Effectiveness of Mobile Technology in Managing Fatigue: Balert App

Ricardo De La Vega 1*, Héctor Anabalón 2, Cristian Jara 2, Eduardo Villamil-Cabello 3, Miguel Chervellino 2 and Álvaro Calvo-Rodríguez 4

1 Department of Physical Education, Sport & Human Movement, Autonomous University of Madrid, Madrid, Spain, 2 AlertPlus, Santiago, Chile, 3 Escuela Universitaria Don Bosco, Madrid, Spain, 4 University of A Coruña, A Coruña, Spain

The performance of professional tasks with a high cognitive, emotional, and even physiological demand, can cause a state of mental fatigue, which implies attentional alterations, greater errors in the tasks performed and a decrease in personal and work productivity caused by a deterioration of the cognitive control processes. The present study presents a mobile phone application named BAlert that allows monitoring and controlling the body’s fatigue processes based on the scores obtained in the Stroop effect and the heart rate variability. A pilot study has been carried out with a sample of 63 adults who have used the application a total of 942 times. The results allow us to classify the subjects, by logistic regression analysis, in their fatigue levels in 74% of the occasions. These results highlight the importance of this mobile application to control work fatigue processes in different possible scenarios (military, health, sports, business, etc.).

Keywords: mental fatigue, heart rate variability, stroop, laboral fatigue, APP

INTRODUCTION

Fatigue, either central or peripheral, has become one of the most important research topics throughout decades (De la Vega et al., 2021). Besides, the analysis of psychophysiological fatigue is considered very important in different contexts (Lohani et al., 2019) resulting in studies related to human’s response to external and internal loads (Wijesuriya et al., 2007; Wilson et al., 2007). The external loads exerted on the individual are added to their skills and coping strategies, leading to a level of tolerance and adaptation to each situation (Folkman and Lazarus, 1988). Along the last decades, there has been shown differences between physical and mental fatigue roles, highlighting clear limitations in the study of central fatigue due to its indirect ways of measuring it, which emphasizes the importance of developing new central fatigue analysis procedures (Bittner et al., 2000). On the contrary, in order to measure physiological fatigue, there have been indicated clear methodologies (De la Vega et al., 2021).

One of the main and most studied areas involving fatigue is physical activity and sports performance (Rampinini et al., 2008). Those physically fatigueing activities, such as team sports, also demand cognitive efforts, as it requires sustained attention to make accurate and quick decisions sustained by processing information from a changing environment (Walsh, 2014). According to these demands, sports may also produce mental fatigue (understood as a psychobiological state characterized by tiredness and lack of energy produced by demanding and sustained periods of cognitive activity (Marcora et al., 2009; Smith et al., 2016). This mental fatigue combined with the implicit physical fatigue usually results in a reduction of performance (Pageaux et al., 2015; Smith et al., 2016). It is also important to assess the impact of mental fatigue on decision-making skills because it is well-known in the literature that decreases cognitive performance resulting in a negative impact during
laboratory-controlled tests on decision-making (Pageaux and Leppers, 2018). Stoop effect can be defined as the inference between two stimuli with different levels of relevance that are happening at the same time (Banich, 2019) which has no single locus on the brain. This could be useful to measure fatigue.

In the occupational and labor field, different technological applications are being developed that allow analyzing the response of workers to demands and requirements. Given the important role of sleep debt and the effect on cognitive performance as a relevant factor of fatigue (Belenky and Akersted, 2011), mathematical models, techniques and procedures have been developed, measuring performance in the operating environment and clinical practice to reduce risks of poor performance, lost productivity, errors, incidents and accidents in the workplace and play a crucial role in managing the risk of fatigue to improve performance, productivity, safety, and in the long run, term, improve the health and well-being of workers (Thomas et al., 2006; Van Dongen et al., 2010). Certain detection technologies may be helpful in determining both types of fatigue, but technology aimed at increasing alertness may only be suitable for countering TR fatigue. Technologies currently exist which enable detection of driver fatigue and interventions that have the potential to dramatically reduce crash probability (May and Baldwin, 2009). The successful implementation of these technologies depends on the cause and type of fatigue experienced (Begum et al., 2012; He et al., 2013).

