A neurophysiological training evaluation metric for Air Traffic Management

G. Borghini-*IEEE Student Member*, P. Aricò-*IEEE Student Member*, F. Ferri, I. Graziani, S. Pozzi, L. Napoletano, J. P. Imbert, G. Granger, R. Benhacene, and F. Babiloni-*IEEE Member*

Abstract— **The aim of this work was to analyze the possibility to apply a neuroelectrical cognitive metrics for the evaluation of the training level of subjects during the learning of a task employed by Air Traffic Controllers (ATCos). In particular, the Electroencephalogram (EEG), the Electrocardiogram (ECG) and the Electrooculogram (EOG) signals were gathered from a group of students during the execution of an Air Traffic Management (ATM) task, proposed at three different levels of difficulty. The neuroelectrical results were compared with the subjective perception of the task difficulty obtained by the NASA-TLX questionnaires. From these analyses, we suggest that the integration of information derived from the power spectral density (PSD) of the EEG signals, the heart rate (HR) and the eye-blink rate (EBR) return important quantitative information about the training level of the subjects. In particular, by focusing the analysis on the direct and inverse correlation of the frontal PSD theta (4 –7 (Hz)) and HR, and of the parietal PSD alpha (10-12 (Hz)) and EBR, respectively, with the degree of mental and emotive engagement, it is possible to obtain useful information about the training improvement across the training sessions.**

I. INTRODUCTION

The rapid growth in worldwide air travel dramatically increased the demand for air traffic services. This demand increases loading on already burdened air traffic control systems and operators (Air Traffic Controllers; ATCos) that are at or near their designated maximum handling capacities [1]. Anecdotal evidence suggests that ATCos increasingly speak not of the difficulty of a given traffic density, but of the associated traffic complexity [2]. It is being recognized that complexity factors can interact in nonlinear ways and that individual differences between ATCos can mean that different controllers respond differently to the same constellation of complexity factors [3]. Factors such as skills, training, experience, fatigue and other "stressors" all mediate the relationship between task demands, safety and performance of the AT-controllers. Hence, it is easy to understand how quantitative information about the skill level of the controllers could help to evaluate and to decide if the ATCos might need more training before working onto real scenarios. Several studies described the correlation of spectral power of the EEG bands with the complexity of the task that the subjects are performing [4]. In fact, an increase of electroencephalographic (EEG) power spectral density (PSD) especially over the frontal cortex in the theta band (4 - 7 (Hz)) and a EEG PSD decrease in the alpha band (8-12 (Hz)) over the parietal cortex have been observed when the required mental workload, the task's complexity, the amount of information processing increase. Furthermore, it has been suggested that an increased Heart Rate (HR) could be related with an increased mental workload and engagement, while the Eyeblinks Rate (EBR) and duration are inversely correlated with the increase of the mental workload and attention [4]. The hypothesis at the base of this study is that the variations of EEG PSD in theta frequency band, over frontal areas, and in alpha band, over the parietal ones, together with the variation of the HR and EBR could be taken as indexes related to the training level of the subjects. To validate such hypothesis the EEG PSD gathered in a group of students during a daily training along a week must correlate with the levels of expertise of the task (behavioral data) and with the levels of the subject's perceived workload, as assessed by the NASA-TLX questionnaire.

G. Borghini is with IRCCS Fondazione Santa Lucia, Rome, Italy (corresponding author to provide phone: +39 3493112548; e-mail: gianluca.borghini@gmail.com).

P. Aricò, I. Graziani and F. Ferri are with the Dept. Physiology and Pharmacology, University "Sapienza" of Rome, Italy (e-mails: arico@dis.uniroma1.it, ilenia.graziani@gmail.com, ibfedericoferri@gmail.com).

S. Pozzi and L. Napoletano are with Deep Blue Research and Consulting, Rome, Italy (e-mails: simone.pozzi@dblue.it, linda.napoletano@dblue.it).

J. P. Imbert, G. Granger and R. Benhacene are with Ecole Nationale de l'Aviation Civile, Toulouse, France (e-mails: jean-paul.imbert@recherche.enac.fr, geraud.granger@recherche.enac.fr, railane.benhacene@recherche.enac.fr).

