillary light reflexes of the eye to the environmental lighting during the baseline recording. Baseline pupil diameter for each condition and time-on-task interval was then exported to SPSS for further analysis.

For the recording of the EEG, we used a BioSemi Active-Two with Ag/AgCl active electrodes at 32 + 2 scalp sites (International 10–20 system). There were six additional electrodes attached. Two electrodes were placed on the left and right mastoids as reference. To allow correction for ocular movement artifacts, we placed two electrodes next to the outer side of the eyes for horizontal electrooculogram (HEOG) and two above and below the left eye for vertical electrooculogram (VEOG). Online signals were recorded with a sample rate of 512 Hz and 24-bit A/D conversion. Extensive research of the P3 shows the distinction between the P3a and P3b potential. The P3a is linked to novelty detection and best seen at the Cz and Fz electrodes, while the P3b is linked to salience processing and is best seen at the Pz electrode (Polich, 2007). Because we were interested in the latter, we analyzed the EEG signal at the Pz

electrode. Reviewing the voltage maps confirmed that the amplitude of the P3 was indeed largest at Pz. The EEG data were analyzed in Brain Vision Analyzer (Brain Products). After rejection of out-of-range and eye movement artifacts, using the Gratton and Coles method (Gratton, Coles, & Donchin, 1983), the ERPs were averaged offline. Segments with amplitudes higher than 200 μ V and lower than -200μ V (0.122% of the data) and voltage steps above 50 μ V/ms (0.004% of the data) were removed. The data were also inspected on low activity (below 0.1 μ V) and filtered (low cutoff at 0.1 Hz and high cutoff at 40 Hz). After baseline correction for the 200 ms before the stimulus onset, we aggregated the data per condition and measured the positive peak between 300 and 450 ms after the onset of the stimulus. Trials in which performance errors occurred were excluded. The mean P3 peak activity for each condition and time-on-task interval was then exported to SPSS for further analysis.

Statistical Analysis

The subjective behavioral and psychophysiological data were exported to SPSS and statistically analyzed using repeated measures analysis of variance (ANOVA) using the Greenhouse-Geisser correction for degrees of freedom. First, main and interaction effects of time-on-task and task difficulty were tested. Then, significant effects were further qualified by examining changes from block 1 through 6, in which the fatigue manipulation occurred, and changes from block 6 to 7, in which the reward/motivation manipulation occurred.

In addition to the repeated measures ANOVA, we also analyzed the data using a multilevel approach with Mplus statistical software (Muthén & Muthén, 1998). The multilevel method takes into account that in some designs measures may not be fully independent from each other, but are nested on various levels. Repeated measures data can be treated as multilevel data, because the repeated measures (i.e., the time-on-task blocks) are nested within individuals. The multilevel analyses take this information into account when calculating associations between variables, because it partitions the variance to each level. Using this multilevel method, we calculated the associations between the various outcome measures (i.e., pupil diameter, P3, subjective measures, performance) on the individual level with the nested structure of the data taken into account (i.e., blocks are nested within persons). We used a two-level model with time-on-task block at the first level (Level 1; $N = 126$), and individuals at the second level (Level 2; $N = 18$). In this operationalization, a high correlation between the outcome variables implies that a change in one variable corresponds with a similar change in another variable within individuals, taking into account that each individual has been measured during multiple blocks. For more information about multilevel analysis, see Snijders and Bosker (1999).

Results

Subjective Measures

Pre- and posttask analysis of the RSME confirmed that our fatigue manipulation was successful as participants reported significantly higher levels of subjective fatigue after the experiment than before, $t(19) = -9.5$, $p < .001$. In a repeated measures analysis of the fatigue measure after each block, we found that, with increasing time-on-task, participants felt more fatigued from block 1 through 6, $F(2.4, 33.2) = 28.13$, $p < .001$, $\eta_p^2 = .67$. After the reward

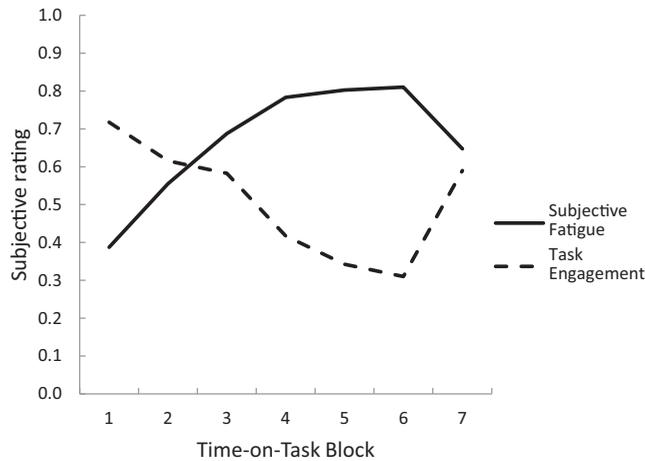


Figure 2. Subjective fatigue and engagement ratings with time-on-task. Responses are given ranging from 0 to 1 with 0.05 increments.

manipulation, subjective mental fatigue significantly decreased from block 6 to 7, $F(1,17) = 7.52, p < .05, \eta_p^2 = .96$.

Subjective task engagement significantly decreased with increasing time-on-task from block 1 through 6, $F(2.8,47.1) = 28.23, p < .001, \eta_p^2 = .62$. After the reward manipulation from block 6 to 7, however, there was a significant increase in subjective engagement, $F(1,18) = 18.16, p < .001, \eta_p^2 = .50$. The latter finding shows that this manipulation effectively increased subjective task engagement during block 7 (see Figure 2).

Behavioral Measures

Repeated measures analysis showed significant main effects for time-on-task, $F(3.3,61.8) = 4.6, p < .001, \eta_p^2 = .20$, and task difficulty, $F(1.3,23.9) = 16.7, p < .001, \eta_p^2 = .47$, on d prime. The main effects were further qualified by a significant interaction between task difficulty and time-on-task from block 1 through 6, $F(5.6,107) = 4.6, p < .001, \eta_p^2 = .20$. The interaction revealed that the 1-back and 2-back tasks showed a significant decline in d prime from block 1 to 6 (1-back, $F(4,75.8) = 6.5, p < .001, \eta_p^2 = .26$; 2-back, $F(2.5,48.2) = 3.3, p < .05, \eta_p^2 = .15$), whereas there was no such change in performance on the 3-back task ($\eta_p^2 = .08$; see Figure 3). After the reward manipulation, performance significantly increased from block 6 to 7 on all tasks, $F(1,19) = 17, p < .01, \eta_p^2 = .47$; 1-back, $F(1,19) = 28.8, p < .001, \eta_p^2 = .60$; 2-back, $F(1,19) = 4.3, p < .05, \eta_p^2 = .19$; 3-back, $F(1,19) = 4.8, p < .05, \eta_p^2 = .20$. Note that these results are in line with the results on the subjective measures and indicate a relation between mental fatigue and task engagement on the one hand and task performance on the other hand.

