Eye-gaze independent EEG-based brain–computer interfaces for communication

2012 J. Neural Eng. 9 045001

(http://iopscience.iop.org/1741-2552/9/4/045001)

J. Neural Eng. 9 (2012) 045001 (15pp) doi:10.1088/1741-2560/9/4/045001

Eye-gaze independent EEG-based brain–computer interfaces for communication

A Riccio[1,2], D Mattia[1], L Simione[2,3,4], M Olivetti[2,3] and F Cincotti[1,5]

[1] Neuroelectrical Imaging and BCI Lab, Fondazione Santa Lucia IRCCS, Rome, Italy
[2] Deparment of Psychology, University of Rome 'Sapienza', Rome, Italy
[3] Interuniversity Centre for Research on Cognitive Processing in Natural and Artificial Systems (ECONA), Rome, Italy
[4] Institute of Cognitive Science and Technologies, National Research Council, Rome, Italy
[5] Department of Computer, Control and Management Engineering 'Antonio Ruberti', Sapienza University of Rome, Italy

E-mail: a.riccio@hsantalucia.it

Received 15 November 2011
Accepted for publication 12 April 2012
Published 25 July 2012
Online at stacks.iop.org/JNE/9/045001

Abstract

The present review systematically examines the literature reporting gaze independent interaction modalities in non-invasive brain–computer interfaces (BCIs) for communication. BCIs measure signals related to specific brain activity and translate them into device control signals. This technology can be used to provide users with severe motor disability (e.g. late stage amyotrophic lateral sclerosis (ALS); acquired brain injury) with an assistive device that does not rely on muscular contraction. Most of the studies on BCIs explored mental tasks and paradigms using visual modality. Considering that in ALS patients the oculomotor control can deteriorate and also other potential users could have impaired visual function, tactile and auditory modalities have been investigated over the past years to seek alternative BCI systems which are independent from vision. In addition, various attentional mechanisms, such as covert attention and feature-directed attention, have been investigated to develop gaze independent visual-based BCI paradigms. Three areas of research were considered in the present review: (i) auditory BCIs, (ii) tactile BCIs and (iii) independent visual BCIs. Out of a total of 130 search results, 34 articles were selected on the basis of pre-defined exclusion criteria. Thirteen articles dealt with independent visual BCIs, 15 reported on auditory BCIs and the last six on tactile BCIs, respectively. From the review of the available literature, it can be concluded that a crucial point is represented by the trade-off between BCI systems/paradigms with high accuracy and speed, but highly demanding in terms of attention and memory load, and systems requiring lower cognitive effort but with a limited amount of communicable information. These issues should be considered as priorities to be explored in future studies to meet users' requirements in a real-life scenario.

1. Introduction

A brain–computer interface (BCI) is a technology that utilizes brain signals to control external devices (Wolpaw and Wolpaw 2012). Under this condition, a BCI system provides the human body with an alternative artificial channel that can substitute, restore or enhance the natural outputs (i.e. peripheral nerves, muscles) which have been lost because of a disease or injury. Up to now, the first two BCI applications have been the most explored with the aims of allowing/improving the communication and interaction with the environment of people with severe motor disability (Sellers *et al* 2010, Zickler *et al* 2011, Aloise *et al* 2006, 2011, Vanacker *et al* 2007, Tonin *et al* 2011, Neuper *et al* 2009).

The brain signals directly subserving as inputs to BCI systems can be recorded in a variety of electrophysiological and metabolic methods (Hinterberger *et al* 2004, Weiskopf *et al* 2004, Coyle *et al* 2004). Experimental evidence on which recorded signal is the best to measure the user's intent has still to be provided. Although many studies are currently available on the advantages and disadvantages of each recording method (Birbaumer 2006, Mak and Wolpaw 2009), the most extensively investigated and applied BCI systems rely on a non-invasive electrophysiological recording such as the electroencephalogram (EEG). The EEG-based BCIs will be the focus of this review.

1.1. EEG signal features for BCIs

The non-invasive electrical signals are measured by means of the EEG and are by far the most extensively applied in BCI system control. A variety of EEG signals have been used as measures of the brain activity: event-related potentials (ERPs; Farwell and Donchin 1988, Nijboer *et al* 2008a, Sellers and Donchin 2006, Piccione *et al* 2006, Riccio *et al* 2011), frequency oscillations (particularly the EEG sensorimotor rhythms; SMRs) (Wolpaw *et al* 2000, Pfurtscheller *et al* 2000), slow cortical potentials (SCPs) (Birbaumer *et al* 1999, Neumann *et al* 2003) and steady-state responses (SSRs) (Cheng *et al* 2002).

The SCPs are slow (range of seconds) voltage changes recorded over the sensorimotor cortex, phase- and time-locked to specific sensorimotor events (Kübler *et al* 2001). The SCPs typically consist of negative potential shifts that precede actual or imagined movements or other cognitive tasks. Birbaumer *et al* (1999, 2000) have shown that people can learn to perform mental tasks to produce SCP changes and thereby control the movement of an object on a computer screen (Elbert *et al* 1980, Birbaumer *et al* 1999, 2000). This paradigm was at the base of the implementation of the thought translation device (TTD) that has been tested in people with late-stage amyotrophic lateral sclerosis (ALS), proving that it can supply basic communication capability (Kübler *et al* 2001).

The steady-state evoked potentials (SSEPs) are stable oscillations in voltage that can be elicited by rapid repetitive stimulation conveyed through visual, auditory and somatosensory modality. In the steady-state visual evoked potential (SSVEP)-based BCIs, stimuli flickering at different frequencies are visually presented to subjects, who direct their attention to one of the stimuli. The attended stimulus elicits enhanced SSVEP responses at the corresponding frequency, recorded over occipital brain areas (for a review see Vialatte *et al* (2010)). This increase in SSVEP amplitude can be detected at the level of single trial, classified and translated into control commands (Middendorf *et al* 2000, Liu *et al* 2010). Detection of SSPs has been also documented in auditory (auditory steady-state evoked potential (ASSEP); see Plourde 2006 for a review) and somatosensory systems (steady-state somatosensory evoked potentials (SSSEPS) (Namerow *et al* 1974).

The ERPs, embedded within the EEG background activity, are manifestations of neural activity that is triggered by, and involved in the processing of, specific events. The ERP-based BCIs are implemented with an *oddball paradigm*, wherein a rare target (oddball event) is intersected with frequent non-target events. These BCIs usually exploit an endogenous ERP component, known as P300, as input signal. The P300 is a positive deflection that occurs in the scalp recorded EEG approximately 300 ms after the presentation of the rare visual, auditory or somatosensory—task relevant—stimulus (Sutton *et al* 1965). By focusing attention on the rare target (e.g., by keeping a mental count of its occurrence), the P300 amplitude can be increased and therefore its detection and classification improves. For a review about the visual P300-based BCIs see Kleih *et al* (2011).

The BCIs based on sensorimotor rhythms (SMRs) are operated by voluntary modulation of such rhythms recorded over scalp sensorimotor areas, within a frequency range of 8 to 30 Hz (mu and beta band). The SMRs have a long-standing history related to motor behavior (Berger 1930, Jasper and Andrew 1938, Jasper and Penfield 1949). In fact it has been repeatedly shown that execution or imagination of limb movements induces changes in this rhythmic activity. Pfurtscheller and Aranibar (1979) and Pfurtscheller and Neuper (1992) have further elucidated this phenomenon and demonstrated that SMRs decrease and/or increase during motor behavior (event-related desynchronization, ERD, and event-related synchronization, ERS).

All the above-described EEG signals carry on relevant features for BCI control and a number of neural computational methods have been developed and implemented to extract those features and then classify them on a single trial basis, to ultimately translate them into an external command (Lotte *et al* 2007, Krusienski *et al* 2011). From a 'pure' application point of view, many BCI systems for communication have been exploiting features extracted from ERP components; these latter being mainly (but not only) induced by visual stimulation (Mugler *et al* 2010, Kleih *et al* 2011, Kaufmann *et al* 2011).

1.2. Dependent and independent BCIs

In 2002, Wolpaw *et al* (2002) distinguished 'present-day' EEG-based BCIs as dependent and independent systems. Accordingly, independent EEG-based BCIs are those which do not exploit in any way the normal brain output channels (muscles and/or nerves) to generate the EEG activity carrying the relevant information to control the system. As communication devices, such BCIs can use slow cortical potentials, P300-evoked potentials or sensorimotor rhythms. In contrast, dependent BCIs are systems which indirectly depend on muscles and cranial nerve activity to generate the control features. Examples of such BCIs are represented by BCIs based on the visual-evoked potential (VEP) and their stable oscillatory components (i.e. SSVEP) recorded from the scalp over visual cortex, whose amplitude depends on the gaze direction (Vidal *et al* 1973, Sutter 1992, Gao *et al* 2003).

Over 12 years of development and publications in this field, it is important to mention that this categorization of BCI systems still remains to some extent based on the assumption that current systems are neither purely dependent nor purely independent. A clear example of this latter statement is represented by the visual P300-based BCI systems as is evident from recent studies which have challenged their definition as independent BCIs.