F. Babiloni, is with the Dept. Physiology and Pharmacology, University "Sapienza" of Rome, Italy and BrainSigns srl (e-mail: fabio.babiloni@uniroma1.it).

II. MATERIAL AND METHODS

A. Experimental group and ATM simulation task.

A group of six healthy volunteers has been selected in terms of age (21±4 years) and previous computer game skills and experience. The subjects have been asked to learn to execute correctly an ATM task (LABY), that never did before, under easy (E), medium (M) and hard (H) conditions, randomly selected and proposed. A reference condition (NEUTRAL), in which the subjects watched the stimuli's tasks without responding to them, has been defined for evaluate the variations of the neurophysiological parameters. The LABY microworld is a functional simulation of Air Traffic Control (ATC) that captures the underlying processes involved in electronic air traffic management (ATM) with a simplified version of the operational human-machine interface. Microworlds are computer-based human-in-the-loop simulation environments that offer testing, behavioural/physiological measurement, and training capabilities, with the flexibility to build various scenarios [5]. The LABY microworld is based upon the main task of guiding *N* plane(s) around a predetermined route, indicated by a green path (Fig. 1). Participants must input numerical values such as heading, flight level, speed, etc., in order to direct flight around the trajectory and to avoid any conflicts or obstacles which may occur during the flight-route. Penalties are applied if the aircrafts deviate off the route or if other constraints are not met. The difficulty of the task can be altered according to how many aircrafts the participant have to control, the number and type of clearances required over the time and the number/trajectory of other interfering flights. The subjects trained daily (almost for an hour a day) for 5 days (SESSIONS T1÷T5) and their neurophysiological signals have been recorded in the first (T1), in the third (T3) and in the fifth (T5) session, while the behavioral and performance data have been collected every day. At the end of each experimental condition the subjects filled the NASA-TLX questionnaire for the evaluation of the perceived workload of the proposed task.

B. *Signal analysis*

Electroencephalogram (EEG) and physiological signals, including vertical electrooculogram (EOG) and electrocardiogram (ECG), have been recorded by the digital monitoring *BEmicro* system (EBNeuro system). The sixteen EEG channels, the ECG and the EOG channels have been collected simultaneously with a sampling frequency of 256 (Hz). All the EEG electrodes have been referenced to both earlobes, and the impedances of the electrodes were kept below 10 (kΩ).

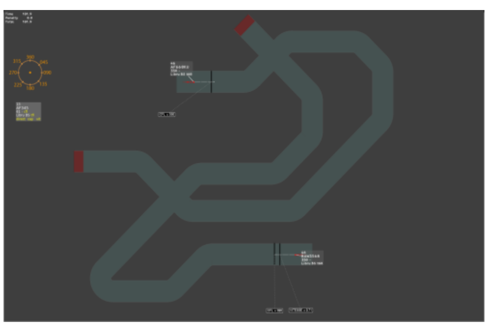

Figure 1. The LABY is a dynamic environment whereby an ATC must issue directional commands to guide *N* airplane(s) around a predetermined route, indicated by a green path, in order to avoid any conflicts or obstacles which may occur during the flight-route.

The bipolar electrodes for the heart activity have been placed on the Erb's point, while the bipolar electrodes for the EOG have been positioned vertically on the left eye. The acquired EEG signals have been digitally band-pass filtered by a 4^{th} order Butterworth filter (low-pass filter cut-off frequency: 30 (Hz), high-pass filter cut-off frequency: 1 (Hz)) and then segmented in epochs of 4 (sec), 2 (sec)-overlapped. The EOG signal has been used to remove eyes-blink artifacts from the EEG data by using the Gratton method [6]. For other source artifacts, a specific procedure for artifact removal, based on the approach involving the *Riemman* geometry theory has then been applied [7]. For each EEG-epoch, the *Power Spectral Density* (PSD) was calculated using a periodogram with *Hanning* window (2 seconds length) in the EEG frequency bands defined, for each subject, by the estimation of the *Individual Alpha Frequency* (IAF) value [8]. The PSDs have then been analyzed by estimating the *Coefficient of Determination* (r^2), or *r-square* [9], between the considered experimental condition and the reference condition. As $0 < r^2 < 1$ by definition, a signed r^2 has been derived by multiplying the coefficient of determination by the sign of the slope of the corresponding linear model of the regression analysis. In this way, it has been possible to obtain information not only about if the two datasets were different, but also about the direction of such difference. The HR and the EBR have been estimated by calculating the distance between consecutive peaks occurring in the ECG and in the EOG signals. In particular it have been used the R-peaks and the eyeblinks peaks and then they have been normalized by the *z-score* transformation with respect to the reference condition (NEUTRAL) [10].

