Analysis of Human Performance as a Measure of Mental Fatigue

André Pimenta, Davide Carneiro, Paulo Novais, and José Neves

CCTC/DI - Universidade do Minho
Braga, Portugal
{apimenta,dcarneiro,pjon,jneves}@di.uminho.pt

Abstract. In our daily life, we often have the feeling of being exhausted due to mental or physical work, and a sense of performance degradation in the execution of simple tasks. The maximum capacity of operation and performance of an individual, whether physical or mental, usually also decreases gradually as the day progresses. The loss of these resources is linked to the onset of fatigue, which is particularly noticeable in long and demanding tasks or repetitive jobs. However, good management of the working time and effort invested in each task, as well as the effect of breaks at work, can result in better performance and better mental health, delaying the effects of fatigue. This paper details a non-invasive approach on the monitoring of fatigue of a human being, based on the analysis of the performance of his interaction with the computer.

Keywords: Fatigue, Mental Fatigue, Performance, Behavioural Analysis, Classification

1 Introduction

In today's busy society, people push their limits in order to find a balance between ever ambitious goals and time to dedicate to enjoyable moments, family or personal projects [1]. This balance, when it comes to exist, is generally obtained at the expenses of less time for rest and sleep, with an impact on the well-being and health of the individual. Fatigue, above all others, must be considered here. The effects of fatigue may not be felt immediately. Nonetheless, when prolonged it may have serious consequences on health and safety at the workplace as well as on labour productivity [2]. Ultimately, fatigue can put people who have safety-sensitive occupations at risk, as any mistake can lead to the loss of their own lives or those of others [3].

This paper presents a new approach on the problem of fatigue detection and monitoring, that can be included in the umbrella of Hybrid Artificial Intelligent Systems [4,5]. It looks at the effects of fatigue on the performance of the individual's interaction with the computer. This approach can be deemed both non-invasive and non-intrusive as it relies on the observation of the individual's use of the mouse and keyboard. It builds on a previous study, now to address the evolution of fatigue throughout the day and the role that task complexity and Circadian Rhythm play in the degradation of performance.

1.1   Circadian Rhythm, Environment and their Role on Fatigue

In studying fatigue one must acknowledge the major role of the biological clock, i.e., one cannot expect to beat one's rhythm and biological need to rest. Indeed, the role of our biological clock goes beyond compelling the body to fall asleep and to wake up again. It also modulates our hour-to-hour waking behaviour, which in turn affects our sense of fatigue, alertness and performance, generating circadian rhythmicity in almost all neuro-behavioural variables [6]. This means that there is a natural fluctuation of the level of fatigue we experience during the day, that is independent of the intensity of the tasks we perform. It can be seen as a base-level of fatigue.

The context and environment in which the individual carries out the activity are also very important in determining the level of fatigue during the day. Indeed, factors such as workload, distractions, periods of boredom, motivational factors, stress, food intake, posture, ambient temperature, background noise, lighting conditions or drug intake (e.g., caffeine) are examples of contextual variables that can have a positive or negative influence on the level of fatigue of a human being [7]. Smith and Miles conducted a study which analyses the effects of lunch on cognitive vigilance tasks [8]. The authors analysed the performance of individuals who did not take lunch before performing tasks compared to those who took, to show some differences between morning and afternoon performance in subjects who abstained from eating lunch.

Figure 1 details the influence of several environment variables and external factors in task performance, as well as their relationship. Black boxes depict task characteristics and while red ones depict characteristics of the Human operator. The dotted line indicates a positive correlation (measured with the Pearson product-moment correlation coefficient) between actual, measurable task workload and the worker's perception of that workload (estimated with a mental workload tool such as NASA-TLX [9] or SWAT [10]).

1.2   Human Performance as a Measure of Fatigue

When under the effect of prolonged excessive mental workload, inadequate sleep patterns, or circadian effects, Humans may exhibit slowed cognitive processes or may not respond at all to the stimuli and information received. This is generally due to the amount of information exceeding the brain's processing capacity at that moment. In contrast, when the mental workload is significantly below the adequate level, the individual tends to get tired and bored, which will likely result in an increase in the number of errors in the tasks being performed [12]. Performance as a measure of fatigue has been used previously as an objective measure of fatigue in humans. Morris & Miller, for example, conducted a study which analysed the performance of pilots during simulated flight as a means of detecting and measuring fatigue [13]. The authors have shown that the decrease in flight performance occurs due to fatigue.

