Psicológica (2020), 41, 103-126
doi: 10.2478/psicolj-2020-0006

Inconsistencies between mental fatigue measures under compensatory control theories

Enrique Muñoz-de-Escalona[1*], José J. Cañas[1], Paulo Noriega[2]

[1]*Cognitive Ergonomics Group, Mind, Brain and Behavior Research Centre (CIMCYC), Department of Experimental Psychology, University of Granada, Spain*

[2]*ErgoUX Lab, CIAUD, Faculdade de Arquitetura, Universidade de Lisboa, Portugal*

Mental fatigue has traditionally been defined as a condition of reduced cognitive efficiency and performance, accompanied by a subjective feeling of fatigue. Even though we could expect to find associations between the three defining characteristic of mental fatigue (performance impairment, physiological deactivation and subjective fatigue), research has shown that the emergence of inconsistencies between measures is more frequent than one might expect: people proved capable of maintaining adequate performance levels even after having declared themselves fatigued. This could be explained under the compensatory control mechanism models, which state that humans are able to provide additional resources under demanding conditions, but only at the expense of psychophysiological cost and subjective fatigue. We tested this explanation by manipulating task complexity and time performing a simulated air-traffic control task. We collected psychophysiological, performance and subjective data. A decrease in pupil size was seen in the low-aircraft-density condition, while pupil size remained constant in the high-aircraft-density condition. Participants' task performance was optimal in both conditions, though they showed an increase in subjective feelings of fatigue, especially in the high-complexity task condition. Thus, complexity seemed to trigger compensatory mechanisms, which reallocated extra resources that physiologically activated participants in order to deal with a higher complexity task, whereas subjective fatigue could be acting as a signal to the organism of impending resource depletion. Our findings support compensatory control theories and offer an explanation of inconsistencies between fatigue measures. Further research on compensatory mechanisms is needed to enable better management of fatigue effects to prevent work-related accidents.

Abbreviations. TOT: Time On Task, ATCo: Air Traffic Controller.

---

[*]**Corresponding author:** Enrique Muñoz-de-Escalona, Universidad de Granada, Department of Experimental Psychology, Campus de Cartuja s/n, 18071, Granada, Spain. E-mail: enriquemef@ugr.es. **Acknowledgements.** This research has been funded by Project AIRPORTS from the Spanish Ministry of Industry, Grant number: PI-1461/2015.

Human behaviour and mental activity require energy. In the same way that a car needs petrol and oxygen to move through streets, human beings need energy from food to perform mental and motor activities. In a sense, we could say that in the life sciences, research has followed a mechanistic paradigm according to which human machinery function depends on supplied energy. Thus, for instance, it is assumed that performance on a task will improve or deteriorate depending, among other things, on the quantity and quality of the energy (resources) supplied (Kahneman, 1973).

An important feature of the mechanistic paradigm is the assumption that resources are limited and can be depleted. As a result, performance could be affected by the emergence of mental fatigue (Gopher and Donchin, 1986; Kahneman, 1973). Mental fatigue lowers mental alertness levels, and thus performance deteriorates (Grandjean, 1989) after long periods of 'demanding cognitive activity which requires sustained mental efficiency' (Lorist et al., 2005). Therefore, mental fatigue has been considered responsible for performance decrements as well as the main cause of work-related accidents (Grandjean, 1989; Dawson, et al., 2011; Hopstaken et al, 2016).

Many studies have tried to define and establish neuropsychological, psychosocial and even bio-mathematical models of fatigue (Dawson et al, 2011). However, almost all models agree that fatigue has three well-established components, two of which are observable (performance and physiological) and one of which is a non-observable subjective feeling. Thus, based on conventional definition of fatigue, it can be defined as the deterioration of task performance with Time on Task (TOT), accompanied by a decrease in physiological activation and the emergence of subjective feelings of fatigue. For this reason, mental fatigue has been approached with a methodology that includes three kinds of measures corresponding to the three defining characteristics of mental fatigue (Hockey, 2013):

- **Performance measures:** this type of methodology is composed of primary task measures (e.g. nº of errors, reaction time, etc.) and secondary task measures (e.g. choice reaction-time tasks, time estimation, memory-search tasks etc.). The aim is to measure objective performance indexes in order to quantify the quantity and the quality of performed tasks. Mental fatigue would result in task performance impairment (Chalder et al., 1993; Smets et al., 1995; Van der Linden, et al., 2003; Barker and Nussbaum, 2011; Vrijkotte, et al., 2018).
- **Subjective measures:** this sort of methodology includes all self-reported measures (e.g. NASA Task Load Index or Multidimensional Fatigue Inventory (MFI) (Smets et al., 1995)). The aim is to obtain easy and low-cost subjective data about perceived mental fatigue.

Mental fatigue would induce a subjective feeling of fatigue (DeLuca, 2005; Kanfer, 2011).

- **Physiological measures**: this type of methodology comprises every physiological response to mental activation (e.g. pupil diameter, heart rate variability (HRV), electroencephalogram (EEG), etc.). The aim is to collect objective physiological data, which reacts to mental fatigue. Literature research has shown that physiological changes arise when a person is experiencing mental fatigue (Kohl, 2016). EEG measurement techniques are considered one of the most effective methods for assessing mental fatigue, as reflected by numerous studies. For example, Zhang et al. (2006) found that Approximate Entropy of EEG can serve as an index for detecting mental fatigue level (Zhang et al., 2008). Moreover, it has also be found that relative wavelet packet energy in β, β/α, (α+θ)/ β frequency bands and wavelet packet entropy of EEG are strongly correlated with mental fatigue (Nie, 2005; Zhang and Yu, 2010). Shen et al. (2008) tested an EEG-based mental fatigue monitoring system using a probabilistic-based support vector-machines (SVM) method, which provided a valuable estimation of confidence in the prediction of mental fatigue. On the other hand, eye parameters such as pupil dilation, blinks, fixations and saccades have also been widely used for estimating mental workload (Hess & Polt,1964; Beatty,1982; Nakayama et al., 2002; Bailey & Iqbal, 2008; Di Stassi et al., 2010; Jainta and Baccino, 2010; Benedetto et al. 2011; Lemonnier, Roland and Baccino, 2014; De Alwis, 2017; Yokoyama et al. 2018) as well as for mental fatigue (Roy et al., 2014; Horiuchi, 2017). Researchers have found high positive correlations between pupil diameter changes indicative of decreased physiological activation and the amount of resources used to perform certain tasks (Kahneman and Beatty, 1966; Beatty and Lucero-Wagoner, 2000; LeDuc, et al., 2005; Van Orden et al., 2001; Ahlstrom and Friedman-Berg, 2006). For example, Hopstaken et al. (2015) found that increasing mental fatigue coincided with diminished stimulus-evoked pupil dilation; similarly, Hopstaken et al (2016) found that baseline pupil diameter significantly decreased with time on task throughout the experiment as fatigue increased, which was in line with the former findings.

Presumably, since the three types of measures reflect the same concept, we would expect them to be interrelated. In other words, we would expect correlations within measurements (associations), in such a manner that we could be able to detect fatigue with each kind of measure independently.

However, research has found inconsistencies (disassociations and insensitivities) between measures more frequently than one might expect (Johnson, et al. 1997; Johnson, et al.1998; Paul et al., 1998; Krup and Elkins, 2000; Christodoulou, 2005; DeLuca, 2005; Bailey, et al. 2007; Bruce et al., 2010; Walker et al., 2012; Ishii et al., 2013; Hockey, 2013, Hancock, 2017). An insensitivity occurs when a workload measure does not reflect any change with a change in task load, whereas a dissociation takes place when that measure contradict the change in task load (Hancock, 2017). For example, it has frequently been found that people performing prolonged difficult tasks have exhibited no discernible perfomance decrements even after having declared that they are fatigued (for a review see Hockey, 2013). Additionally, as Gilbert (2009) has stated, most inconsistencies have involved the subjective component of the individual. At this point, it is unclear why such discrepancies between the results for different types of measures which are supposed to be measuring the same construct are so common, and the issue of how and why in certain situations fatigue does not affect performance or physiological activation when a person claims feeling fatigued requires explanation and further investigation.

