

PAPER • OPEN ACCESS

# Enhancing clinical communication assessments using an audiovisual BCI for patients with disorders of consciousness

To cite this article: Fei Wang *et al* 2017 *J. Neural Eng.* **14** 046024

View the [article online](#) for updates and enhancements.

## Related content

- [Detecting awareness in patients with disorders of consciousness using a hybrid brain-computer interface](#)  
Jiahui Pan, Qiuyou Xie, Yanbin He *et al.*
- [A P300 auditory brain-computer interface based on mental repetition](#)  
Alessandro Marassi, Riccardo Budai and Luca Chittaro
- [An independent SSVEP-based brain-computer interface in locked-in syndrome](#)  
D Lesenfants, D Habbal, Z Lugo *et al.*

## Recent citations

- [The Self-Face Paradigm Improves the Performance of the P300-Speller System](#)  
Zhaohua Lu *et al*
- [The N400 for brain computer interfacing: complexities and opportunities](#)  
K V Dijkstra *et al*
- [Motor imagery EEG recognition with KNN-based smooth auto-encoder](#)  
Xianlun Tang *et al*



The Department of Bioengineering at the University of Pittsburgh Swanson School of Engineering invites applications from accomplished individuals with a PhD or equivalent degree in bioengineering, biomedical engineering, or closely related disciplines for an open-rank, tenured/tenure-stream faculty position. We wish to recruit an individual with strong research accomplishments in Translational Bioengineering (i.e., leveraging basic science and engineering knowledge to develop innovative, translatable solutions impacting clinical practice and healthcare), with preference given to research focus on neuro-technologies, imaging, cardiovascular devices, and biomimetic and biorobotic design. It is expected that this individual will complement our current strengths in biomechanics, bioimaging, molecular, cellular, and systems engineering, medical product engineering, neural engineering, and tissue engineering and regenerative medicine. In addition, candidates must be committed to contributing to high quality education of a diverse student body at both the undergraduate and graduate levels.

[CLICK HERE FOR FURTHER DETAILS](#)

**To ensure full consideration, applications must be received by June 30, 2019. However, applications will be reviewed as they are received. Early submission is highly encouraged.**

# Enhancing clinical communication assessments using an audiovisual BCI for patients with disorders of consciousness

Fei Wang<sup>1,3,6</sup>, Yanbin He<sup>2,4,6</sup>, Jun Qu<sup>1,3</sup>, Qiuyou Xie<sup>2</sup>, Qing Lin<sup>2</sup>, Xiaoxiao Ni<sup>2</sup>, Yan Chen<sup>2</sup>, Jiahui Pan<sup>1,3</sup>, Steven Laureys<sup>5</sup>, Ronghao Yu<sup>2,7</sup> and Yuanqing Li<sup>1,3,7</sup>

<sup>1</sup> School of Automation Science and Engineering, South China University of Technology, Guangzhou 510640, People's Republic of China

<sup>2</sup> Coma Research Group, Centre for Hyperbaric Oxygen and Neurorehabilitation, Guangzhou General Hospital of Guangzhou Military Command, Guangzhou 510010, People's Republic of China

<sup>3</sup> Guangzhou Key Laboratory of Brain Computer Interface and Applications, Guangzhou 510640, People's Republic of China

<sup>4</sup> Southern Medical University, Guangzhou 510515, People's Republic of China

<sup>5</sup> Coma Science Group, Cyclotron Research Centre and Neurology Department, University and University Hospital of Liège, B-4000 Liège, Belgium

E-mail: [auyqli@scut.edu.cn](mailto:auyqli@scut.edu.cn) and [gesund@139.com](mailto:gesund@139.com)

