

Passive Brain-Computer Interfaces for Robot-Assisted Rehabilitation

Domen Novak, Benjamin Beyeler, Ximena Omlin and Robert Riener

Abstract Stroke patients must exercise intensely with rehabilitation robots to achieve satisfactory rehabilitation outcome, but ensuring appropriate exercise difficulty is a challenging task. Brain-computer interfaces would be suitable for such difficulty adaptations since they capture both conscious and subconscious aspects of workload, but have seen little use in rehabilitation. This chapter reviews previous work on passive brain-computer interfaces and highlights the practical challenges of applying the technology to motor rehabilitation. Preliminary results of a study on workload estimation in a rehabilitation robot with healthy subjects are then presented. Adaptive stepwise regression is used to estimate different types of workload from electroencephalography signals recorded at different sites. Results show that electroencephalography can achieve more accurate workload estimation than autonomic nervous system responses and that adaptive estimation methods can further improve accuracy. However, the number of electrode sites needs to be reduced and issues such as motion artefacts must be resolved before passive brain-computer interfaces can be used in motor rehabilitation.

Keywords Machine learning · Passive brain-computer interface · Psychophysiology · Rehabilitation · Robotics

1 Introduction

1.1 Robot-Assisted Rehabilitation

Stroke affects about 1 million people in Europe each year (Brainin et al. 2000; Thorvaldsen et al. 1995). Though a stroke often causes severe impairment, it is possible to regain lost motor functions and improve quality of life through

D. Novak (✉) · B. Beyeler · X. Omlin · R. Riener
Sensory-Motor Systems Lab, ETH Zurich, Zurich, Switzerland
e-mail: domen.novak@hest.ethz.ch

appropriate therapy. Successful therapy is characterized by intensive, repetitive exercises of long duration (Bütefisch et al. 1995; Kwakkel et al. 2002; Nelles 2004). With respect to these criteria, the normal manually-assisted therapy has several limitations: it is labour-intensive, time-consuming, and expensive. By contrast, robot-assisted rehabilitation can reduce the number of therapist hours and increase the duration and number of training sessions. Furthermore, the robot provides multimodal feedback and supports the assessment of impairment score and functional ability (Guidali et al. 2011).

Despite its advantages, robotic guidance alone is not sufficient to guarantee positive rehabilitation outcome. The patient's motivation is also an important determinant of the outcome (Maclean 2002). Motivation has been highlighted as an additional advantage of robot-assisted rehabilitation, which can be enhanced with virtual environments that are viewed as more fun, engaging and motivating than conventional therapy (Colombo et al. 2007; Mihelj et al. 2012).

1.2 Exercise Difficulty Adaptation

One way to improve patient motivation is to ensure an appropriate exercise difficulty level: the patient should be challenged in a moderate but engaging way without causing undue boredom or stress. Difficulty can be adapted based on the patient's exercise performance (Cameirão et al. 2010; Zimmerli et al. 2012), but this ignores subjective factors such as perceived workload. For example, the patient may be successfully completing the task but only with excessive effort that quickly leads to frustration. An unobtrusive alternative for dynamic difficulty adaptation in motor rehabilitation was proposed in the form of psychophysiological measurements.

Psychophysiological measurements are defined as measurements of the body's responses to psychological factors such as workload, engagement and stress. The first such measurements used in motor rehabilitation were autonomic nervous system (ANS) responses, as the sensors are relatively cheap and can be quickly attached to the patient. Closed-loop rehabilitation difficulty adaptation systems based on ANS responses were first presented in 2011 for both upper limb (Novak et al. 2011a) and lower limb (Koenig et al. 2011) rehabilitation. Since then, other authors have proposed alternative solutions based on ANS responses (Badesa et al. 2012; Guerrero et al. 2013; Shirzad and Van der Loos 2013), but results have been mixed. Since ANS responses are heavily affected by physical workload, which is an integral part of motor rehabilitation, it can be difficult to extract psychological aspects (Novak et al. 2011b). Examples of existing rehabilitation robots that have been used with ANS-based difficulty adaptation are shown in Fig. 1.

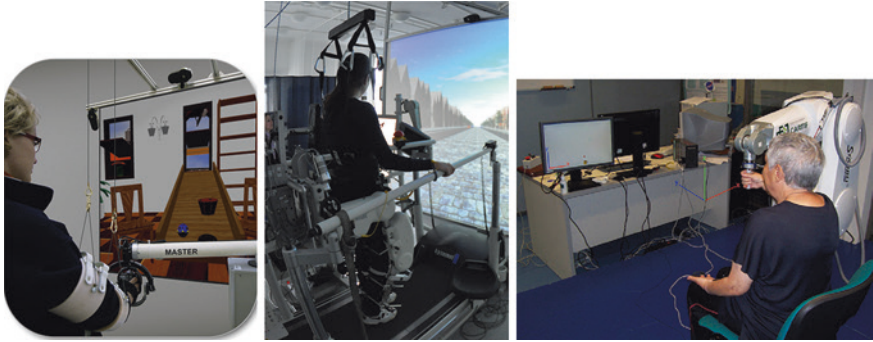


Fig. 1 Existing rehabilitation systems that have been tested with closed-loop difficulty adaptation based on autonomic nervous system responses. *Left to right* HapticMaster (Novak et al. 2010), Lokomat (Koenig et al. 2011), PhysioBot (Guerrero et al. 2013)

1.3 Passive Brain-Computer Interfaces

A promising alternative or complement to ANS responses are passive brain-computer interfaces (BCIs). Unlike classic ‘active’ BCIs, which measure intentionally generated brain activity (e.g. due to motor imagery), passive BCIs measure brain activity that occurs in response to, for example, stress or workload without conscious user effort (Zander and Kothe 2011). They have been used to classify workload, emotions and attention in many applications, including computer games (Chanel et al. 2011; Girouard et al. 2009), simulated flight (Wilson and Russell 2007) and driving (Zhao et al. 2011). Passive BCIs may represent a more promising and practical application of BCI than active ones since the required temporal resolution is much lower (Van Erp et al. 2012).

Recently, passive BCIs were used in physical human-robot interaction to detect mental workload and adapt robot behaviour accordingly (George et al. 2012). As brain activity should be less vulnerable to physical workload, passive BCI could offer an alternative to ANS responses in motor rehabilitation applications. It may even be possible to combine central and autonomic nervous system responses to obtain the optimal amount of information for exercise difficulty adaptation. Such a combination of passive BCI and ANS responses has already been tested in, for example, computer games (Chanel et al. 2011).

Passive BCIs, however, have their own weaknesses. They can require a significant time to apply, making them problematic in rehabilitation where the goal is to maximize the amount and intensity of exercise in a limited time period. They are vulnerable to motion artefacts, both due to sensor movement and due to electrical activity caused by muscle activation. Even without motion artefacts, inferring useful information from brain signals is not trivial and generally requires advanced machine learning techniques.

In the following sections, we shall discuss the problems of introducing passive BCIs to motor rehabilitation, suggest potential solutions, and finally show our own implementation of a passive BCI in the ARMin rehabilitation robot.

