34th Annual International Conference of the IEEE EMBS
San Diego, California USA, 28 August - 1 September, 2012

A Subjective Assessment of a P300 BCI System for Lower-Limb Rehabilitation Purposes

Matthieu Duvinage, Thierry Castermans, Mathieu Petieau,
Karthik Seetharaman, Thomas Hoellinger, Guy Cheron and Thierry Dutoit

Abstract— **Recent research has shown that a P300 system can be used while walking without requiring any specific gait-related artifact removal techniques. Also, standard EEG-based Brain-Computer Interfaces (BCI) have not been really assessed for lower limb rehabilitation/prosthesis. Therefore, this paper gives a first baseline estimation (for future BCI comparisons) of the subjective and objective performances of a four-state P300 BCI plus a non-control state for lower-limb rehabilitation purposes. To assess usability and workload, the System Usability Scale and the NASA Task Load Index questionnaires were administered to five healthy subjects after performing a real-time treadmill speed control. Results show that the P300 BCI approach could suit fitness and rehabilitation applications, whereas prosthesis control, which suffers from a low reactivity, appears too sensitive for risky and crowded areas.**

I. Introduction

Since its beginning, Brain-Computer Interface (BCI) research has focused on increasing performance without really considering the patient himself [1]. Actually, classification accuracy or information transfer rates have often been used to show the superiority of one paradigm or one method against another. For a long time, new world records were emphasized and some of them were presented as a revolution for heavily disabled patients or for potential gamers that could use BCI as an additional input.

However, although information transfer rate and classification accuracy are important, they only represent a part of a reality. As a proof of that, BCI applications are known for having a very high abandonment rate, i.e. users quickly give up using the system. As a proposition towards a better acceptance rate, some subject-based feedback questionnaires were suggested. The most famous ones are the System Usability Scale (SUS), the NASA Task Load Index (NASA-TLX) and for games, the Game Engagement Questionnaire, which measure the usability, the cognitive load and the engagement, respectively [1].

One of the most noble BCI purposes is rehabilitation. Developments in that field have allowed several main steps. Hand grasping was made possible thanks to detection of Steady-State Visual Evoked Potential (SSVEP) that arises when looking at a flickering image at a higher frequency than 6 Hz [2]. This allows one to grasp or release an object. In [3], a patient could control a wheelchair by modulating his EEG signals thanks to three different mental states.

Recently, interesting locomotion-related studies have appeared using Electroencephalography (EEG). Several researchers have focused on the analysis of spontaneous brain signals during the gait cycle [4], [5]. Other researchers have even claimed that a decoding of EEG signals is easy to obtain using linear decoding scheme [6]. However, as reviewed in [7], those studies could somehow show incoherent and/or subjective results. Actually, when performing a time-frequency analysis [4], similar results can be obtained with an accelerometer signal positioned on the head as with the claimed EEG-based cortical signals. Additionally, no distinction between descending (command) and ascending (feedback) signals has been done making it impossible to know whether a prosthesis could be controlled that way. On top of that, these studies have only been conducted at artificially low-speeds to attenuate artifacts (under 2.5 km/h, a range of walking speeds for which the gait style is more erratic and quite different from standard walk [8]) or without considering accelerations. All these weaknesses suggest that a lot of research has to be done before understanding all aspects of locomotion cortical control.

This is why a more robust and close-to-market but less natural scheme using gait modeling and a standard BCI has been proposed in [9]. Inspired from biology, this approach uses Central Pattern Generators (CPGs) widely used in robotics. These CPGs can model automatic gait patterns based on kinematics. When the patient wants to modify the gait pattern frequency, i.e. the gait speed, he uses a standard BCI with high-level commands. A proof of concept with a non natural P300 paradigm was given in [10]. Actually, the P300 evoked potential is an involuntary positive potential that arises around 300 ms after the user has perceived a relevant and rare stimulus [11]. One strong advantage of this approach is the spontaneous aspect of P300 implying very low requirements to manage the interface. Furthermore, although the P300 is known to be less efficient than the SSVEP, this study has a double interest as argued in [12]: P300 is more suitable for people suffering from epilepsy or people having difficulties in accurately controlling eye muscles. On the other hand, an external screen is needed to produce the potential and solutions like the augmented reality eyewear (Vuzix, Rochester, NY, USA) have to be envisaged to display stimuli on a semi-transparent module.