With that in mind, some researchers have proposed new resource theories which postulate the existence of a compensatory control mechanism that regulates performance under stress and high mental workload conditions. Specifically, the function of this mechanism is to provide additional resources dynamically, depending on the level of stress and mental workload experienced. Thus, task performance will not necessarily be impaired in demanding and fatiguing conditions (e.g., a high-complexity task) due to this provision of extra resources needed to perform the task successfully, but at the expense of psychophysiological cost and subjective fatigue. Hockey (1997) proposed the compensatory control model of fatigue, which asserts that 'the maintenance of performance stability under demanding conditions is an active process under the control of the individual, requiring the management of cognitive resources through the mobilisation of mental effort' (p.78).

The compensatory control model of fatigue makes a distinction between two levels of control (Figure 1): a lower control loop, which manages routine regulatory activity, and an upper loop, which is needed to deal with harder challenges to performance. Even initially, executive control is exerted to the extent that the achievement of a selected task goal (G) is prioritized over many other possible contender goals. However, the other possible goals (g1, g2, etc) remain relevant to motivational needs of the individual (rest, change task, etc.). But if strain is detected in the lower control loop, the upper control loop kicks in. A choice is made between two possible options: 1) increasing effort (reallocating extra resources) to protect performance from impairment, with an increase in costs or 2) maintaining the current level of effort with minimal costs, despite the risk of performance decline (either maintaining the current goal with a lower performance standard or displacing the goal with one of the competing goals). In either case, the feeling of fatigue is expected to disappear as strain becomes extinct (Hockey, 2011).

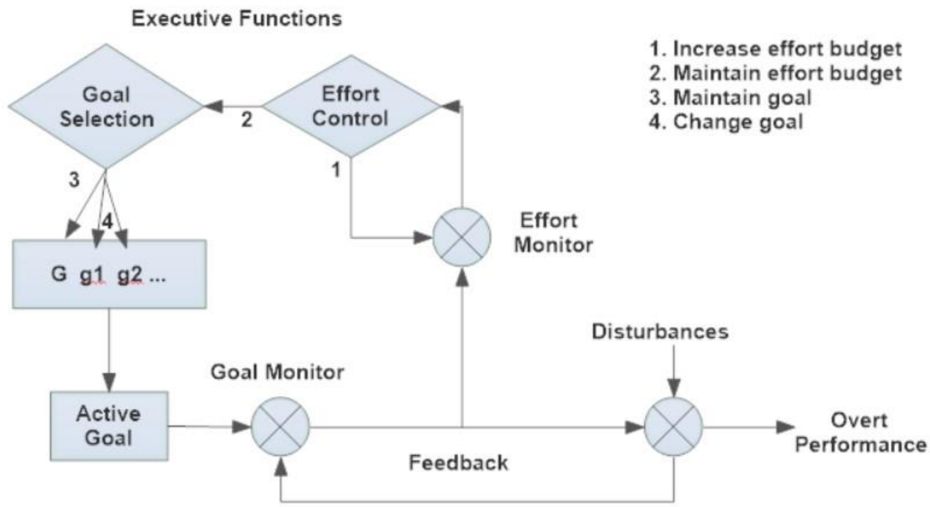


Figure 1. Compensatory control model of fatigue (after Hockey 1997)

Therefore, as Hockey (2011) suggests, 'the subjective fatigue state may be identified broadly with the outcome of the monitoring process, which detects control problems and the need for greater effort'(p.18), hence 'within the context of an effort-fatigue compensatory loop a sensed need for greater effort reflects the same affective state as a sensed increase in discomfort of fatigue'(p.18).

Even though there are some studies whose results are consistent with Hockey's compensatory control model (Sauer et al 2003; Venables and Fairclough 2009), we think further research is needed to strengthen compensatory control theories, particularly in relation to inconsistencies between fatigue measures. To address that need, and to provide greater knowledge about the phenomenon of mental fatigue, we conducted the experiment reported herein.

The main objective of this study is to test whether measurements of the aforementioned three aspects of fatigue correlate or not and, if the latter happens, to gather evidence regarding how and why such inconsistencies between measures occur, using the compensatory mechanism approach as our conceptual framework.

We manipulated two independent variables in this experiment. The first independent variable was 'Time on Task' (TOT), whereas the second independent variable was task complexity. An intuitive approach would lead us to believe that both of them are factors that can cause mental fatigue (Stern, 1997): Resources will run out as time is spent on task performance, so the more time is spent, the fewer resources will be available and therefore the level of mental fatigue will be greater. Moreover, it could also be assumed that task complexity determines the quantity of resources dedicated to a task, so that more complex tasks require more resources. Hence, the more complex the task is, the greater resource depletion will be with TOT and thus, the greater the level of mental fatigue. For example, we may think that the end of the working day would be a more likely time for work-related accidents than the beginning of the day, due to fatigue caused by resource depletion, especially with high complexity tasks.

Our hypothesis is that, in line with recent empirical findings (see a review in Hockey, 2013), we will find dissociations between the three kinds of fatigue measures when testing the effects of these two factors, TOT and complexity, that can be explained in terms of compensatory control mechanisms. Moreover, we predict that particularly subjective fatigue will not correlate with the physiological (pupil size) or performance indices because subjective fatigue functions as an alarm signal regarding resource depletion that kicks in even while there are still enough resources to perform the task properly, as high physiological and performance index scores would be reflecting.

The effects of TOT and complexity were tested in an air-traffic control simulation experiment in which participants were trained and instructed to avoid conflicts between aircraft. This complex and dynamic task allowed us to observe the possible effects of our manipulated independent variables, as well as our measured dependent variables. Participants performed the task for 120 minutes while their performance, pupil diameter and feelings of fatigue were measured.

MATERIALS AND METHOD

^{ATC}Lab-Advanced software. The software used for simulating ATCo tasks was an air traffic simulator called ^{ATC}Lab-Advanced, which is available for free public download. (Fothergill et al., 2009) (see Figure 2). It provided a high level of realism as well as a simplified and easy handling, which allowed it to be used by all participants in several learning sessions. Additionally, it allowed strong experimental control of air traffic scenarios parameters.

For scenario development, first, the static characteristics of the simulation environment (control sector size and possible pathways through which aircraft could travel) were defined. Next, aircraft quantity (density) and initial aircraft parameters (altitude, assigned altitude, speed, time of appearance on stage and planned route) were defined for each aircraft presented in the scenario.

Finally, we note that the ^{ATC}Lab-Advanced simulator provided participants with every needed tool to carry out the ATCo task, such as the route (the aircrafts' fixed route was displayed), distance scale (which allows horizontal aircraft distance measurement) and altitude and speed change tools. Once the structural and dynamic scenario parameters were established, a file that could be launched by the simulator was obtained. This file recorded a '.log' file with performance data (tools used during the task and aircrafts' parameters) for each participant during simulation.

Scenarios. Scenarios used in the study varied according to whether the participant was in the training or the experimental stage. Thus, during the training sessions, the standard scenarios provided by the software creator software were used, but a specific scenario was programmed by the experimenters to achieve experimental session goals. This specific scenario was programmed with two different experimental conditions.

The structural features of both experimental conditions were identical, that is, identical control sectors and routes, as well as the same number of total initial number of aircraft (9) and number of aircraft under participant control (6). For a better understanding, refer to Figure 2, which represents the initial simulator screen presented to participants; the capital black letters (starting route spots) do not appear on the radar screen.