In 1988, Farwell and Donchin (1988) were the first who introduced a visual P300-based paradigm as the basis for a BCI (P300 speller). The P300 speller consists of a 6 × 6 symbol matrix wherein letters and symbols are arranged within rows and columns. Throughout the course of a trial, rows and columns are intensified one after the other, in a random order. The user is asked to attend (not necessarily fixate) the desired letter and mentally count the number of times it flashes. This stimulus is the 'rare event' in an oddball paradigm (Fabiani *et al* 1987) and elicits the P300. The computer identifies the attended item as the intersection of the row and column eliciting the largest P300. In this first P300-based BCI, the assumption was that attention can be focused away from the gaze fixation point (i.e. the target letter in this context) according to Posner's theory (Posner 1980).

Since then, the endogenous ERP component P300 has been one of the major features in BCI research and application for communication and control purpose. Many studies were indeed conducted by using the P300 speller (Donchin 2000, Nijboer *et al* 2008a, Kleih *et al* 2010) and many others had addressed the optimization of the presentation paradigm, all involving both able- and not able-bodied subjects (Sellers and Donchin 2006, Piccione *et al* 2006).

Recent studies have disclosed how the performance of P300-based BCIs that use matrix format may depend to some extent on the user's ability to gaze (i.e. to control his own eyes) and furthermore, how some short-latency features (e.g. VEP such as P100, N200) elicited by visual P300-based BCIs may play a role in the classification accuracy achieved by such systems in addition to the P300.

Brunner *et al* (2010) and Treder and Blankertz (2010) have shown how the performance of a P300-based BCI, that uses a speller matrix format, deteriorates when switching from an overt to a covert attention condition. In both studies, the overt attention condition was realized by instructing the participants (healthy volunteers) to gaze at the target item whereas in the covert attention condition the participants had to fixate on a central point while paying attention to the intensification of the target item.

According to the finding of significantly higher performance accuracy when the target was fixated (overt attention; gaze control), the two research groups concluded that the P300-based BCI performance partially depends on the subject's ability to fixate and therefore the system also relies on eye-gaze control. Moreover, they both found that early components of ERPs (such as VEPs occurring within a time interval comprises within 300 ms from the stimulus) may have accounted for the classification accuracy obtained under overt attention condition (target fixation). On the other hand, in a 'pure' covert attention modality (central fixation) performance accuracy was mostly depending on the P300 component of ERPs (occurring after 300 ms from the stimulus).

In conclusion, the P300 was acknowledged without any doubt as informative in overt attention modality (eye gaze) but as less informative than the earlier VEPs measured at occipital and parieto-occipital sites.

This latter finding raises the issue that most of the P300-based BCIs which use a speller matrix format could also elicit early visual components and the accuracy of classification they achieve may depend to some degree on these occipital VEPs, in addition to the P300. Taken as a whole, these BCI systems should be therefore better referred to as ERP-based BCIs (Treder and Blankertz 2010). Also, the evidence that exogenous ERP components which are elicited and modulated by overt attention processes play a role in the real-time performance of such BCIs indicates that they cannot be defined as purely independent BCIs.

Many of the potential BCI users may show several degrees of impairment of their oculomotor control or other visual impairments, like in the case of ALS or acquired brain injury; hence, the development of new independent ERP-based BCI paradigms is a crucial challenge in modern BCI research (Nijboer *et al* 2008b). In addition, severe motor disabilities can often be associated with various cognitive, sensorial and functional peculiarities. This encourages the BCI community to develop a range of BCI paradigms in order to cope with a broad spectrum of user's needs and with the heterogeneity of potential users.

Overall, these theoretical and real considerations bring us to the central core of this review that is illustrated in the next section.

1.3. Scope of the review

The impairment of vision (including the eye gaze control) can virtually affect the performance of any of the EEG-based BCIs that rely on the user's visual abilities. The BCI systems can indeed exploit the visual channel in two different ways: as the modality to convey to the user instant feedback of user and system interaction performance (SMR-based BCIs; SCP-based BCIs) or as the sensory modality through which the user attends the stimulus whose elicited brain responses actually drive the BCI system (i.e. ERP-based BCIs; SSVEP-based BCIs).

Considering this distinction and what has been previously discussed on the possible dependence of visual ERP-based BCI performance on overt attention processes funneled by muscular activity, we will review primarily those studies that address purely eye-gaze independent EEG-based BCIs, namely those operated by means of non-visual paradigm. As a second body of the review, studies regarding visual EEG-BCI systems developed specifically to overcome the impasse of gaze dependence will be discussed. Only BCI systems developed with the aim to restore the communication in severely disable people and requiring an external sensory stimulation will be considered.

The present review is therefore structured as follows: (i) first, the auditory BCI paradigms for communication will be systematically reviewed, the auditory modality being purely independent; (ii) for the same motivation, somatosensory (tactile) BCIs will be shortly treated in a separate session. Finally, studies investigating the non-gaze dependent visual BCIs will be reviewed.

For each domain, the essential elements of the tasks required to control the independent BCIs for communication, the accuracy, defined as the ratio of the number of correct selections over the number of total selections, and when available the information transfer rate (ITR) will be discussed. Table 1 summarizes the performances in terms of accuracy, bits per symbol (Wolpaw formula; Wolpaw *et al* 2002), number of symbols per minute and bits per minute (ITR) reported in each considered study.

2. Methods

2.1. Search

The searching and selection of papers was performed through PubMed, PsychINFO and IEEE databases. This procedure was also integrated with cross-checking references cited in relevant and recent papers published in the issue.

Three different semantic areas were at the base of search, depending on the sensory modality employed in the different BCI studies:

- First area: auditory BCIs.
  The relevant papers were collected searching with the following keywords: [(Brain–Computer Interface OR Brain Machine Interface) AND (Auditory)]. The first run of the search identified 96 articles. This first collection was then filtered by reading the articles' titles and abstracts, so that all articles not matching the scope of the present review were excluded from further discussion. This procedure of selection provided us with 25 articles (out of 96). We further excluded, by reading the full articles of the database, one conference paper reporting the preliminary results of a study successively presented as a journal article (Höhne *et al* 2011) and nine theoretical studies which did not have communication as a target BCI application.
- Second area: tactile BCIs.
  The keywords [(Brain–Computer Interface OR Brain Machine Interface) AND (Tactile)] were used to search articles in this area. Twelve articles were identified and five were excluded, based on the same procedure as described above. Among the remaining six articles, one was excluded because it was a conference proceeding, reporting preliminary findings which were then reported in a full-length paper (Cincotti *et al* 2007).
- Third area: gaze independent visual BCI.
  Articles were collected if they contained in their title, abstract or keywords the following terms [(Brain–Computer Interface OR Brain Machine Interface) AND (Covert Attention OR Gaze Independent)].
  According to the criteria and procedures described in and used for the previous two researches, 15 articles were finally selected (out of a total of 28 articles appearing online). The complete reading of these remaining 15 papers yielded to a further exclusion of two conference papers which presented preliminary data successively included in full-length papers (Treder *et al* 2011b, Zhang *et al* 2010).

3. Results

- First area: *auditory BCIs*.
  The final database included 15 papers published from 2004 and February 2012.
  Among the selected articles, 12 dealt with ERP-based BCI for communication whereas 1 reported about a BCI system based on auditory steady-state responses (ASSRs). In addition, an auditory feedback was applied in a BCI operated via a motor imagery task and in a BCI based on the modulation of slow cortical potentials in two articles. These latter studies are not discussed in depth being beyond the scope of the present review which is focused on sensory stimulation-driven BCIs. All the selected papers were journal papers or full-length peer-reviewed conference papers.
- Second area: *tactile BCIs*.
  Six papers were selected after the exclusion criteria procedure. They were published between 2004 and September 2010. Three investigated the feasibility of a BCI based on steady-state somatosensory-evoked potentials (SSSEPs). One study was about a somatosensory ERP-based BCI. Finally in two studies the tactile modality was used as a BCI feedback channel for sensorimotor tasks. These latter studies are not extensively discussed being sensorimotor-based BCIs beyond the scope of the present review. All the selected papers were journal papers or full-length peer-reviewed conference papers.
- Third area: *gaze independent visual BCIs*.
  After applying the above-mentioned exclusion criteria, the final database included 13 papers published from 2004 and December 2011. Among those articles, six dealt with ERP-based BCIs for communication and four reported on SSVEP-based BCI systems. Finally three studies investigated modulation of EEG alpha band oscillation by covert visual attention: these studies are reported but not fully discussed being beyond the scope of the present review. All the selected papers were journal papers or full-length peer-reviewed conference papers.

3.1. Auditory BCIs

3.1.1. Methodological studies on auditory BCIs. Up to now, the majority of studies on auditory BCIs have exploited the ERPs as a brain control signal. To operate an auditory ERP-based BCI system, participants are presented with a series of acoustic stimuli (events) each of which falls into two classes: target stimuli (or a stream of stimuli) and non-target stimuli. The subject is asked to concentrate on a target stimulus. Both target and non-target acoustic stimuli can be represented by tones, environmental sounds or words.

Hill *et al* (2004) developed a paradigm for an auditory BCI, starting from the assumption that ERPs can be modulated by the selective attention of a subject who is focusing on one of two sequences of auditory stimuli dichotically presented. In order to allow a binary choice, the authors asynchronously presented the subjects with two different auditory sequences

Table 1. Comparison of the accuracy, bits per symbol, symbols per minute and bits per minute of the BCIs reviewed in this paper.