M. Duvinage (FNRS Research Fellow), T. Castermans and T. Dutoit are with Faculty of Electrical Engineering, TCTS Lab, University of Mons, 7000 Mons, Belgium

M. Petieau, K. Seetharaman, T. Hoellinger and G. Cheron are with LNMB Lab, Université Libre de Bruxelles, CP 168, 50 Av F Roosevelt, Brussels, Belgium

Matthieu Duvinage is the corresponding author: Matthieu.Duvinage@umons.ac.be

In order to compare this approach to other BCI pipelines, the main objective of this paper is to propose a first baseline estimation of the P300 BCI objective and subjective performances in the framework of lower-limb fitness, rehabilitation and prostheses. To simulate such applications, a strong feedback was introduced by controlling the speed of the treadmill subjects were walking on.

In this paper, section 2 describes the P300 approach, the pipeline and the experiment description. Section 3 details the subjective and objective performance measures. Section 4 presents and discusses the results.

II. P300 SYSTEM

This section first details the P300 approach. Then, the acquisition system and the pipeline are explained. Finally, the experiment and its purpose are presented.

A. P300 Approach

Following previous work [10], we are interested in a four-speed BCI managing a non-control state, which does not send any instruction to the orthosis control system. The screen is composed of two rows and two columns representing Low, Medium and High speed states and the Stop state as depicted in Figure 1. This arrangement could mimic the corner of the Vuzix eyewear. The different speeds respectively correspond to 1.5, 3 and 4.5 km/h whereas the Stop state simulates the standing state (0 km/h).

To modify the current speed, the subject has to focus on the corresponding letter during a whole trial. After randomly flashing each row/column, the different P300 responses and the assumed letter can be detected. On the other hand, when the subject does not want to modify the speed, i.e. during the non-control state, he does not look at the screen and no P300 response is elicited.

Fig. 1: P300 visualization is divided into four states: Low-speed, Medium-speed, High-speed and Stop. A fifth state is detected by the system when the user is not looking at the screen.

B. P300 Pipeline

EEG was recorded using a 32-electrode cap connected to the ANT acquisition system (Advanced Neuro Technology, ANT, Enschede, The Netherlands) digitizing the signals at 512 Hz. Left ear was chosen as reference. Mastoid was not used because of possible pollution from EMG signals of the neck while walking. Electrode impedance was measured and maintained under 20 kΩ for each channel using electrode gel.

Following [10], providing the EEG signals downsampled at 32 Hz, the pipeline was composed of several main components: a temporal high-pass filter, an xDAWN-based spatial filter [13], an epoch averaging and a Linear Discriminant Analysis (LDA) classifier using a voting rule for the final decision sent to a Virtual Reality Peripheral Network (VRPN) server [14].

The frequency band of interest was obtained by high-pass filtering the EEG signals (1 Hz cutoff) through a 4th order Butterworth filter. Thus, after the downsampling, the undesired slow drift in the measurement and high-frequency noise such as power line interference were removed [15].

Afterwards, a spatial filter was designed using an xDawn algorithm [13]. By linearly combining EEG channels, this algorithm defines a P300 subspace. When projecting EEG signals into this subspace, P300 detection is enhanced.

Then, the resulting signal was epoched using a time window of 600 ms starting immediately after the stimulus. Groups of two epochs corresponding to a specific row/column were averaged and magnitude features were extracted over the windows. The flash, no flash and inter-repetition durations were respectively 0.2 s, 0.1 s and 1 s.