Consequently, the measurement of the performance of the individual while performing a task results in an objective way of measuring fatigue in the human being as a relationship between the two exists. Usually, tasks that imply
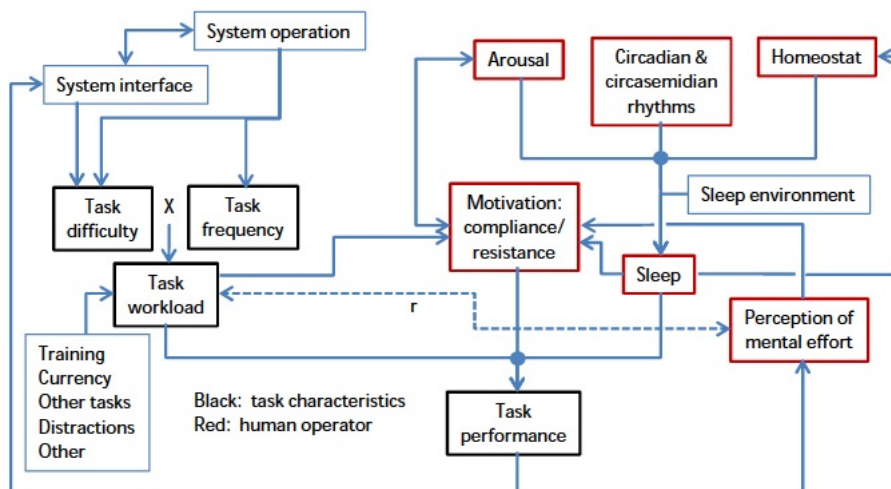
Fig. 1. Environment variables and external factors on Mental Performance Task (Source: [11]).

excessive mental and physical workload as well as very low workload (observed in monotonous and repetitive activities) result in significant performance decrements. Thus being, the loss of performance due to fatigue can be studied as a function that takes as input the duration of the task and the quantity and/or quality of the work carried out [2,14].

The challenge lies in the fact that a decrease in the performance is often not "seen" immediately, which may be explained by the concept of *reverse capacity* [15]. The theory of reverse capacity states that an individual rarely works at his maximum capacity, and can thus temporarily absorb the load of cognitive or physical additional work while still maintaining performance indicators, as long as their inverse capacity is not exceeded. For this reason, the effects of fatigue may not be immediately visible [15]. The effects of fatigue on performance, in particular, may only be felt in the course of time, when it may be to late to avoid them.

2   A non-invasive Approach to Fatigue Monitoring

This paper introduces a system for detecting and monitoring mental fatigue through the analysis and measurement of human performance indicators. As previously postulated in [16] it uses the keyboard and mouse as sensors of user performance. Specifically, the user is monitored through particular Operating System events fired from the use of the computer's mouse and keyboard:

- MOV, timestamp, posX, posY - an event describing the movement of the mouse, in a given time, to coordinates (posX, posY) in the screen;

- MOUSE_DOWN, timestamp, [Left|Right], posX, posY - this event describes the first half of a click (when the mouse button is pressed down), in a given time. It also describes which of the buttons was pressed (left or right) and the position of the mouse in that instant;
- MOUSE_UP, timestamp, [Left|Right], posX, posY - an event similar to the previous one but describing the second part of the click, when the mouse button is released;
- MOUSE_WHEEL, timestamp, dif - this event describes a mouse wheel scroll of amount dif, in a given time;
- KEY_DOWN, timestamp, key - identifies a given key from the keyboard being pressed down, at a given time;
- KEY_UP, timestamp, key - describes the release of a given key from the keyboard, in a given time;

From these events, that fully describe the interaction of the user with the mouse and keyboard, it is possible to extract the following features:
KEY DOWN TIME - the timespan between two consecutive KEY_DOWN and KEY_UP events, i.e., for how long was a given key pressed.
UNITS - milliseconds