Study	Modality	Classes	Accuracy (%)	Bits per symbol	Symbols per minute	Bits per minute (ITR)	Analysis	Population studied	Number of subjects	Comments
Hill *et al* (2004)	Auditory	2	72.6	0.15	9.23	1.41	Offline*	Healthy	15	*Analysis performed without ICA and per-channel normalization
			82.4	0.32		3.03	Offline **			**With ICA and per-channel normalization performed
Hill and Schölkopf (2012)	Auditory	2	84.8	0.36	12	4.38*	Online	Healthy	13	*In the study is reported an ITR of $\mathbf{4.98 \pm 2.3\ bits/min^{-1}}$ computed for each subject separately before averaging
Halder *et al* (2010)	Auditory	2	78.5	0.25	7.93	1.98	Offline *	Healthy	20	The accuracy was calculated using the task that was optimal for each subject
			93.3	0.64	0.63	0.41	Offline **			*2 sequences of stimuli were averaged for classification; the ITR reported in the study is of $\mathbf{2.46\ bits/min^{-1}}$ obtained computing ITR for each subject separately before averaging **25 sequences averaged for classification; $\mathbf{0.39\ bits/min^{-1}}$ reported, computing ITR for each subject separately before averaging
Guo *et al* (2010)	Auditory	8	85.6	2.00	4.2	8.4	Offline*	Healthy	14	*Classification was performed averaging 5 sequences of stimuli; authors reported $\mathbf{8.77\ bits/min^{-1}}$ computing ITR for each subject separately before averaging
			90.2	2.26	1.4	3.16	Offline**			**Classification was performed averaging 15 sequences of stimuli; authors obtained $\mathbf{3.43\ bits/min^{-1}}$ computing ITR for each subject separately before averaging
Kim *et al* (2011)	Auditory	2	86.3	0.42	4.28**	1.78	Offline	Healthy	6	The accuracy is calculated averaging the best score of each subject **Obtained averaging the 'analysis window size' resulting in the highest accuracy for each subject
Seller and Donchin (2006)	Auditory	4	65.2*	0.52	0.86–3.46	0.43–1.80**	Offline	Healthy	3	*Obtained averaging the maximum accuracy for each subject with different methods of classification; the authors employed a bootstrapping approach to estimate the smallest number of trials that would yield specified detection accuracies
		4	70.6*	0.64	0.67–2.81	0.43–1.80**	Offline	End-users (ALS)	3	**The ITR reported by the authors is not differentiated between healthy subjects and end users

Table 1. (Continued.)

Study	Modality	Classes	Accuracy (%)	Bits per symbol	Symbols per minute	Bits per minute (ITR)	Analysis	Population studied	Number of subjects	Comments
Furdea *et al* (2009)	Auditory	25	65.0	2.1	0.61	1.28*	Online	Healthy	15	*The value does not correspond with the ITR reported by Furdea and colleagues. The authors calculated the ITR with a different metric. **The number of symbols per minute includes the ITI (3.75 s)**
Klobassa *et al* (2009)	Auditory	36	59.3	2.11	*	1.86**	Online	Healthy	5	*Depending on offline classification accuracy the number of *stimulus presentation sequences* was reduced in the online condition. For this reason we were not able to report the number of symbols selected per minute
	Auditory + visual	36	68.7	2.67	*					**Value reported by the authors, the condition is not specified
Schreuder *et al* (2010)	Auditory	5	94.0*	1.87	5.91**	11.08	Offline	Healthy	5	The accuracy is calculated using the maximum selection score reached for each subject
			93.6**	1.85	3.07****	5.69				*Stimulus onset asynchrony (SOA): 175 ms ** SOA: 300 ms Value obtained using ***11.6 trials and ****13 trials on average for classification
Schreuder *et al* (2011)	Auditory	6	77.4*	1.42	1.89	2.72	Online	Healthy	21	*First session: fixed number of iterations
		6	84.3**	1.64	2.80	4.68			15	**Second session: only the subjects who performed well during the first session were involved. Dynamic stopping included
		6	86.1***	1.74	2.87	5.13			15	***Second session: only the subjects who performed well during the first session were involved. No visual support. **The metric now includes the inter trial interval and the feedback time. The ITR was computed for each subject separately before averaging**.
Höhne *et al* (2011)	Auditory	9	89.4	2.34	1.97	4.61*	Online	Healthy	20	The ITR reported by the authors (3.4 bits min^{-1}) includes the ITI and all pauses taken by the subjects
Kübler *et al* (2009)	Auditory	25	12.1	0.08	0.66	0.05	Online	End users (ALS)	4	**The number of symbols per minute includes the ITI (1.87 s)**

Table 1. (Continued.)

Study	Modality	Classes	Accuracy (%)	Bits per symbol	Symbols per minute	Bits per minute (ITR)	Analysis	Population studied	Number of subjects	Comments
Müller-Putz *et al* (2006)	Tactile	2	68.0*	0.09	18.75	1.79	Online	Healthy	4	*Obtained averaging the accuracy achieved by each subject in the best session
Zhang *et al* (2007)	Tactile	2	63.00*	0.05	12	0.6	Offline	Healthy	8	*SSSEP
		2	83.00**	0.34		4.11			8	**SSSEP+SSVEP+ mu-band power
		3	61.00***	0.23		2.76			8	***SSSEP+SSVEP+ mu-band power
Brouwer and van Erp (2010)	Tactile	6	*		2.72	3.71**	Online	Healthy	11	*Performances are not clearly reported **Best condition; **the ITR was computed for each subject separately before averaging**
Kelly *et al* (2005)	Visual	2	70.3	0.12	7.5	0.91	Offline	Healthy	10	
Zhang *et al* (2010)	Visual	2	72.6	0.14	15	2.17	Online	Healthy	18	
Treder and Blankertz (2010)	Visual	36	60	2.15	2	4.29	Offline	Healthy	13	
Treder *et al* (2011)	Visual	36	91.3*	4.29	2	8.59	Online	Healthy	13	*Hex-o-Spell condition
			88.2**	4.04	2	8.08				**Cake Speller condition
			97.1***	4.83	2	9.66				***Center Speller condition
Liu *et al* (2010)	Visual	36	94.4	4.57*	1.25	5.71	Online	Healthy	8	*Random condition
			96.3	4.72**	1.25	6.02				**Searching condition
Aloise *et al* (2011)	Visual	36	77.82	3.34	2.5	8.35*	Online	Healthy	10	***The ITR was computed for each subject separately before averaging**
Acqualagna *et al* (2010)	Visual	30	90	3.9*	1.23	4.8	Offline	Healthy	9	*Best condition (SOA = 133 ms; 10 sequences)
Acqualagna and Blankertz (2011)	Visual	30	94.0*	4.29	1.2	5.14	Online	Healthy	12	*Condition no-color with SOA = 116 ms
			94.7**	4.35	1.2	5.22				**Condition color with SOA = 116 ms
			93.6***	4.25	2.4	10.20				***Condition color with SOA = 83 ms

Modality refers to the sensory modality through which the user attends the stimulus*; classes* refers to the number of possible targets; *accuracy* is the number of correct selections divided by the total amount of selections; *bits per symbol* is calculated with the Wolpaw *et al* (2002) formula; *symbols per minute* refers to the number of selections performed in 1 min; *bits per minute (ITR)* is obtained by multiplying the *bits per symbol* and the *symbols per minute; analysis* refers to whether the data analysis is online or offline; *population studied* column reports whether the subjects involved in the study were healthy subjects or end-users; *number of subjects* refers to the number of subjects involved in the study.
Data not explicitly reported in the reviewed studies were calculated by the authors.
In order to provide comparable values the employed metrics do not take into account the inter stimulus interval (ITI).
For the same reason we reported the *bits per minute* calculated with the average accuracy of the subjects.
When differently computed, procedures have been specified in the comments.

of stimuli, one in each ear, and the subject was required to direct his attention to one of the presented auditory streams by counting the number of target beeps contained in it. The offline estimation of classification accuracy in 15 healthy participants showed a high inter-subject variability and it was on average 82.6% after applying the independent component analysis (ICA) followed by per-channel normalization of data.

In a more recent study, Hill and Schölkopf (2012) reported on the assessment of the online performance of such acoustic streaming approach. The acoustic ERP-based BCI was tested online with 13 healthy participants and an online accuracy of 84.8% on average was reported. The bits/min^{-1} was 4.98, computing ITR for each subject separately before averaging and without considering the inter-trial interval duration. In table 1 we report an ITR calculated by averaging the accuracy across all the experimental subjects, in order to allow the comparison with other studies, where the accuracy of the single subject is not reported.

Kanoh *et al* (2010) investigated the auditory stream segregation phenomenon as a possible paradigm to develop a binary choice BCI system. In their study, subjects were presented with two auditory *oddball* sequences to the right ear, set so that the presented tones were perceived by subjects as two segregated auditory streams. Subjects were required to pay attention to one of the two streams. The methodology adopted for the offline estimation of the accuracy was not clearly reported by Kanoh *et al* (2010). Accordingly, as also pointed out in Hill and Schölkopf (2012), it can be interpreted as the performance was calculated using the same data used to train the classifier. Therefore the performances were not reported in table 1.