Several differences in dynamic scenario parameters were established according to the scenario used in each condition:

1. **Low aircraft density scenario (low-complexity task condition):** A total of 35 aircraft were presented, with 25 coming from external locations A,D,E,F,W,P,N,L,M and Z and 10 from inner locations C and J. More specifically, 3 aircraft came from A,D,E,F and W; 2 from P,N,L,M and Z; 6 from J (4 coming to P and 2 to A); and 4 from C (coming to A,D,E and F respectively).
2. **High aircraft density scenario (high-complexity task condition):** A total of 70 aircraft were presented, thus double the density of the low aircraft density condition. 50 came from external locations

A,D,E,F,W,P,N,L,M and Z and 20 from inner locations. Specifically, 6 each came from A,D,E,F and W; 4 from P,N,L,M and Z; 12 from J (8 of them coming to P and 4 to A); and 8 from C (coming to A,D,E and F respectively).

It should be noted that, traffic density is always variable in the sense that the number of flights cannot be keep exact during the whole 2 hours session, as this would not be realistic. In this sense, there is a small intragroup variation, which was designed to be equivalent in both conditions.

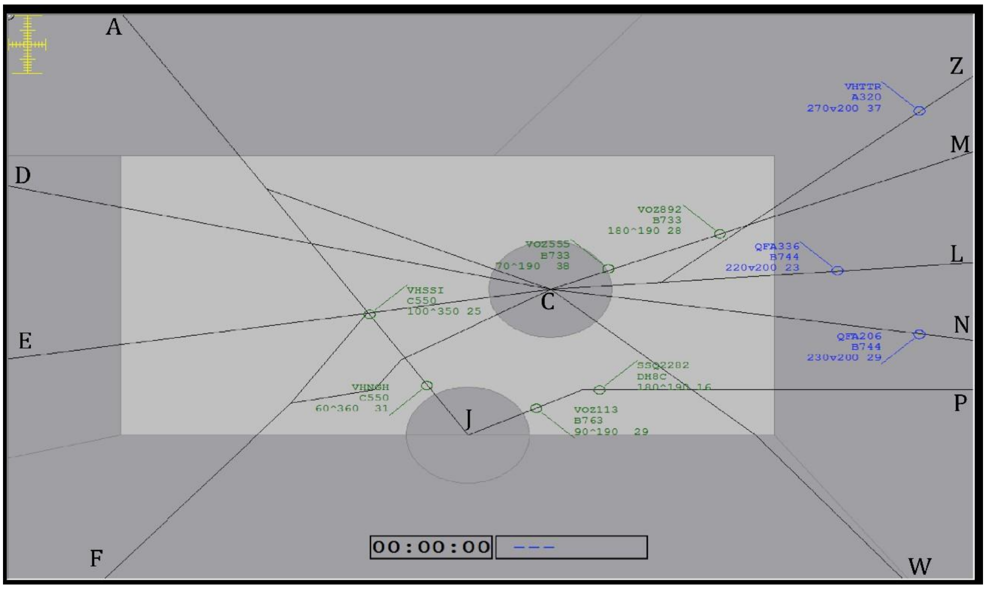


Figure 2. ^{ATC}Lab-Advanced software initial screenshot presented to participants

Procedure. Participants had to perform the ATCo tasks with the previously described ^{ATC}Lab-Advanced software, so they had to learn how to use it before proceeding through the experimental stage in which the performance data were collected. Thus, we established 2 distinct stages:

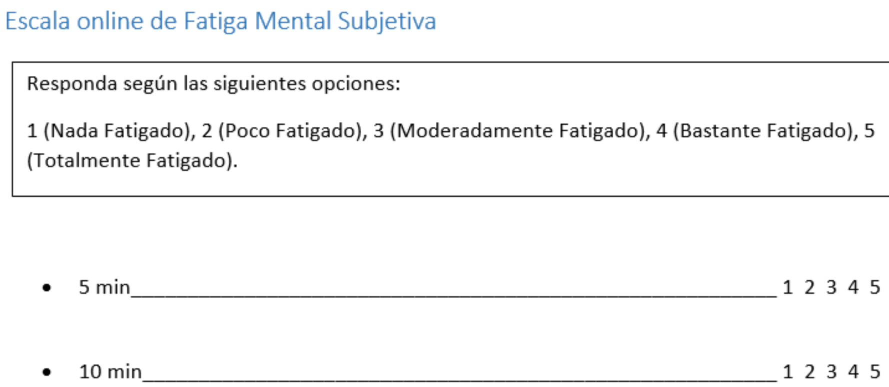
1. Training stage: This took place for a total of 2 hours over 2 consecutive days, 1 hour of training each day. The main objective of this first stage was for participants to familiarise themselves with the software so that they could handle it comfortably during the experimental stage. The training stage procedure was as follows: during the first day, researcher explained the participants their main task goal (maintaining air traffic security and preventing potential conflicts between aircraft), they started reading a short manual about the operation of the simulator

for about 20 minutes and were asked to call the researcher once they had finished. Later, the participants sat in front of the running simulator while the researcher reviewed the manual in detail with the participants to ensure both correct understanding of the task and the assimilation of knowledge through content review. Participants then started using the simulator on their own while the researcher executed different air traffic control scenarios in order of difficulty (3 scenarios). The participants had free access to both the manual and researcher at all times in case of doubts or questions. The researcher also periodically checked the participants' performance to monitor their learning. Through the second day, participants would continue working through scenarios (4 in total) in order of increasing difficulty, under the same conditions as the previous day. Once the training period concluded, participants were ready for the experimental session, which took place the next day.

2. Data collection stage: The aim of the data collection stage, which lasted a total of 2 total hours, was to collect experimental performance data from participants while they performed ATCo tasks. Both objective (physiological and execution data) and subjective (mental fatigue subjective index) data were collected. The participants were told the differences between the training and experimental stages, which were as follows: First, they would perform ATCo tasks in front of an eye-tracker system that had been previously calibrated. Secondly, participants were instructed to fill in the mental fatigue scale every 5 minutes, when a scheduled alarm sounded. The whole session lasted 2 hours. Once the differences were explained, participants began the task, with participants working in the low-aircraft-density condition (35 aircraft) or in the high-aircraft-density condition (70 aircraft). At the end of the session, the participants were thanked and given extra study credits.

Tobii T120 eyetracker. To obtain the pupil diameter variable, an infrared-based eye tracker system, the Tobii T120 Video System was used. This system achieves a minimum accuracy of .01 mm, is completely non-intrusive and provides high precision and an excellent head compensatory movement mechanism, which ensures high-quality data collection. In addition, the freedom of movement it offers participants allows them to act naturally in front of the screen, as if it were an ordinary computer display.

Subjective mental fatigue scale. We needed an on-line subjective mental fatigue measure, thus we developed an easy and intuitive subjective instant self-assessment fatigue scale with which participants could write down how much fatigue they were currently experiencing. The scale was designed to elicit a response at every 5-minute interval throughout the 2 hours of the experimental stage. The scale ranged from 1 'no mental fatigue' to 5 'maximum mental fatigue', presented from left to right in ascending amount of fatigue experienced. Participants were taught to use the scale just before experimental stage begin and were instructed about mental fatigue concept in order to avoid possible confusion with similar concepts such as workload or frustration (See Figure 3).


Escala online de Fatiga Mental Subjetiva

Responda según las siguientes opciones:

1 (Nada Fatigado), 2 (Poco Fatigado), 3 (Moderadamente Fatigado), 4 (Bastante Fatigado), 5 (Totalmente Fatigado).

- 5 min____ 1 2 3 4 5
- 10 min____ 1 2 3 4 5

Figure 3. Instantaneous self-assessment subjective fatigue scale

Participants. Seventy psychology students at the University of Granada participated in the study for extra course credit. Participants were divided into 2 groups, 34 participants in the low complexity condition and 36 participants in the high complexity condition. Participants' ages ranged from 18 to 27, with an average of 21.6 and a median of 20. A total of 56 women (80%) and 14 men (20%) participated. This gender imbalance is likely the result of the fact that psychology students at University of Granada are mostly women. A further requirement was that none of the participants had any previous experience in ATCo tasks. Informed consent was obtained from each participant.