TIME BETWEEN KEYS - the timespan between two consecutive KEY_UP and KEY_DOWN events, i.e., how long did the individual took to press another key.
UNITS - milliseconds

MOUSE VELOCITY - The distance travelled by the mouse (in pixels) over the time (in milliseconds). The velocity is computed for each interval defined by two consecutive MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, mup and mdo, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$, that took place respectively in the instants $time_1$ and $time_2$. Let us also assume two vectors $posx$ and $posy$, of size n, holding the coordinates of the consecutive MOUSE_MOV events between mup and mdo. The velocity between the two clicks is given by $r\_dist/(time_2 - time_1)$, in which r_dist represents the distance travelled by the mouse and is given by equation 1.
UNITS - pixels/milliseconds

$$r\_dist = \sum_{i=0}^{n-1} \sqrt{(posx_{i+1} - posx_i)^2 + (posy_{i+1} - posy_i)^2} \qquad (1)$$

MOUSE ACCELERATION - The velocity of the mouse (in pixels/milliseconds) over the time (in milliseconds). A value of acceleration is computed for each interval defined by two consecutive MOUSE_UP and MOUSE_DOWN events, using the intervals and data computed for the Velocity.
UNITS - pixels/milliseconds^2

TIME BETWEEN CLICKS - the timespan between two consecutive MOUSE_UP and MOUSE_DOWN events, i.e., how long did it took the individual to perform

another click.
UNITS - milliseconds

DOUBLE CLICK DURATION - the timespan between two consecutive MOUSE_UP events, whenever this timespan is inferior to 200 milliseconds. Wider timespans are not considered double clicks.
UNITS - milliseconds

AVERAGE EXCESS OF DISTANCE - this feature measures the average excess of distance that the mouse travelled between each two consecutive MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, mup and mdo, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$. To compute this feature, first it is measured the distance in straight line between the coordinates of mup and mdo as depicted in equation 2.

$$s\_dist = \sqrt{(x2 - x1)^2 + (y2 - y1)^2} \tag{2}$$

Then, it is measured the distance actually travelled by the mouse by summing the distance between each two consecutive MOUSE_MV events. Let us assume two vectors $posx$ and $posy$, of size n, holding the coordinates of the consecutive MOUSE_MV events between mup and mdo. The distance actually travelled by the mouse, $real\_dist$ is given by equation 1. The average excess of distance between the two consecutive clicks is thus given by r_dist/s_dist.
UNITS - pixels

AVERAGE DISTANCE OF THE MOUSE TO THE STRAIGHT LINE - in a few words, this feature measures the average distance of the mouse to the straight line defined between two consecutive clicks. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, mup and mdo, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$. Let us also assume two vectors $posx$ and $posy$, of size n, holding the coordinates of the consecutive MOUSE_MOV events between mup and mdo. The sum of the distances between each position and the straight line defined by the points $(x1, y1)$ and $(x2, y2)$ is given by equation 3, in which $ptLineDist$ returns the distance between the specified point and the closest point on the infinitely-extended line defined by $(x1, y1)$ and $(x2, y2)$. The average distance of the mouse to the straight line defined by two consecutive clicks is this given by s_dists/n.
UNITS - pixels

$$s\_dists = \sum_{i=0}^{n-1} ptLineDist(posx_i, posy_i) \tag{3}$$

DISTANCE OF THE MOUSE TO THE STRAIGHT LINE - this feature is similar to the previous one in the sense that it will compute the s_dists between two consecutive MOUSE_UP and MOUSE_DOWN events, mup and mdo, according to equation 3. However, it returns this sum rather than the average value during

the path.