Halder *et al* (2010) proposed an auditory three stimuli paradigm to operate a BCI for communication purposes (binary selection). The authors enrolled 20 healthy subjects and presented them with three different tasks, related to three stimuli *oddball* sequences with two targets and a standard tone. The two targets differed in one physical property between the three tasks: pitch, loudness or direction of tones. The participants had to focus on the predefined target and ignore the other stimuli. The offline analysis showed that the best accuracy (i.e. the best discrimination between target and non-target) was obtained with different tasks in different subjects: 14 subjects obtained the best performance in the task with the target and non-target differing for pitch; 6 subjects best performed in the direction task and the last 2 in the task where target and non-target differed for amplitude (loudness). The authors concluded that an individual preliminary screening session to determine the optimal task is recommended to provide each potential user with the optimal solution. The mean of the best accuracy scores (considering the optimal task) achieved offline for each volunteer was 78.54% with 2.46 bits/min^{-1}; this latter feature is calculated separately for each subject without considering the inter-trial interval duration. In table 1 we also reported the ITR calculated with the accuracy averaged across the 20 subjects.

Guo *et al* (2010) proposed an auditory ERP-based device presenting eight spoken digits (1–8) as stimuli. In order to select the target the subject had to focus on the desired number discriminating two dimensions of the stimuli: the location (left–right) and the gender of the voice (male–female). The offline classification accuracy obtained relying on the N200 and the P300 components recorded from 14 healthy subjects, was 85% on average, with an ITR of 8.77 bit/min^{-1} (average of individual ITR). As for the papers mentioned above, we reported in table 1 the ITR calculated with the accuracy averaged across the subjects in two conditions: when (i) 15 trials and (ii) 5 trials were averaged for the offline analysis. A trial is considered a sequence of eight stimuli.

The possibility of utilizing the modulation of the ASSEP to control a BCI has been discussed by Lopez *et al* (2009) and recently demonstrated as feasible by Kim *et al* (2011). In this latter study, the subjects were presented with two pure tone burst trains with different beat frequency, to the left and to the right sound field. The offline classification analysis, performed in six healthy volunteers, revealed different accuracies for each subject depending on the analysis window size, which varied from 2 to 20 s after the auditory stimulus onset. More specifically they reported that the best accuracy, estimated at different time windows for each subject, ranged from 80% to 92% with an average of 86.33% (mean of the analysis window size with the best accuracy was 14 s). The authors also reported an accuracy of 71.4% achieved by one participant during a pilot online test.

Other attempts to implement visual modality independent BCIs have used the auditory channel as feedback modality to inform subjects on their instant performance in modulating the SMRs (Nijboer *et al* 2008b) or SCPs (Pham *et al* 2005). These studies will not however be discussed and performances were not reported in table 1, being beyond the scope of the present review.

Overall these studies have established the feasibility of auditory stimuli eliciting ERPs or modulating the ASSEP to be used in gaze-independent BCIs. They indeed represent an important theoretical basis for the BCI paradigms described in the next session and should encourage their further development.

3.1.2. Auditory spellers. One of the first studies exploring the auditory modality in an ERP-based BCI system implemented for basic communication utilized a set of words such as ‘Yes’, ‘No’, ‘Stop’, ‘Pass’ as auditory stimuli; YES and NO were the two possible targets (Sellers and Donchin 2006). The authors compared the auditory, visual and auditory plus visual stimulation modalities across ten experimental sessions. The system was not linked to a speller device but subjects were asked to affirmatively and negatively reply to some questions by focusing their attention on the ‘Yes’ or ‘No’ stimuli. The three different stimulation modalities were investigated in three healthy and three ALS participants. An offline accuracy of 65.2% on average was achieved averaging the maximum accuracy obtained by each healthy subject in the auditory modality. A comparable accuracy of 70.6% was achieved averaging the best accuracies of the three ALS participants. The authors reported an ITR varying between 0.43 and 1.80 bits/min^{-1}. Unfortunately we were unable to calculate separately ITR for healthy and ALS users due to a lack of reported information (see table 1).

In an effort to improve the bitrate, Furdea *et al* (2009) used words acoustically presented that map onto a visual matrix. In this paradigm, numbers were assigned to identify rows and columns of a 5 × 5 speller matrix and such numbers were presented for the selection (i.e. the spoken words were 'one' to 'ten'). In order to select a letter the users were asked to attend to the two numbers representing the coordinates of the target character in the matrix. The authors compared the auditory and visual modality in 15 healthy volunteers who achieved an online accuracy of 65% in the auditory modality and 94.6% in the visual one. The ITR was 1.54 bits/min^{-1} and 6.8 bits/min^{-1} when including the interval between the characters' presentation, respectively. Subsequently such a device was tested by four ALS patients (Kübler *et al* 2009) who reached an ITR of 0.05 and were indeed not able to control the device.

In a similar design, Klobassa *et al* (2009) coded different environmental sounds within a 6 × 6 speller matrix. In this study, five healthy subjects (out of ten) were challenged (across 11 sessions) with the auditory modality speller alone, whereas five tested the speller initially with the combined auditory and visual modalities and successively the visual stimuli were gradually withdrawn until only the auditory stimuli remained. The online accuracies achieved in the two conditions (auditory only versus auditory plus visual) were comparable between the first and the last session. The mean accuracies achieved in the last session, respectively in the first and second conditions, were 59.38% (SD ± 22.1) and 68.78% (SD ± 13.9) and the authors reported higher accuracies in simultaneous auditory and visual stimulus presentation than in auditory stimulus presentation. Similar findings were also reported in another study by Belitski *et al* (2011).

Other speller-like auditory BCI paradigms exploited the spatial features of auditory stimuli to improve classification performance. A multi-class auditory paradigm using spatially distributed cues (AMUSE) was first proposed and tested offline (Schreuder *et al* 2010) and then online as reported in a more recent study (Schreuder *et al* 2011). In the final version of the system (Schreuder *et al* 2011) the BCI drove a speller interface using six spatial locations. The subjects were surrounded by six speakers arranged in a circle at ear height; each speaker presented always the same tone stimulus. The spelling of a letter needed two steps: the first step allowed the selection of a group of letters, each location corresponding to a group of characters; in the second step, the letters of the selected group were divided over the locations and assigned to a speaker and its corresponding tone. The subjects were asked to focus on the stimulus coming from the speaker which corresponded to the desired group of letters first and then to the desired letter. Such spatial organization of stimuli was inspired by the Hex-o-Spell, a speller device implemented in a BCI system operated via the sensorimotor rhythm modulation (Blankertz *et al* 2007, Williamson *et al* 2009; see also the visual paragraph, Treder and Blankertz 2010, 2011b). Sixteen out of the 21 healthy subjects who tested AMUSE were able to write a full German sentence during the first BCI session and 14 out of 15 subjects wrote two full German sentences in the second session. The average accuracy was of 77.4%, 84.3% and 86.1% for the first, the second and the third written sentences respectively. Moreover, as for the spelling of the last two sentences a dynamic stopping was introduced and the visually presented support labels (Hex-o-Spell) were switched off while writing the last sentence. Therefore, such a speller could be used without any visual support and the vocalization of the labels was also omitted in order to improve the interface efficiency. The average of the ITR for the first session was 2.84 bits/min^{-1}; this value increased to 5.26 bits/min^{-1} during the second session, and without considering the subjects unable to successfully write in the first session. In the same way, the writing efficiency (i.e. characters per minute) varied from an average of 0.62 char/min for the spelling of the first sentence to an average of 0.91 and 0.94 char/min for the second and the third sentence spelling, respectively. The ITR reported by the authors is calculated with the method proposed by Schlögl *et al* (2007) and it is not comparable with the ITR reported in the studies previously cited. However for allowing a comparison with them we reported in table 1 the ITR calculated with the Wolpaw formula (Wolpaw *et al* 2002).

The same research group (Höhne *et al* 2011) developed a multiclass P300-based BCI paradigm where the auditory stimuli were presented via headphones and varied in two dimensions: pitch—three tones differing in hertz and thus resulting in high, medium, and low tones—and location—the tones were presented either to the left ear, or to the right or to both, thus resulting in left, right and middle location. Differently from Schreuder *et al* (2010, 2011) the information coded by the two dimensions was independent and not redundant. Accordingly, it resulted in nine different classes of three different tones for each of the three locations embedded in a modified T9 predictive text system for mobile phones (PASS2D speller). To test such a speller device, 20 healthy subjects were asked to spell two German sentences focusing on the target stimuli representing the desired selection and ignoring the non-target ones. They had a visual support showing a screen with the representation of the 3 × 3 matrix with the nine numbers corresponding to the nine stimuli. A visual support was also used for reporting the selections made and the words available with the T9 system. The online spelling accuracy was 89.37% on average, with an ITR of 3.4 bits/min^{-1} (0.89 characters per minute). This latter also included all pauses taken by the subjects for relaxation during the spelling of the sentences. The authors also investigated the possible influence of the similarities in the stimulus dimensions (pitch and direction) on the accuracy value, and pointed out how the pitch dimension was more discriminative than the direction dimension. This observation is in line with the findings in Halder *et al* (2010).