2 Hardware Selection and Setup

The crucial requirements for BCI hardware in motor rehabilitation are non-invasiveness and ease of use. A typical rehabilitation session lasts approximately one hour and should be spent exercising as intensively as possible. To be effective, the BCI should therefore require as little additional setup time as possible. Furthermore, since the added benefit of BCI is simply a more appropriate exercise difficulty (rather than, for example, allowing exercise to be performed at all), patients and therapists are likely unwilling to deal with great inconveniences in applying the equipment. Given these requirements, two promising technologies are electroencephalography (EEG) and functional near infrared spectroscopy (fNIRS), as they are both noninvasive and portable.

2.1 EEG Hardware

EEG is by far the most studied physiological signal for noninvasive BCIs. It measures the brain's electrical activity using electrodes placed on the scalp. Preparation time, however, can be up to 30 min when using a cap with ~15 gelled signal electrodes, a reference, ground, and electrooculography (EOG) electrodes. This is not suitable for motor rehabilitation, so we should aim to minimize the number of electrode sites while making each individual electrode quick and easy to apply.

2.1.1 Electrode Locations

There is no clear agreement on where to place electrodes for passive BCIs, possibly since each passive BCI measures a different aspect of the user's psychological state and thus requires a different electrode placement. Table 1 shows example electrode placements from various passive EEG studies.

We can see from the table that frontal sites are strongly represented, which is convenient as these sites are generally not covered by hair and allow easier electrode application. They are, however, more susceptible to EOG interference. Other popular sites are central (C3, C4, Cz) and parietal (P3, P4, Pz). However, authors often do not make a distinction between workload types, which may range from visual processing to short-term memory recall to decision making under temporal pressure.

Table 1 Signal electrode locations in different passive EEG studies, as well as the psychological variable(s) of interest

Study	Electrode locations	Psychological variable
Missonnier et al. (2003)	F3, F4, Fz, C3, C4, Cz, P3, P4, Pz	Working memory load
Wilson and Russell (2003)	F7, Fz, Pz, T4, T5, Oz	Mental workload
Berka et al. (2004)	Fz, Cz, POz	Vigilance
Ryu and Myung (2005)	Fz, Pz, O1, O2	Mental workload
Fairclough and Venables (2006)	Cz, P3, P4, Pz,	Engagement, distress, worry
Pesonen et al. (2006)	Fp1, Fp2, F3, F4, F7, F8, Fz, C3, C4, Cz, P3, P4, Pz, T3, T4, T5, T6, O1, O2	Auditory memory load
Berka et al. (2007)	F3, F4, Fz, C3, C4, Cz, POz	Engagement, workload
Wilson and Russell (2007)	F7, Fz, Pz, T5, O2	Mental workload
Venables and Fairclough (2009)	F3, F4, C3, C4, P3, P4	Mood, workload
Antonenko et al. (2010)	5 frontal, 4 temporal, 3 central, 2 parietal, 2 occipital	Cognitive load
Heger et al. (2010)	Fp1, Fp2, F3, F4, F7, F8, C3, C4, Cz, P3, P4, Pz, T3, T4, O1, O2	Workload
Wu et al. (2010)	Fp1, Fp2, Fz, Cz, Pz, O1, O2	Arousal
Knoll et al. (2011)	AF3, AF4, F3, F4, F7, F8, FC5, FC6, T7, P7, P8, T8, O1, O2	Cognitive workload
Brouwer et al. (2012)	F3, F4, Fz, FCz, C3, C4, Pz	Workload
Coffey et al. (2012)	F3, F4, Fz, FCz, C3, C4, Cz, Pz	Workload
George et al. (2012)	Fp1, Fp2, F3, F4, F7, F8, C3, C4, Cz, T8, P3, P4, Pz, T7, O1, O2	Mental workload

Robot-assisted rehabilitation involves many brain activities, including motion planning and visual processing. In order to identify the optimal electrode placement, we recommend an initial study of frontal, central and parietal sites with actual rehabilitation tasks. If possible, we should aim to minimize the setup to frontal sites. Though some studies place ground and reference electrodes in spots such as Cz or FCz, we recommend placing them in convenient spots such as the forehead or the ears/mastoids (Berka et al. 2004; Coffey et al. 2012; Ryu and Myung 2005; Wilson and Russell 2007).

2.1.2 Electrode Types

The most convenient device for measuring EEG in rehabilitation would be a low-cost device with integrated electrodes such as the Emotiv EPOC headset (Emotiv Systems, Australia). The Emotiv has previously been used for workload measurement and has successfully shown correlations between frontal signals and task difficulty (Knoll et al. 2011). However, in an evaluation with active BCI, it achieved significantly worse results than a medical EEG device (Mayaud et al. 2013).

As an alternative to low-cost systems, we can consider dry (ungelled) or even noncontact electrode systems, as patients are unlikely to accept gel on the scalp simply for automated difficulty adaptation. Such electrodes have already been shown to be comparable to classic electrodes in active BCI (Chi et al. 2012; Guger et al. 2012).

2.2 fNIRS Hardware

fNIRS has begun gaining ground in passive BCI due to the relatively quick and simple setup (Solovey and Girouard 2009). It provides a measure of blood oxygen concentration indicative of brain activity using one or more infrared light source-detector pairs that probe tissue up to depths of 1–3 cm. Since the light from the source is absorbed by hemoglobin and deoxygenated hemoglobin in the blood, changes in light intensity at the detector can be related to changes in relative concentrations of hemoglobin. The main weakness of fNIRS is that, due to the use of light, ambient lighting or dark hair can easily distort measured signals (Coyle et al. 2004).

2.2.1 Probe Locations

There are many possible placements for fNIRS probes, with the most common being the motor cortex (Sitaram et al. 2007) and the frontal/prefrontal cortex (Ayaz et al. 2012; Ong et al. 2013). The prefrontal cortex has been recommended for passive brain-computer interfaces since it deals with high-level processing such as working memory and problem solving (Solovey and Girouard 2009). fNIRS measurements from the prefrontal cortex are taken by placing the probe on the forehead, which is not covered by hair. This is both user-friendly and prevents problems with dark hair affecting measurements.

Since almost all passive BCI research with fNIRS has been performed with forehead measurements, this is also the best current candidate for motor rehabilitation, though tests with different locations on the forehead should be performed to find the area with the best response to rehabilitation tasks.

2.2.2 Probe Types

The basic fNIRS technology of multiple source-detector pairs is common to all existing probes. The main requirements for practical use are to block out ambient light, which distorts the signal, and to tightly fix the probe to the head. Previous studies have used probes embedded in black hats to block ambient light, and suggestions have been made that probes for applied studies should be embedded in hats or helmets (Solovey and Girouard 2009). At the moment, this appears more

user-friendly for rehabilitation than dimming the light in the room, and is therefore a prime concern in hardware selection.

Another potentially important feature of the probe is the short separation channel. This is an additional source-detector pair whose light does not penetrate deeply enough to measure brain activity, but does measure the same physiological noise as the other channels (Gagnon et al. 2012). It is frequently used to reduce noise, and is practical since it can be built into the same probe as the other channels and therefore does not increase the setup time.

2.3 Hybrid BCIs

It should be reiterated that there is no serious barrier to combining multiple sensor types, creating so-called hybrid BCIs. Hybrid passive BCIs such as EEG combined with ANS responses (Chanel et al. 2011) or even EEG combined with fNIRS (Coffey et al. 2012) have already achieved good results. The only obstacles are increased setup time and possible physical overlapping between sensors (e.g. fNIRS and frontal EEG). However, both can be mitigated with appropriate equipment.