Received 16 January 2017

Accepted for publication 10 April 2017

Published 14 June 2017



## Abstract

**Objective.** The JFK coma recovery scale-revised (JFK CRS-R), a behavioral observation scale, is widely used in the clinical diagnosis/assessment of patients with disorders of consciousness (DOC). However, the JFK CRS-R is associated with a high rate of misdiagnosis (approximately 40%) because DOC patients cannot provide sufficient behavioral responses. A brain–computer interface (BCI) that detects command/intention-specific changes in electroencephalography (EEG) signals without the need for behavioral expression may provide an alternative method. **Approach.** In this paper, we proposed an audiovisual BCI communication system based on audiovisual ‘yes’ and ‘no’ stimuli to supplement the JFK CRS-R for assessing the communication ability of DOC patients. Specifically, patients were given situation-orientation questions as in the JFK CRS-R and instructed to select the answers using the BCI. **Main results.** Thirteen patients (eight vegetative state (VS) and five minimally conscious state (MCS)) participated in our experiments involving both the BCI- and JFK CRS-R-based assessments. One MCS patient who received a score of 1 in the JFK CRS-R achieved an accuracy of 86.5% in the BCI-based assessment. Seven patients (four VS and three MCS) obtained unresponsive results in the JFK CRS-R-based assessment but responsive results in the BCI-based assessment, and 4 of those later improved scores in the JFK CRS-R-based assessment. Five patients (four VS and one MCS) obtained unresponsive results in both assessments. **Significance.** The experimental results indicated that the audiovisual BCI could provide more sensitive results than the JFK CRS-R and therefore supplement the JFK CRS-R.

**Keywords:** audiovisual brain-computer interface, disorders of consciousness (DOC), coma recovery scale-revised (CRS-R), communication, assessment

(Some figures may appear in colour only in the online journal)



Original content from this work may be used under the terms of the [Creative Commons Attribution 3.0 licence](https://creativecommons.org/licenses/by/3.0/). Any further distribution of this work must maintain attribution to the author(s) and the title of the work, journal citation and DOI.

<sup>6</sup> These authors contributed equally to the manuscript.

<sup>7</sup> Authors to whom any correspondence should be addressed.

## 1. Introduction

Disorders of consciousness (DOC), such as vegetative state (VS) and minimally conscious state (MCS), are clinically diagnosed on the basis of behavior scales such as the JFK coma recovery scale-revised (JFK CRS-R) and the Glasgow coma scale (GCS), which rely on motor responses to external stimulation [1]. People in a vegetative state may awaken but show no awareness of themselves or their environment, whereas patients in a minimally conscious state may demonstrate inconsistent but reproducible signs of awareness [2]. Furthermore, emergence from MCS (EMCS) is characterized by a reliable and consistent demonstration of functional interactive communication or the functional use of two different objects [3]. Patients in locked-in state (LIS) similarly show very limited signs of awareness due to profound sensory and motor deficits but have retained self-awareness and normal or near-normal cognitive capacities [4]. Several studies have not included LIS as a DOC [3, 5].

Among these behavior assessment methods, the JFK CRS-R is suggested to be reliable and widely applied in clinical assessments [6]. The JFK CRS-R was first introduced by Giacino *et al* in 1991 and was revised in 2004 [7]. The JFK CRS-R contains six subscales, including auditory, visual, motor, oromotor, communication and arousal functions. The score in each subscale depends on whether the DOC patient has specific behavior responses to sensory stimuli. For example, in the communication subscale, if clear, discernible and accurate responses occur for all situational orientation questions, the patient receives a score of 2. If clear, discernible responses occur for at least two but not all questions, the patient receives a score of 1. Otherwise, a score of 0 is given. However, behavior-based scales, such as the JFK CRS-R, have some problems [8]. First, many patients with DOC cannot maintain a stable state during the evaluation. Second, these patients are usually unable to make normal physical movements [9]. As a consequence, high rates of clinical misdiagnosis can occur. For example, recent studies have observed that 37% to 43% of patients diagnosed as being in a VS actually show signs of awareness [6, 10, 11].