The only two studies using auditory ERP-based BCIs conducted on potential end-users reported different results (Sellers and Donchin 2006, Kübler *et al* 2009).

Sellers and Donchin (2006) tested the four-choice auditory-based BCI with the participation of three ALS end-users (see also the beginning of this section) who displayed, in the best condition, the offline classification accuracy of 79.6%, 73.2% and 59.1% over ten recording sessions.

Kübler *et al* (2009) tested the same 5 × 5 auditory matrix presented in Furdea *et al* (2009) with four ALS end users in one experimental session. The achieved online accuracies were 25%, 0%, 0% and 23.53% for each participant who 'performed' offline with classification accuracies of 58.33%, 25%, 25% and 41.17%, respectively. These performance values are still relatively low next to those previously achieved by the same patients using a visual P300-based BCI (Nijboer *et al* 2008a). Generally, the performance values of ALS participants testing the multiclass BCI (Kübler *et al* 2009; 25 classes) in comparison with the four classes device presented by Sellers and Donchin (2006) show a lower ITR (0.05 bits/min^{-1}), probably due to the working memory load required in the multiclass task.

Nonetheless the problem for some people who need basic communication supported by ERPs-based BCIs and cannot rely on the visual channel definitely encourages further studies with auditory BCIs.

3.2. Tactile BCIs

The somatosensory system may be another alternative pathway through which stimuli can be delivered to operate a BCI system for communication. This alternative channel may be an option in the case of people suffering from ALS or other neurological disorders that also impair vision.

Müller-Putz *et al* (2006) investigated the usability of the steady-state somatosensory evoked potential (SSSEP) paradigm in which a vibratory (tactile) stimulation is used to operate a binary choice BCI. Four subjects tested the system participating in five sessions over five different days: each subject was stimulated with two specific stimulation frequencies simultaneously delivered on the right and on the left index finger. Before the BCI experiments a frequency screening was indeed performed, in order to identify the subject-specific 'resonance-like' frequency. The BCI task consisted in focusing attention either to the left or right finger tips. Two subjects were not able to focus on the target stimulation during entire online sessions. The two remaining subjects who were able to concentrate on the target stimulation during the overall sessions displayed an accuracy of 71.7% and 83.1% (best values), respectively. The average of the best accuracy values across subjects was 68.0, with an ITR of 1.79 bits/min^{-1}, calculated taking into account only the stimulation time and without considering the inter trials interval besides the duration of cue presentation and feedback (see table 1).

Zhang *et al* (2007) presented an initial work using tactile spatial attention paradigms with two or three targets to control a BCI system. They showed that the amplitude of SSSEP could be modulated by switching attention between two stimulations delivered either with the same modality (tactile) and/or with different modalities (tactile and visual). Tactile stimuli were delivered by means of two Braille elements attached to the distal segments of both index fingers of eight healthy subjects. At the same time, visual stimuli consisting of five capital letters were presented at the center of an LCD monitor. Different features were considered for the estimation of classification accuracy and the results were: (i) 63% of accuracy when considering only the SSSEP feature to distinguish left and right tactile stimuli; (ii) 83.2% of accuracy when SSSEP, SSVEP and mu-band power (identified over peri-central cortex) features were considered to discriminate tactile versus visual class; and (iii) 61.7% of accuracy by considering all three features to distinguish three classes, two in the tactile modality and one in the visual modality. To allow a comparison with the other BCIs the performances in the three conditions are reported in table 1.

Recently Brouwer and van Erp (2010) evaluated the feasibility of a tactile P300-based BCI in an online study carried out with 11 healthy volunteers. The tactile stimuli were delivered by means of several devices (tactors) placed around the waist. The authors performed a series of experiments varying the number of tactors and the stimulus onset asynchrony (SOA, defined as the time interval between the onset of two different stimuli). The achieved online accuracy was above chance level (50% in the condition with two tactors; 25% in the condition with four tactors; 17% in the conditions with six tactors) in all the experimental conditions, ranging from an average of 58% for six tactors to an average of 73% for two tactors (table 1 reports the performance value of the best condition).

The feasibility of the tactile BCIs was also demonstrated in studies where the vibrotactile stimulation was employed to convey the instant feedback of the user's performance in a SMR-based BCI operated via motor imagery tasks (Cincotti *et al* 2007, Chatterjee *et al* 2007).

Overall these preliminary pieces of evidence are in favor of the usage of the somatosensory pathways to funnel information relevant for a (ERP-based) BCI paradigm. More systematic studies are required to establish the real contribution of this stimulation (or even feedback) modality in the BCI technology development and application.

3.3. Gaze-independent visual BCIs

3.3.1. Methodological studies on gaze-independent visual BCI. Covert attention can modulate SSVEP amplitude (Kelly *et al* 2005a, Zhang *et al* 2010). How such modulation could be used to control a BCI system without the need for gaze control is an intriguing issue recently receiving substantial attention.

In a pioneer study (Kelly *et al* 2004) it was shown that the transition from overt to covert attention in a SSVEP-based paradigm, allowing a binary decision, resulted in a reduction of classification accuracy by about 20% on average. In light of this performance decrease, the same authors redesigned the paradigm modifying the bilaterally displayed stimuli (visual angle) and obtained an average of binary accuracy of 70.3% (Kelly *et al* 2005a). Accordingly, only the data of Kelly *et al* 2005a are therefore reported in table 1.

Allison *et al* (2008) investigated the hypothesis that the superimposition of visual stimulation patterns could evoke classifiable changes in SSVEP. They presented the subjects with two images each oscillating at a different frequency. The oscillating images could be presented either in a superimposed or separated condition, in order to explore the role of gaze function on system performance. In half of the 14 enrolled healthy subjects the overlaid condition induced differences in SSVEP activity elicited by the visual stimulation patterns that were robust enough to predict an online BCI control. The authors demonstrated that such SSVEP differences depend on the selective attention paid to one of two superimposed stimulation patterns. However, the authors simulated the online control using the R^2 and did not clearly report accuracy values.

Zhang *et al* (2010) proposed a covert non-spatial visual selective attention paradigm to operate a SSVEP-based BCI. Two sets of dots with different colors and flickering frequencies were used to induce the perception of two superimposed transparent surfaces. A group of 18 healthy subjects was asked to selectively attend to one of the two surfaces in order to control the BCI system to perform a binary decision task during a three day training program. An average of accuracy of 72.6% was achieved in the last training session. As reported in table 1 the system would allow 15 selections per minute, without considering the inter trial intervals in the analysis, with an ITR of 2.17 bits/min^{-1}.

Modulation of the EEG alpha band oscillation by covert visual attention has also been investigated as a mechanism to operate a BCI (Kelly *et al* 2005b, Van Gerven *et al* 2009, Treder *et al* 2011a). In these studies, it was shown that covert attention-based paradigms could induce changes in the alpha frequency oscillations both in the presence and even in the absence of delivered flickering stimuli. These studies will not however be discussed and performances are not be reported in table 1, being beyond the scope of the present review, which is focused on sensory stimulation-driven BCIs.

As reported in table 1 the bits per minute achieved in the studies discussed in the present session remain below 2.17 bits/min^{-1} (Zhang *et al* 2010). Nevertheless the low cognitive effort required to perform the 'BCI' tasks could reasonably make these BCIs appropriated for users with reduced cognitive capabilities.

3.3.2. Gaze-independent visual spellers. Treder and Blankertz (2010) proposed an ERP-based BCI with a speller interface inspired by the Hex-o-Spell (Blankertz *et al* 2007) to be used in a covert attention modality. The interface was arranged in such a way that six vertices of a hexagon corresponded to six circles, each of which contained five letters of the alphabet and was randomly highlighted. When a given circle was chosen, the five contained letters were expanded into the six circles, with the sixth one intended as a backdoor for returning to the group level in case the wrong group was selected. Under this modality of stimulation, an offline classification accuracy up to 60% was obtained.

In a successive study, Treder *et al* (2011b) proposed three alternatives of the previously presented two stages speller. Such spellers, based on covert attention and non-spatial feature attention modalities, were tested on 13 healthy volunteers. In the first variant, the stimulation consisted in intensifying the Hex-o-Spell circles by flashing them with colors. In the second variant, the 'Cake Speller', triangular shapes were substituted for the previous circles and intensified by flashing them in a single color. For selecting letters in the latter two interfaces participants used covert spatial attention and color attention to focus on the target. In the 'Center Speller', the third variant, the elements were presented as single geometric shapes each having a unique color, one after the other in the center of the screen. Each element corresponded to a group of letters first and then to one letter belonging to the selected group.

Accuracy was higher for the 'Center Speller' with respect to the other two spellers: the average accuracies for Hexo-Spell, Cake Speller and Center Speller were 90.4%, 88.0% and 97.0%, respectively. The spelling speed reported by the authors was on average about 2 char/min with an ITR of 9.8 bits/min^{-1}.