C. *Statistical analysis*

The results derived from the different analysis have been then validated by the statistical analysis performed by using the STATISTICA software (Statsoft). The one-way repeated measures ANOVA (Confidence Interval, CI = .95) was used for all the data with the factor SESSIONS. Such factor has three levels, one for each day of the week in which the EEG recording was made (T1, T3 and T5). Duncan post-hoc tests have also been performed.

III. RESULTS

A. *Performance analysis*

Throughout the training sessions, the performance of the subjects increased continuously in terms of mean performance level and accuracy. Figure 2 shows the performance's index adopted across the different training days. By the inspection of Fig. 2 it is easy to note the simultaneous increase of the performances level and the decrease of the amplitude of the standard deviations in the learning curve. On the second day of training, all the subjects reached a good level of performance (almost the 90%) and since the third day, they could reach performance level higher than 95%. The one-way ANOVA performed on the global LABY score showed significant differences across the sessions ($F(4, 180) = 34.74$ with a $p < 10^{-5}$). The post-hoc Duncan test showed that the first two sessions (T1 and T2) were statistically different from all the others ($p < 10^{-4}$) while the last three ones (T3, T4 and T5) were not statistically different to each other.

B. *Frontal PSD theta*

The ANOVA results reported in Figure 3 show a significant modulation of the of EEG PSD in theta band over the frontal areas (EEG channels: AF3, AF4, F3, Fz, and F4) across the different training sessions ($F(2, 400) = 43.45$), $p < 10^{-5}$ and also the Duncan's post-hoc test confirmed these differences $p < 10^{-4}$.

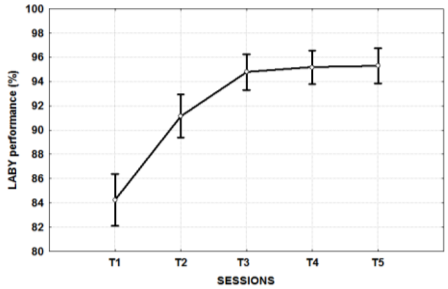

Figure 2. The trend of the global LABY score across the five different training sessions (T1÷T5). The figure reports the mean performance value and the standard deviations for the sessions. A statistical significant increase of the performance was obtained at the end of the period when compared to the first day of training.

It is evident that in the central session (T3), when the subjects have been supposed to have learnt how to execute correctly the task and focused the cognitive resources for improve their performances, the frontal PSD theta reached the highest increment respect all the other sessions.

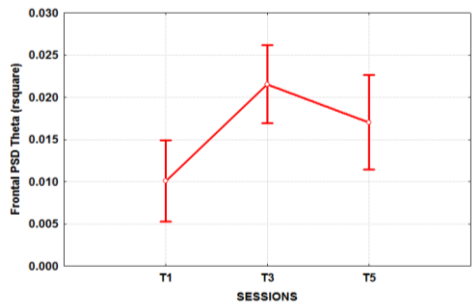

Figure 3. Signed r-square of the frontal EEG PSD in theta band over the frontal EEG channels AF3, AF4, F3, Fz and F4 across the training sessions T1, T3 and T5. At T3, the frontal PSD theta reached the highest increment ($p < 10^{-5}$).