There were no significant time-on-task changes in RTs for the 1-back ($\eta_p^2 = .05$) and 2-back ($\eta_p^2 = .03$) task. These results show that the decrease in d prime in the 1-back and 2-back tasks were not due to decreased reaction times, so there was no sign of a speed/accuracy tradeoff with increasing time-on-task (Wickelgren, 1977). However, participants displayed significantly shorter reaction times with increasing time-on-task on the 3-back task, $F(3,57.6) = 3.18, p < .05, \eta_p^2 = .14$. This indicates that performance on the 3-back task became more efficient over time. It is likely that the difficult 3-back task is more prone to learning effects during this 2-1/2 h

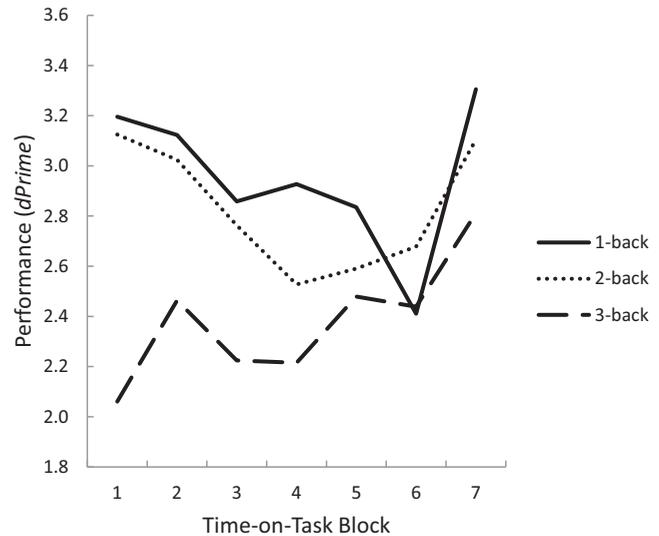


Figure 3. Performance on the n -back task with time-on-task for each difficulty level.

experiment, whereas the 1-back and 2-back tasks reached maximum performance levels relatively early.

Physiological Measures

Pupil diameter. As hypothesized, with increasing time-on-task, baseline pupil diameter significantly decreased from block 1 through 6, $F(2.5,42.2) = 4.7, p < .01, \eta_p^2 = .22$, and significantly increased again after the manipulation from block 6 to 7, $F(1,17) = 10.6, p < .01, \eta_p^2 = .38$ (see Figure 4). The baseline diameter also displayed a strong task difficulty main effect, $F(1.9,32.6) = 10.8, p < .001, \eta_p^2 = .39$, in which the diameter was significantly larger at more difficult tasks. In Figure 4, it can be observed that the initial diameter on block 1 is much lower for lower difficulty levels of the n -back task. This lower initial value

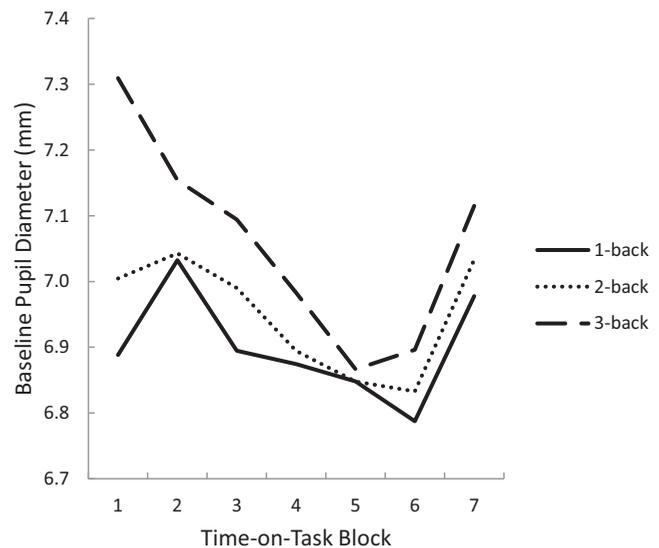


Figure 4. Baseline pupil diameter with time-on-task during the 1-back, 2-back, and 3-back task.

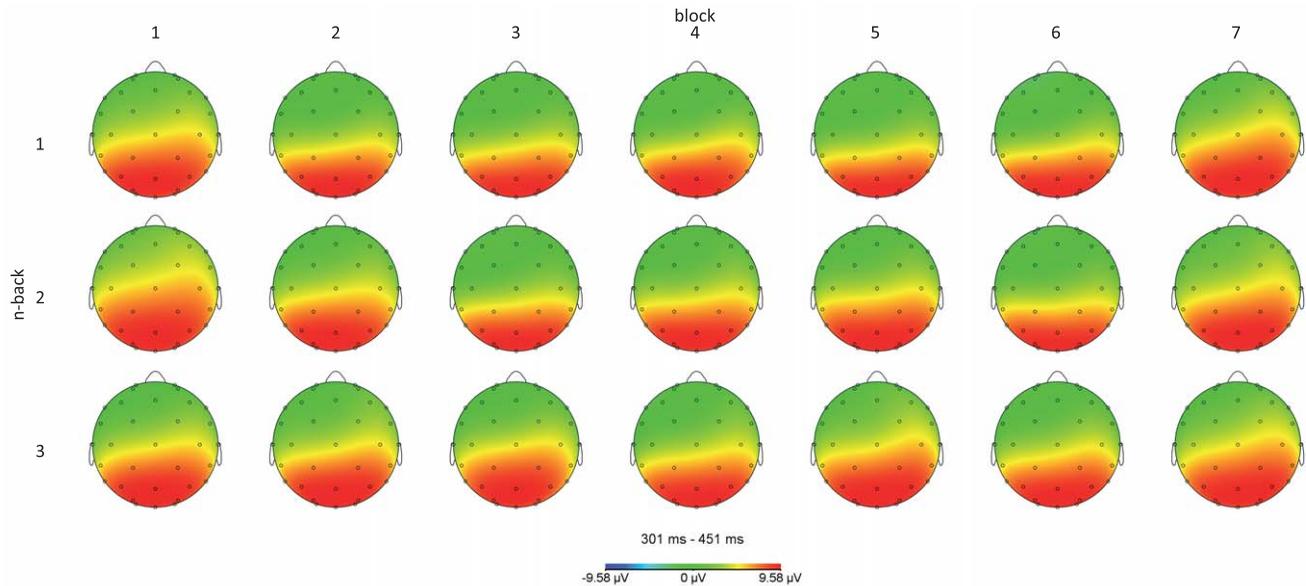


Figure 5. Localization of the P3 amplitudes depict a clear maximum at the parietal electrodes.