Another visual interface with unusual organization of the stimuli and feedback was described in Liu *et al* (2011). The authors adapted the classical visual speller in a new covert attention-based paradigm, with characters clustered and presented around the gaze point in a small near-central visual field. In this paradigm the active visual search was the cognitive process used to elicit a P300 component. The subjects had to decide where the target letter was presented. The authors assumed that the covert searching and localization of the attended target would have elicited classifiable features and investigated two conditions: (i) a *random position (RP)* where the subjects could not anticipate the position of the target and (ii) a *fixed position (FP)* where each character could only appear at one or two specific positions. For the eight healthy participants who tested the system, the online average accuracy was 94.4% for the RP and 96.3% for the FP. The authors reported the written symbol rate (WSR, symbols/min; Furdea *et al* 2009) with a pick at 1.38 symbols/min, which correspond to 5.71 bits/min^{-1} for the RP and 6.02 bits/min^{-1} for the FP calculated with the metric adopted in table 1.

A similar approach was adopted by Aloise *et al* (2012) who tested the device (geospell) usability in terms of effectiveness (*accuracy*), efficacy (*workload*) and satisfaction. The geospell was tested by ten healthy subjects who reached an accuracy of 77.82% and an ITR of 8.35 bits/min^{-1}, higher than the one achieved in the Liu *et al* (2011) paper, because of a shorter SOA, but still lower than the ITR obtained with the testing of the Center Speller in Treder *et al* (2011). Moreover, while driving the geospell in covert attention the subjects perceived a workload not significantly different to that perceived when they operated the p300 speller (Farwell and Donchin 1988) in overt attention.

Acqualagna *et al* (2010) presented a paradigm based on a central rapid serial visual presentation (RSVP speller) of stimuli. The novel BCI paradigm was tested by nine healthy volunteers who, in an offline study, were presented with fast bursts of stimuli presented successively at a single central location. The study investigated two different time (SOA 83 ms and 133 ms) and color (black letters and letters divided in three color groups) conditions. The mean classification performance was up to 90% with ten repetitions and the best condition was the no-color, 133 ms condition.

In a successive online study, Acqualagna and Blankertz (2011) reported a mean online classification accuracy in 12 healthy subjects operating the RSVP speller under three conditions. They reached an accuracy of 93.6% in the best condition with colored stimuli and an SOA of 83 ms with an ITR of 10.20 bits/min^{-1} (Wolpaw formula, see table 1).

4. Conclusion

A BCI is a device that allows a direct interaction between the brain and external applications, bypassing peripheral output (i.e. nerves and muscles). The general distinction between dependent and independent BCI systems, i.e. BCIs which do or do not exploit normal brain output channels to generate EEG control signals, still remains a debatable issue. In fact, recent studies have clearly demonstrated that the performance of visual P300-based BCIs may also rely on early components of visual ERPs (Brunner *et al* 2010, Treder and Blankertz 2010, Bianchi *et al* 2010). As a consequence, most of the visual ERP-based BCIs, tested over the last decade with healthy users and potential end-users, should be regarded as eye-gaze dependent BCIs.

Recently many research groups have deployed and tested a range of BCI systems, independent from eye-gaze with communication purposes. As such these BCIs have been expressly designed for people in clinical conditions like complete locked-in, or other neurological conditions with sensorial (mainly visual) impairments. In the present review, BCI studies published between 2004 and 2012, aiming to provide communication devices independent from eye-gaze and based on external sensorial stimulation (auditory, tactile and visual), were discussed. Accordingly, we first introduced the studies on auditory and tactile BCI systems and paradigms (first and second area in section 3), and second the visual BCIs relying on a range of covert and overt attentional paradigms were reviewed (third area in section 3). Two of the three areas—auditory BCIs and gaze independent, visual BCIs—were further organized into two subsections dealing with theoretical studies and studies focused on speller device applications, respectively.

The overall considered studies were summarized in table 1, where we specifically addressed their reported performances in term of accuracy, bits per symbol (Wolpaw formula; Wolpaw *et al* 2002), number of symbols per minute and bits per minute (ITR).

To date, in spite of the growing number of publications a systematic approach to gaze-independent BCIs is still lacking, as there exist many different BCIs using different paradigms. Likewise, the comparison between the available independent BCI paradigm in terms of system performances, resources and capability required to use BCIs is limited (if not possible at all) due to the lack of common, standard metrics needed for a consistent comparison between BCI systems (Quitadamo *et al* 2009). As the various BCI designs are moving toward clinical use, it has become increasingly important to reach a general consensus on how they should be compared. Nevertheless, the current available findings on gaze-independent BCIs are without any doubt promising. The speller devices developed during the last two years, both in the visual area (Treder *et al* 2011b, Acqualagna and Blankertz 2011, Aloise *et al* 2012) and in the auditory research area (Höhne *et al* 2011, Schreuder *et al* 2011) have considerable potentiality, as indicated by the reported accuracies and ITR obtained in healthy volunteers.

The availability of studies testing such independent BCI approaches exclusively on healthy users brings about another relevant issue. As a counterpart of the reported relatively high ITR, it stands out to the level of subject's attentional and other cognitive (i.e. working memory) demands required to accomplish the related BCI tasks.

In this regard, it is noteworthy that up to 50% of patients with ALS has clinical evidence of a cognitive impairment as revealed by neuropsychological testing (Woolley *et al* 2010). Several studies have also shown that the primary non-motor deficit in ALS occurs in the domain of attention and cognitive flexibility (Massman *et al* 1996, Frank *et al* 1997). Many other acquired neurological disorders causing severe motor disability are often accompanied by cognitive impairments (such as traumatic brain injury, stroke, etc); hence, future research efforts within the BCI community should address this aspect as one of the priorities in the design of new BCI paradigms and protocols. In an online study, Kübler *et al* (2009) have reported a very low level in performances of severe ALS users testing an auditory BCI system. The authors attributed this unsuccessful clinical trial to the excessive cognitive workload in using such independent BCI.

This point stresses the relevance of developing flexible and adaptable BCI-based devices for different end-users, in accordance to a user-centered design principle. For example, the ITR reported in studies evaluating few-choices-BCI devices is, on the one hand, inherently limited as only a little information can be communicated, but on the other hand the cognitive effort required to accomplish the 'BCI' tasks would be reasonably suitable for users with reduced (or even impaired) cognitive resources.

Therefore, system performances expressed within a common metrics framework should be evaluated also by taking into account the cognitive requirements/resources needed to perform the various BCI tasks. Both these aspects should be explored in future studies to meet users' requirements in a real-life scenario.

One emergent BCI research line regards the development of *hybrid* BCI systems (Pfurtscheller *et al* 2010). A *hybrid* BCI term identifies two kinds of BCI systems: (i) a BCI that combines two different types of brain signals (e.g. VEPs and SMRs) to provide a control/output signal and (ii) a system that combines a BCI channel with a muscle-based channel (i.e. EMG signals) (Allison *et al* 2012). In this respect, it has been recently demonstrated that to combine visual attention and motor imagery to operate a BCI is feasible (Allison *et al* 2010, Brunner *et al* 2010, Pfurtscheller *et al* 2010).

The development and validation of this hybrid approach may have a substantial impact on the further implementation of gaze-independent BCIs. As shown by Schmidt *et al* (2012), the performance of a BCI-driven speller can be enhanced by incorporating an automatic selection verification based on the so-called error potential (ErrPs). The authors implemented the ErrPs in the Center Speller (Treder *et al* 2011b) pointing out that the approach could be adapted to other gaze-independent spellers.

Another way to increase efficiency and its related selection speed of the gaze-independent BCIs may be represented by the 'dynamic stopping' modality, a functionality that is currently under exploitation in an independent BCI system (Schreuder *et al* 2011). The advantage of this BCI operation modality resides in the statistical evaluation of the confidence interval of the classifier selection performed after each stimulation sequence and thus *adapting* the number of stimulations needed (and delivered) to achieve the best performance of a given BCI system and subject interaction (Schreuder *et al* 2011).

A similar *autoadaptative* modality combined with a *no-control state* (an automatic pause in the classification output) has been also evaluated in different dependent BCI systems for communication and environmental interaction purposes (Zhang *et al* 2008, Aloise *et al* 2011).

The transfer of such an asynchronous approach to the no-gaze dependent BCIs might represent another opportunity for future development research aiming at a substantial increase in the efficiency of these BCI systems and their usability.

Acknowledgments

The work was partially supported by the EU FP7-224631 'TOBI' (tools for brain–computer interaction) project and by the Italian Agency for Research on ALS–ARiSLA project 'Brindisys'. This paper only reflects the authors' views, and funding agencies are not liable for any use that may be made of the information contained herein.