As we know, fatigue is a function of the interaction of multiple factors, including the background of the sleep/wake cycle, the circadian rhythm phase and the workload, being modulated by individual differences in response to these factors (Van Dongen et al., 2005). A fatigue-inducing factor is one that shifts the distribution of fatigue risk in the direction of increasing the risk of error, incident or accident as is the case with the workload and cognitive performance People vary from each other in their sensitivity to these factors. This relative variability in sensitivity to sleep loss appears to be a lasting individual trait (Belenky and Akersted, 2011).

Fatigue measurements is operationally defined by subjective self-assessment and objectively by decreased alertness that reflects lower performance. The fatigue self-report consists of a verbal response (for example, the subject says “I’m tired”) or a written response responding to a fatigue scale. The lower performance can also be measured by a variety of forms such as cognitive test and others (Thomas et al., 2006).

There are individual differences in response to factors that cause fatigue which could have an association with genetic markers. There are also age-related differences as older people perform less than younger people when both are in rest, but perform better than younger people when both have sleep restrictions. In addition, there are individual differences which are likely to affect fatigue measured by self-perception and objective measures.

As we have presented, detection of fatigue could be very useful. However, there are very few reliable Apps that can provide a suitable tool for that purpose. Even some cutting edge technologies, such us eye-tracker, have proven that are useful classifying mental fatigue (Li et al., 2020), but it demands much more materials than a simple smartphone.

Certain detection technologies may be helpful in determining both types of fatigue, but technology aimed at increasing alertness may only be suitable for countering TR fatigue.

In this sense, virtual platforms and applications for evaluation and research will provide a useful tool to obtain individual subjects and collect all the information (Hernández-Mendo and y González-Ruiz, 2012; Huberty et al., 2019). Technologies currently exist which enable the detection of driver fatigue and interventions that have the potential to dramatically reduce the crash probability. The successful implementation of these technologies depends on the cause and type of fatigue experienced.

The main goal of this research is to analyze the usefulness of the BeAlert application for the detection of fatigue processes through the measurement of heart rate variability and the values obtained in the Stroop test.

**Materials and Methods**

This study followed a quasi-experimental design (Montero and León, 2007) and it received the approval of the University Ethical Commission in compliance with the Helsinki Declaration. All subjects were informed about the procedure and gave their written consent to participate. This study was carried out complying with the Standards for Ethics in Sport and Exercise Science Research (Harris et al., 2019).

**Participants**

The participants included 63 individuals from Madrid, Spain, 20 of whom were female and 43 male. The participants were aged between 18 and 42 years ($M_{\text{years}} = 23.62$, $SD_{\text{years}} = 4.87$). All of the participants regularly engaged in physical activity, between 4 and 14 h per week ($M_{\text{hours/week}} = 8.15$, $SD_{\text{hours/week}} = 3.02$), BMI between 16.4 and 31.4 ($M_{\text{BMI}} = 23.18$, $SD_{\text{BMI}} = 2.4$); and number of hours aware between 0 and 10 ($M_{\text{hours aware}} = 4.4$, $SD_{\text{hours aware}} = 3.01$). Six participants were excluded from the study for not completing the measurements correctly. Intentional sampling methods were used (Montero and León, 2007). Participants were asked to complete the mobile app task at least once a day for a period of 1 week. The total data collected was $N = 942$ ($N_{\text{female}} = 285$; $N_{\text{male}} = 657$).

**Instrumentation and Study Variables**

The number of hours of physical activity per week and the scores obtained on the DT test were used as controlled variables. This allows us to have a relatively similar population in the variable “level of physical activity,” which can affect the fatigue results obtained in the mobile app. Therefore, only the subjects in which there were no statistically significant differences in their weekly level of physical exercise were used in our pilot study (de la Vega et al., 2021).