also results in a time-on-task curve that is less steeply declining. This is also reflected in the significant interaction effect between time-on-task and task difficulty, $F(5.4,92.2) = 2.8$, $p < .05$, $\eta_p^2 = .14$. A subsequent analysis also tested the time-on-task effects for each task individually. The results were in line with the observation that the more difficult the task, the stronger the time-on-task effect (1-back: *ns*, $\eta_p^2 = .09$; 2-back: $F(3,59.2) = 4.0$, $p < .05$, $\eta_p^2 = .17$; 3-back: $F(3.2,64.1) = 7.6$, $p < .01$, $\eta_p^2 = .27$).

P3 amplitude. We confirmed that the P3b was largest at electrode Pz. The voltage maps in Figure 5 show the localization of the P3 ERP during each time-on-task block. The P3b amplitude showed a strong significant main effect for time-on-task, $F(3.3,63) = 4.8$, $p < .01$, $\eta_p^2 = .20$. Figure 6 shows the P3 with increasing time-on-task for the different *n*-back tasks. Subsequent analysis revealed that the 1-back and 2-back task showed a significant decrease in P3 amplitude from block 1 through 6: 1-back: $F(3.8,72) = 3.3$, $p < .05$, $\eta_p^2 = .15$; 2-back: $F(3.4,64) = 4.4$, $p < .01$, $\eta_p^2 = .19$. The 3-back task did not show a significant change in P3 on this interval ($\eta_p^2 = .03$). This divergent progression of the 3-back task was expressed in a significant interaction effect between task difficulty and time-on-task, $F(26.8,139.3) = 2.9$, $p < .01$, $\eta_p^2 = .13$. Similar to the subjective and behavioral data, the P3 amplitude showed a reverse in pattern after the manipulation and increased significantly from block 6 to 7, $F(1,19) = 26.2$, $p < .001$, $\eta_p^2 = .58$, as was expected in the case of a reengagement effect. Raw P3 ERPs are displayed in Figure 7. By looking at the mean amplitudes, the steepest decrease can be observed during the first hour on the task, whereas the decrease of the pupil diameter seems more gradual over the whole experiment.

Multilevel Analysis

In line with our predictions, we found all measures to change congruently with increasing time-on-task. That is, during the first six blocks, subjective fatigue increased and performance, baseline diameter, and the P3 amplitude decreased. These findings were

obtained using ANOVAs, which is an approach adopted in the majority of studies in the field of behavioral and psychophysiological sciences. However, this method does not provide direct insight into the association between the different measures we used in the present study. Therefore, we also tested the associations between measures using multilevel analysis with Mplus statistical software (Muthén & Muthén, 1998). Using the multilevel approach, we were able to correlate the various measures within individuals while taking the nested structure (i.e., time-on-task blocks are nested within individuals) of the data into account. The use of multilevel analyses is justified when there is sufficient variance explained at two or more levels of analysis. The intraclass