3 Signal Processing

3.1 Artefact Removal

Motor rehabilitation is a noisy environment for passive brain-computer interfaces, and numerous artefacts must be considered. The most problematic ones are motion artefacts due to movement of the head or entire body. While commonly minimized in BCI, motion is an integral part of motor rehabilitation.

Motion affects measured signals either directly (e.g. by causing electrode/probe movement) or indirectly via human physiology. For EEG, the main indirect problem is that the electrodes also measure head and neck muscle activity as well as eye movement and blinking. Neck muscle artefacts are prominent toward the back of the head while eye artifacts are prominent toward the forehead. They cannot be removed by simple bandpass filtering, as frequency bands of the EEG, electromyogram (EMG) and EOG partially overlap (Vaughan et al. 1996). For fNIRS, motion can increase blood flow through the scalp, and head orientation can affect the signal due to gravity's effect on blood (Matthews et al. 2008). fNIRS is notably less vulnerable to eye artefacts than EEG.

Motion artefacts can be reduced using secondary sensors. For instance, eye artefacts can be removed from the EEG by using the EOG as a reference for noise removal algorithms (Croft and Barry 2000). Larger artefacts such as head movement can be detected using accelerometers and reduced in both EEG and fNIRS using

e.g. adaptive finite impulse filtering (Matthews et al. 2008). An alternative approach is to remove artefacts without a secondary sensor using computational methods such as principal or independent component analysis. This has been successfully performed to remove motion artefacts from EEG during walking (Gwin et al. 2010).

Besides motion artefacts, additional noise is caused by cardiorespiratory activity, which is visible in both the EEG (due to e.g. ECG or electrode movement as a result of respiration) and the fNIRS (affecting the blood flow). This noise is commonly removed by measuring cardiorespiratory activity using additional sensors and including this information as an input to adaptive filtering. Most notably, physiological noise could be measured in fNIRS using the short separation channel (Sect. 2.2.2), which may be a simple and convenient solution.

We should, however, consider to what degree physiological noise should be removed at all. Changes in heart rate or respiration also reflect psychological changes, so brain signals containing such physiological ‘noise’ may actually allow more accurate inference of the subject’s psychological state. Similarly, EOG artefacts seen in the EEG reflect eye movement and may provide useful information about visual processing. We believe that physiological noise removal should depend on the research goal. If the goal is to show that brain activity alone can be used to infer workload in rehabilitation, physiological noise should be minimized. However, if the goal is to obtain the most accurate psychophysiological inference, a passive BCI should be evaluated both with and without physiological noise.

3.2 Feature Extraction

In the context of psychophysiology and passive BCI, feature extraction refers to extracting a number of psychologically relevant features from raw physiological signals. They are generally calculated over a certain time period (window) and then fed to the psychophysiological inference algorithms. The length of this window depends on the measured signal and application, with values between 30 s and 5 min being common in psychophysiology (Novak et al. 2012). While EEG responds faster to stimuli than ANS signals and theoretically allows shorter windows, this is probably unnecessary. Feature extraction in closed-loop psychophysiological systems is generally done every time an action is taken by the system. As we should not adapt the rehabilitation task difficulty more than once a minute (or even less frequently), shorter windows are not needed.

3.2.1 EEG Feature Extraction

With regard to signal analysis, passive BCIs differ significantly from active ones. While active BCIs tend to focus on event-related potentials, passive BCIs generally measure brain activity over the entire time period of interest. This activity is examined in multiple frequency bands: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz), beta (13–30 Hz) and gamma (30–70 Hz). The overwhelmingly popular

EEG features for passive BCIs are total power in a particular band (e.g. Antonenko et al. 2010; Wilson and Russell 2007; Wu et al. 2010) and total power in a particular band divided by total power in all bands (e.g. Fairclough and Venables 2006). These features are commonly normalized with respect to a baseline (rest) condition in order to reduce intra- and intersubject variability (e.g. Antonenko et al. 2010).

Not all frequency bands are equally contaminated by motion artefacts. Particularly, frequencies above 20 Hz are significantly affected by EMG (Whitham et al. 2007), reducing their usefulness unless artefact removal methods are applied. This may be problematic since beta and gamma bands are connected to aspects of attention and mental workload (e.g. Herrmann et al. 2004). Though alpha and theta bands still contain a large amount of information about mental workload tasks (Klimesch 1999), this problem should be kept in mind.

3.2.2 fNIRS Feature Extraction

The first step in fNIRS feature extraction is to calculate oxygenated and deoxygenated hemoglobin concentrations using the modified Beer-Lambert law (Villringer and Chance 1997). The commonly extracted features are then simply the mean values of the two concentrations over the time period of interest (Ayaz et al. 2012; Girouard et al. 2009; Ong et al. 2013). As with EEG, the concentrations are often normalized by expressing them as a percentage of change from the baseline level.

4 Psychophysiological Inference

Once a set of potentially relevant features has been extracted from the EEG and/or fNIRS signals, the set should be assigned a label. For motor rehabilitation, this label can be categorical such as “task is too easy/too hard” (Novak et al. 2011a) or “workload is low/high” (George et al. 2012; Koenig et al. 2011). Alternatively, the label can be a continuous number that represents perceived task difficulty or workload (Badesa et al. 2012; Guerrero et al. 2013). The label type affects the actions that can be taken by the robot. Categorical labels are used to trigger discrete actions such as “change difficulty by one level” or “activate/deactivate robotic assistance” while continuous labels can be used for smoother, continuous control such as changing the gain of the robotic assistance.

4.1 Categorical Inference

Categorical labels in psychophysiology and passive BCI are inferred almost exclusively with classifiers based on supervised machine learning. A popular example of such a classifier is linear discriminant analysis (LDA), which has been used in ANS- and EEG-based motor rehabilitation systems despite its simplicity (George et al. 2012; Koenig et al. 2011; Novak et al. 2011a). Other classifiers that have been

previously used for psychophysiological inference include support vector machines, nearest-neighbour classifiers, Bayesian networks and neural networks (Novak et al. 2012). Most of these classifiers are also used in active BCIs (Lotte et al. 2007), with one significant difference. Active BCIs frequently employ hidden Markov models, which take temporal dynamics into account (Lotte et al. 2007; Zimmermann et al. 2013). These are uncommon in passive BCIs and ANS-based psychophysiological systems where dynamics within a time period are not very important.

The best classifier to use in a particular application is uncertain, and our recent review of psychophysiological measures (Novak et al. 2012) did not find a systematic advantage of any specific classifier, though we do not recommend nearest-neighbour classifiers since they are not robust to irrelevant features and features with different numerical ranges. Furthermore, dimensionality reduction methods such as principal component analysis or sequential feature selection are recommended to remove irrelevant input features (Novak et al. 2012). Dimensionality reduction is more relevant for EEG-based passive BCIs, which have a large number of input features compared to fNIRS-based BCIs.