Considering these limitations of behavioral assessment scales, brain-computer interfaces (BCIs) could be used as an assistance tool for clinical evaluation because they can detect command/intention-specific changes in brain signals, such as electroencephalography (EEG), without requiring any behavioral expression [9, 12]. Several BCI systems have recently been applied in awareness detection for patients with DOC [13–16]. Cruse *et al* accessed 16 VS patients by having them complete a motor imagery (MI) task in which they were instructed to imagine movements of their right hand and toes in response to commands [13]. Three of the patients could repeatedly and reliably generate appropriate EEG responses to two distinct commands. Pan *et al* reported a visual hybrid BCI that combined P300 and SSVEP to detect awareness in four VS, three MCS and one LIS patients [14]. They successfully demonstrated command following in three patients (one VS, one MCS and one LIS). In our previous study, we applied an audiovisual BCI system to detect the awareness of seven DOC patients; the patients were instructed to selectively

attend a target of two different number stimuli [15]. Five of the patients exhibited command following. These studies focused on DOC patients' overall awareness levels and did not evaluate the patients' functional consciousness in detail, as might be done with the JFK CRS-R. In addition, several studies have reported that BCIs could detect functional communication in LIS patients. Lulé *et al* tested a four-choice (Yes, No, Stop and Go) auditory oddball BCI with 18 patients (13 MCS, 3 VS and 2 LIS) [17]. Their results showed that one LIS patient had a correct response rate of 60% and therefore was able to communicate using the BCI. Overall, the performance of these BCIs designed for DOC patients is generally poor. This is mainly because DOC patients' recognition levels, such as attention, are substantially lower than those of healthy subjects, and recognition levels are associated with the performance of a BCI system. Furthermore, to our knowledge, no BCI has been designed to directly assist in behavioral scale-based assessments. It is difficult to obtain sensitive results for several items in traditional behavioral scales. BCI methods are superior to clinical assessment because the former does not depend on the patients' behavioral responses. However, the performance of BCIs in DOC patients is generally poor. A statistical test can be performed to show whether BCI results are significant. Furthermore, we need to develop novel BCIs that are suitable for DOC patients. Previous studies have validated the effectiveness of audiovisual BCIs [15, 18]. In this study, we used an audiovisual paradigm to improve BCI performance.

In this paper, we propose a new protocol to assist JFK CRS-R for assessing the communication ability of DOC patients. The new protocol was designed following JFK CRS-R communication subscale-based assessment. Specifically, the new audiovisual BCI system was based on semantically congruent audiovisual stimuli, i.e. 'Yes' and 'No'. We imitated the JFK CRS-R communication subscale by administering situation-orientation questions to DOC patients. After a question was proposed, two flashing buttons with the Chinese words 'Yes' and 'No' appeared on the computer screen. When a button was flashing, the corresponding spoken word 'Yes' or 'No' (in Chinese) was heard simultaneously. The patients were instructed to selectively focus on the flashing button corresponding to the correct answer ('Yes' or 'No') to the question and the corresponding spoken word. The BCI system determined the patient's choice by detecting event-related potentials (ERPs), such as P300. The detection result was presented as online feedback. Based on the feedback results, the clinical examiners could make a determination of the communication ability of the patient. Thirteen patients participated in our experiments with BCI- and JFK CRS-R-based communication assessments. The experimental results demonstrated the efficacy of our BCI approach in supplementing the JFK CRS-R assessment of communication in DOC patients.

## 2. Methods

### 2.1. Subjects

Thirteen patients (eight VS and five MCS; mean  $\pm$  SD,  $37 \pm 12$  years of age; twelve males; see table 1) from a local

**Table 1.** Summary of patients' clinical status. TBI, traumatic brain injury; NTBI, non-traumatic brain injury; VS, vegetative state; MCS, minimally conscious state; EMCS, emergence from MCS; LIS, locked-in state. JFK CRS-R subscales: auditory, visual, motor, oromotor, communication, and arousal functions.