To carry out this research, a mobile application (AP Móvil), has been created based on the application of two independent tests: (i) The Stroop effect, which allows the evaluation of the speed in making decisions simple (identify "same color—same word" -e.g.: “the green word written in green letters); and complex (identify "color—word of another color “e.g.: “the green word written in red”). The number of correct responses
and the reaction time to the appearance of the stimuli on the screen are taken into account; and (ii). The temporal parameters of the Heart Rate Variability (AVNN, RMSSD, SDNN, and PNN50). The app is available for download on Android and IOS systems (https://play.google.com/store/apps/details?id=net.alertplus.amfsmovil&hl=es_419&gl=US). The engineering team implemented both tests with the previous considerations offered to be able to properly record the values of both tests. The application allows the registration of all the values that, later, are used to offer the results on fatigue to the participants.

The duration of this test was 5 min approximately. In addition, the values of the body mass index and the number of hours awake were taken.

**Procedure**

All participants voluntarily participated in the research. An informational meeting was held where participants completed their informed consent. Instructions were offered to all volunteers so that they could download the application on their mobile phones, explaining its use. In order to carry out this pilot study, they were asked to carry out the test at least once a day for a period of 2 weeks.

**Statistics**

We have used the Multinomial Logistic Regression Analysis (MLRA), which allows us to analyze the relationship between multiple explanatory variables and different categorical variables (Schulte et al., 2003). It has been decided that group 1 (Fast STROOP and low SDNN) is the reference group on which the statistics of the rest of the groups will be analyzed (Fast STROOP and High SDNN; Slow STROOP and Low SDNN; Slow STROOP and High SDNN). The dependent variable used is composed of deviation in SDNN score, as a result of subtracting the SDNN of each person with respect to the theoretical SDNN that they should have according to their age (Nunan et al., 2010). From this variable, we can therefore interpret values >0 as High SDNN, and values <0 as low SDNN. On the other hand, we have the variable “Average STROOP Milliseconds” as the average time it takes each person to click for each of the 35 stimuli that were presented. For the purpose of this research, values <500 ms were interpreted as fast STROOP, and values ≥500 ms were interpreted as slow STROOP.

The classification variables based on the dependent variable are the scores offered in the mobile app in terms of stress level (S), RMSSD, AVNN, perceived quality of food (PQF), gender (male, female), body mass index (BMI) and the number of hours awake (HA). The global fit of the model was checked with the Nagelkerke R-squared statistic (Nagelkerke, 1991). All computations are carried out using the R Commander statistical language (see R Project http://cran.r-project.org). The significance level was set at \( p < 0.05 \).

**RESULTS**

The global fit of the model was checked with the Nagelkerke R-squared statistic with a value of 0.81 and with the classification table that shows a classificatory success rate of 74%. Both indicators reflect a very good fit of our logistic regression model with the chosen variables.

Descriptive statistics are presented in Tables 1, 2. The descriptive statistics values of each of the study variables are presented, as well as the grouping variables used.

\[
\text{SD} = \text{standard deviation; SDNN} = \text{standard deviation of normal to normal intervals; RMSSD} = \text{root mean square}
\]
TABLE 3 | Results from multivariate logistic regression model containing all explanatory variables (full model).