4.2 *Continuous Inference*

Continuous inference is less common than categorical inference in both ANS-based psychophysiological systems (Novak et al. 2012) and in BCIs (Lotte et al. 2007), but has gained attention in motor rehabilitation since it allows smoother tuning of different parameters (Badesa et al. 2012; Guerrero et al. 2013; Mihelj et al. 2009). While continuous inference can also be based on machine learning techniques such as regression and neural networks (Novak et al. 2012), motor rehabilitation studies have preferred to use fuzzy logic (Badesa et al. 2012; Guerrero et al. 2013; Mihelj et al. 2009).

Fuzzy logic defines the relationship between physiological input features and the output label using if-then rules. Unlike classical logic, fuzzy rules and definitions have degrees of truth. For instance, while a fuzzy rule may state “if blood oxygenation is high, workload is high”, blood oxygenation can be 70 % ‘high’ and 30 % ‘low’ at a certain time. The if-then rules are manually defined by experts and are appropriate for noisy systems where a precise mathematical model does not exist, but experts can identify general rules underlying the system—which is the case in passive BCIs. Furthermore, fuzzy logic does not require training data, which can potentially simplify the design of the system.

5 Preliminary Implementation

This section describes our own preliminary experiment with a passive BCI in the ARMin arm rehabilitation robot. To illustrate the principle, we present first results; a more detailed analysis is planned for the future as a journal publication.

5.1 Goal

Passive BCIs have potential in rehabilitation robotics, but several application-specific issues remain. Outstanding questions include:

1. Passive BCIs commonly involve regular rest periods to allow physiological activity to return to a baseline state. These must be avoided at all costs in motor rehabilitation, as they decrease the amount of exercise performed by the patient. However, can passive BCIs still provide useful information without them?
2. Given the high levels of physical activity in motor rehabilitation, how heavily contaminated by motion artefacts are the measured signals?
3. What measurement locations are needed to obtain useful information? Specifically, are frontal locations (user-friendly, dominant in fNIRS) sufficient?
4. What psychological quantities do we wish to infer from the BCI data? Previous work has focused on workload (George et al. 2012; Koenig et al. 2011) or arousal/valence (Badesa et al. 2012; Guerrero et al. 2013; Mihelj et al. 2009). However, workload has many aspects [e.g. physical, temporal, mental (Hart and Staveland 1988)] that may be correlated with each other in a rehabilitation task. Furthermore, perceived workload may not always positively correlate with the effort the user puts into the task; for instance, users may become frustrated and give up as workload becomes too high.
5. What level of accuracy can the psychophysiological inference achieve? In one of our previous studies, for instance, using physiological measurements for closed-loop adaptation of a rehabilitation task was less accurate than simply using the task performance, but combining both sources of information gave the best accuracy (Novak et al. 2011a). Can a passive BCI outperform or at least complement task performance information and ANS responses?

Though we naturally cannot satisfactorily answer all questions at once, we designed a first study to obtain exploratory information.

5.2 Study Protocol

Ten healthy subjects (8 males, 2 females, 27.6 ± 3.7 years of age) were asked to perform a “whack a mole” game with the ARMin III rehabilitation robot. The ARMin III (Nef et al. 2009) has an exoskeletal structure with seven actuated degrees of freedom, including a hand module. The subject’s dominant arm is connected to the robot with cuffs on the upper arm and forearm. The hand is fixed to the hand module with elastic straps. The dimensions of the device are adjustable to the individual subject, and gravity and friction compensation allow the arm to be moved in all directions without resistance. A photo of the subject performing the task is shown in Fig. 2.

Fig. 2 A subject performing the task with the ARMin robot while monitored with sensors

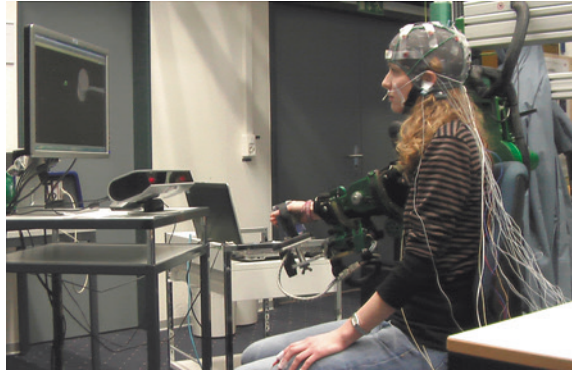


Fig. 3 A screenshot of the “whack a mole” task



The principle of the game is to hit monsters with a hammer before they disappear (Fig. 3). The hammer is moved around the screen with the end-effector of the robot, and a ‘hitting’ movement is performed by turning the forearm. Monsters appear at one of nine locations (3×3 layout) and disappear if not hit within a certain amount of time. Each monster has a mathematical equation attached to it, and the subject should only hit a monster if the equation is correct. 50 % of all equations are correct.

The task has two adjustable parameters: the equation difficulty and the frequency with which monsters are spawned. A new monster can spawn every 1.5, 2.5, 4 or 6 s. An individual monster remains on the screen 2.5 times the spawn interval, so there are at most three monsters on the screen at any time. The equation difficulty has five possible levels, from very easy (e.g. $2 + 5 = 7$) to very difficult (e.g. $45 + 33 + 63 = 141$). With 5 equation difficulty levels and 4 temporal difficulty levels, there are 20 possible conditions in total.

The study protocol began with a practice round where the subject played until he/she was comfortable and understood the task. The questionnaire was demonstrated, and the sensors were applied and calibrated. There was then a 60-s baseline

period during which subjects were asked to relax, not move, keep their eyes open, and remain silent. After the baseline period, subjects performed 19 60-s task periods, with each task period having a different combination of equation difficulty and monster spawn difficulty. Of the 20 possible combinations, only the easiest one (level 1 equations, one monster per 6 s) was omitted as it was found to be extremely boring for the subjects. The 19 combinations were presented in a random order that was generated differently for each subject.

Physiological measurements were taken continuously during the study. Each task period was followed by the questionnaire (Sect. 5.3) before the next task period began. Subjects were told to answer it for the preceding task period, not the entire task to then. After the 19th task period and questionnaire, the experiment was concluded.

5.3 *Measurements*

5.3.1 Questionnaire

The NASA-TLX (Hart and Staveland 1988) was used to obtain reference self-report values of workload during the task. It has been extensively used in human workload studies, including previous closed-loop psychophysiological work (Wilson and Russell 2007). It consists of six scales: mental workload, temporal workload, physical workload, performance, effort and frustration. Subjects rate each on a visual scale from ‘very low’ to ‘very high’. A computerized version was presented, with the subjects moving a slider along the visual scale by pronating/supinating their forearm in the robot. The selections were saved as numerical values from 0 to 100.

The task difficulty levels should affect the different NASA-TLX scales. Equation difficulty should affect mental and temporal workload, as subjects have to perform more complex mental arithmetic in the same amount of time. The monster spawn difficulty should affect both temporal and physical workload, as subjects must perform calculations faster and move their arm more often. Effort should increase with all workload types, though only up to a point; excessive workload may lead to a decrease in effort as subjects give up. Similarly, performance should decrease and frustration should increase with increasing workload.

5.3.2 Physiology

The primary measurement was EEG, which was measured with the g.GAMMAcap (g.tec Medical Engineering GmbH, Austria) and g.Butterfly active electrodes. Electrodes were placed at 14 locations of the International 10–20 system: Fz, F3, F4, F8, F7, Cz, C3, C4, Pz, P4, POz, O1 and O2. All signals were referenced to an electrode at position FPz and grounded with an electrode on the left earlobe.