This study was carried out in accordance with the recommendations of the local ethical guidelines of the committee of the University of Granada institution called "Comité de Ética de Investigación Humana". The protocol was approved by the "Comité de Ética de Investigación Humana". All subjects gave written informed consent in accordance with the Declaration of Helsinki.

Experimental room conditions. Sessions were held in several different rooms, depending on whether the participant was in the training or data collection stage: During the training stage, participants could work in one of three different rooms equipped for training with the simulator, and no special attention to room conditions was needed. However, during the data collection stage, standardizing room conditions was essential. Pupillary responsiveness is related to interaction of sympathetic and parasympathetic components and depends on three kind of reflexes: light, proximity and psychosensory reflexes (Beatty & Lucero-Wagoner, 2000). The parasympathetic component is highly sensitive to lighting conditions, thus lighting conditions were kept constant (55 lx measured at the eye by a calibrated luxmeter), with constant artificial lighting and no natural light in the rooms. Moreover, participants always sat into the same place, a comfortable chair spaced 60 cm from the eye-tracker system and the testing room was temperature controlled to 21ºC.

Experimental design. The experimental design was a 22 X 2 Mixed Factorial. One of the independent variables, TOT, was a within-subjects factor. This variable had 24 intervals of 5 minutes, though the first 5-minute interval was taken as a baseline measure of the pupil diameter of each subject. In addition, we discarded the last interval from our analysis due to a high level of experimental noise during the last 5 minutes, which coincided with a researcher being present in the experimental room to oversee and end the session and would unintentionally increase participants' arousal. The second independent variable was Task Complexity, which was measured by the number of aircraft (traffic density) entering into the sector. This variable had two levels, Low Complexity task (low aircraft density scenario) and High Complexity task (high aircraft density scenario). The dependent variables were task performance, pupil diameter, and subjective feeling of fatigue.

Dependent measures. We operationalized the dependent variable 'performance' by dividing the number of conflicts by total present aircraft number at a given time, since we thought that using only the number of conflicts as a performance measurement wouldn't be an appropriate performance indicator, since it largely reflects air traffic density, so that way, it wouldn't be possible to establish comparison between the two traffic density conditions. It is worth noting that a conflict between 2 aircraft takes place when the safe distance between both aircrafts, either horizontal (5 nautical miles) or vertical (1 thousand feet), is violated. In order to quantify the number of conflicts per interval, a specific software was designed to analyze .log files provided by the ATC simulator.

For the physiological measure (pupil diameter), while our eye tracking system allows continuous sampling rate recording at 120Hz, we set a total of

24 intervals lasting 5 minutes each to facilitate subsequent analyses. However, given high inter-individual variability relative to average pupil size, we followed the recommendations provided by Sebastiaan Mathôt et al (2018) about base-line correction of pupil-size data: we took the first interval (5 minutes) as a reference of standard individual average pupil size, which was then subtracted from the obtained value in each of the remaining 23 intervals, thereby giving a differential standardized value that allowed us to compare participants. Analyses were carried out both for the left and right pupils. A negative value meant that the pupil was contracting while a positive value meant that it was dilating.

Subjective mental fatigue was obtained on-line by the instant self-assessment fatigue scale explained above

RESULTS

We used a two-way, mixed factorial ANOVA to analyse obtained results. First, we can see in Figure 4 that air traffic control task performance was optimal; thus, participants were able to avoid the vast majority of conflicts. However, in the ANOVA, the main effect of traffic density turned out to be significant ($F(1,68) = 89.5$, $MSe = .002$, $p < .05$; $\eta 2 = .57$). As expected, average performance was better in the low-aircraft density condition than in the high-aircraft density condition, which shows that our independent variable manipulation of task complexity proved to be effective. Furthermore, the main effect of interval also proved to be significant: $F(21,1428) = 12.51$, $MSe = .001$, $p < .05$; $\eta 2 = .16$). Thus, we noted a lower performance peak, which coincides with the maximum air traffic intensity peak on both conditions. This effect could not be considered a fatigue effect since it happens at approximately half of the total performance time but not at the end as would be expected if fatigue were the cause. We also found that the interaction of traffic density x TOT was significant: $F(21,1428) = 4.22$, $MSe = .006$, $p < .05$; $\eta 2 = .05$. However, we consider this interaction effect very small and difficult to interpret. Baseline measurements between groups were very similar (low complexity: .11; high complexity: .14) and presented no significant differences ($t(1,68) = -.51$, $p > .05$).

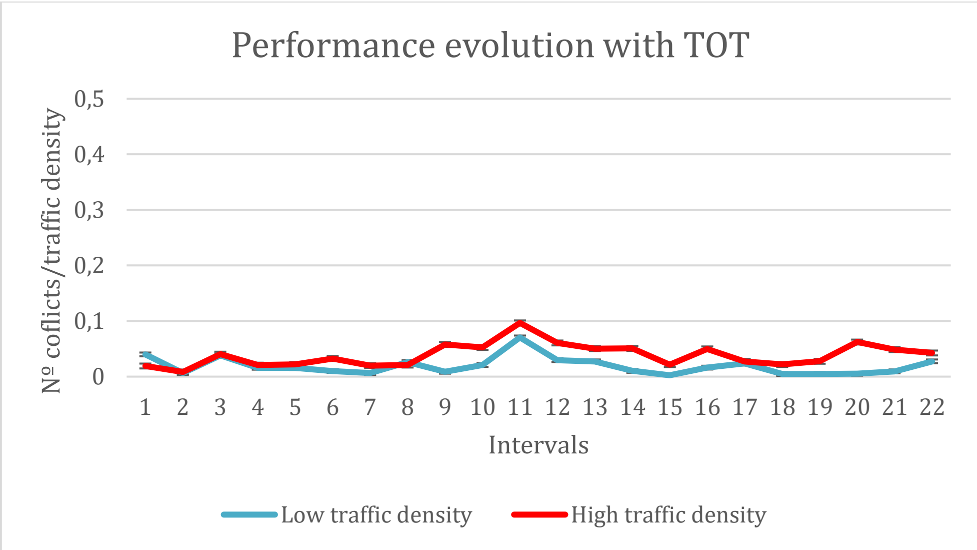


Figure 4. Participants´ performance during task development

For pupil diameter, considering that all participants presented no ocular abnormalities, and left and right pupil measurements were highly correlated both for low ($r = .92$; $p<.01$) and high-complexity ($r = .79$; $p<.01$) conditions, only data recorded from the left pupil were used for further analyses. The main effect of traffic density was found to be significant for pupil diameter ($F (1,68) = 5.95$, $MSe = 0.47$, $p <.05$; $\eta2 = .08$). In other words, average pupil size was smaller in the low complexity condition than in the high complexity condition. We also found a significant main interval effect for pupils ($F (21,1428) = 6.22$, $MSe =.13$, $p < .05$; $\eta2 = .08$). Figure 5 shows a decrease in the pupil diameter variable until the fifth interval, from which it increases from interval 5 to 6, stabilizing from then until the last interval. However, this interval effect must be considered in light of its interaction with the traffic density variable, which also was significant ($F (21,1428) = 2.49$, $MSe =.013$, $p < .05$; $\eta2 = .04$). We can see how changes occur in the variable depending on air traffic density, that is, depending on task complexity level. As shown in Figure 5, an identical decrease in pupil diameter occurred in both experimental conditions until interval 2. Nevertheless, from then on, pupil diameter continued to decrease in the low aircraft density condition, stabilizing in the sixth interval and then slightly increasing again, while it remained constant in the high aircraft density condition. Trend analysis revealed a quadratic trend in pupils ($F (1,33) = 27.25$, $MSe = .91$, $p <.01$; $\eta2 = .45$) in the low-aircraft density condition. Baseline measurements between groups were very similar (low complexity: 3.55mm; high complexity: 3.53mm) and presented no significant differences ($t (1,68) = .176$, $p >.05$).