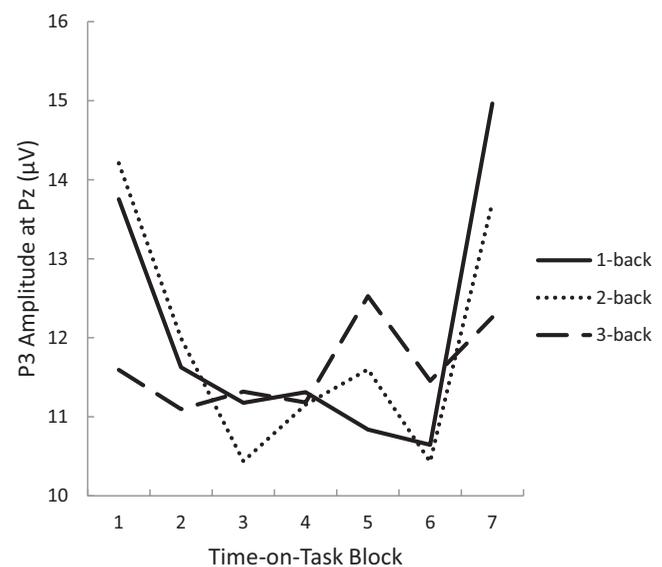
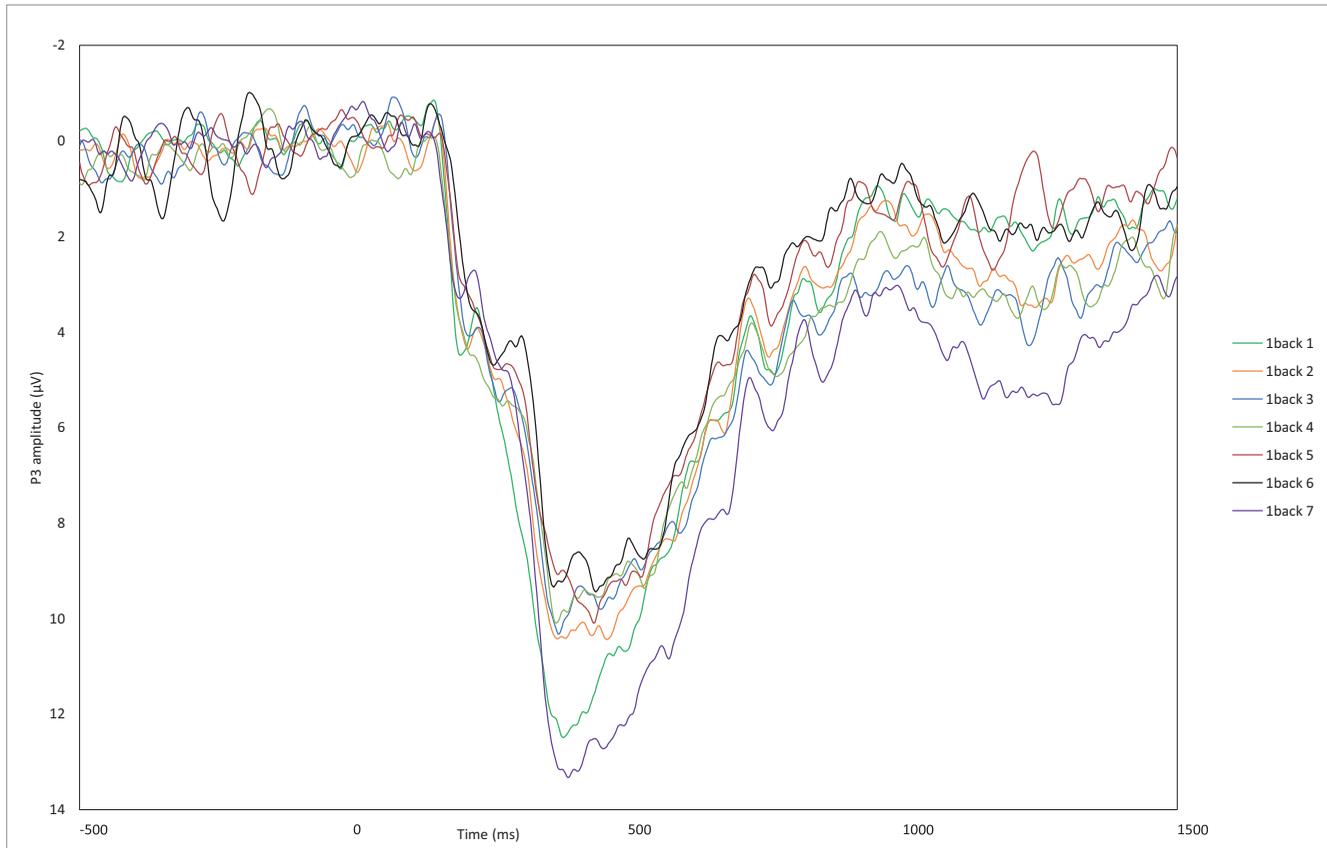
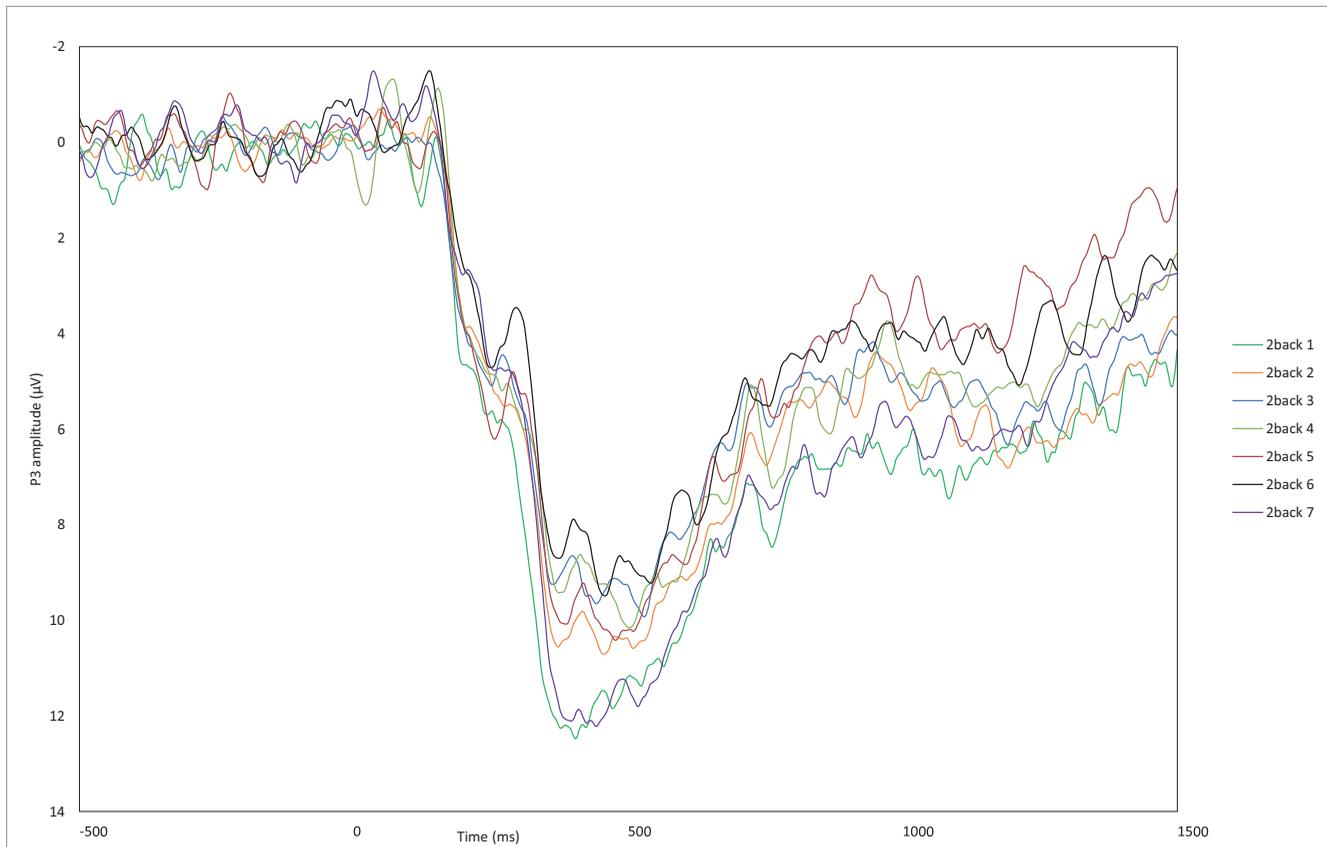


Figure 6. P3 amplitude with time-on-task during the 1-back, 2-back, and 3-back task.