Finally, a 12-fold LDA classifier was applied to each two-grouped averaged time windows giving a value which represents the distance to a hyperplane separating at best the target/non-target classes. For a given trial, in a voting classifier, the row/column, which had been activated was determined by summing six consecutive LDA outputs (12 repetitions) and by choosing the maximum value. The decision was sent to a VRPN server to finally automatically control the treadmill speed [14].

In this pipeline, neither standard or gait-related artifact removal methods were used. Regarding gait artifact, it was shown that it is relatively small and not significant at low speeds in the waveform [16]. Moreover, on the basis of classification rates, current gait-related artifact removal methods are not significantly efficient compared to raw data in the context of a low-cost embedded system [17].

C. Experiment Description

Following [10], the experiment was twofold: a two-step training session and a realistic test session. The first part of the training session aimed at training the classifiers to detect the P300 response. It consisted of 25 trials of random letters (around 12 minutes). Then, a session of 10 trials of non-control state was recorded while the subject was not looking at the screen. This aimed at determining a threshold (by a Receiver Operating Characteristic (ROC) analysis) from which the voting rule result was significant. A practical application should not make mistakes while the subject is not looking at the screen given that this state has a higher probability of occurrence (except in city center). Moreover, this kind of mistakes would force the subject to frequently re-adjust the current speed due to misclassification of non-control states and would destabilize him due to unexpected speed modifications. In other words, the False Positive Rate (FPR), i.e. the number of non-target elements classified as target ones divided by the total number of non-target, should be as low as possible. In the ROC analysis, the threshold was set around FPR=1%.

Then, the test session aimed at controlling the overall system working in real-time. It consisted of a 60-target session (around 30 minutes) in order to get sufficient data to obtain a good overview of the system performance. Before each trial, a beep was emitted to indicate that a trial would start within three seconds. At this point, the subject declared what his target was and focused on it until it was properly recognized. This simulated a real use scenario and allowed an assessment of the objective performance measures. After this correct recognition of a target letter, the subject did not look at the screen during one trial to simulate the non-control state. Additionally, a uniform distribution of commands was requested and checked during the experiment. In order to assess the need for concentration, after the middle of the experiment, the experimenter spoke with the subject to detect potential influence on control failures.

Four male and one female subjects (with age between 22 and 34 years old) participated in this experiment. Subject 3 never had any experience with a BCI system before the experiment. A 20-inch screen was placed at about 1 meter in front of the subject. Subjects were healthy and did not have any known locomotion-related or P300 disturbing diseases or handicaps. After the experiment, each subject was administered SUS and NASA-TLX questionnaires described here below (they were familiarized with them before doing the experiment). Finally, an interview about the answers was performed to identify the system strengths and weaknesses.

III. Performance Measures

In this section, the measures of performance are described. First, following [18], the main focus is on subjective feedback measures including SUS and NASA-TLX questionnaires. Then, objective measures, composed of classification rate, non-decision rate and error rate, are introduced.

A. Subjective Feedback Measures

The System Usability Scale (SUS) questionnaire has been proven to be a reliable, robust and low-cost usability evaluation tool that can be used for global assessment of system usability. Moreover, it has become a standard survey among thousands of studies as reported in [19]. This allows one to grade the studied interface and compare it with other surveys.

The SUS questionnaire is composed of ten items using a Likert scale. Each item corresponds to a statement and the respondent has to indicate the degree of agreement or disagreement on a 5 (7) point scale. Here, on top of that, a global single item evaluation was asked as proposed in [19]. This is graded from the worst imaginable (0) to the best imaginable (6) through OK (3). After properly scoring and weighting the different ten statements, a SUS score can be derived. The higher the SUS value, the better the system is.

Another widely used questionnaire is the NASA Task Load Index (NASA-TLX) [20]. This is a brief and powerful questionnaire for workload evaluation.