Units - pixels

SIGNED SUM OF ANGLES - with this feature the aim is to determine if the movement of the mouse tends to "turn" more to the right or to the left. Let us assume three consecutive MOUSE_MOVE events, $mov1, mov2$ and $mov3$, respectively in the coordinates $(x1, y1)$, $(x2, y2)$ and $(x3, y3)$. The angle α between the first line (defined by $(x1, y1)$ and $(x2, y2)$) and the second line (defined by $(x2, y2)$ and $(x3, y3)$) is given by $degree(x1, y1, x2, y2, x3, y3) = \tan(y3 - y2, x3 - x2) - \tan(y2 - y1, x2 - x1)$. Let us now assume two consecutive MOUSE_UP and MOUSE_DOWN events, mup and mdo. Let us also assume two vectors $posx$ and $posy$, of size n, holding the coordinates of the consecutive MOUSE_MOV events between mup and mdo. The signed sum of angles between these two clicks is given by equation 4.

Units - degrees

$$s\_angle = \sum_{i=0}^{n-2} degree(posx_i, posy_i, posx_{i+1}, posy_{i+1}, posx_{i+2}, posy_{i+2}) \quad (4)$$

ABSOLUTE SUM OF ANGLES - this feature is very similar to the previous one. However, it seeks to find only how much the mouse "turned", independently of the direction to which it turned. In that sense, the only difference is the use of the absolute of the value returned by function $degree(x1, y1, x2, y2, x3, y3)$, as depicted in equation 5.

Units - degrees

$$s\_angle = \sum_{i=0}^{n-2} \mid degree(posx_i, posy_i, posx_{i+1}, posy_{i+1}, posx_{i+2}, posy_{i+2}) \mid \quad (5)$$

DISTANCE BETWEEN CLICKS - represents the total distance travelled by the mouse between two consecutive clicks, i.e., between each two consecutive MOUSE_UP and MOUSE_DOWN events. Let us assume two consecutive MOUSE_UP and MOUSE_DOWN events, mup and mdo, respectively in the coordinates $(x1, y1)$ and $(x2, y2)$. Let us also assume two vectors $posx$ and $posy$, of size n, holding the coordinates of the consecutive MOUSE_MOV events between mup and mdo. The total distance travelled by the mouse is given by equation 1.

From these features it is possible to obtain a measure of the user's performance (e.g. an increased distance between clicks or sum of angles represents decreased performance). Once information about the individual's performance exists in these terms, it is possible to start monitoring his fatigue, in real-time, and without the need for any explicit or conscious interaction. This makes this approach especially suited to be used in work environments in which people use computers as it requires no change in their working routines. This is the main advantage of this work, especially when compared to more traditional approaches

that still rely on questionnaires (with issues concerning wording or question construction), specific hardware (that has additional costs and is frequently intrusive) or the availability of human experts.

3 Case Study

In order to assess the validity of the approach described in the previous section a case study was implemented aimed at collecting data over a relatively long period of time. This allowed us to determine if and how people are affected by fatigue in each of the mentioned features.

3.1 Participants

The participants, twenty seven in total (20 men, 7 women) were students from the University of Minho who volunteered to participate. Their age ranged between 18 and 30. The following requirements were established to select, among all the volunteers, the ones that participated: (1) familiarity and proficiency with the use of the computer; (2) use of the computer on a daily basis and throughout the day; (3) owning at least one personal computer. These restrictions constituted no problem as a large slice of the population of out institution, and particularly the ones affiliated to the Department of Informatics, fulfil the requirements.

3.2 Methodology

The methodology followed to implement the study was devised to be as minimally intrusive as the approach it aims to support. Participants were provided with an application for logging the previously mentioned events of the mouse and keyboard. This application, which maintained the confidentiality of the keys used, needed only to be installed in the participant's primary computer and would run on the background, starting automatically with the Operating System. The only explicit interaction needed from the part of the user was the input of very basic information on the first run, including the identification and age.

This application was kept running for approximately one month, collecting interaction data whenever there was interaction with the computer. During the whole process, participants did not need to perform any additional task and were asked to perform their activities regularly, whether they were work or leisure-related, as they would if they were not participating on the study.

Finally, when the period of one month ended, participants were asked to send in the file containing the log of their interaction during the duration of the study. The data collected was used to build the features described previously. The resulting dataset as well as the process by means of which data were analysed is described in the following sub-sections.

3.3   Data Selection

As previously stated, the data collected spans a period of approximately 30 complete days of computer use, in a work context as well as in a leisure one. From this data, we extracted particular moments for analysis.