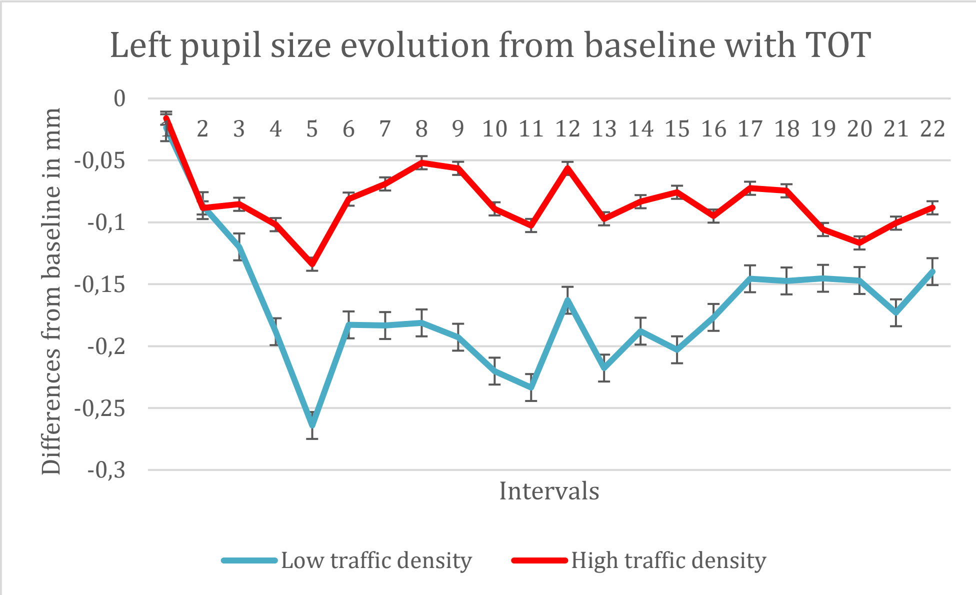


Figure 5. Participants´ left pupil diameter differences from baseline in mm, during task development

As shown in Figure 6, a linear increase in subjective experienced fatigue occurs over time, although we found a greater increase in the high-complexity task condition. The main effect of traffic density on task was not significant (F (1,68) = .94, MSe = 11.96, p >.05; η2 = .01). However, the interval main effect turned out to be significant (F (21,1428) = 80.5, MSe .47, p <.05; η2 = .54). Thus, a significant variation in subjective fatigue occurred over the intervals, assuming an upward linear and quadratic trend, given a linear (F (1,68) = 230.64, MSe = 3.24, p <.05; η2 = .78) and a quadratic (F (1,68) = 21.3, MSe = 1.9, p <.05; η2 = .24) contrast test. Lastly, the interaction between the two variables was significant, (F (21,1428) = 1.99, MSe .47, p <.05; η2 = .03), maintaining linear and quadratic trends on both low-aircraft (F (1,33) = 77.07, MSe = 279.14, p <.01; η2 = .7 & F (1,33) = 11.37, MSe = 6.41, p <.05; η2 = .16) and high-aircraft (F (1,35) = 168.55, MSe = 485.36, p <.01; η2 = .83 & F (1,35) = 15.91, MSe = 31.72, p <.01; η2 = .31) density conditions, respectively. The quadratic trend could be interpreted as a ceiling effect: participants reached the maximum level of fatigue. Regarding subjective fatigue baseline measures, there exist a difference of .18 between groups (low complexity: 1.24; high complexity: 1.06), which represents less than 1% of difference. Nevertheless, we found this difference statistically significant (t (1,68) = 2.19, p <.05) and despite it could be due to the low variance error in the ANOVA, this partly limits any inference on this variable.

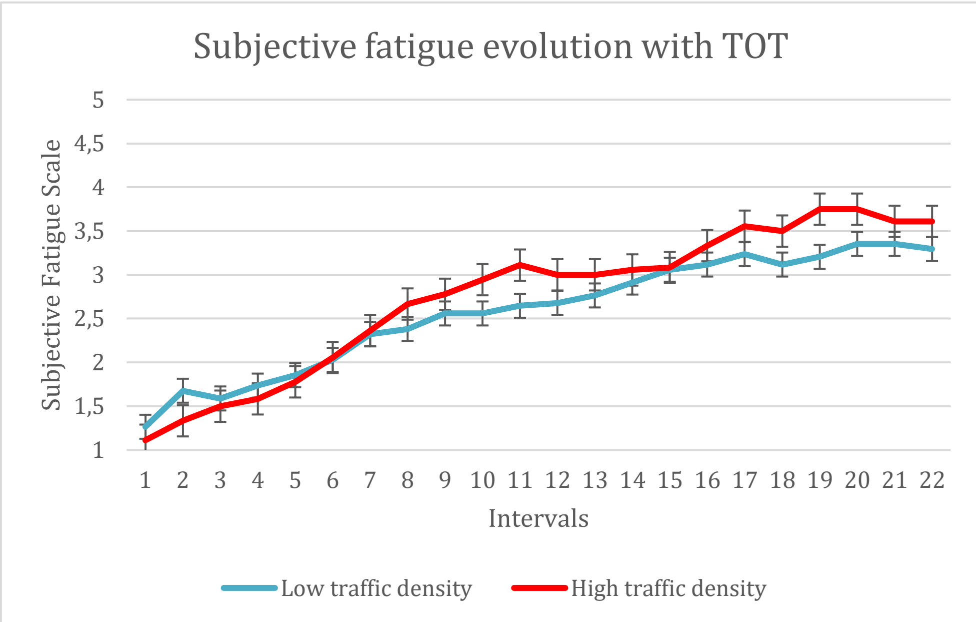


Figure 6. Participants´ subjective mental fatigue during task development

Finally, we performed a bayesian correlation analysis in order to assess the probability of the null hypothesis being true (it contrasts the likelihood of the data fitting under the null hypothesis with the likelihood of fitting under the alternative hypothesis). In Bayesian statistics, a Bayesian Factor B01=1 means no evidence in favour of either the null hyphothesis or the alternative. Bayesian Factors < 1 denote evidence inclined toward the alternative hypothesis, while Bayesian Factors > 1 denote evidence inclined toward the null hypothesis (Jarosz and Wiley, 2014). As depicted in Table 1, we could not find correlations between pupil diameter and subjective fatigue measures for either low ($r = -.13$; $p>.05$; B01=3.28) or high-complexity ($r = -.12$; $p>.05$; B01=.3) conditions. Similarly, we could not find correlations between subjective fatigue and performance measures for either low ($r = -.26$; $p>.05$; B01=1.95) or high-complexity ($r = .38$, $p>.05$; B01=1.19) conditions.

Table 1. Correlation chart between measures

	Low-Complexity Task	High-Complexity Task

		Pupil	Subjective Fatigue	Performance	Pupil	Subjective Fatigue	Performance
Pupil	Pearson	1	-,130 B_{01}=3.28	-,019 B_{01}=3.86	1	-,122 B_{01}=.3	-,307 B_{01}=.67
	Sig. (bilateral)		,553	,931		,581	,154
	N	23	23	23	23	23	23
Subjective Fatigue	Pearson	-,130 B_{01}=3.28	1	-,261 B_{01}=1.95	-,122 B_{01}=.3	1	,382 B_{01}=1.19
	Sig. (bilateral)	,553		,229	,581		,072
	N	23	23	23		23	23
Performance	Pearson	-,019 B_{01}=3.86	-,261 B_{01}=1.95	1	-,307 B_{01}=.67	,382 B_{01}=1.19	1
	Sig. (bilateral)	,931	,229		,154	,072	
	N	23	23	23	23	23	23

DISCUSSION

Mental fatigue is defined as a condition of reduced cognitive efficiency and performance caused by excessive mental workload sustained over time, accompanied by a subjective feeling of being exhausted and physical discomfort. Hence, **we could expect that a person experiencing mental fatigue would show task performance impairment, lower physiological activation and higher subjective mental fatigue.** In other words, we could expect to find associations between performance, physiological activation and subjective feeling of fatigue.