This questionnaire follows a two-step procedure. First, subjects have to rate six different item subscales which assess mental demand, physical demand, temporal demand, effort, frustration and performance. Each item is rated using a 20-temporal step bipolar scale resulting in a score between 0 and 100. In addition to detailed instructions, a pair of words are written at each extremity of the scale to help subjects. Given that there are two very different tasks (control and non-control), the evaluation was done 1) on a global basis and 2) focused on control commands.

The second step aims at determining the source of loads. By performing pair-wise comparisons based on the most contributive load item, the subject can determine the weight of each item in the overall workload. By computing the weighted average, a NASA-TLX score is obtained. But, in this study, this weighting was not used.

Some questions are really dependent on the application. Thus, subjects were asked to separately consider fitness and rehabilitation on a treadmill (which are quite close to the experiment considerations). They were also asked to consider the case of a prosthesis in a general context. Obviously, given that this latter application is quite different from the actual experiment, subjective results are less reliable but can give an indication. Of course, in case of small handicaps that are not considered for this latter application, very efficient techniques based on purely mechanical systems are available (for instance, for ankle prostheses [21]).

B. Objective Performance Measures

Given the specific design of the system, three objective measures were used. The classification accuracy is defined as the ratio between the number of correctly recognized targets and the total number of targets.

The non-decision rate is computed as the ratio between the number of non-control states and the number of trials when a command control is supposed to be emitted.

The error rate is defined as the resulting rate when a command control is emitted.

IV. Discussion

In this section, results are firstly discussed relying on objective measures. Then, subjectives measures and results of interviews are exposed.

Considering objective measures, results indicate desired functioning. Indeed, as depicted in Table I, no error occurred over all the experiments, which represent a real-time test of two and a half hours. As explained in [10], the price to pay for this "reliability" is a non-zero non-decision rate. This rate sometimes nearly reached 10% (no difference was observed by subjects while speaking during the test). This indicates that decreasing the number of repetitions to speed up the system is quite risky. To obtain the same behavior of providing commands when the system is quite certain, this would require a higher non-decision rate. On the other hand, being less conservative in the ROC analysis would provoke more unwanted speed modifications that could force the user to re-adjust his current speed more frequently, which could discourage him. In an end-user application, a trade-off should be made depending on the subject profile/feelings. However, all the subjects said that they preferred the conservative approach, which gives the impression of a sufficiently reliable system (even if they did not strictly test another approach). Moreover, it avoids some fears of suddenly modifying the speed when the subject was not prepared to.

TABLE I: Objective results show that the system is working as desired. It makes no error at the expense of a non-zero non-decision rate that can nearly reach 10%. However, all the non-control states are perfectly recognized.

Subject	Classification	Non-decision	Non-control classification
1	93.5%	6.5%	100%
2	90.6%	9.4%	100%
3	90.6%	9.4%	100%
4	96.8%	3.2%	100%
5	100%	0%	100%
Mean	94.3%	5.7%	100%
STD	4.1%	4.1%	0%

Considering questionnaires completed by the five subjects, results depicted in Table II are roughly just passable in terms of usability but subjects need a relatively low cognitive load. Based on SUS questionnaires, for fitness and rehabilitation, the system passes the *acceptable* threshold of 70 as defined in [22]. Consistently with the one-item global SUS score, the system can be defined as *good*. On the other hand, such a P300-based prosthesis system is on the edge of reaching this *acceptable* usability but can only be defined as *marginal (high)*. Given the specific framework of daily use, and, as stressed on by several subjects, the slow reactivity and the risk of non-recognition in case of emergency stop are highly damaging. Additionally, this is not really adapted for crowded places, i.e. typically a place where many speed modifications have to be done. However, some subjects indicated that this approach could be suited for leisurely walk in large areas. According to them, this could be a way for heavily disabled people to walk outside again. More interestingly, all subjects agreed that, excluding those drawbacks, the global approach is a nice way to control. This suggests that if the system can be increased in reactivity, it would be more broadly accepted.