Specifically, four distinct periods of one our each were selected in order to analyze the effect of Circadian rhythm and the effect of acute fatigue in a work day: (1) the start of the day, when the user is mentally fresh; (2) immediately before lunch break; (3) after lunch break; and (4) the end of the day, when the individual is most fatigued. This selection stems from the willingness to study the differences between the beginning and end of the day as well the potential effect of a lunch break. Lunch break is automatically inferred from the lack of interaction during a relative long time in a given period.

3.4   Statistical Data Analysis

To determine the effects of acute fatigue on the interaction patterns of an individual it is necessary to determine that the distributions of the data collected in each moment are statistically and significantly different. To this aim, the following approach was implemented.

First, it was determined, using the Pearson's chi-squared test, that most of the distributions of the data collected are not normal. Given this, the Kruskal-Wallis test was used in the subsequent analysis. This test is a non-parametric statistical method for testing whether samples originate from the same distribution. It is used for comparing more than two samples that are independent, and thus prove the existence of distinct behaviours. The null hypothesis considered is: H_0 : all samples have identical distribution functions against the alternative hypothesis that at least two samples have different distribution functions. For each set of samples compared, the test returns a $p-value$, with a small $p-value$ suggesting that it is unlikely that H_0 is true. Thus, for every Kruskal-wallis test whose $p-value < \alpha$, the difference is considered to be statistically significant, i.e., H_0 is rejected. In this work a value of $\alpha = 0.05$ is considered, a standard value generally accepted by research.

This statistical test is performed for each of the features considered, with the intention of determining if there are statistically significant differences between the several distributions of data, which will in turn confirm the existence of effects of fatigue on the interaction patterns.

Table 1 details, for each feature, the percentage of participants whose interaction patterns were significantly affected by fatigue and the average value of the $p-value$.

After having determined that a large majority of the participants were indeed significantly affected by acute fatigue, it was analysed the evolution of user performance throughout the day. Figure 2 shows these effects for a particular participant, in two different features and in each of the four previously mentioned moments. It is possible to see how, during the day, the excess of distance between

Table 1. Percentage of participants whose interaction patterns were significantly affected by fatigue, in each feature, and the average value of the $p - value$.

Metric	significant	p-value
Time between keys	100%	9.88e-74
Key down time	88.0%	7.42e-16
Mouse Acceleration	100%	3.83e-284
Mouse Velocity	100%	1.02e-126
Average Excess of Distance	94.4%	1.36e-13
Distance Between Clicks	94.4%	2.83e-07
Click Durations	94.4%	4.46e-143
Errors per key	77.7%	6.54e-21

clicks and the time between keys increases gradually. Both examples essentially depict a significant loss of performance.

In order to understand the effects of fatigue during the day it is useful to look at the evolution of the features. Figures 3 and 4 depict, respectively, the evolution of the velocity of the mouse and the key down time during the day, for a particular participant. They essentially depict the same trend of decreasing efficiency: the velocity of the mouse decreases from nearly 0.9 pixels/ms to 0.6 pixels/ms while the key down time increases from 86 ms to more than 90 ms. Both pictures belong to the same participant and the same day, between 9:27 AM to 5:50 PM.

Both the statistical differences and the trend identified show that there is indeed a gradual loss of performance over the day. Moreover, it is also proved that this loss of performance can be detected through non-invasive methods, specifically by looking at how the user interacts with the computer.

3.5   Classification Results

Having proved the existence of a change on the behaviours over the day due to fatigue, as well as the loss of performance in the use of mouse and keyboard, the final step is to train a classifier that can accurately quantify the level of fatigue of an individual.

To this aim, the performance of several classifiers using the data collected was assessed. Table 2 depicts the performance of these classifiers, with the KNN holding the best results although all of them behave fairly good. These tests were performed with a dataset with a total of 2514 instances and 10-fold cross method.