C. *Parietal PSD alpha*

Figure 4 shows the trend of the parietal EEG PSD in alpha band over the EEG channels P3, Pz and P4, represented as variation of signed r-square. Repeated measures ANOVA showed significant differences of the parietal PSD alpha ($F(2, 240)=43.27$ with an associated p value $< 10^{-5}$) and a decreasing trend of the spectral PSD from T1 to T5 has been found out across the training sessions.

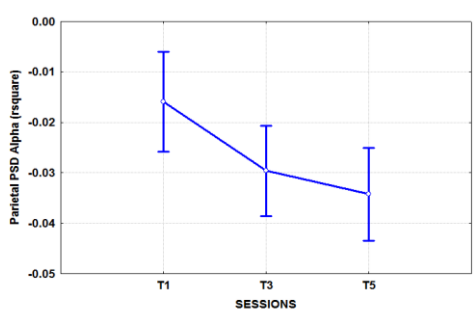

Figure 4. Signed r-square of the parietal EEG PSD in alpha frequency band during the training sessions (T1, T3 and T5). The continuous decrement is significant across all the training sessions ($p<10^{-5}$).

D. *Heart and Eyeblinks rates*

Figure 5 and 6 show the results of the statistical analysis of the autonomic parameters of HR and of EBR. The HR shows that the subjects were emotively engaged in correspondence of the central training session (T3), as the HR in T3 was the highest one, and that at the end of the training period they were more confident with the experimental task, as both the HR and the EBR decreased and increased, respectively.

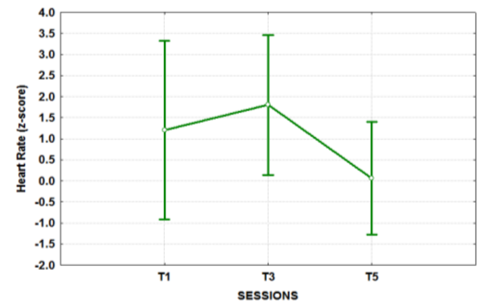

Figure 5. Heart Rate (z-score) values across the trianing sessions. The trend shows how in the central part of the training period (T3) the subjects showed an high emotive engagment, as the HR got the highest value.

In fact, the Duncan's post-hoc tests reported significant ($p<.01$) differences between the HR and EBR values of the first (T1) and last (T5) training session. In addition, the EBR z-score shows how the subjects kept to pay attention to the task, as it was negative even at the end of the training.

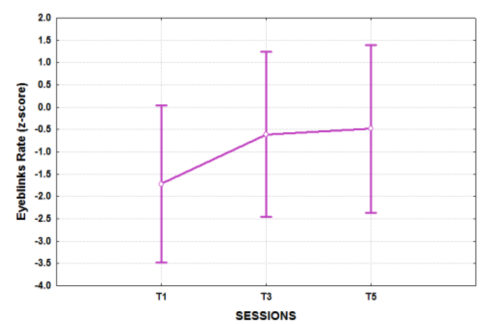

Figure 6. Eyesblink rate (z-score) values across the training sessions. The values are all negative because the subjects paid attention to the task for the whole training period and it shows how the subjects got more confident with task session after session.

F. *Perceived workload*

The one-way ANOVA for the NASA-TLX data shows significant differences among the training sessions ($F(4, 180)=19.39$ and $p< 10^{-5}$). A post-hoc test allowed to check out that the average scores of the NASA-TLX were statistically different until the fourth session (T4), whereas the T4 and T5 sessions were perceived as similar in terms of workload.

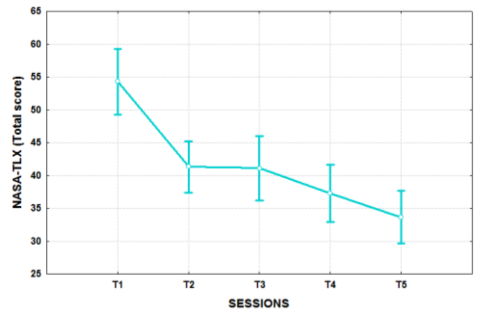

Figure 7. Average NASA – TLX scores of the training sessions. After each training session the subjects perceived the difficulty of the experimental task easier than the previous one.