Regarding NASA-TLX values, workload is below average. Indeed, workload during control seems to be slightly less important than a previous P300 study about an interface control (browser, word processing, software configuration, etc.) but close to average [23]. However, when considering the entire system, workload drastically drops. This shows that the entire system is perceived as a low workload demanding system with peaks during control commands that are relatively easily manageable. For a prosthesis application, a more important frustration is noticed again because of the context that increases risks and the need for reactivity and reliability.

Several subjects pointed out some weak points that could be enhanced to significantly increase the usability. Firstly, a technical support is generally desired, at least, at the beginning. This could be highly improved by using a user-friendly EEG cap with dry electrodes if similar performances are provided. This will also contribute to getting a less cumbersome device. Secondly, the loss of control is one fear of subjects. In addition to a recognition error in risky situations, the stress can provoke another error, etc, until a crash/an abandonment. Thirdly, as expected, the synchronous functioning is quite penalizing and an asynchronous system would be more appreciated. Fourthly, one subject proposed a system with acceleration/deceleration command instead of fixed speeds in order to avoid too important unwanted speed variations and to provide a more flexible system.

V. Conclusions and Future Work

In this paper, a first estimation of the subjective performance of a P300 BCI-based for lower-limb rehabilitation purposes is provided. In this experiment, the treadmill speed was used as a feedback. Performances were assessed on an objective basis (classification, error and non-decision rates) and on a subjective basis (SUS and NASA-TLX questionnaires). Five healthy subjects participated in this experiment.

Regarding objective measures, the system is working as designed. Over the five subjects, no error occurred at the expense of the non-decision rate that could sometimes nearly reach up to 10%. Generally, this specific working allowing not to detect a command state when information appears to be too uncertain was well appreciated. Even if they did not test another approach during the experiment, subjects found the tested approach rather reliable. Although they could feel uncomfortable with the non-decision rate, they prefer avoiding an unexpected speed modification.

Regarding subjective measures, mixed results appear. According to subjects, this approach could suit fitness and rehabilitation but not really prostheses. More precisely, they consider that such a prosthesis control system could be used in large areas, e.g. for promenades, but not in crowded and risky areas, e.g. city center. This is mainly due to the lack of reactivity and to the non-decision rate at critical time.

Future work will be devoted to refine these results on a larger population, to study the longitudinal performances and to enhance the P300 system based on user feedback. This includes asynchronous control [12], [24].

Then, it will focus on the study of several BCI paradigms. Typically, following a within-subject experimental design, P300, SSVEP and EOG-based eye movements interfaces will be compared [25]. Moreover, when a better understanding of spontaneous gait signals will be available, this assumed more natural but less stable interface will also be considered.

TABLE II: Comparing SUS and Global SUS scores, results are consistent with [22]. Moreover, it indicates that this P300 approach seems to be more suitable for fitness and rehabilitation than prosthesis mainly due to the lack of reactivity. Considering the overall workload, it appears to be quite small in average and the overload peaks reached during control states are smaller than [23].

	SUS (Global SUS)			NASA-TLX Global (Command)		
	Fitness	Rehabilitation	Prosthesis	Fitness	Rehabilitation	Prosthesis
Subject 1	87.5 (4)	87.5 (4)	77.5 (4)	20 (39.16)	20 (39.16)	25 (44.16)
Subject 2	57.5 (2)	67.5 (4)	62.5 (3)	10 (43)	10 (41.6)	10 (42.16)
Subject 3	80 (5)	75 (5)	70 (5)	25.8 (46.6)	25.83 (46.33)	27.5 (48.33)
Subject 4	77.5 (4)	75 (4)	67.5 (2)	20.8 (33.3)	24.16 (36.6)	27.5 (42.5)
Subject 5	77.5 (4)	72.5 (3)	67.5 (3)	14.16 (26.66)	18.33 (30.83)	21.66 (34.16)
Mean	76 (3.8)	75.4 (4)	69 (3.4)	18.15 (37.7)	19.66 (39)	22.33 (42.262)
STD	11.1 (1.1)	7.43 (0.71)	5.47 (1.14)	6.15 (7.9)	6.19 (5.85)	7.3 (5.15)

ACKNOWLEDGMENT

M. Duvinage is a FNRS (Fonds National de la Recherche Scientifique) Research Fellow. The authors want to express their gratitude for the financial support provided by Fonds Européen de Développement Régional (FEDER) and the Région Wallonne and by the European Commission through the MINDWALKER project (FP7 - 2007-2013).