To build the dataset it was used part of the data collected in the case study described. Specifically, we selected instances from the two most different distributions: instances resulting from periods where the performance is lower(M4)
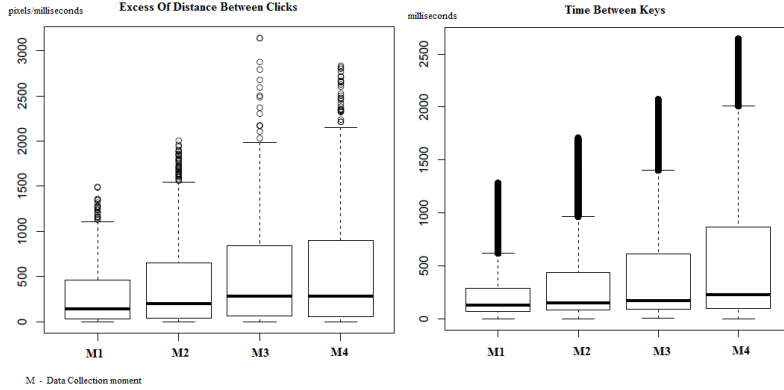
Fig. 2. The effects of fatigue on the excess of distance (left) and on the time between keys (right) for a participant, in each of the four moments.
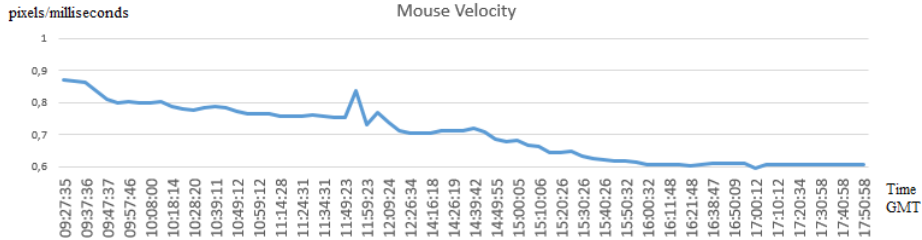
Fig. 3. Fatigue causes a gradual and consistent decrease in the mouse velocity over the day.

were labelled as *Fatigue*, and instances resulting from periods of better performance (M1) were labelled as *Normal*. The dataset describes 10 attributes: key down time, time between keys, error per key, mouse acceleration, mouse velocity, distance between clicks, click duration and average excess of distance. These attributes were selected to represent the features that revealed, in the overall, to be more significantly affected by fatigue in the moments considered.

Given this, we are currently using the k-Nearest Neighbor algorithm (k-NN) for classifying the interaction of a user with the computer in terms of fatigue. This is a non-parametric method that consists of a majority vote of its neighbours, with the object being assigned to the most common class among its k nearest neighbours. In this specific case, a value of k = 10 was used.

4   Conclusion

This paper presented an approach for measuring the level of fatigue of an individual through the analysis of his performance. This is achieved by analysing
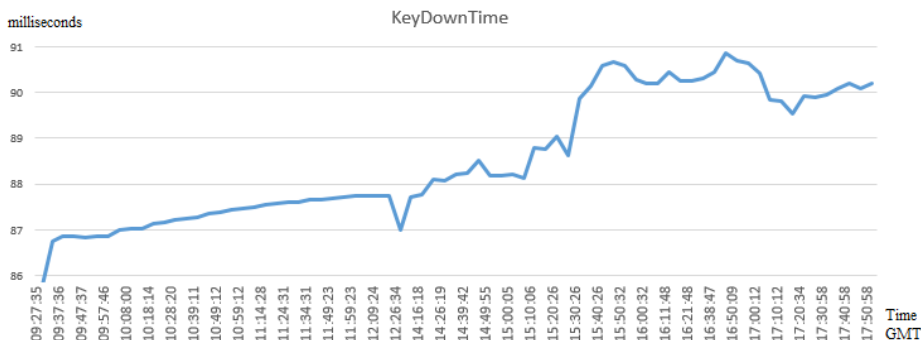
Fig. 4. Fatigue causes a gradual and consistent increase in the key down time over the day.