Patients	Age	Gender	Aetiology	Time since onset (months)	JFK CRS-R (Subscale scores) clinical diagnosis		
					Before the experiment	Shortly after the experiment	Two months after the experiment
P1	29	M	NTBI	8.5	4(1-0-1-0-0-2) VS	4(1-0-1-0-0-2) VS	4(1-0-1-0-0-2) VS
P2	37	M	NTBI	2	5(0-0-2-1-0-2) VS	5(0-0-2-1-0-2) VS	5(0-0-2-1-0-2) VS
P3	38	M	TBI	1	7(1-1-2-1-0-2) VS	7(2-1-3-1-0-2) VS	7(2-1-3-1-0-2) VS
P4	33	M	TBI	2	7(1-0-2-2-0-2) VS	7(1-0-2-2-0-2) VS	7(1-0-2-2-0-2) VS
P5	40	M	NTBI	2	5(1-0-2-0-0-2) VS	6(1-0-2-1-0-2) VS	7(1-1-2-1-0-2) VS
P6	52	M	NTBI	4.5	5(1-1-0-1-0-2) VS	8(2-3-0-1-0-2) MCS	9(2-3-1-1-0-2) MCS
P7	42	M	NTBI	4	7(1-1-2-1-0-2) VS	11(3-2-3-1-0-2) MCS	11(3-2-3-1-0-2) MCS
P8	26	M	NTBI	1	7(1-1-2-1-0-2) VS	15(4-5-2-1-1-2) MCS	15(4-5-2-1-1-2) MCS
P9	19	M	NTBI	1.5	8(1-1-3-1-0-2) MCS	8(1-1-3-1-0-2) MCS	8(1-1-2-2-0-2) VS
P10	17	M	TBI	2	8(1-1-3-1-0-2) MCS	11(1-3-3-2-0-2) MCS	17(2-4-6-2-1-2) EMCS
P11	46	F	TBI	1.5	7(1-0-3-1-0-2) MCS	19(3-5-6-2-1-2) EMCS	20(3-5-6-2-1-3) EMCS
P12	46	M	NTBI	2	9(1-1-4-1-0-2) MCS	11(1-1-5-2-0-2) MCS	20(4-5-6-2-1-2) EMCS
P13	59	M	NTBI	2	14(3-4-2-1-1-2) MCS	18(4-5-3-1-2-3) LIS	18(4-5-3-1-2-3) LIS

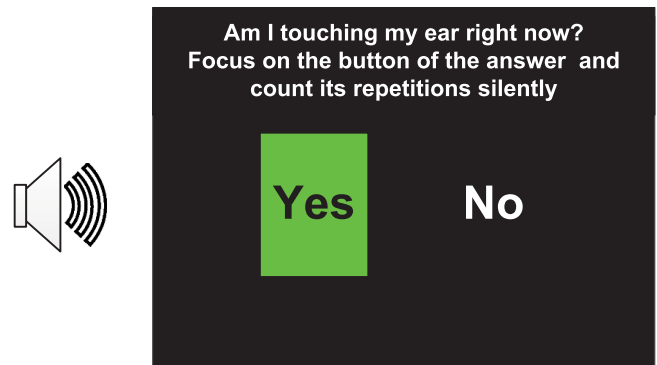
hospital participated in the experiment. None of the patients had a history of impaired visual or auditory acuity. The Ethics Committee of the General Hospital of the Guangzhou Military Command of PLA in Guangzhou, China, which complies with the Code of Ethics of the World Medical Association (Declaration of Helsinki), approved the experimental procedures. Each patient's legal guardian provided written informed consent for the experiment and for the publication of their individual details in this manuscript. The clinical diagnoses were based on the JFK CRS-R, which comprises six subscales addressing auditory, visual, motor, oromotor, communication and arousal functions [7]. Details, including the JFK CRS-R scores of the thirteen patients, are shown in table 1.

## 2.2. Data acquisition

EEG data were amplified using a SynAmps2 device (Compumedics, Neuroscan, Inc., Australia) sampled at 250 Hz and filtered between 0.01 and 30 Hz. The EEG signals were recorded from 30 electrodes using an EEG cap (LT 37) based on the international 10–20 system and referenced to the right mastoid. All electrode impedances were maintained below 5 k $\Omega$  during data collection.