However, many authors have found that in certain situations where a decrease in performance was expected due to a highly demanding and/or lengthy task, no such decrease appeared (performance insensitivity). In these cases, people proved capable of maintaining adequate performance levels, even after having declared themselves fatigued (for a review see Hockey, 2013).

These results have led some authors to propose new resource theories that defend the existence of a **compensatory control mechanism responsible for the ability to maintain high levels of performance and physiological activation, even when the individual feels fatigued**.

The results from our study supported the existence of such a compensatory control mechanism. Our findings revealed relative stable performance in both traffic density conditions, as participant's performance was optimal; however, we did find differences in activity levels: Participants performing the low-complexity task experienced higher pupil diameter decrement, specifically during the first 5 intervals, than participants in the high-complexity task. From then on, in the high-complexity condition, pupil size remained stable, whereas a slight increase can be noted in the low-complexity condition until the end of the task (quadratic trend). On the other hand, we observed that subjective fatigue linearly increased in both groups (as trend analysis showed up) through the intervals, although we found a greater increase with TOT in the high-complexity task condition. Hence, our results showed:

- **Performance and physiological insensibilities:** we would expect performance to get worse and physiological activation to get lower with TOT, especially in the high-complexity task. However, **both remained stable at different levels** throughout the whole task, and the latter even slightly increased in the low-complexity condition.
- **Dissociations between subjective and physiological/performance measures**: taking into account the stability found on performance and physiological measures (at different levels), we would not expect subjective fatigue to appear with TOT. Nevertheless, our results revealed a steady subjective fatigue increase in both groups with TOT, being greater in the high-complexity condition. In other words, **we did not find correlations between subjective fatigue measures and physiological/performance measures**.

We think these insensitivities and dissociations can be explained by the compensatory control mechanism approach. **The compensatory mechanism modulates human performance, by reallocating resources when necessary, in order to successfully cope with the task**: participants in the high-complexity task maintained a steady higher level of activation over time to deal with a higher complexity task. In contrast, participants in the low-complexity task did not require that high level of activation and their levels decreased during the first intervals until they reached the appropriate activation levels for that lower task difficulty, slightly increasing from then until the last interval. Although we do not have a clear explanation about that slight increase in activation, we believe that the participant might be fighting against task disengagement and boredom by slightly increasing their activation level throughout the rest of the task. We must bear in mind that a 2-hour low-complexity monitoring task can be very exasperating. Although good performance and stable physiological activation levels would not be in principle indicators of mental fatigue, subjective fatigue linearly increased in both conditions with TOT, especially in the high-complexity condition. **This subjective fatigue increase would be higher in the high-complexity condition because a higher level of resources are being reallocated (effort) by the compensatory control mechanism in order to cope with a higher difficulty task**. Hence, we could say that as Hockey (2013) has noted,

subjective fatigue acts as an organismic alarm signal against resource depletion (similarly to thirst or hunger), even when there are still enough available resources to tackle the task.

In summary, **the compensatory mechanism allows the individual to maintain good performance on task development with TOT, even under high difficulty conditions, and this will be reflected by higher physiological activation. But we can also state that this need for greater effort would result in subjective fatigue feelings, which act as an alarm signal against resource depletion in such a manner that the more resources are being reallocated (higher effort), the more subjective fatigue increases**. Thus, for example, a worker may be maintaining an appropriate level of performance and high physiological activation in a task with a high level of mental demand, yet still experience high levels of subjective fatigue. This indicates the use of extra resources, but because of the risk of resource depletion, the risk of accidents is increased. Therefore, returning to the car metaphor used in the introduction, subjective fatigue acts as a low fuel warning light that is activated when there is still enough fuel to continue while warning us that our fuel is nearly running out.

Several limitations must be addressed before concluding. First, we only had one physiological measure, pupil diameter. Future research should also utilize **other physiological measures**, such as electrodermal activity, heart-rate variability and EEG. We hypothesize that every physiological measurement would reflect activation changes produced by the compensatory control mechanism. That is, higher activation would be reflected in higher electrodermal activity, lower hear-rate variability and higher EEG activity, and the opposite for low activation. However, inconsistencies have been found not only between performance, psychological and subjective measures, but **also within different indicators of each kind of measure; thus, further research is needed to untangle this issue**. It also needs to be noted that we conducted this research with students under simulated conditions; hence, we think it would be necessary to **replicate this experiment under real working conditions** to improve validity. On the other hand, the design aspect involved use of two different subject groups who were assigned either the low or high density task, but no subjects were tested in both tasks, which means that any differences could be in part attributed to group difference. In any case, we think that possible inter-group differences may have been in part solved statistically through randomization procedure (N=70). We also checked baseline values (1st interval) of both groups in dependent variables and found they are very similar. However, despite performance and pupil size presented no statistical differences, we found statistical differences in subjective experienced fatigue. This baseline similarities between groups gives us certain confidence that no differences could be in part attributed to group differences, except in the case of subjective experienced fatigue. Finally, participants were alerted by an alarm every 5 minutes to assess subjective mental fatigue scale and it may be possible that these breaks could have influenced the impact of time-on-task and, consequently, the accumulated mental fatigue. However, we think that as the fatigue scale was designed to be very easy handling and to be answered within 1 or 2 seconds, we believe that these breaks to assess mental fatigue are not very determining and do not brake the situational awareness of participants but, must be taken into consideration in future research.

In conclusion, we could say that, based on our results, **some inconsistencies between fatigue measurements would be explained according to the compensatory control mechanism action. Moreover, our results support the current resource theories** (i.e. Hockey, 1997, 2011 2013) that replaced outdated classical resource theories (i.e. Kahneman, 1973).

Our results suggest a need for **caution about considering psychophysiological indexes such as pupil size as the only indicator of mental fatigue** (Cañas, 2017; Hopstaken et al, 2016; Sakamoto et al, 2009). Fatigue recognition systems have been introduced in some cars which claim to recognize drivers' fatigue by only taking into account drivers' eye behaviour. According to our results, ocular behaviour could indicate high levels of physiological activation, which would not be, in principle, an indicator of mental fatigue, but could instead indicate the use of extra resources to maintain optimal driving performance even though the driver has already started to feel fatigued.