References

Acqualagna L and Blankertz B 2011 A gaze independent spelling based on rapid serial visual presentation *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **2011** 4560–3

Acqualagna L, Treder M S, Schreuder M and Blankertz B 2010 A novel brain–computer interface based on the rapid serial visual presentation paradigm *Conf. Proc. IEEE Eng. Med. Biol. Soc.* **1** 2686–9

Allison B Z, Brunner C, Kaiser V, Müller-Putz G R, Neuper C and Pfurtscheller G 2010 Toward a hybrid brain–computer interface based on imagined movement and visual attention *J. Neural Eng.* **7** 26007

Allison B Z, McFarland D J, Schalk G, Zheng S D, Jackson M M and Wolpaw J R 2008 Towards an independent brain–computer interface using steady state visual evoked potentials *Clin. Neurophysiol.* **119** 399–408

Aloise F, Aricò P, Schettini F, Riccio A, Salinari S, Mattia D, Babiloni F and Cincotti F 2012 A covert attention P300-based brain–computer interface: Geospell *Ergonomics* **55** 538–51

Aloise F, Schettini F, Arico P, Leotta F, Salinari S, Mattia D, Babiloni F and Cincotti F 2011 P300-based brain–computer interface for environmental control: an asynchronous approach *J. Neural Eng.* **8** 025025

Aloise F *et al* 2006 The ASPICE project: inclusive design for the motor disabled *AVI '06: Proc. of the Working Conf. on Advanced Visual Interfaces* pp 360–3

Belitski A, Farquhar J and Desain P 2011 P300 audio-visual speller *J. Neural Eng.* **8** 025022

Berger H 1930 Über das Elektrenkephalogramm des Menschen, 2nd report *J. Psychol. Neurol. Leipzig.* **40** 60–79

Bianchi L, Sami S, Hillebrand A, Fawcett I P, Quitadamo L R and Seri S 2010 Which physiological components are more suitable for visual ERP based brain–computer interface? A preliminary MEG/EEG study *Brain Topogr.* **23** 180–5

Birbaumer N 2006 Breaking the silence: brain–computer interfaces (BCI) for communication and motor control *Psychophysiology* **43** 517–32

Birbaumer N *et al* 1999 A spelling device for the paralysed *Nature* **398** 297–8

Birbaumer N, Kübler A, Ghanayim N, Hinterberger T, Perelmouter J, Kaiser J, Iversen I, Kotchoubey B, Neumann N and Flor H 2000 The thought translation device (TTD) for completely paralyzed patients *IEEE Trans. Rehabil. Eng.* **8** 190–3

Birbaumer N, Murguialday A R and Cohen L 2008 Brain–computer interface in paralysis *Curr. Opin. Neurol.* **21** 634–8

Blankertz B, Krauledat M, Dornhege G, Williamson J, Murray-Smith R and Müller K-R 2007 A note on brain actuated spelling with the Berlin brain–computer interface *Universal Access in Human-Computer Interaction. Ambient Interaction* vol 4555 ed C Stephanidis (Berlin: Springer) pp 759–68

Brouwer A M and van Erp J B 2010 A tactile P300 brain–computer interface *Front. Neurosci.* **4** 036003

Brunner P, Joshi S, Briskin S, Wolpaw J R, Bischof H and Schalk G 2010 Does the 'P300' speller depend on eye gaze? *J. Neural Eng.* **7** 056013

Chatterjee A, Aggarwal V, Ramos A, Acharya S and Thakor N V 2007 A brain–computer interface with vibrotactile biofeedback for haptic information *J. Neuroeng. Rehabil.* **4** 40

Cheng M, Gao X, Gao S and Xu D 2002 Design and implementation of a brain–computer interface with high transfer rates *IEEE Trans. Bio-Med. Eng.* **49** 1181–6

Cincotti F *et al* 2007 Vibrotactile feedback for brain–computer interface operation *Comput. Intell. Neurosci.* **2007** 48937

Coyle S, Ward T, Markham C and McDarby G 2004 On the suitability of near-infrared (NIR) systems for next-generation brain–computer interfaces *Physiol. Meas.* **25** 815–22

Donchin E, Spencer K M and Wijesinghe R 2000 The mental prosthesis: assessing the speed of a P300-based brain–computer interface *IEEE Trans. Rehabil. Eng.* **8** 174–9

Elbert T, Rockstroh B, Lutzenberger W and Birbaumer N 1980 Biofeedback of slow cortical potentials: I. *Electroencephalogr. Clin. Neurophysiol.* **48** 293–301

Fabiani M, Gratton G, Karis D and Donchin E 1987 The definition, identification, and reliability of measurement of the P3 component of the event related brain potential *Advances in Psychophysiology* vol 2 ed P K Ackles, J R Jennings and M G H Coles (Greenwich: JAI Press)

Farwell L A and Donchin E 1988 Talking off the top of your head: toward a mental prosthesis utilizing event-related brain potentials *Electroencephalogr. Clin. Neurophysiol.* **70** 510–23

Frank B, Haas J, Heinze H J, Stark E and Münte T F 1997 Relation of neuropsychological and magnetic resonance findings in amyotrophic lateral sclerosis: evidence for subgroups *Clin. Neurol. Neurosurg.* **99** 79–86

Furdea A, Halder S, Krusienski D J, Bross D, Nijboer F, Birbaumer N and Kübler A 2009 An auditory oddball (P300) spelling system for brain–computer interfaces *Psychophysiology* **46** 617–25

Gao X, Xu D, Cheng M and Gao S 2003 A BCI-based environmental controller for the motion-disabled *IEEE Trans. Neural Syst. Rehabil. Eng.* **11** 137–40

Guo J, Gao S and Hong B 2010 An auditory brain–computer interface using active mental response *IEEE Trans. Neural Syst. Rehabil. Eng.* **18** 230–5

Halder S *et al* 2010 An auditory oddball brain–computer interface for binary choices *Clin. Neurophysiol.* **121** 516–23

Hill N J, Lal T N, Bierig K, Birbaumer N and Scholkopf B 2004 Attention modulation of auditory event-related potentials in a brain–computer interface *IEEE Int. Workshop on Biomedical Circuits and Systems* pp S3/5/INV–S3/17–20

Hill N J and Schölkopf B 2012 An online brain–computer interface based on shifting attention to concurrent streams of auditory stimuli *J. Neural Eng.* **9** 026011

Hinterberger T, Weiskopf N, Veit R, Wilhelm B, Betta E and Birbaumer N 2004 An EEG-driven brain–computer interface combined with functional magnetic resonance imaging (fMRI) *IEEE Trans. Biomed. Eng.* **51** 971–4

Höhne J, Schreuder M, Blankertz B and Tangermann M 2011 A novel 9-class auditory ERP paradigm driving a predictive text entry system *Front. Neurosci.* **5** 99

Jasper H H and Andrew H L 1938 Electro-encephalography: III. Normal differentiation of occipital and precentral regions in man *Arch. Neurol. Psychiat.* **39** 96–115

Jasper H H and Penfield W 1949 Electrocorticograms in man: effect of the voluntary movement upon the electrical activity of the precentral gyrus *Arch. Psychiat. Z. Neurol.* **183** 163–74

Kanoh S, Miyamoto K and Yoshinobu T 2010 A brain–computer interface (BCI) system based on auditory stream segregation *J. Biomech. Sci. Eng.* **5** 32–40

Kaufmann T, Schulz S M, Grünzinger C and Kübler A 2011 Flashing characters with famous faces improves ERP-based brain–computer interface performance *J. Neural Eng.* **8** 056016

Kelly S P, Lalor E, Finucane C and Reilly R B 2004 A comparison of covert and overt attention as a control option in a steady-state visual evoked potential-based brain–computer interface *Conf. Proc. IEEE Eng. Med. Biol Soc.* **7** 4725–8

Kelly S P, Lalor E C, Reilly R B and Foxe J J 2005a Visual spatial attention tracking using high-density SSVEP data for independent brain–computer communication *IEEE Trans. Neural Syst. Rehabil. Eng.* **13** 172–8

Kelly S P, Lalor E, Reilly R B and Foxe J J 2005b Independent brain–computer interface control using visual spatial attention-dependent modulations of parieto-occipital alpha *2nd Proc. Conf. on Int. IEEE EMBS (2005)* pp 667–70

Kim D-W, Hwang H-J, Lim J-H, Lee Y-H, Jung K-Y and Im C-H 2011 Classification of selective attention to auditory stimuli: toward vision-free brain–computer interfacing *J. Neurosci. Methods* **197** 180–5

Kleih S C *et al* 2011 Out of the frying pan into the fire—the P300-based BCI faces real-world challenges *Prog. Brain Res.* **194** 27–46

Kleih S C, Nijboer F, Halder S and Kübler A 2010 Motivation modulates the P300 amplitude during brain–computer interface use *Clin. Neurophysiol.* **121** 1023–31

Klobassa D S, Vaughan T M, Brunner P, Schwartz N E, Wolpaw J R, Neuper C and Sellers E W 2009 Toward a high-throughput auditory P300-based brain–computer interface *Clin. Neurophysiol.* **120** 1252–61

Krusienski D J, Grosse-Wentrup M, Galán F, Coyle D, Miller K J, Forney E and Anderson C W 2011 Critical issues in state-of-the-art brain–computer interface signal processing *J. Neural Eng.* **8** 025002

Kübler A, Furdea A, Halder S, Hammer E M, Nijboer F and Kotchoubey B 2009 A brain–computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients *Ann. NY Acad. Sci.* **1157** 90–100

Kübler A, Neumann N, Kaiser J, Kotchoubey B, Hinterberger T and Birbaumer N P 2001 Brain–computer communication: self-regulation of slow cortical potentials for verbal communication *Arch. Phys. Med. Rehabil.* **82** 1533–9

Liu T, Goldberg L, Gao S and Hong B 2010 An online brain–computer interface using non-flashing visual evoked potentials *J. Neural Eng.* **7** 036003

Liu Y, Zhou Z and Hu D 2011 Gaze independent brain–computer speller with covert visual search tasks *Clin. Neurophysiol.* **122** 1127–36