<table>
<thead>
<tr>
<th>Fast stroop—Low SDNN</th>
<th>β estimate</th>
<th>Standard error</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress (β_1)</td>
<td>−0.048</td>
<td>0.95</td>
<td>0.000</td>
</tr>
<tr>
<td>Age 17–21 (β_2)</td>
<td>−10.454</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>Age 22–30 (β_3)</td>
<td>−8.515</td>
<td>0</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSSD (β_4)</td>
<td>0.065</td>
<td>1.07</td>
<td>0.000</td>
</tr>
<tr>
<td>HR (β_5)</td>
<td>0.417</td>
<td>1.52</td>
<td>0.000</td>
</tr>
<tr>
<td>AVNN (β_6)</td>
<td>0.019</td>
<td>1.02</td>
<td>0.000</td>
</tr>
<tr>
<td>Alimentation (β_7)</td>
<td>0.431</td>
<td>1.54</td>
<td>0.143</td>
</tr>
<tr>
<td>Female (β_8)</td>
<td>−0.303</td>
<td>0.74</td>
<td>0.503</td>
</tr>
<tr>
<td>Epworth (β_9)</td>
<td>−0.162</td>
<td>0.85</td>
<td>0.009</td>
</tr>
<tr>
<td>BMI (β_10)</td>
<td>0.335</td>
<td>1.40</td>
<td>0.000</td>
</tr>
<tr>
<td>Hours awake (β_11)</td>
<td>0.064</td>
<td>1.06</td>
<td>0.373</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Slow stroop—Low SDNN</th>
<th>β estimate</th>
<th>Standard error</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress (β_1)</td>
<td>−0.001</td>
<td>1.00</td>
<td>0.999</td>
</tr>
<tr>
<td>Age 17–21 (β_2)</td>
<td>−4.208</td>
<td>0.01</td>
<td>0.000</td>
</tr>
<tr>
<td>Age 22–30 (β_3)</td>
<td>−4.75</td>
<td>0.01</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSSD (β_4)</td>
<td>0.009</td>
<td>1.01</td>
<td>0.529</td>
</tr>
<tr>
<td>HR (β_5)</td>
<td>0.227</td>
<td>1.25</td>
<td>0.002</td>
</tr>
<tr>
<td>AVNN (β_6)</td>
<td>0</td>
<td>1.00</td>
<td>0.857</td>
</tr>
<tr>
<td>Alimentation (β_7)</td>
<td>−0.922</td>
<td>0.40</td>
<td>0.000</td>
</tr>
<tr>
<td>Female (β_8)</td>
<td>0.227</td>
<td>1.25</td>
<td>0.535</td>
</tr>
<tr>
<td>Epworth (β_9)</td>
<td>0.196</td>
<td>1.22</td>
<td>0.000</td>
</tr>
<tr>
<td>BMI (β_10)</td>
<td>0.12</td>
<td>1.13</td>
<td>0.062</td>
</tr>
<tr>
<td>Hours awake (β_11)</td>
<td>0.051</td>
<td>1.05</td>
<td>0.302</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Slow stroop—High SDNN</th>
<th>β estimate</th>
<th>Standard error</th>
<th>Sig</th>
</tr>
</thead>
<tbody>
<tr>
<td>Stress (β_1)</td>
<td>−0.043</td>
<td>0.96</td>
<td>0.000</td>
</tr>
<tr>
<td>Age 17–21 (β_2)</td>
<td>−10.317</td>
<td>0.00</td>
<td>0.000</td>
</tr>
<tr>
<td>Age 22–30 (β_3)</td>
<td>−10.222</td>
<td>0.00</td>
<td>0.000</td>
</tr>
<tr>
<td>RMSSD (β_4)</td>
<td>0.069</td>
<td>1.07</td>
<td>0.000</td>
</tr>
<tr>
<td>HR (β_5)</td>
<td>0.377</td>
<td>1.48</td>
<td>0.000</td>
</tr>
<tr>
<td>AVNN (β_6)</td>
<td>0.016</td>
<td>1.02</td>
<td>0.000</td>
</tr>
<tr>
<td>Alimentation (β_7)</td>
<td>−0.843</td>
<td>0.43</td>
<td>0.004</td>
</tr>
<tr>
<td>Female (β_8)</td>
<td>1.325</td>
<td>3.87</td>
<td>0.003</td>
</tr>
<tr>
<td>Epworth (β_9)</td>
<td>−0.033</td>
<td>0.97</td>
<td>0.585</td>
</tr>
<tr>
<td>BMI (β_10)</td>
<td>0.29</td>
<td>1.34</td>
<td>0.001</td>
</tr>
<tr>
<td>Hours awake (β_11)</td>
<td>−0.065</td>
<td>0.94</td>
<td>0.303</td>
</tr>
</tbody>
</table>

of successive differences; BMI = body mass index; Amo = mode amplitude—% intervals corresponding to Mode; VR = variational range—difference between min and max r-r intervals; Mo = mode—most frequent R-R interval value.

In the Table 3, all the independent explanatory variables of the model are presented in the first column. The order, from highest to lowest importance for the model as a whole, is as follows: (1). Stress; (2). Age; (3). RMSSD; (4). HR; (5). AVNN; (6). Food; (7). Female; (8). Epworth; (9). BMI; and (10). Hours awake.

The interpretations summarize the meaning of each exponent of the β: how much the probability of occurrence of the event increases or decreases given a score in each independent variable (ceteris paribus).