RESUMEN

Inconsistencias entre medidas de fatiga mental bajo las teorías del control compensatorio. Tradicionalmente, la fatiga mental ha sido definida como una condición de reducción en los niveles de eficiencia cognitiva y rendimiento, acompañada de una sensación subjetiva de fatiga mental. A pesar de que podríamos esperar encontrar asociaciones (convergencia) entre las tres características definitorias de la fatiga mental (deterioro en el rendimiento, reducción en los niveles de activación fisiológica y surgimiento de fatiga subjetiva), la literatura ha revelado que la aparición de inconsistencias (divergencia) entre las medidas de fatiga es más frecuente de lo esperado: la gente se muestra capaz de mantener niveles adecuados de rendimiento a pesar de haber declarado encontrarse fatigados. Esto puede explicarse a partir de los modelos del mecanismo de control compensatorio,

los cuales afirman que los seres humanos son capaces de proveerse con recursos adicionales bajo condiciones de elevada demanda cognitiva, únicamente a expensas de un coste psicofisiológico y del surgimiento de la sensación subjetiva de fatiga mental. En el presente estudio, ponemos a prueba esta explicación manipulando el tiempo y la complejidad de una tarea de simulación de control de tráfico aéreo. Recabamos datos psicofisiológicos, de rendimiento y subjetivos. Nuestros resultados desvelan una disminución del diámetro pupilar en la condición de baja densidad de tráfico aéreo, en tanto que se mantiene constante en la condición de alta densidad. El nivel de rendimiento de los participantes resultó ser óptimo en ambas condiciones, a pesar de que se aprecia un incremento lineal en los niveles de fatiga subjetiva, especialmente en la condición de alta complejidad. Así, la complejidad parece activar el mecanismo compensatorio, el cual provee al organismo con recursos adicionales que mantienen fisiológicamente activados a los participantes de la condición de alta densidad de tráfico al objeto de hacer frente a una tarea de mayor dificultad, mientras que la fatiga subjetiva podría estar actuando como una señal del organismo para impedir el agotamiento de los recursos cognitivos. Nuestros hallazgos apoyan las teorías del control compensatorio y ofrecen una posible explicación sobre algunas inconsistencias entre las medidas de fatiga mental. La ciencia necesita seguir investigando el fenómeno del mecanismo compensatorio para favorecer la gestión de los efectos de la fatiga mental y prevenir los accidentes laborales.

REFERENCES

Ahlstrom, U., & Friedman-Berg, F. J. (2006). Using eye movement activity as a correlate of cognitive workload. *International Journal of Industrial Ergonomics, 36*(7), 623-636. https://doi.org/10.1016/j.ergon.2006.04.002

Bailey, A., Channon, S., & Beaumont, J. G. (2007). The relationship between subjective fatigue and cognitive fatigue in advanced multiple sclerosis. *Multiple Sclerosis*, *13*(1), 73-80. http://dx.doi.org/10.1177/1352458506071162

Bailey, B. P.; Iqbal, S. T. (2008) Understanding changes in mental workload during execution of goal-directed tasks and its application for interruption management. *ACM Trans. Comput. Interact*. *14*, 1–28. https://doi.org/10.1145/1314683.1314689

Barker, L. M., & Nussbaum, M. A. (2011). The effects of fatigue on performance in simulated nursing work. *Ergonomics*, *54*(9), 815-829. https://doi.org/10.1080/00140139.2011.597878

Beatty, J. (1982). Task-evoked pupillary responses, processing load, and the structure of processing resources. *Psychol. Bull*., *91*, 276–292. https://doi.org/10.1037/0033-2909.91.2.276

Beatty, J., & Lucero-Wagoner, B. (2000). The pupillary system. *Handbook of Psychophysiology*, *2,* 142-162.

Benedetto, S., Pedrotti, M., Minin, L., Baccino, T., Re, A., & Montanari, R. (2011). Driver workload and eye blink duration. *Transportation research part F: traffic psychology and behaviour, 14*(3), 199-208. https://doi.org/10.1016/j.trf.2010.12.001

Bruce, J. M., Bruce, A. S., & Arnett, P. A. (2010). Response variability is associated with selfreported cognitive fatigue in multiple sclerosis. *Neuropsychology, 24*(1), 77-83. https://doi.org/10.1037/a0015046

Cañas, J. J. (2017). Fatigue and Theories of Resources: The Energetic Metaphor of the Mind Functioning Might Be Reconsidered. In: *Advances in Neuroergonomics and Cognitive Engineering* (pp. 269-277). Springer, Cham. https://doi.org/10.1007/978-3-319-41691-5_23

Chalder, T., Berelowitz, G., Pawlikowska, T., Watts, L., Wessely, S., Wright, D., & Wallace, E.P. (1993). Development of a fatigue scale. *Journal of Psychosomatic Research*, *37*(2), 147-153. https://doi.org/10.1016/0022-3999(93)90081-P

Christodoulou, C. (2005). The assessment and measurement of fatigue. In: J. DeLuca (Ed.), *Fatigue as a window to the brain* (pp. 19-36). Cambridge, MA: MIT Press.

Dawson, D., Ian Noy, Y., Härmä, M., Åkerstedt, T., & Belenky, G. (2011). Modelling fatigue and the use of fatigue models in work settings. Accident Analysis & Prevention, *43*(2), 549-564. https://doi.org/10.1016/j.aap.2009.12.030

De Alwis Edirisinghe, V. (2017). Estimating Mental Workload of University Students using Eye Parameters (Master's thesis, NTNU).

DeLuca, J. (2005). 19 Fatigue: Its Definition, Its Study, and Its Future. *Fatigue as a window to the brain*, 319.

Di Stasi, L. L., Marchitto, M., Antolí, A., Baccino, T., & Cañas, J. J. (2010). Approximation of on-line mental workload index in ATC simulated multitasks. *Journal of Air Transport Management*, *16*(6), 330-333. https://doi.org/10.1016/j.jairtraman.2010.02.004

Fothergill, S., Loft, S., & Neal, A. (2009). ATC-labAdvanced: An air traffic control simulator with realism and control. *Behavior Research Methods, 41*(1), 118-127. https://doi.org/10.3758/BRM.41.1.118

Gilbert, D. (2009). *Stumbling on happiness*. Vintage: Canada.

Gopher, D., & Donchin, E. (1986). Workload: An examination of the concept. In K. R. Boff, L. Kaufman, & J. P. Thomas (Eds.), *Handbook of perception and human performance, Vol. 2*. Cognitive processes and performance (p. 1–49). John Wiley & Sons.

Grandjean, E. (1989). Fitting the task to the man: A textbook of occupational ergonomics (4th ed.) Taylor & Francis/Hemisphere.

Hancock, P. A. (2017). Whither Workload? Mapping a Path for Its Future Development. In: *International Symposium on Human Mental Workload: Models and Applications* (pp. 3-17). Springer, Cham. https://doi.org/10.1007/978-3-319-61061-0_1

Hess, E. H.; Polt, J.M. (1964) Pupil Size in Relation to Mental Activity during Simple Problem-Solving. *Science*, *143*, 1190–1192. https://doi.org/10.1126/science.143.3611.1190

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*(1), 73-93. https://doi.org/10.1016/S0301-0511(96)05223-4

Hockey, G. R. J. (2011). A motivational control theory of cognitive fatigue. Cognitive fatigue: Multidisciplinary perspectives on current research and future applications, 167-87. https://doi.org10.1037/12343-008

Hockey, R. (2013). *The psychology of fatigue: Work, effort and control* Cambridge University Press.

Hopstaken, J. F., van der Linden, D., Bakker, A. B., & Kompier, M. A. (2015). The window of my eyes: Task disengagement and mental fatigue covary with pupil dynamics. *Biological psychology, 110*, 100-106. https://doi.org/10.1016/j.biopsycho.2015.06.013

Hopstaken, J. F., van der Linden, D., Bakker, A. B., Kompier, M. A., & Leung, Y. K. (2016). Shifts in attention during mental fatigue: Evidence from subjective, behavioral, physiological, and eye-tracking data. *Journal of Experimental Psychology.Human Perception and Performance,* https://doi.org/10.1037/xhp0000189

Horiuchi, R., Ogasawara, T., & Miki, N. (2017). Fatigue evaluation by detecting blink behaviour using eyeglass-shaped optical sensor system. *IET Micro & Nano Letters, 12*(8), 554-559. https://doi.org/10.1049/mnl.2017.0136

Ishii, A., Tanaka, M., Shigihara, Y., Kanai, E., Funakura, M., & Watanabe, Y. (2013). Neural effects of prolonged mental fatigue: A magnetoencephalography study. *Brain Research, 1529,* 105-112. https://doi.org/10.1016/j.brainres.2013.07

Jainta, S., & Baccino, T. (2010). Analyzing the pupil response due to increased cognitive demand: An independent component analysis study. *International Journal of Psychophysiology*, *77*(1), 1-7. https://doi.org/10.1016/j.ijpsycho.2010.03.008

Jarosz, A. F., & Wiley, J. (2014). What are the odds? A practical guide to computing and reporting Bayes factors. *The Journal of Problem Solving*, *7*(1), 2. https://doi.org/10.7771/1932-6246.1167

Johnson, S. K., DeLuca, J., Diamond, B. J., & Natelson, B. H. (1998). Memory dysfunction in fatiguing illness: Examining interference and distraction in short-term memory. *Cognitive Neuropsychiatry,* *3*(4), 269-285. https://doi.org/10.1080/135468098396099

Johnson, S. K., Lange, G., DeLuca, J., Korn, L. R., & Natelson, B. (1997). The effects of fatigue on neuropsychological performance in patients with chronic fatigue syndrome, multiple sclerosis, and depression. *Applied Neuropsychology, 4*(3), 145-153. https://doi.org/10.1207/s15324826an0403_1

Kahneman, D. (1973). Attention and effort (Vol. 1063). Englewood Cliffs, NJ: Prentice-Hall.