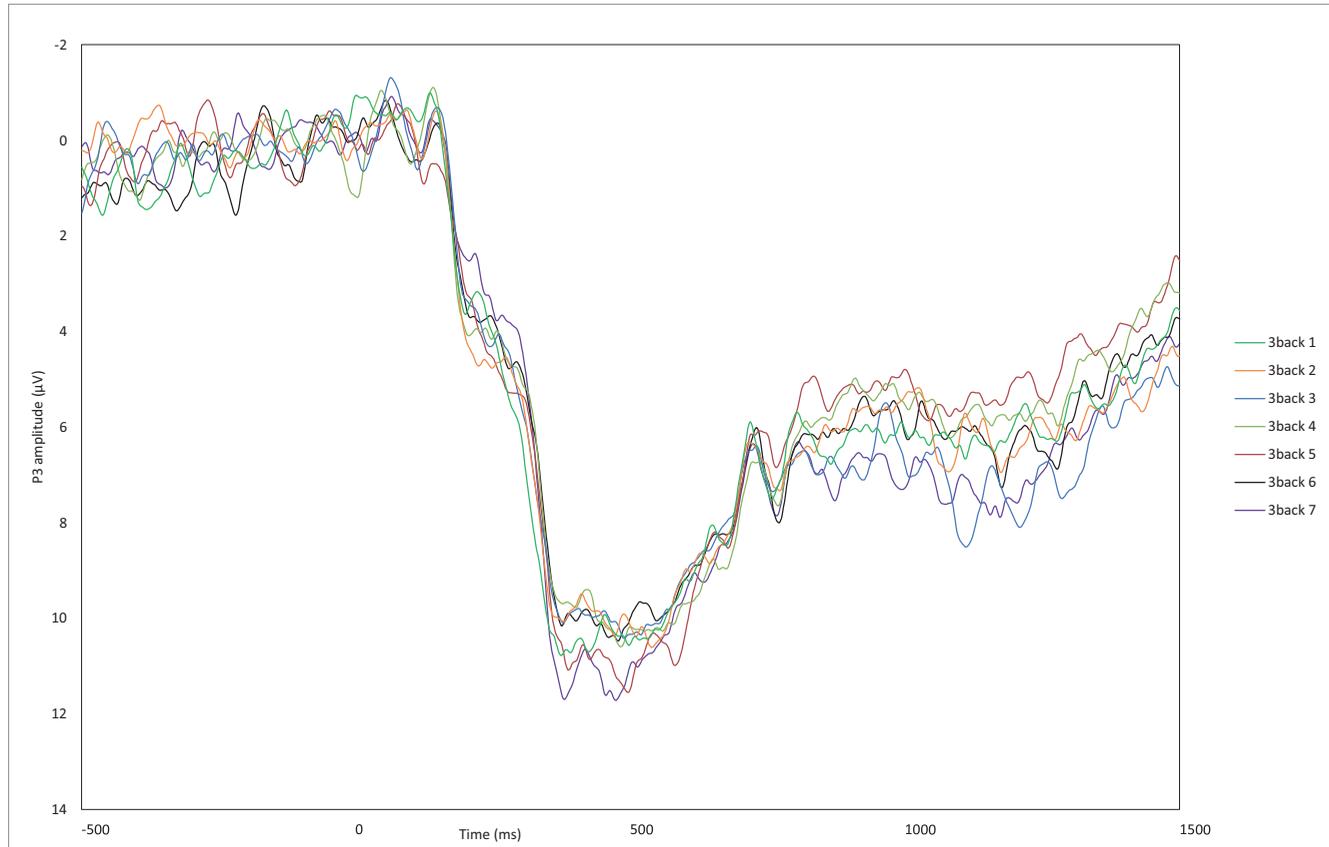


(a)



(b)

Figure 7. Average P3 amplitudes for each time-on-task block during the (a) 1-back, (b) 2-back, and (c) 3-back task at electrode Pz.



(c)

Figure 7. Continued.

correlation (ICC), displayed in Table 1, indicated that there indeed was sufficient variance explained on both levels for each observed variable. Table 1 also shows the correlations for each pair of observed variables.

We found strong correlations between the P3 amplitude, *d* prime, and both subjective measures. This statistically confirms that, in general, a change in one type of measure (e.g., P3 amplitude) was accompanied by a change in another measure (e.g., *d* prime), underlining the link between these variables regardless of when this change took place. We also found correlations between baseline pupil diameter and both of the subjective measures. This correlation increased in strength with higher task difficulty. The correlations between the baseline pupil diameter with *d* prime and P3 did not reach significant levels.

Table 1a. Multilevel Correlations Between Measures on the 1-Back Task

	ICC	1	2	3	4
1. Subjective fatigue	0.29	–	–	–	–
2. Subjective task engagement	0.42	–0.70***	–	–	–
3. Performance	0.25	–0.41***	0.49***	–	–
4. P3 amplitude	0.54	–0.28***	0.31**	0.25*	–
5. Baseline pupil diameter	0.89	–0.24	0.35***	0.32*	0.11

p* < .05. *p* < .01. ****p* < .001.

The multilevel findings are important because they directly support the relatedness of the P3 amplitude, performance on the task, and subjective levels of fatigue and engagement. Our findings also suggest that the baseline (baseline pupil diameter) and stimulus-evoked effects (P3 amplitude) both moved in the hypothesized direction (i.e., declined over time), but their time trajectories differed, leading to nonsignificant correlation.

Discussion

The results of the present study confirm the link between mental fatigue, task disengagement, and impaired performance on cognitively demanding tasks. An important asset of the study was that we simultaneously measured subjective, behavioral, and

Table 1b. Multilevel Correlations Between Measures on the 2-Back Task

	ICC	1	2	3	4
1. Subjective fatigue	0.29	–	–	–	–
2. Subjective task engagement	0.42	–0.70***	–	–	–
3. Performance	0.22	–0.41***	0.48***	–	–
4. P3 amplitude	0.41	–0.27**	0.35***	0.36**	–
5. Baseline pupil diameter	0.86	–0.28*	0.33**	0.05	0.12

p* < .05. *p* < .01. ****p* < .001.

Table 1c. Multilevel Correlations Between Measures on the 3-Back Task

	ICC	1	2	3	4
1. Subjective fatigue	0.30	–	–	–	–
2. Subjective task engagement	0.42	–0.69***	–	–	–
3. Performance	0.47	0.02	0.07	–	–
4. P3 amplitude	0.64	0.05	–0.09	0.33**	–
5. Baseline pupil diameter	0.84	–0.61***	0.51***	0.11	0.06

* $p < .05$. ** $p < .01$. *** $p < .001$.

physiological responses and directly tested the association between these measures. We found that, with increasing time-on-task, subjective mental fatigue increased and the participants' task engagement, as measured by P3 and baseline pupil diameter, decreased. At the behavioral level, this disengagement was accompanied by a decline in cognitive performance. We also found that the detrimental effects of fatigue on the subjective, physiological, and performance measures could be reversed by increasing the task rewards. Increasing the rewards led to task reengagement in spite of previous signs of fatigue. Importantly, this reengagement was accompanied by increased pupil diameter and P3 amplitude. The pattern of results in the first six time-on-task blocks suggests that motivation decreased to a point where resources were no longer fully invested. An explanation could be that disengagement occurs to preserve resources for the possibility of encountering more rewarding tasks in the future. By increasing the motivation (i.e., increasing the task rewards) for engagement, it becomes worthwhile to reengage in the task to prevent cognitive failures and keep up task performance.