IV. Discussion

The neurophysiological parameters, the task performance score and the experienced workload describe a story in which the training improvement of the subjects could be analyzed and quantitatively described. As it is possible to see from the results, after a couple of training sessions the subject started feeling more confident and learnt how to execute the task correctly, and in the central part of the training period (T3) the cognitive and emotive engagements became the highest. The frontal PSD theta and the HR reached the highest values, whereas the parietal PSD alpha and the EBR decreased significantly respect the first sessions. The continuous decrement of the parietal PSD alpha and of the EBR also showed how the subjects kept to pay attention to the execution of the task. From a perception point of view, the NASA-TLX scores demonstrated that session after session the subjects experienced less workload, especially at the end of the training period (T5), respect to the beginning of it (T1).

V. Conclusion

The integration of information derived by the EEG, ECG and EOG signals could be used as possible “cognitive metric” of the learning process and the training progress of learners throughout their periods of professional formation. After a fixed period of training it could be possible to compare subjects’ cognitive performances by estimating the neurophysiological parameters presented in this study, and by the High Resolution EEG should be also possible to investigate which are the cortical areas mainly involved during the execution of the task and in the different difficulty conditions [11].

Acknowledgment

This work is co-financed by EUROCONTROL acting on behalf of the SESAR Joint Undertaking (the SJU) and the EUROPEAN UNION as part of Work Package E in the SESAR Programme. Opinions expressed in this work reflect the authors' views only and EUROCONTROL and/or the SJU shall not be considered liable for them or for any use that may be made of the information contained herein. This work is also supported by BrainSigns srl, with a grant from Regione Lazio, FILAS, named “BrainTrained”. CUP F87I12002500007.

References

[1] B. D. Nordwall. FAA programs advance, with caution flags flying. *Aviation Week and Space Technology*, vol. 148(5), 43–47, 1998.

[2] B. Kirwan, M. Rodgers, e D. Schäfer, *Human Factors Impacts in Air Traffic Management*. Ashgate, 2005.

[3] R. H. Mogford, K. Harwood, E. D. Murphy, and R. J. Roske-Hofstrand. Review and evaluation of applied research techniques for documenting cognitive processes in air traffic control, 1994.

[3] W. Pawlak, C. Brinton, K. Crouch, and K. Lancaster. A framework for the evaluation of air traffic control complexity, in: *Guidance, Navigation, and Control Conference*, American Institute of Aeronautics and Astronautics, 1996.

[4] G. Borghini, L. Astolfi, G. Vecchiato, D. Mattia, and F. Babiloni. Measuring neurophysiological signals in aircraft pilots and car drivers for the assessment of mental workload, fatigue and drowsiness, *Neurosci. Biobehav. Rev.*, 2012.

[5] C. Gonzalez, P. Vanyukov, and M. K. Martin. The use of microworlds to study dynamic decision-making. *Computers in Human Behavior*, vol. 21, 273–286, 2005.

[6] G. Gratton, M. G. Coles, and E. Donchin. A new method for off-line removal of ocular artifact, *Electroencephalogr. Clin. Neurophysiol.*, vol. 55, n. 4, pagg. 468–484, 1983.

[7] A. Barachant, A. Andreev, and M. Congedo. The Riemannian Potato: an automatic and adaptive artifact detection method for online experiments using Riemannian geometry, in: *Proceedings of TOBI Workshop IV*, pp. 19–20, 2013.

[8] W. Klimesch. EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res. Brain Res. Rev.*, vol. 29, n. 2–3, pagg. 169–195, 1999.

[9] N. R. Draper. Applied regression analysis bibliography update 1994-97. *Commun. Stat. - Theory Methods*, vol. 27, n. 10, pp. 2581–2623, 1998.

[10] G. W. Milligan and M. C. Cooper. A study of standardization of variables in cluster analysis. *J. Classif.*, vol. 5, n. 2, pp. 181–204, 1988.

[11] M. Zavaglia, L. Astolfi, F. Babiloni, M. Ursino. A neural mass model for the simulation of cortical activity estimated from high resolution EEG during cognitive or motor tasks, *Journal of neuroscience methods*, vol. 157 (2), pp. 317-329, 2006.