Kahneman, D., & Beatty, J. (1966). Pupil diameter and load on memory. *Science, 154*, 1583-1585.

Kanfer, R. (2011). Determinants and consequences of subjective cognitive fatigue. In: L. P. Ackerman (Ed.), *Cognitive fatigue: Multidisciplinary perspectives on current research and future applications* (pp. 189-208). Washington, DC: American Psychological Association. https://doi.org/10.1037/12343-009

Krupp, L. B., & Elkins, L. E. (2000). Fatigue and declines in cognitive functioning in multiple sclerosis. *Neurology,* *55*(7), 934-939. https://doi.org/10.1212/WNL.55.7.934-a

Kohl, A. D. (2016). Understanding mental fatigue: A grounded theory study (Doctoral dissertation, Saybrook University).

LeDuc, P. A., Greig, J. L., & Dumond, S. L. (2005). Involuntary eye responses as measures of fatigue in US army apache aviators. *Aviation, Space, and Environmental Medicine, 76*(Supplement 1), C91.

Lemonnier, S., Brémond, R., & Baccino, T. (2014). Discriminating cognitive processes with eye movements in a decision-making driving task. *Journal of Eye Movement Research*, *7*(4). https://doi.org/10.16910/jemr.7.4.3

Lorist, M. M., Boksem, M. A., & Ridderinkhof, K. R. (2005). Impaired cognitive control and reduced cingulate activity during mental fatigue. *Cognitive Brain Research, 24*(2), 199-205. https://doi.org/10.1016/j.cogbrainres.2005.01.018

Mathôt, S., Fabius, J., Van Heusden, E., & Van der Stigchel, S. (2018). Safe and sensible preprocessing and baseline correction of pupil-size data. *Behavior research methods*, 50(1), 94-106. https://doi.org/10.3758/s13428-017-1007-2

Nakayama, M.; Takahashi, K.; Shimizu, Y. (2002). The act of task difficulty and eye-movement frequency for the "Oculo-motor indices.". In: *Proceedings of the Symposium on Eye Tracking Research & Applications—ETRA'02*, New Orleans, Louisiana, ACM Press: New York, USA, pp. 37–42. https://doi.org/10.1145/507072.507080

Nie,Nie Xin Xin (聂欣 nie xin). (2005). Quantitative assessment of mental fatigue for human body based on nonlinear analysis of physiological signals (Doctoral dissertation).

Paul, R. H., Beatty, W. W., Schneider, R., Blanco, C. R., & Hames, K. A. (1998). Cognitive and physical fatigue in multiple sclerosis: Relations between self-report and objective performance. *Applied Neuropsychology, 5*(3), 143-148. https://doi.org/10.1207/s15324826an0503_5

Roy, R. N., Charbonnier, S., & Bonnet, S. (2014). Eye blink characterization from frontal EEG electrodes using source separation and pattern recognition algorithms. *Biomedical Signal Processing and Control, 14*, 256-264. https://doi.org/10.1016/j.bspc.2014.08.007

Sakamoto, K., Aoyama, S., Asahara, S., Mizushina, H., & Kaneko, H. (2009, July). Relationship between emotional state and pupil diameter variability under various types of workload stress. In *International Conference on Ergonomics and Health Aspects of Work with Computers* (pp. 177-185). Springer, Berlin, Heidelberg.

Sauer, J., Wastell, D. G., Robert, G., Hockey, J., & Earle, F. (2003). Performance in a complex multiple-task environment during a laboratory-based simulation of occasional night work. *Human Factors*, *45*(4), 657-670. https://doi.org/10.1518%2Fhfes.45.4.657.27090

Shen, K. Q., Li, X. P., Ong, C. J., Shao, S. Y., & Wilder-Smith, E. P. (2008). EEG-based mental fatigue measurement using multi-class support vector machines with confidence estimate. Clinical Neurophysiology, *119*(7), 1524-1533. https://doi.org/10.1016/j.clinph.2008.03.012

Smets, E. A., Garssen, B., Bonke, B., & de Haes, J. M. (1995). The Multidimensional Fatigue Inventory (MFI) psychometric qualities of an instrument to assess fatigue. *Journal of Psychosomatic Research, 39*(3), 315-325. https://doi.org/10.1016/0022-3999(94)00125-O

Stern, J. A. (1997). The pupil of the eye: What can it tell us about 'mental processes'. *Human Engineering for Quality Life (HQL) Quarterly, 8*, 1-2.

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*(1),45-65. https://doi.org/1010.1016/S0001-6918(02)00150-6

Van Orden, K. F., Limbert, W., Makeig, S., & Jung, T. P. (2001). Eye activity correlates of workload during a visuospatial memory task. *Human Factors, 43*(1), 111-121. https://doi.org/10.1518%2F001872001775992570

Venables, L., & Fairclough, S. H. (2009). The influence of performance feedback on goal-setting and mental effort regulation. *Motivation and Emotion, 33*(1), 63-74. https://doi.org/10.1007/s11031-008-9116-y

Vrijkotte, S., Meeusen, R., Vandervaeren, C., Buyse, L., Cutsem, J. V., Pattyn, N., & Roelands, B. (2018). Mental Fatigue and Physical and Cognitive Performance During a 2-Bout Exercise Test. *International journal of sports physiology and performance, 13*(4), 510-516. https://doi.org/10.1123/ijspp.2016-0797

Walker, L. S., Berard, J. A., Berrigan, L. I., Rees, L. M., & Freedman, M. S. (2012). Detecting cognitive fatigue in multiple sclerosis: Method matters. *Journal of The Neurological Sciences, 316*(1-2), 86-92. https://doi.org/10.1016/j.jns.2012.01.021

Yokoyama, H., Eihata, K., Muramatsu, J., & Fujiwara, Y. (2018). Prediction of Driver's Workload from Slow Fluctuations of Pupil Diameter. In: 21st International Conference on Intelligent Transportation Systems (ITSC) (pp. 1775-1780). IEEE. https://doi.org/10.1109/ITSC.2018.8569279

Zhang, C., Zheng, C., Ouyang, Y., & Yu, X. (2008). Analysis of mental fatigue basing on power spectrum features of EEG. *Space Medicine and Medical Engineering, 21*(1), 35.

Zhang, C., Zheng, C., Yu, X., & Li, X. (2006). Analysis of Mental Fatigue Based on Approximate Entropy of EEG. *Space Medicine and Medical Engineering, 19*(6), 459.

Zhang, C., & Yu, X. (2010). Estimating mental fatigue based on electroencephalogram and heart rate variability. *Polish Journal of Medical Physics And Engineering, 16*(2), 67-84. https://doi.org/10.2478/v10013-010-0007-7

(Manuscript received: 24 May 2019; accepted: 16 February 2020)