Additionally, the EOG was recorded with two electrodes (Grass Technologies, USA): one to the upper right of the right eye and one to the lower left of the left eye. EOG was used only to correct ocular artefacts in the EEG. Both EEG and EOG were sampled at 600 Hz using a g.USBamp signal amplifier (g.tec).

In addition to EEG, four ANS responses were measured: the electrocardiogram (ECG), skin conductance, respiration and skin temperature. ECG was measured with three surface electrodes placed on the trunk. Respiration was measured using a thermistor flow sensor beneath the nose. Skin conductance was measured using a g.GSR sensor (g.tec). Electrodes were placed on the medial phalanges of the second and third fingers of the nondominant hand. Peripheral skin temperature was measured using a g.TEMP sensor (g.tec) attached to the distal phalanx of the fifth finger of the nondominant hand. All ANS signals were sampled at 600 Hz using a second g.USBamp amplifier.

Finally, eye tracking was performed using the SMI RED 250 (SensoMotoric Instruments, Germany), a remote eye tracker placed underneath and slightly in front of the screen. Though it is more commonly mounted directly below the screen, it was moved forward to ensure that the distance between the eyes and tracker is within the optimal operating conditions. The sampling frequency was 250 Hz.

5.4 Feature Extraction

Several features were extracted from the raw physiological signals for the baseline period and the 19 task periods. Each feature was calculated over the entire 60-s period.

5.4.1 EEG

The EEG was first bandpass-filtered between 1 and 30 Hz. Eye movement and blink artifacts were then removed using a recursive least-squares filter with EOG as the noise reference. The power spectral density of each EEG channel was then calculated using Welch's method. For each EEG channel, we calculated four features used in previous studies:

- alpha power divided by total power,
- theta power divided by total power,
- alpha power divided by theta power,
- $1/(\text{alpha power} + \text{theta power})$.

These features can optionally be individualized for each subject by using the peak frequency method to set the borders of the alpha band (Goljahani et al. 2012). Additional features based on the beta and gamma band were considered but later omitted due to concerns over data quality (see Sect. 5.6.1).

5.4.2 ANS Responses

From the ECG, intervals between two normal heartbeats (NN intervals) were extracted. Then, *mean heart rate* as well as three standard measures of heart rate variability (HRV) were calculated (Task Force, 1996): the *standard deviation of NN intervals (SDNN)*, the *square root of the mean squared differences of successive NN intervals (RMSSD)*, and the *number of interval differences of successive NN intervals greater than 50 ms divided by the total number of NN intervals (pNN50)*.

From the skin conductance signal, we detected all skin conductance responses (SCRs). A SCR is a transient increase in skin conductance whose amplitude exceeds 0.05 microsiemens and whose peak occurs less than 5 s after the beginning of the increase. *SCR frequency* and *mean SCR amplitude* were calculated.

From the respiration signal, we calculated *mean respiratory rate* and *standard deviation of respiratory rate*.

From the temperature signal, we calculated the *final skin temperature* as the mean temperature during the last 5 s of each period. Additionally, the *mean derivative of skin temperature* was calculated over the entire period.

5.4.3 Eye Tracking

Eye tracker feature extraction was done by the manufacturer's provided software, BeGaze 3.1, which first segments the recorded signals into blinks, saccades (rapid gaze shifts from point to point) and fixations. It then outputs the *number of blinks*, *number of saccades*, and *number of fixations* as well as the *mean blink duration*, *mean saccade duration*, and *mean fixation duration*.

For saccades, BeGaze outputs the *mean saccade velocity*, saccade velocity variability and *mean saccade amplitude*. For fixations, it outputs the *mean pupil diameter*, the *standard deviation of pupil diameter* and *mean gaze dispersion* (the amplitude of small movements performed by the eyes during a fixation). Finally, it outputs the *ratio of total fixation time and total saccade time*.

All of these features can optionally be individualized for each subject by setting different thresholds for fixations, saccades and blinks in BeGaze 3.1.

5.5 Psychophysiological Inference

Though we have previously worked extensively with classification algorithms (Koenig et al. 2011; Novak et al. 2011a), we chose to assign continuous values to each task period as suggested by other authors in rehabilitation robotics (Badesa et al. 2012; Guerrero et al. 2013). This was partially also why we selected the NASA-TLX as a reference—it measures each workload scale as a value between 0 and 100.

Stepwise linear regression was used to predict NASA-TLX reported values from the extracted physiological features. Since data were analyzed offline, cross-validation was used to obtain the results. The regression algorithm was trained with three approaches:

- Leave period out: Trained with data from all but one task period of one subject, then tested on the subject's remaining task period. Repeat for all subjects. In leave-period-out crossvalidation, EEG and eye tracking features were individualized to each subject as described in Sect. 5.4.
- Leave subject out: Trained with data from all but one subject, then tested on the remaining subject. Repeat for all subjects. In leave-subject-out crossvalidation, features were normalized for each task period. This was done for a period by subtracting the feature's baseline value (obtained during the initial baseline period) from the current value and dividing the result by the baseline value.
- Adaptive leave subject out: Same as leave subject out, but after each task period, the regression weights are updated using information from that task period through Kalman filtering. It thus gradually adapts to the current subject. The approach is computationally the same as in our previous adaptive LDA (Koenig et al. 2011; Novak et al. 2011a), except with a regression rather than classification function.

The measure of regression quality was the mean absolute error between the reported and predicted workload; the lower the error, the better. Regression functions were created separately for EEG, ANS and eye tracking data. Furthermore, to evaluate what accuracy would be achieved by a completely random regression function, regression functions were also created using twelve randomly generated features. These features' values were generated randomly for each time period from either a normal (6 features) or a uniform (6 features) distribution.

5.6 Initial Results and Discussion

5.6.1 EEG Data Quality

An examination of the EEG data found major motion artefacts: regardless of the task difficulty, measured power in the beta and especially gamma bands was much higher during any task period than during rest (up to nearly triple the baseline value). Tests before and after the official measurement protocol showed that high gamma activity is present even when no task is displayed and subjects simply move their arm inside the robot in a circular motion. Power in alpha and theta bands did not significantly increase when performing circular motions, and sometimes actually decreased during task periods.

From these observations, we conclude that beta and gamma band features cannot be used in motor rehabilitation without extensive artefact removal. It is currently unclear how much effort this would require. Rehabilitation robots already

measure arm movement, which could be used as an input to a noise reduction algorithm. However, this would only help with arm motions, not with head motions, which likely have a larger effect. For our first investigation, we chose to only utilize alpha and theta band information.

5.6.2 Correlations Between Game Difficulty and NASA-TLX

Pearson correlation coefficients were calculated separately for each subject, then averaged across subjects to obtain the final result. The mean correlation coefficient between *equation difficulty and mental workload* was 0.65 (range: 0.49–0.81) while the mean correlation coefficient between *monster spawn frequency and temporal workload* was 0.67 (range: 0.49–0.82). Workload was thus indeed induced by the task as desired.