Table 2. Results of several classifiers on WEKA.

data set	Correctly Classified	Incorrectly Classified	Kappa s	MAE	ROC Area	RMSE
MultilayerPerceptron	92,97%	7.03%	0.83	0.09	0.939	0.240
KNN (k=10)	**96,85%**	**3.15%**	**0.92**	**0.04**	**0.992**	**0.155**
BayesNet	88,68%	11.31%	0.71	0.13	0.887	0.292
Regression	96,89%	3.10%	0.92	0.07	0.981	0.167

the interaction patterns of the individual with the mouse and keyboard. Thus being, it relies on behavioural analysis rather on traditional physiological sensors or questionnaires, being non-invasive, non-intrusive and objective.

The results achieved from the implementation of the case study detail prove that it is indeed possible to analyse and quantify human performance through the use of the mouse and keyboard, and this is a valid way to measure fatigue. It was also observed that the trend is towards the decrease in the performance of the interaction, visible through a generalized slower use of both the mouse and the keyboard or an increased distance between consecutive clicks.

This work opens the door to the development of real-time systems for fatigue monitoring through performance indicators. These can contribute to the development of healthier habits and even improve the quantity and quality of the work produced. All in all, such approaches can have a positive effect on both the worker and the institution. Nonetheless, we acknowledge that more accurate approaches can be developed when considering more specific context information not yet considered, such as the type of the task or workload, as well as environment information. This will be essential to develop accurate approaches to very specific domains in which the interaction with the peripherals is necessarily different than the one we studied in this paper. Moreover, we will also validate this approach against other existing ones, namely self-report mechanisms and EEG.

Acknowledgments

References

1. Williamson, R.J., Purcell, S., Sterne, A., Wessely, S., Hotopf, M., Farmer, A., Sham, P.C.: The relationship of fatigue to mental and physical health in a community sample. Social psychiatry and psychiatric epidemiology **40**(2) (2005) 126–132
2. Tucker, P.: The impact of rest breaks upon accident risk, fatigue and performance: a review. Work & Stress **17**(2) (2003) 123–137
3. Jaber, M., Neumann, W.: Modelling worker fatigue and recovery in dual-resource constrained systems. Computers & Industrial Engineering **59**(1) (2010) 75–84
4. Woźniak, M., Graña, M., Corchado, E.: A survey of multiple classifier systems as hybrid systems. Information Fusion **16**(0) (2014) 3 – 17 Special Issue on Information Fusion in Hybrid Intelligent Fusion Systems.
5. Borrajo, M., B.B.C.E.B.J.C.J.: Hybrid neural intelligent system to predict business failure in small-to-medium-size enterprises. International Journal of Neural Systems **21**(04) (2011) 277–296 PMID: 21809475.
6. Harris, W.: Fatigue, circadian rhythm, and truck accidents. In: Vigilance. Springer (1977) 133–146
7. Folkard, S., Tucker, P.: Shift work, safety and productivity. Occupational medicine **53**(2) (2003) 95–101
8. Smith, A.P., Miles, C.: The effects of lunch on cognitive vigilance tasks. Ergonomics **29**(10) (1986) 1251–1261
9. Hart, S.G., Staveland, L.E.: Development of nasa-tlx (task load index): Results of empirical and theoretical research. Advances in psychology **52** (1988) 139–183
10. Reid, G.B., Nygren, T.E.: The subjective workload assessment technique: A scaling procedure for measuring mental workload. Advances in Psychology **52** (1988) 185–218
11. Turner, R.: The handbook of operator fatigue. Ergonomics **56**(9) (2013) 1486–1486
12. McClernon, C.K., Miller, J.C.: Variance as a measure of performance in an aviation context. The International Journal of Aviation Psychology **21**(4) (2011) 397–412
13. Morris, T., Miller, J.C.: Electrooculographic and performance indices of fatigue during simulated flight. Biological psychology **42**(3) (1996) 343–360
14. Tanabe, S., Nishihara, N.: Productivity and fatigue. Indoor Air **14**(s7) (2004) 126–133
15. Kahneman, D.: Remarks on attention control. Acta Psychologica **33** (1970) 118–131
16. Pimenta, A., Carneiro, D., Novais, P., Neves, J.: Monitoring mental fatigue through the analysis of keyboard and mouse interaction patterns. In: Hybrid Artificial Intelligent Systems. Springer (2013) 222–231