The global fit of the model was checked with the Nagelkerke R-squared statistic with a value of 0.81 and with the classification table that shows a classificatory success rate of 74%. Both indicators reflect a very good fit of our logistic regression model with the chosen variables. Based on the odd ratios and Wald’s statistic, the order of importance of the independent variables is: (i). stress; (ii). Age; (iii). RMSSD; (iv). HR; (v). AVNN; (vi). Food; (vii). gender (female); (vii). Epworth; (viii). BMI.

The profiles that maximize the probability for each group are: (i). Person older than 30 years, low STRESS, low RMSSD, high HR, high AVNN, low EPWORTH, and high BMI for fast STROOP and low SDNN; (ii). Person older than 30 years, with high HR, low alimentation and high EPWORTH for slow STROOP and low SDNN. (iii). Women, older than 30 years, low STRESS, high RMSSD, high HR, high AVNN, low alimentation and high BMI for slow STROOP and high SDNN; and (iv). Men between 17 and 21 years, high STRESS, low RMSSD, low HR, low AVNN, high alimentation, and low BMI for fast STROOP and low SDNN.

DISCUSSION

Given the poor subjective perception of the level of fatigue, it is advisable to emphasize the establishment of objective parameters in the assessment (Shahid et al., 2010). For a more objective profile, various psychometric evaluations have been used, such as the PVT, the WAFA and the Stroop test, among others. In addition, the incorporation of biometric measurements that evaluate the psychophysiological status such as heart rate variability (HRV) has been very useful as a marker of fatigue and has been an element used in the evaluation of stress and acute mental load in workers from various areas and also as a relevant element in the detection of driver fatigue (Castaldo et al., 2015; Fallahi et al., 2016; Abtahi et al., 2018; Digiesi et al., 2020). The relationship between the variability of the heart rate of the people evaluated with sleep debt has also been described (Heinze et al., 2011). We use this evaluation together with objective elements such as the evaluation of the cognitive level with the Stroop Test and in this way the structure of the Balert device is organized to include the fundamental variables responsible for the development of fatigue.

Early detection of mental fatigue could be very practical in order to prevent occupational hazard and it has also importance in productivity (Li et al., 2020) in different sectors such as education, sports, the military context or the professions that must manage fatigue situations such as drivers. In the present study, a Stroop task had to be carried out with a large proportion of congruent trials. This large proportion of congruent trials paradoxically increases working memory load (Kane and Engle, 2003), since congruent trials encourage goal neglect such that the goal of naming the color has to be actively maintained in working memory and cannot be automatized during the time course of the task. Due to this complication, the task was more cognitively demanding than tasks used previously.
in the analysis of mental fatigue; in addition, since the Stroop task has also been employed in studies of self-control loss, an analysis of the time evolution of this loss of cognitive control was possible.

Finally, it should be noted that it would be of interest to weight all these average time results in the Stroop taking into account the success/error of the person clicking on each color, and check if the results change by weighing these successes/errors in the average time of the Stroop. A study on Stroop should be sought and consulted with which to replicate the construction of an indicator that summarizes the temporal information of the Stroop relativized with the rate of successes/errors of people in each click.

**DATA AVAILABILITY STATEMENT**

The datasets presented in this article are not readily available because right now the development company owns the rights and is in the development phase. Requests to access the datasets should be directed to hector.anabalon@alertplus.cl.

**ETHICS STATEMENT**

The studies involving human participants were reviewed and approved by Autonomous University of Madrid. The patients/participants provided their written informed consent to participate in this study.

**AUTHOR CONTRIBUTIONS**

RD: planning of the research design, statistical analysis, analysis of results, and writing of the article. HA: app development, research design planning, and article writing. CJ and MC: design and app development. EV-C: planning of the research design, analysis of results, and writing of the article. AC-R: statistical analysis. All authors contributed to the article and approved the submitted version.

**REFERENCES**


**Conflict of Interest:** HA, CJ, and MC were employed by company AlertPlus.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

*Copyright © 2021 De La Vega, Anabalón, Jara, Villamil-Cabello, Chervellino and Calvo-Rodríguez. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.*