However, there were also significant correlations between the different NASA-TLX scales. The mean correlation coefficient between *mental* and *temporal workload* was 0.41 (range: 0.07–0.71) while the mean correlation coefficient between temporal and physical workload was 0.42 (range 0.01–0.86). Effort was significantly correlated with all three workload types, though interestingly the correlation coefficients were negative in some subjects. The mean absolute correlation coefficients were 0.55 for *effort* and *mental workload* (range: –0.59 to 0.87), 0.46 for *effort* and *physical workload* (range: –0.36 to 0.74) and 0.54 for *effort* and *temporal workload* (range: –0.60 to 0.90). Finally, the mean absolute correlation coefficient between *effort* and *frustration* was 0.52 (range: –0.62 to 0.85). The same subjects who have negative correlations between effort and workloads (3 out of 10) also have negative correlations between effort and frustration.

While these results depend on the task, they suggest two things. First of all, it is not necessary to try and measure all types of workload in rehabilitation robotics, as they are correlated with each other. Second, subjects do sometimes respond to high workload by giving up and no longer putting as much effort into the task, as evidenced by negative correlation coefficients between effort and workload in some subjects.

For a motor rehabilitation task such as ours, we therefore suggest inferring two psychological dimensions from physiological measurements: the workload the subject is experiencing and how he/she is coping with it (actively or passively). This reinforces the suitability of the arousal/valence emotion model, which was used by previous studies (Badesa et al. 2012; Guerrero et al. 2013; Mihelj et al. 2009), but was suggested to be suboptimal due to the inability of ANS responses to accurately measure valence (Novak et al. 2010). An EEG-based passive BCI could measure valence more accurately than ANS responses, making this model more suitable. As different workload dimensions are difficult to separate in haptic and rehabilitation robotics (Novak et al. 2011b), such a two-dimensional model would likely be sufficient in most cases. A model with more dimensions, however, would be suitable for rehabilitation scenarios that consist of alternative mental and physical challenges (e.g. Koenig et al. 2011; Mihelj et al. 2012). A promising

candidate in such a case would be the proposed but untested arousal/valence/physical workload model of Mihelj et al. (2009).

5.6.3 Accuracy of Psychophysiological Inference

Since the NASA-TLX scales are significantly correlated, we present first results for estimation of *mental workload* and *effort* in Fig. 4. An example of reported and estimated (through leave-period-out linear regression) workload is shown for two subjects in Fig. 5.

All three physiological modalities estimated *mental workload* significantly better than random in leave-period-out cross validation where the regression function is trained on other data from the same subject. The accuracy of both EEG and eye tracking was significantly better than that of ANS responses. However, no modality provided significantly better than random results in leave-subject-out cross validation where the regression function is trained on data from other subjects.

ANS and EEG estimated *effort* significantly better than random, but again only in leave-period-out cross validation. ANS achieved a slightly better result than EEG, though the difference between the two was not significant. Leave-subject-out results were poor and actually significantly worse than random estimation in the case of EEG and eye tracking. However, the adaptive algorithm was able to greatly decrease leave-subject-out error, reaching approximately the same accuracy as in the leave-period-out case.

These results suggest that both mental workload and effort can be estimated better than random with EEG or other physiological data. The estimation can be done in the presence of physical activity and with only a single initial baseline, though user-specific models are needed. Furthermore, they demonstrate that EEG has advantages over previously used ANS responses in a rehabilitation robot: it

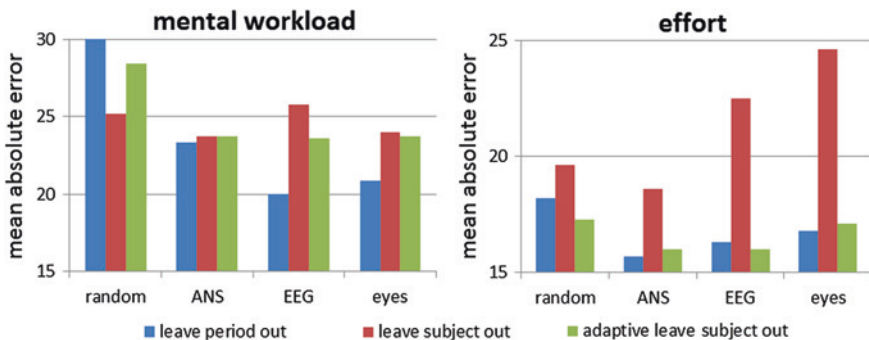


Fig. 4 Mean absolute error (difference between estimated and self-reported value in questionnaire units) for regression of mental workload (*left*) and effort (*right*) using autonomic nervous system responses, electroencephalography and eye tracking. The error when using random input data is shown for comparison

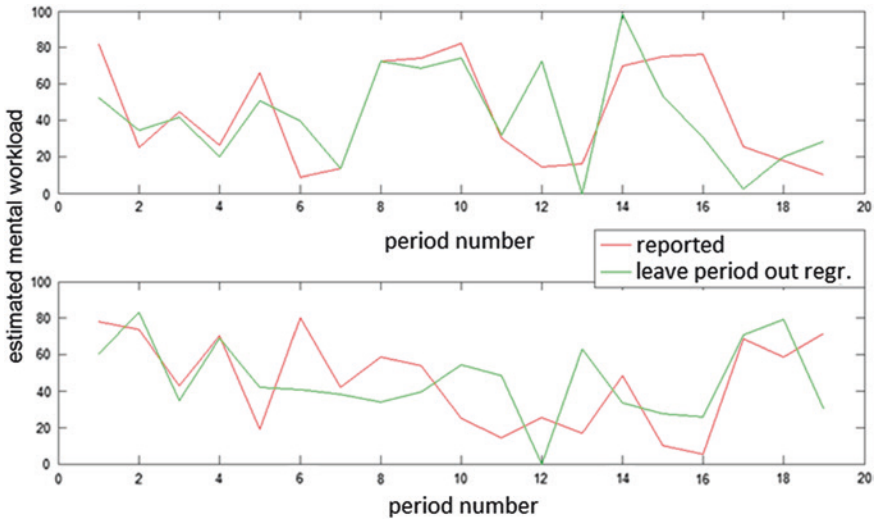


Fig. 5 Reported and estimated workload (through leave-period-out linear regression) for two subjects over the entire study. The *top* graph represents a subject with relatively low estimation error while the *bottom* graph represents a subject with high error

is able to estimate mental workload significantly better than ANS responses. If subject-specific models are not available, the adaptive algorithm can be used to greatly increase the error. However, the current offline implementation assumes that the algorithm can always adapt perfectly after each task period, which would not be the case in reality.

6 Conclusion and Outlook

In our review of the state of the art, we identified both EEG and fNIRS as promising passive BCI modalities for rehabilitation robotics. As the sensors need to be set up quickly in a rehabilitation environment, we should aim to minimize the number of electrodes/probes, use only frontal sites (not covered by hair), and use only dry (non-gelled) sensors, though not all of these goals may be achievable in practice. The two main practical problems in rehabilitation are the high level of physical activity, which results in motion artefacts, and the lack of baseline periods due to the need to maximize rehabilitation intensity.