Lopez M-A, Pomares H, Pelayo F, Urquiza J and Perez J 2009 Evidences of cognitive effects over auditory steady-state responses by means of artificial neural networks and its use in brain–computer interfaces *Neurocomputing* **72** 3617–23

Lotte F, Congedo M, Lécuyer A, Lamarche F and Arnaldi B 2007 A review of classification algorithms for EEG-based brain–computer interfaces *J. Neural Eng.* **4** R1–R13

Mak J N and Wolpaw J R 2009 Clinical applications of brain–computer interfaces: current state and future prospects *IEEE Rev. Biomed. Eng.* **2** 187–99

Massman P J, Sims J, Cooke N, Haverkamp L J, Appel V and Appel S H 1996 Prevalence and correlates of neuropsychological deficits in amyotrophic lateral sclerosis *J. Neurol. Neurosurg. Psychiatr.* **61** 450–5

Middendorf M, McMillan G, Calhoun G and Jones K S 2000 Brain–computer interfaces based on the steady-state visual-evoked response *IEEE Trans. Rehabil. Eng.* **8** 211–4

Mugler E M, Ruf C A, Halder S, Bensch M and Kübler A 2010 Design and implementation of a P300-based brain–computer interface for controlling an internet browser *IEEE Trans. Neural Syst. Rehabil. Eng.* **18** 599–609

Müller-Putz G R, Scherer R, Neuper C and Pfurtscheller G 2006 Steady-state somatosensory evoked potentials: suitable brain signals for brain–computer interfaces? *IEEE Trans. Neural Syst. Rehabil. Eng.* **14** 30–7

Namerow N S, Sclabassi R J and Enns N F 1974 Somatosensory responses to stimulus trains: normative data *Electroencephalogr. Clin. Neurophysiol.* **37** 11–21

Neumann N, Kübler A, Kaiser J, Hinterberger T and Birbaumer N 2003 Conscious perception of brain states: mental strategies for brain–computer communication *Neuropsychologia* **41** 1028–36

Neuper C, Scherer R, Wriessnegger S and Pfurtscheller G 2009 Motor imagery and action observation: modulation of sensorimotor brain rhythms during mental control of a brain–computer interface *Clin. Neurophysiol.* **120** 239–47

Nijboer F, Furdea A, Gunst I, Mellinger J, McFarland D J, Birbaumer N and Kübler A 2008b An auditory brain–computer interface (BCI) *J. Neurosci. Methods* **167** 43–50

Nijboer F *et al* 2008a A P300-based brain–computer interface for people with amyotrophic lateral sclerosis *Clin. Neurophysiol.* **119** 1909–16

Pfurtscheller G *et al* 2010 The hybrid BCI *Front. Neurosci.* **4** 30

Pfurtscheller G and Aranibar A 1979 Evaluation of event-related desynchronization (ERD) preceding and following voluntary self-paced movement *Electroencephalogr. Clin. Neurophysiol.* **46** 138–46

Pfurtscheller G, Guger C, Müller G, Krausz G and Neuper C 2000 Brain oscillations control hand orthosis in a tetraplegic *Neurosci. Lett.* **292** 211–4

Pfurtscheller G and Neuper C 1992 Simultaneous EEG 10 Hz desynchronization and 40 Hz synchronization during finger movements *Neuroreport* **3** 1057–60

Pham M *et al* 2005 An auditory brain–computer interface based on the self-regulation of slow cortical potentials *Neurorehabil. Neural Repair* **19** 206–18

Piccione F, Giorgi F, Tonin P, Priftis K, Giove S, Silvoni S, Palmas G and Beverina F 2006 P300-based brain–computer interface: reliability and performance in healthy and paralysed participants *Clin. Neurophysiol.* **117** 531–7

Plourde G 2006 Auditory evoked potentials *Best Pract. Res. Clin. Anaesthesiol.* **20** 129–39

Posner M I 1980 Orienting of attention *Q. J. Exp. Psychol.* **32** 3–25

Quitadamo L R, Mattia D, Cincotti F, Babiloni F, Cardarilli G C, Marciani M G and Bianchi L 2009 Standard model, file formats and methods in brain–computer interface research: why? *Biomedical Engineering* ed Carlos Alexandre and Barros de Mello (Rijeka: Intech)

Riccio A *et al* 2011 Workload measurement in a communication application operated through a P300-based brain–computer interface *J. Neural Eng.* **8** 025028

Schlögl A, Kronegg J, Huggins J and Mason S G 2007 Evaluation criteria for BCI research *Towards Brain–Computer Interfacing* ed G Dornhege, J del, R Millán, T Hinterberger, D McFarland and K-R Müller (Cambridge, MA: MIT Press) pp 297–312

Schmidt N M, Blankertz B and Treder M S 2012 Online detection of error-related potentials boosts the performance of mental typewriters *BMC Neurosci.* **13** 19

Schreuder M, Blankertz B and Tangermann M 2010 A new auditory multi-class brain–computer interface paradigm: spatial hearing as an informative cue *PloS One* **5** e9813

Schreuder M, Rost T and Tangermann M 2011 Listen, you are writing! Speeding up online spelling with a dynamic auditory BCI *Front. Neuroprosthetics* **5** 112

Sellers E W and Donchin E 2006 A P300-based brain–computer interface: initial tests by ALS patients *Clin. Neurophysiol.* **117** 538–48

Sellers E W, Vaughan T M and Wolpaw J R 2010 A brain–computer interface for long-term independent home use *Amyotroph. Lateral Scler.* **11** 449–55

Sutter E E 1992 The brain response interface: communication through visually-induced electrical brain responses *J. Microcomput. Appl.* **15** 31–45

Sutton S, Braren M, Zubin J and John E R 1965 Evoked-potential correlates of stimulus uncertainty *Science* **150** 1187–8

Tonin L, Carlson T, Leeb R, del R and Millán J 2011 Brain-controlled telepresence robot by motor-disabled people *Conf. Proc. IEEE Eng. Med. Biol. Soc. 2011* pp 4227–30

Treder M S, Bahramisharif A, Schmidt N M, van Gerven M A and Blankertz B 2011a Brain–computer interfacing using modulations of alpha activity induced by covert shifts of attention *J. Neuroeng. Rehabil.* **8** 24

Treder M S and Blankertz B 2010 (C)overt attention and visual speller design in an ERP-based brain–computer interface *Behav. Brain Functions* **6** 28

Treder M S, Schmidt N M and Blankertz B 2011b Gaze-independent brain–computer interfaces based on covert attention and feature attention *J. Neural Eng.* **8** 066003

Vanacker G, del R Millán J, Lew E, Ferrez P W, Moles F G, Philips J, Van Brussel H and Nuttin M 2007 Contex-based filtering for assisted brain actuated wheelchair driving *Comput. Intell. Neurosci.* 1–12

van Gerven M, Bahramisharif A, Heskes T and Jensen O 2009 Selecting features for BCI control based on a covert spatial attention paradigm *Neural Netw.* **22** 1271–7

Vialatte F B, Maurice M, Dauwels J and Cichocki A 2010 Steady-state visually evoked potentials: focus on essential paradigms and future perspectives *Prog. Neurobiol.* **90** 418–38

Vidal J J 1973 Towards direct brain–computer communication *Annu. Rev. Biophys. Bioeng.* **2** 157–80

Weiskopf N, Scharnowski F, Veit R, Goebel R, Birbaumer N and Mathiak K 2004 Self-regulation of local brain activity using real-time functional magnetic resonance imaging (fMRI) *J. Phys.* Paris **98** 357–73

Williamson J, Murray-Smith R, Blankertz B, Krauledat M and Müller K-R 2009 Designing for uncertain, asymmetric control: interaction design for brain–computer interfaces *Int. J. Hum.-Comput. Stud.* **67** 827–41

Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Brain–computer interfaces for communication and control *Clin. Neurophysiol.* **113** 767–91

Wolpaw J R, McFarland D J and Vaughan T M 2000 Brain–computer interface research at the Wadsworth Center *IEEE Trans. Rehabil. Eng.* **8** 222–6

Woolley S C, Moore D H and Katz J S 2010 Insight in ALS: awareness of behavioral change in patients with and without FTD *Amyotroph. Lateral Scler.* **11** 52–6

Wolpaw J R and Wolpaw E W 2012 *Brain–Computer Interfaces: Principles and Practice* (Oxford: Oxford University Press)

Zhang H, Guan C and Wang C 2008 Asynchronous P300-based brain–computer interfaces: a computational approach with statistical models *IEEE Trans. Bio-Med. Eng.* **55** 1754–63

Zhang D, Maye A, Gao X, Hong B, Engel A K and Gao S 2010 An independent brain–computer interface using covert non-spatial visual selective attention *J. Neural Eng.* **7** 16010

Zhang D, Wang Y, Maye A, Engel A K, Gao X, Hong B and Gao S 2007 A brain–computer interface based on multi-modal attention *3rd Int. IEEE/EMBS Conf. on Neural Engineering (2007 CNE '07)* pp 414–7

Zickler C *et al* 2011 A brain–computer interface as input channel for a standard assistive technology software *Clin. EEG Neurosci.* **236** 236–44