This paper presents research results of the Belgian Network DYSCO (Dynamical Systems, Control, and Optimization), funded by the Interuniversity Attraction Poles Programme, initiated by the Belgian State, Science Policy Office. The scientific responsibility rests with its author(s).

REFERENCES

[1] H. Gürkök, G. Hakvoort, and M. Poel, "Evaluating user experience in a selection based brain-computer interface game: A comparative study," in *Entertainment Computing - ICEC 2011*, ser. Lecture Notes in Computer Science, J. Coutinho Anacleto, S. Fels, N. Graham, B. Kapralos, M. Saif El-Nasr, and K. Stanley, Eds., vol. 6972. Berlin, Germany: Springer Verlag, October 2011, pp. 77–88.

[2] R. Ortner, B. Allison, G. Korisek, H. Gaggl, and G. Pfurtscheller, "An SSVEP BCI to control a hand orthosis for persons with tetraplegia," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 19, no. 1, pp. 1 –5, feb. 2011.

[3] J. d. R. Millán, F. Galn, D. Vanhooydonck, E. Lew, J. Philips, and M. Nuttin, "Asynchronous Non-Invasive Brain-Actuated Control of an Intelligent Wheelchair," in *31st IEEE EMBC.*, 2009.

[4] J. Gwin, K. Gramann, S. Makeig, and D. Ferris, "Electrocortical activity is coupled to gait cycle phase during treadmill walking," *NeuroImage*, vol. 54, no. 2, pp. 1289 – 1296, 2011.

[5] M. Wieser, J. Haefeli, L. Bütler, L. Jäncke, R. Riener, and S. Koeneke, "Temporal and spatial patterns of cortical activation during assisted lower limb movement," *Exp. Brain Res.*, vol. 203, pp. 181–191, 2010.

[6] A. Presacco, R. Goodman, L. W. Forrester, and J. L. Contreras-Vidal, "Neural decoding of treadmill walking from non-invasive, electroencephalographic (EEG) signals." *Journal of Neurophysiology*, vol. 106, no. 305, pp. 1875–1887, 2011.

[7] T. Castermans, M. Duvinage, G. Cheron, and T. Dutoit, "EEG and human locomotion descending commands and sensory feedback should be disentangled from artifacts thanks to new experimental protocols," in *Biostec conference*, 2012.

[8] H. van Hedel, L. Tomatis, and R. Müller, "Modulation of leg muscle activity and gait kinematics by walking speed and bodyweight unloading," *Gait and Posture*, vol. 24, no. 1, pp. 35 – 45, 2006.

[9] M. Duvinage, T. Castermans, T. Hoellinger, G. Cheron, and T. Dutoit, "Modeling human walk by PCPG for lower limb neuroprosthesis control," in *Neural Engineering (NER), 2011 5th International IEEE/EMBS Conference on*, 2011, pp. 317 –321.

[10] M. Duvinage, T. Castermans, R. Jiménez-Fabiàn, T. Hoellinger, C. D. Saedeleer, M. Petieau, K. Seetharaman, G. Cheron, O. Verlinden, and T. Dutoit, "A five-state p300-based foot lifter orthosis: Proof of concept," in *3rd IEEE Biosignals and Biorobotics conference*, 2012.

[11] L. Farwell and E. Donchin, "Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials," *Electroencephalography and Clinical Neurophysiology*, vol. 70, no. 6, pp. 510 – 523, 1988.