In our first implementation with the ARMin III, we showed that EEG in the beta and especially gamma bands is strongly contaminated by motion artefacts, to the degree where such artefacts would be difficult to remove even with reference motion sensors. We therefore only used alpha and theta bands. Nonetheless, we were able to show that information from these two bands allows both mental

workload and effort to be estimated significantly better than random, with EEG outperforming ANS responses (previously used for task adaptation in rehabilitation robotics) in mental workload estimation. The estimation algorithms are computationally inexpensive and suitable for real-time use. However, subject-specific models or an adaptive (learning) algorithm are required. This may be due to the fact that subjects respond differently to workload, with some actually decreasing their effort as workload increases.

The immediate next step of our study is to compare the accuracy of EEG-based workload inference with the accuracy that can be achieved using nonphysiological data such as task score and movement information. Furthermore, we will attempt to identify the EEG channels that contribute the most to workload inference and thus attempt to minimize the number of channels used. We will develop algorithms to try and reduce motion artefacts in the EEG using either the robot's built-in position sensors or additional sensors to measure head movement. At the same time, we will conduct a second study to test whether fNIRS could provide a more convenient alternative to EEG. Finally, once an optimal, minimum-configuration setup is available, we will test it with actual patients undergoing motor rehabilitation to test both accuracy compared to healthy subjects and acceptance by the target population.

Acknowledgments This work was supported by the Swiss National Science Foundation through the National Centre of Competence in Research Robotics and by the Clinical Research Priority Program “NeuroRehab” University of Zurich.

References

- Antonenko P, Paas F, Grabner R, Gog T (2010) Using electroencephalography to measure cognitive load. *Educational Psychol Rev* 22(4):425–438
- Ayaz H, Shewokis PA, Bunce S, Izzetoglu K, Willems B, Onaral B (2012) Optical brain monitoring for operator training and mental workload assessment. *NeuroImage* 59(1):36–47
- Badesa F, Morales R, Garcia-Aracil N, Sabater J, Perez-Vidal C, Fernandez E (2012) Multimodal interfaces to improve therapeutic outcomes in robot-assisted rehabilitation. *IEEE Trans Syst Man Cybern Part C Appl Rev* 42(6):1152–1158
- Berka C, Levendowski D, Cvetinovic M, Petrovic M, Davis G, Lumicao M et al (2004) Real-time analysis of EEG indexes of alertness, cognition, and memory acquired with a wireless EEG headset. *Int J Hum Comput Interact* 17(2):151–170
- Berka C, Levendowski DJ, Lumicao MN, Yau A, Davis G, Zivkovic VT, Craven PL (2007) EEG correlates of task engagement and mental workload in vigilance, learning, and memory tasks. *Aviat Space Environ Med* 78(5 Suppl):B231–B244
- Brainin M, Borstein N, Boysen G, Demarin V (2000) Acute neurological stroke care in Europe: results of the European stroke care inventory. *Eur J Neurol* 7:5–10
- Brouwer A-M, Hogervorst M, Van Erp JBF, Heffelaar T, Zimmerman PH, Oostenveld R (2012) Estimating workload using EEG spectral power and ERPs in the n-back task. *J Neural Eng* 9(4):045008
- Bütefisch C, Hummelsheim H, Denzler P, Mauritz K (1995) Repetitive training of isolated movements improves the outcome of motor rehabilitation of the centrally paretic hand. *J Neurol Sci* 130:59–68

- Cameirão MS, Badia SBI, Oller ED, Verschure PFMJ (2010) Neurorehabilitation using the virtual reality based rehabilitation gaming system: methodology, design, psychometrics, usability and validation. *J Neuroeng Rehabil* 7:48
- Chanel G, Rebetz C, Bétrancourt M, Pun T (2011) Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Trans Syst Man Cybern Part A Syst Hum* 41(6):1052–1063
- Chi YM, Wang Y-T, Wang Y, Maier C, Jung T-P, Cauwenberghs G (2012) Dry and noncontact EEG sensors for mobile brain-computer interfaces. *IEEE Trans Neural Syst Rehabil Eng* 20(2):228–235
- Coffey EBJ, Brouwer A-M, van Erp JBF (2012) Measuring workload using a combination of electroencephalography and near infrared spectroscopy. In: *Proceedings of the human factors and ergonomics society 56th annual meeting*, vol 56(1), pp 1822–1826
- Colombo R, Pisano F, Mazzone A, Delconte C, Micera S, Carrozza MC et al (2007) Design strategies to improve patient motivation during robot-aided rehabilitation. *J Neuroeng Rehabil* 4:3
- Coyle S, Ward T, Markham C (2004) Physiological noise in near-infrared spectroscopy: implications for optical brain computer interfacing. In *Proceedings of the 26th annual international conference of the IEEE engineering in medicine and biology society*, pp 4540–4543
- Croft RJ, Barry RJ (2000) Removal of ocular artifact from the EEG: a review. *Clin Neurophysiol* 30(1):5–19
- Fairclough SH, Venables L (2006) Prediction of subjective states from psychophysiology: a multivariate approach. *Biol Psychol* 71(1):100–110
- Gagnon L, Cooper RJ, Yücel MA, Perdue KL, Greve DN, Boas DA (2012) Short separation channel location impacts the performance of short channel regression in NIRS. *NeuroImage* 59(3):2518–2528
- George L, Marchal M, Glondou L, Lecuyer A (2012) Combining brain-computer interfaces and haptics : detecting mental workload to adapt haptic assistance. In: *Proceedings of EuroHaptics 2012*, pp 124–135
- Girouard A, Solovey E, Hirshfield L (2009) Distinguishing difficulty levels with non-invasive brain activity measurements. In: *Proceedings of INTERACT 2009*, pp 440–452
- Goljehani A, D'Avanzo C, Schiff S, Amodio P, Bisiacchi P, Sparacino G (2012) A novel method for the determination of the EEG individual alpha frequency. *NeuroImage* 60(1):774–786
- Guerrero CR, Fraile Marinero JC, Turiel JP, Muñoz V (2013) Using “human state aware” robots to enhance physical human-robot interaction in a cooperative scenario. *Comput Methods Programs Biomed* 112(2):250–259
- Guger C, Krausz G, Allison BZ, Edlinger G (2012) A comparison of dry and gel-based electrodes for P300 BCIs. *Front Neurosci* 6:60
- Guidali M, Duschau-Wicke A, Broggi S, Klamroth-Marganska V, Nef T, Riener R (2011) A robotic system to train activities of daily living in a virtual environment. *Med Biol Eng Comput* 49(10):1213–1223
- Gwin JT, Gramann K, Makeig S, Ferris DP (2010) Removal of movement artifact from high-density EEG recorded during walking and running. *J Neurophysiol* 103(6):3526–3534
- Hart S, Staveland LE (1988) Development of NASA-TLX (task load index): results of empirical and theoretical research. In: Hancock PA, Meshkati N (eds) *Human mental workload*. North Holland Press, Amsterdam
- Heger D, Putze F, Schultz T (2010) Online workload recognition from EEG data during cognitive tests and human-machine interaction. *Lect Notes Comput Sci* 6359:410–417
- Herrmann CS, Munk MH, Engel AK (2004) Cognitive functions of gamma-band activity: memory match and utilization. *Trends Cogn Sci* 8(8):347–355
- Klimesch W (1999) EEG alpha and theta oscillations reflect cognitive and memory performance: a review and analysis. *Brain Res Rev* 29(2–3):169–195
- Knoll A, Wang Y, Chen F, Xu J, Ruiz N, Epps J, Zarjam P (2011) Measuring cognitive workload with low-cost electroencephalograph. In: *Proceedings of INTERACT 2011*, pp 568–571