Task Disengagement and Mental Fatigue

It has been argued that compromised cognitive performance under fatigue might be related to unfavorable tradeoffs between the cost and rewards of task engagement (Boksem & Tops, 2008). Moreover, based on various psychophysiological markers, it has been suggested that the dopamine pathways that are involved in the evaluation of reward information also play an important role in the effects of mental fatigue (e.g., Lorist et al., 2009). Examples of psychophysiological indicators that were used involve the error-related negativity (ERN) and the novelty P3a. In the present study, however, we used a different set of psychophysiological indicators (i.e., pupil diameter and P3b), which have received considerably less attention within the mental fatigue context. These measures seem promising, because there are numerous studies that relate the pupil diameter and the P3 ERP on the one hand, and several attentional and motivational processes that modulate task engagement on the other (e.g., Gilzenrat et al., 2010; Murphy et al., 2011). Based on the present study, we suggest that the time-on-task trajectories of these physiological measures are possible related to the time-on-task trajectories of fatigue and cognitive performance. This is a novel way to measure the role that task disengagement plays in the effects of mental fatigue, and creates opportunities for the use of the P3 and pupil diameter as measures in future fatigue research.

The results of this study also show the flexibility of the P3 and pupil diameter during periods of mental fatigue, and their responsiveness to contingencies in the environment. After we increased the rewards, the P3 amplitude and pupil diameter returned to values that compare to those seen at the start of the experiment, even

though the participants had already been engaged in the task for 2 h and reported high subjective levels of fatigue. Previous studies have shown that, under the right circumstances (i.e., when the rewards are high), participants are able to uphold task performance for a long time, even under high levels of mental fatigue (Boksem et al., 2006; Eccles & Wigfield, 2002; Tops & Boksem, 2010). The effects observed in our reward manipulation confirm this, and show that the P3 and pupil diameter follow the same trajectory as task performance, suggesting reengagement in the task. This specific finding may have broader impact in other fields of research such as self-control. For example, our results seem to overlap with recent findings suggesting that it is unlikely that self-control resources are fundamentally depleted as stated by classical "ego depletion" theory (Baumeister, Bratslavsky, Muraven, & Tice, 1998). Instead, based on our findings, we favor the explanation by Inzlicht and Schmeichel (2012), in which shifts in motivation and attention play an important role in task disengagement. In the present study, disengagement could be caused by diminished predicted rewards (i.e., the task becomes less interesting, but still requires the same level of attention invested). Most importantly, the present findings contradict the theory of depleted resources, because increased motivation leads to reengagement and restored performance on a demanding task. We believe that this is an important contribution to the literature.

Underlying Physiological System

While the present findings provide additional information about the cognitive, motivational, and subjective processes involved in fatigue, they also allow us to further speculate about the involvement of an underlying neuropsychological system. Given that both the P3 and pupil diameter measures were affected by the time-on-task manipulation, there is an interesting possibility that the locus coeruleus norepinephrine (LC-NE) system plays a role in the effects of fatigue. Although there is still much debate about the system underlying the P3 and pupil diameter, an increasing number of recent studies have described that these psychophysiological markers reflect activity of the LC (e.g., Murphy et al., 2011; Murphy, O'Connell, O'Sullivan, Robertson, & Balsters, 2014; Nieuwenhuis, Aston-Jones, & Cohen, 2005). The LC is a nucleus in the brainstem responsible for the release of cortical NE with ascending connections to large parts of the cortex. Nieuwenhuis and colleagues (2005; Nieuwenhuis, 2011) provided an extensive overview of intracranial, pharmacological, and lesion studies with primates suggesting that the P3 may reflect a correlate of stimulus-evoked LC-NE activity. The exact neural pathways that connect the activity of the LC with the pupil and the P3 measures are still under debate and may well be different parallel processes (Nieuwenhuis, de Geus, & Aston-Jones, 2010). To date, there have been relatively few studies that have empirically tested in humans. An early attempt to this comes in the form of a study by Murphy and colleagues (2014), which shows that pupil diameter covaries with blood oxygen level-dependent (BOLD) activity in the human LC.

The possible involvement of the LC-NE system in task engagement is supported by strong projections to the LC from the orbitofrontal and anterior cingulate cortices (Aston-Jones et al., 2002; Rajkowski, Lu, Zhu, Cohen, & Aston-Jones, 2000), which are known to be important in the evaluation of the rewards and costs of a task and interact with the dopamine system (Gottfried, O'Doherty, & Dolan, 2003; Holroyd & Yeung, 2012; McClure, York, & Montague, 2004; O'Doherty, Kringelbach, Rolls, Hornak, & Andrews, 2001; Yeung, Holroyd, & Cohen, 2005). When the

cost/reward tradeoff of a task is favorable, the LC-NE system stimulates exploitation of the rewards by engaging in the task. This enhances cognitive performance and maximizes the benefits of the task. However, when the tradeoff becomes unfavorable, the LC-NE system stimulates disengagement from the task. This is manifested in the tendency to explore the environment for more rewarding tasks and subsequently increases the probability of failures in task-related behavior (e.g., failing to respond to important task-related cues). Aston-Jones and Cohen state that “descending regulation of LC suggests a mechanism for volitional control of waking in the face of fatigue” (Aston-Jones & Cohen, 2005, p. 431). Taking into account that the debate about the extent to which the P3 and pupil diameter are related to the LC-NE system is still ongoing, there is an apparent theoretical and promising psychophysiological explanation of the link between the LC-NE system and task engagement. Therefore, it seems relevant to further explore the possible influence of the LC-NE system on the link between mental fatigue and task engagement.