## 2.3. Experimental procedure

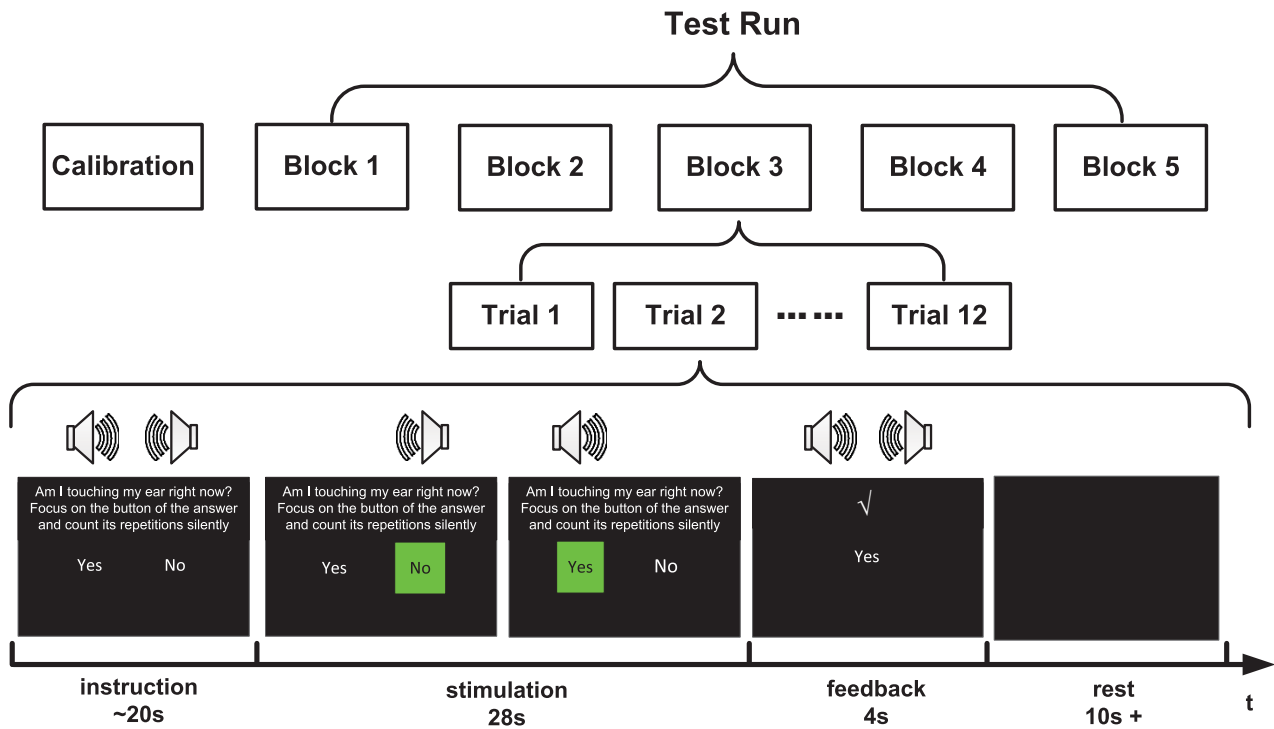
The thirteen patients participated in two assessments: the JFK CRS-R-based behavioral assessment and the BCI-based communication assessment. The JFK CRS-R-based behavioral assessment was conducted by a clinician from the General Hospital of the Guangzhou Military Command of PLA. Following the standard protocol, the clinician administered the following 6 situation-orientation questions to the patient in random order. Questions 1–4: 'Am I touching my ear/nose right now? The clinician touched or did not touch his/her ear/nose'; Questions 5–6: 'Am I clapping my hands right now? The clinician clapped or did not clap'. The communication



**Figure 1.** The GUI of the audiovisual BCI system. The question and instructions are presented in the upper part of the screen. Two buttons ('Yes' and 'No') are randomly presented on the left and right sides of the screen. When a button flashes with the color changing from black to green, the color of the word included in the button changes from white to black, and the corresponding spoken word is simultaneously presented from the headphone on the same side as the button.

scale score depended on the patient's discernible verbal or nonverbal communication response. If clearly discernible and accurate responses were given for all six questions, the patient received a score of 2. If a clearly discernible response (e.g. head nod/shake, thumbs up) was given within 10 s for at least two of the six questions, the patient received a score of 1. If no discernible verbal or nonverbal communication response was given for any question, the patients received a score of 0. The thirteen DOC patients underwent three JFK CRS-R-based assessments shortly before and after the BCI experiment and again two months after the BCI experiment, as shown in table 1.