[12] R. Panicker, S. Puthusserypady, and Y. Sun, "An asynchronous P300 BCI with ssvep-based control state detection," *IEEE Trans. Biomed. Eng.*, vol. 58, no. 6, pp. 1781 –1788, june 2011.

[13] B. Rivet, A. Souloumiac, V. Attina, and G. Gibert, "xDAWN Algorithm to Enhance Evoked Potentials: Application to Brain Computer Interface," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 8, pp. 2035 –2043, 2009.

[14] Russell, T. C. Hudson, A. Seeger, H. Weber, J. Juliano, and A. T. Helser, "VRPN: a device-independent, network-transparent VR peripheral system," in *VRST '01: Proceedings of the ACM symposium on Virtual reality software and technology*. New York, NY, USA: ACM, 2001, pp. 55–61.

[15] U. Hoffmann, J.-M. Vesin, T. Ebrahimi, and K. Diserens, "An Efficient P300-based BrainComputer Interface for Disabled Subjects," *Journal of neuroscience methods*, vol. 167, no. 1, pp. 115–125, Jan. 2008.

[16] K. Gramann, J. T. Gwin, N. Bigdely-Shamlo, D. P. Ferris, and S. Makeig, "Visual evoked responses during standing and walking," *Frontiers in Human Neuroscience*, vol. 4, no. 00202, 2010.

[17] T. Castermans, M. Duvinage, M. Petieau, T. Hoellinger, C. Saedeleer, K. Seetharaman, A. Bengoetxea, G. Cheron, and T. Dutoit, "Optimizing the performances of a p300-based brain-computer interface in ambulatory conditions," *Emerging and Selected Topics in Circuits and Systems, IEEE Journal on*, vol. 1, no. 4, pp. 566 –577, dec. 2011.

[18] P. Emanuele, S. Alessandro, G. Veronica, F. Stefano, , and B. Marta Olivetti, "Toward a usability evaluation of bcis," *International Journal of Bioelectromagnetism*, vol. 13, pp. 121–122, 2011.

[19] A. Bangor, P. T. Kortum, and J. T. Miller, "An empirical evaluation of the system usability scale." *Int. J. Hum. Comput. Interaction*, pp. 574–594, 2008.

[20] S. G. Hart and L. E. Stavenland, "Development of NASA-TLX (Task Load Index): Results of empirical and theoretical research," in *Human Mental Workload*, P. A. Hancock and N. Meshkati, Eds. Elsevier, 1988, ch. 7, pp. 139–183.

[21] R. Jiménez-Fabiàn and O. Verlinden, "Review of control algorithms for robotic ankle systems in lower-limb orthoses, prostheses, and exoskeletons," *Medical Engineering and Physics*, no. 0, pp. –, 2011.

[22] A. Bangor, T. Staff, P. Kortum, and J. Miller, "Determining What Individual SUS Scores Mean : Adding an Adjective Rating Scale," *Journal of Usability Studies*, vol. 4, no. 3, pp. 114–123, 2009.

[23] A. Riccio, F. Leotta, L. Bianchi, F. Aloise, C. Zickler, E.-J. Hoogerwerf, A. Kbler, D. Mattia, and F. Cincotti, "Workload measurement in a communication application operated through a P300-based brain-computer interface," *J. of Neural Eng.*, vol. 8, no. 2, p. 025028, 2011.

[24] H. Zhang, C. Guan, and C. Wang, "Asynchronous P300-based brain–computer interfaces: A computational approach with statistical models," *Biomed. Eng., IEEE Trans. on*, vol. 55, no. 6, pp. 1754 –1763, june 2008.

[25] M. Duvinage, T. Castermans, and T. Dutoit, "Control of a lower limb active prosthesis with eye movement sequences," in *Computational Intelligence, Cognitive Algorithms, Mind, and Brain (CCMB), 2011 IEEE Symposium on*, april 2011, pp. 1 –7.