Limitations

While most of our findings were in line with our hypotheses, a few results remain more open for interpretation. For example, we found that, in contrast to the relatively easier 1-back and 2-back tasks, the performance on the 3-back task did not show a decrease in performance over time, but remained more or less stable. Note, however, that in the 3-back task the level of performance at the start of the task was already much lower. This indicated that the task was more difficult than the other two variants and increases the probability that the observed pattern in performance over time not only reflected a fatigue effect, but also a learning curve. Actually, such a blending of learning and fatigue effects is quite common in the time-on-task studies that use demanding tasks and has been identified in several other studies (Faber et al., 2012; van der Linden, Frese, & Meijman, 2003). In our study, the increase in performance on the 3-back task could also be caused by a learning effect, and may mask the decreased performance due to mental fatigue. This idea is supported by the finding that, after the motivational manipulation, performance levels on the 3-back task were even higher than

in the beginning of the experiment. In contrast, in the 1-back and 2-back versions of the task, performance reverted to levels obtained at the beginning of the experiment. Thus, with the effects of fatigue diminished by the manipulation, possible learning effects became apparent. It should also be noted that the order of the 1-, 2-, and 3-back tasks was fixed within each of the seven time-on-task blocks. Because the tasks alternated relatively often, this should not have a major influence on the general time-on-task effects, but careful interpretation of the task difficulty effects is advised. In a study focused mainly on the task difficulty effect, it may be better to counterbalance the tasks within blocks.

Concluding Remarks

The present study provides evidence for the involvement of the P3 ERP and pupil diameter in the process of task disengagement that coincides with mental fatigue. With a multifaceted approach, we revealed a relation between subjective (i.e., subjective fatigue and engagement), behavioral (i.e., task performance), and physiological measures (i.e., P3 and pupil diameter) of fatigue and task engagement. These findings may help to refine knowledge about neurocognitive mechanisms of fatigue. We also speculated that the LC-NE system may play a role as one of the underlying systems that supports task engagement and disengagement. As such, the present study may also contribute to insight into the ways of dealing with the health and safety issues commonly associated with fatigue. For example, they may be helpful in the development of psychopharmacological interventions that target fatigue symptoms in patient groups. The addition of norepinephrine agents to the predominantly used selective serotonin reuptake inhibitors (SSRIs) may have the potential to further ameliorate fatigue effects in patients with Parkinson’s disease, depression, and burnout symptoms (Stahl, 2002). In addition to health-related issues, knowing which neurocognitive systems drive the tendency to reduce task engagement after sustained performance may also be relevant for developing a workplace environment that prevents mental fatigue or at least minimizes its negative consequences. After all, human factors, and specifically mental fatigue, remain the most important reason for errors and accidents in the workplace (Baker et al., 1994; McCormick et al., 2012).

References

- Aston-Jones, G., & Cohen, J. D. (2005). An integrative theory of locus coeruleus-norepinephrine function: Adaptive gain and optimal performance. *Annual Review of Neuroscience*, *28*, 403–450.
- Aston-Jones, G., Rajkowski, J., Lu, W., Zhu, Y., Cohen, J., & Morecraft, R. (2002). Prominent projections from the orbital prefrontal cortex to the locus coeruleus in monkey. *Society for Neuroscience Abstracts*, *28*, 86–89.
- Baker, K., Olson, J., & Morisseau, D. (1994). Work practices, fatigue, and nuclear power plant safety performance. *Human Factors*, *36*, 244–257.
- Baumeister, R. F., Bratslavsky, E., Muraven, M., & Tice, D. M. (1998). Ego depletion: Is the active self a limited resource? *Journal of Personality and Social Psychology*, *74*, 1252–1265.
- Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychological Bulletin*, *91*, 276–292.
- Begleiter, H., Porjesz, B., Chou, C., & Aunon, J. (1983). P3 and stimulus incentive value. *Psychophysiology*, *20*, 95–101.
- Boksem, M. A. S., Meijman, T. F., & Lorist, M. M. (2006). Mental fatigue, motivation and action monitoring. *Biological Psychology*, *72*, 123–132.
- Boksem, M. A. S., & Tops, M. (2008). Mental fatigue: Costs and benefits. *Brain Research Reviews*, *59*, 125–139.
- Chaudhuri, A., & Behan, P. O. (2004). Fatigue in neurological disorders. *Lancet*, *363*, 978–988.
- Chen, Y. N., Mitra, S., & Schlaghecken, F. (2008). Sub-processes of working memory in the N-back task: An investigation using ERPs. *Clinical Neurophysiology*, *119*, 1546–1559.
- Cohen, J. D., McClure, S. M., & Yu, A. J. (2007). Should I stay or should I go? How the human brain manages the trade-off between exploitation and exploration. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *362*, 933–942.
- Cohen, J. D., Perlstein, W. M., Braver, T. S., Nystrom, L. E., Noll, D. C., Jonides, J., & Smith, E. E. (1997). Temporal dynamics of brain activation during a working memory task. *Nature*, *386*, 604–608.
- Dehaene, S., Kerszberg, M., & Changeux, J.-P. (1998). A neuronal model of a global workspace in effortful cognitive tasks. *Proceedings of the National Academy of Sciences*, *95*, 14529–14534.
- Demyttenaere, K., De Fruyt, J., & Stahl, S. M. (2005). The many faces of fatigue in major depressive disorder. *International Journal of Neuropsychopharmacology*, *8*, 93–105.
- Eccles, J. S., & Wigfield, A. (2002). Motivational beliefs, values, and goals. *Annual Review of Psychology*, *53*, 109–132.
- Faber, L. G., Maurits, N. M., & Lorist, M. M. (2012). Mental fatigue affects visual selective attention. *PLoS One*, *7*, e48073.