- Koenig A, Novak D, Omlin X, Pulfer M, Perreault E, Zimmerli L, Riener R (2011) Real-time closed-loop control of cognitive load in neurological patients during robot-assisted gait training. *IEEE Trans Neural Syst Rehabil Eng* 19(4):453–464
- Kwakkel G, Kollen B, Wagenaar R (2002) Long term effects of upper and lower limb training after stroke: a randomised trial. *J Neurol Neurosurg Psychiatry* 72:473–479
- Lotte F, Congedo M, Lécuyer A, Lamarche F, Arnaldi B (2007) A review of classification algorithms for EEG-based brain-computer interfaces. *J Neural Eng* 4(2):R1–R13
- Maclean N (2002) The concept of patient motivation: a qualitative analysis of stroke professionals' attitudes. *Stroke* 33(2):444–448
- Matthews F, Pearlmutter BA, Ward TE, Soraghan C, Markham C (2008) Hemodynamics for brain computer interfaces. *IEEE Signal Process Mag* 25(1):87–94
- Mayaud L, Congedo M, Van Laghenhove A, Orlikowski D, Figère M, Azabou E, Cheliout-Heraut F (2013) A comparison of recording modalities of P300 event-related potentials (ERP) for brain-computer interface (BCI) paradigm. *Neurophysiol Clin* 43(4):217–227
- Mihelj M, Novak D, Munih M (2009) Emotion-aware system for upper extremity rehabilitation. In: *Proceedings of virtual rehabilitation 2009*, pp 160–165
- Mihelj M, Novak D, Milavec M, Zihel J, Olenšek A, Munih M (2012) Virtual rehabilitation environment using principles of intrinsic motivation and game design. *Presence Teleoperators Virtual Environ* 21(1):1–15
- Missonnier P, Leonards U, Gold G, Palix J, Ibáñez V, Giannakopoulos P (2003) A new electro-physiological index for working memory load in humans. *NeuroReport* 14(11):1451–1455
- Nef T, Guidali M, Riener R (2009) ARMin III—arm therapy exoskeleton with an ergonomic shoulder actuation. *Appl Bionic Biomech* 6(2):127–142
- Nelles G (2004) Cortical reorganization—effects of intensive therapy. *Arch Phys Med Rehabil* 22:239–244
- Novak D, Mihelj M, Zihel J, Olenšek A, Munih M (2011a) Psychophysiological measurements in a biocooperative feedback loop for upper extremity rehabilitation. *IEEE Trans Neural Syst Rehabil Eng* 19(4):400–410
- Novak D, Mihelj M, Munih M (2011b) Psychophysiological responses to different levels of cognitive and physical workload in haptic interaction. *Robotica* 29(03):367–374
- Novak D, Mihelj M, Munih M (2012) A survey of methods for data fusion and system adaptation using autonomic nervous system responses in physiological computing. *Interact Comput* 24:154–172
- Novak D, Zihel J, Olenšek A, Milavec M, Podobnik J, Mihelj M, Munih M (2010) Psychophysiological responses to robotic rehabilitation tasks in stroke. *IEEE Trans Neural Syst Rehabil Eng* 18(4):351–361
- Ong M, Russell PN, Helton WS (2013) Frontal cerebral oxygen response as an indicator of initial attention effort during perceptual learning. *Exp Brain Res* 229(4):571–578
- Pesonen M, Björnberg CH, Hämäläinen H, Krause CM (2006) Brain oscillatory 1–30 Hz EEG ERD/ERS responses during the different stages of an auditory memory search task. *Neurosci Lett* 399(1–2):45–50
- Ryu K, Myung R (2005) Evaluation of mental workload with a combined measure based on physiological indices during a dual task of tracking and mental arithmetic. *Int J Ind Ergon* 35(11):991–1009
- Shirzad N, Van der Loos HFM (2013) Adaptation of task difficulty in rehabilitation exercises based on the user's motor performance and physiological responses. In *Proceedings of the 2013 IEEE international conference on rehabilitation robotics*
- Sitaram R, Zhang H, Guan C, Thulasidas M, Hoshi Y, Ishikawa A, Birbaumer N (2007) Temporal classification of multichannel near-infrared spectroscopy signals of motor imagery for developing a brain-computer interface. *NeuroImage* 34(4):1416–1427
- Solovey E, Girouard A (2009) Using fNIRS brain sensing in realistic HCI settings: experiments and guidelines. In: *Proceedings of UIST 2009*, pp 157–166
- Thorvaldsen P, Asplund K, Kuulasmaa K, Rajakangas AM, Schroll M (1995) Stroke incidence, case fatality, and mortality in the WHO MONICA project. *Stroke* 26(3):361–367

- Van Erp JBF, Lotte F, Tangermann M (2012) Brain-computer interfaces: beyond medical applications. *IEEE Comput* 41:26–34
- Vaughan TM, Wolpaw JR, Donchin E (1996) EEG-based communication: prospects and problems. *IEEE Trans Rehabil Eng* 4(4):425–430
- Venables L, Fairclough SH (2009) The influence of performance feedback on goal-setting and mental effort regulation. *Motiv Emot* 33(1):63–74
- Villringer A, Chance B (1997) Non-invasive optical spectroscopy and imaging of human brain function. *Trends Neurosci* 20:435–442
- Whitham EM, Pope KJ, Fitzgibbon SP, Lewis T, Clark CR, Loveless S, Willoughby JO (2007) Scalp electrical recording during paralysis: quantitative evidence that EEG frequencies above 20 Hz are contaminated by EMG. *Clin Neurophysiol* 118(8):1877–1888
- Wilson GF, Russell C (2003) Real-time assessment of mental workload using psychophysiological measures and artificial neural networks. *Hum Factors* 45(4):635–643
- Wilson GF, Russell C (2007) Performance enhancement in an uninhabited air vehicle task using psychophysiological determined adaptive aiding. *Hum Factors* 49(6):1005–1018
- Wu D, Courtney CG, Lance BJ, Narayanan SS, Dawson ME, Oie KS, Parsons TD (2010) Optimal arousal identification and classification for affective computing using physiological signals: virtual reality Stroop task. *IEEE Trans Affect Comput* 1(2):109–118
- Zander TO, Kothe C (2011) Towards passive brain-computer interfaces: applying brain-computer interface technology to human-machine systems in general. *J Neural Eng* 8(2):025005
- Zhao C, Zheng C, Zhao M, Tu Y, Liu J (2011) Multivariate autoregressive models and kernel learning algorithms for classifying driving mental fatigue based on electroencephalographic. *Expert Syst Appl* 38(3):1859–1865
- Zimmerli L, Krewer C, Gassert R, Müller F, Riener R, Lünenburger L (2012) Validation of a mechanism to balance exercise difficulty in robot-assisted upper-extremity rehabilitation after stroke. *J Neuroeng Rehabil* 9:6
- Zimmermann R, Marchal-Crespo L, Edelmann J, Lambercy O, Fluet M-C, Riener R et al (2013) Detection of motor execution using a hybrid fNIRS-biosignal BCI: a feasibility study. *J Neuroeng Rehabil* 10:4