- Gilzenrat, M. S., Nieuwenhuis, S., Jepma, M., & Cohen, J. D. (2010). Pupil diameter tracks changes in control state predicted by the adaptive gain theory of locus coeruleus function. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 252–269.
- Gottfried, J. A., O'Doherty, J., & Dolan, R. J. (2003). Encoding predictive reward value in human amygdala and orbitofrontal cortex. *Science, 301*, 1104–1107.
- Gratton, G., Coles, M. G. H., & Donchin, E. (1983). A new method for off-line removal of ocular artifact. *Electroencephalography and Clinical Neurophysiology, 55*, 468–484.
- Hockey, G. R. J. (1997). Compensatory control in the regulation of human performance under stress and high workload: A cognitive-energetical framework. *Biological Psychology, 45*, 73–93.
- Holroyd, C. B., & Yeung, N. (2012). Motivation of extended behaviors by anterior cingulate cortex. *Trends in Cognitive Sciences, 16*, 122–128.
- Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Toward a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science, 7*, 450–463.
- Jepma, M., & Nieuwenhuis, S. (2011). Pupil diameter predicts changes in the exploration-exploitation trade-off: Evidence for the adaptive gain theory. *Journal of Cognitive Neuroscience, 23*, 1587–1596.
- Johnson, R. (1993). On the neural generators of the P300 component of the event-related potential. *Psychophysiology, 30*, 90–97.
- Kahneman, D. (1973). *Attention and effort*. Upper Saddle River, NJ: Prentice-Hall Inc.
- Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583–1585.
- Lorist, M. M. (2008). Impact of top-down control during mental fatigue. *Brain Research, 1232*, 113–123.
- Lorist, M. M., Bezdán, E., ten Caat, M., Span, M. M., Roerdink, J. B. T. M., & Maurits, N. M. (2009). The influence of mental fatigue and motivation on neural network dynamics: An EEG coherence study. *Brain Research, 1270*, 95–106.
- Lorist, M. M., Boksem, M. A. S., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research, 24*, 199–205.
- Lorist, M. M., Klein, M., Nieuwenhuis, S., Jong, R., Mulder, G., & Meijman, T. F. (2000). Mental fatigue and task control: Planning and preparation. *Psychophysiology, 37*, 614–625.
- Maslach, C., Schaufeli, W. B., & Leiter, M. P. (2001). Job burnout. *Annual Review of Psychology, 52*, 397–422.
- Massar, S. A., Wester, A. E., Volkerts, E. R., & Kenemans, J. L. (2010). Manipulation specific effects of mental fatigue: Evidence from novelty processing and simulated driving. *Psychophysiology, 47*, 1119–1126.
- McClure, S. M., York, M. K., & Montague, P. R. (2004). The neural substrates of reward processing in humans: The modern role of FMRI. *Neuroscientist, 10*, 260–268.
- McCormick, F., Kadzielski, J., Landrigan, C. P., Evans, B., Herndon, J. H., & Rubash, H. E. (2012). Surgeon fatigue: A prospective analysis of the incidence, risk, and intervals of predicted fatigue-related impairment in residents surgeon fatigue. *Archives of Surgery, 147*, 430–435.
- Meijman, T. F. (1997). Mental fatigue and the efficiency of information processing in relation to work times. *International Journal of Industrial Ergonomics, 20*, 31–38.
- Murphy, P. R., Robertson, I. H., Balsters, J. H., & O'Connell, R. G. (2011). Pupillometry and P3 index the locus coeruleus-noradrenergic arousal function in humans. *Psychophysiology, 48*, 1532–1543.
- Murphy, P. R., O'Connell, R. G., O'Sullivan, M., Robertson, I. H., & Balsters, J. H. (2014). Pupil diameter covaries with BOLD activity in human locus coeruleus. *Human Brain Mapping, 35*, 4140–4154.
- Muthén, L. K., & Muthén, B. O. (1998). *Mplus user's guide* (7th ed.). Los Angeles, CA.
- Nieuwenhuis, S. (2011). Learning, the P3, and the locus coeruleus-norepinephrine system. In R. Mars, J. Sallet, M. Rushworth, & N. Yeung (Eds.), *Neural Basis of Motivational and Cognitive Control* (pp. 209–222). Oxford, UK: Oxford University Press.
- Nieuwenhuis, S., Aston-Jones, G., & Cohen, J. D. (2005). Decision making, the P3, and the locus coeruleus-norepinephrine system. *Psychological Bulletin, 131*, 510–532.
- Nieuwenhuis, S., de Geus, E. J., & Aston-Jones, G. (2010). The anatomical and functional relationship between the P3 and autonomic components of the orienting response. *Psychophysiology, 48*, 162–175.
- O'Doherty, J., Kringelbach, M. L., Rolls, E. T., Hornak, J., & Andrews, C. (2001). Abstract reward and punishment representations in the human orbitofrontal cortex. *Nature Neuroscience, 4*, 95–102.
- Polich, J. (2007). Updating P300: An integrative theory of P3a and P3b. *Clinical Neurophysiology, 118*, 2128–2148.
- Rajkowski, J., Lu, W., Zhu, Y., Cohen, J., & Aston-Jones, G. (2000). Prominent projections from the anterior cingulate cortex to the locus coeruleus in Rhesus monkey. *Society for Neuroscience Abstracts, 26*, 838.15.
- Snijders, T., & Bosker, R. (1999). *Multilevel modeling: An introduction to basic and advanced multilevel modeling*. London, UK: Sage.
- Stahl, S. M. (2002). The psychopharmacology of energy and fatigue. *Journal of Clinical Psychiatry, 63*, 7–8.
- Stern, R. M., Ray, W. J., & Quigley, K. S. (2000). *Psychophysiological Recording*. New York, NY: Oxford University Press.
- Tops, M., & Boksem, M. A. S. (2010). Absorbed in the task: Personality measures predict engagement during task performance as tracked by error negativity and asymmetrical frontal activity. *Cognitive, Affective, & Behavioral Neuroscience, 10*, 441–453.
- van der Linden, D. (2011). The urge to stop: The cognitive and biological nature of acute mental fatigue. In P. L. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 149–164). Washington, DC: American Psychological Association.
- van der Linden, D., & Eling, P. (2006). Mental fatigue disturbs local processing more than global processing. *Psychological Research, 70*, 395–402.
- van der Linden, D., Frese, M., & Meijman, T. F. (2003). Mental fatigue and the control of cognitive processes: Effects on perseveration and planning. *Acta Psychologica, 113*, 45–65.
- van der Linden, D., Frese, M., & Sonnentag, S. (2003). The impact of mental fatigue on exploration in a complex computer task: Rigidity and loss of systematic strategies. *Human Factors, 45*, 483–494.
- Watter, S., Geffen, G. M., & Geffen, L. B. (2001). The n-back as a dual-task: P300 morphology under divided attention. *Psychophysiology, 38*, 998–1003.
- Wickelgren, W. A. (1977). Speed-accuracy tradeoff and information processing dynamics. *Acta Psychologica, 41*, 67–85.
- Wickens, T. D. (2001). *Elementary signal detection theory*. Upper Saddle River, NJ: Oxford University Press.
- Yerkes, R. M., & Dodson, J. D. (1908). The relation of strength of stimulus to rapidity of habit-formation. *Journal of Comparative Neurology and Psychology, 18*, 459–482.
- Yeung, N., Holroyd, C. B., & Cohen, J. D. (2005). ERP correlates of feedback and reward processing in the presence and absence of response choice. *Cerebral Cortex, 15*, 535–544.
- Zijlstra, F. R. H. (1993). *Efficiency in work behaviour: A design approach for modern tools*, (Doctoral dissertation). Delft, The Netherlands: Delft University Press. Retrieved from http://repository.tudelft.nl/assets/uuid:d97a028b-c3dc-4930-b2ab-a7877993a17f/ide_zijlstra_19931123.PDF

(RECEIVED November 25, 2013; ACCEPTED August 21, 2014)