The graphical user interface (GUI) of our audiovisual BCI system is shown in figure 1. First, a situation question and an instruction were presented in the upper part of the screen. The number of questions with an answer of 'yes' was equal to the number of questions with an answer of 'no', and the questions



**Figure 2.** The experimental paradigm of the audiovisual BCI-based communication assessment. The test run contains five blocks, and each block consists of 12 trials. Each trial includes audiovisual instructions (20 s), audiovisual stimulation (20–48 s), feedback of the detection results (48–52 s), and a rest period (not fixed, e.g. 52–62 s). The audiovisual stimulation involves five repetitions of an audiovisual stimulus, such as ‘No’; five repetitions of another audiovisual stimulus, such as ‘Yes’; five repetitions of the first audiovisual stimulus; and five repetitions of the second audiovisual stimulus.

were presented in random order. Two word buttons with ‘Yes’ and ‘No’ in Chinese were randomly presented on the left and right sides of the screen. When an audiovisual stimulus was presented, the color of the corresponding button changed from green to black, and the color of the word included in the button changed from white to black. Simultaneously, the corresponding spoken word (65 dB) was presented in the headphone on the same side as the button.

The experimental paradigm of the audiovisual BCI-based communication assessment is shown in figure 2. Before the online experiment, each patient performed a calibration run of 12 trials. The test run contained five blocks, and each block consisted of 12 trials. Each block was conducted on a separate day because patients were easily fatigued. The test run lasted from one to two weeks.

In the BCI experiment, each trial began with an audiovisual instruction of approximately 20 s. During the instruction, an experimenter administered a situation question to the patient twice; the questions were the same as those used in the clinical evaluation. The patient was asked to focus on the button with the answer and count its repetitions silently. The experimenter and family members explained the instructions repeatedly so that the patients paid attention to the audiovisual target stimuli. There were two rounds of audiovisual stimulations following the instruction period. In the first round, the audiovisual stimulus of one button (randomly chosen from the two buttons, such as ‘No’ in figure 2) was presented five times, and the audiovisual stimulus of the other button, such as ‘Yes’ in figure 1, was then repeated five times. The second round of audiovisual

stimulations was the same as the first round. Each audiovisual stimulus lasted 300 ms. The time interval between every two adjacent audiovisual stimuli was randomly chosen from 700, 900, 1100, 1300, and 1500 ms. If the target was detected by the classification algorithm after two rounds of audiovisual stimulations, the sound of applause and the detected result were presented for 4 s as the feedback; otherwise, a cross appeared on the screen for 4 s. Finally, there was a break at the end of each trial of at least 10 s depending on the patient’s level of arousal. In the BCI-based assessment, if the patients showed sustained eyelid closure during a trial, the recorded trial was discarded, and an arousal facilitation protocol was administered, such as presenting deep pressure stimulation to the shoulder. The next trial began after the patient reawakened. This arousal facilitation protocol was similar to that used in the clinical assessment.

#### 2.4. Data processing

For all trials, the EEG signals recorded from 30 channels were band-pass filtered (0.1–20 Hz). We then extracted epochs corresponding to each stimulus from 0 to 600 ms after stimulus onset for each channel. All epochs were baseline corrected using a baseline of 100 ms before the stimulus onset and down-sampled by a rate of 5. Next, we concatenated the epochs from all 30 channels to obtain a data vector. Finally, a feature vector corresponding to each button was constructed by averaging the vectors from all ten corresponding stimulations. Using EEG data from the calibration run, we first trained an SVM classifier in which the feature vectors corresponding to the target and



**Table 2.** The accuracy of the BCI-based experiment and the scores of the JFK CRS-R communication subscale for all patients. Note that the accuracies that were significantly higher than the chance level (50%) are highlighted in bold.

Patients	Clinical diagnosis	Trials	Hits	Accuracy (%)	<i>p</i> value	JFK CRS-R communication subscale scores		
						Before the experiment	Shortly after the experiment	Two months after the experiment
P1	VS	54	24	44.4	0.2018	0	0	0
P2	VS	60	34	56.7	0.1459	0	0	0
P3	VS	58	35	60.3	0.0538	0	0	0
P4	VS	60	34	56.7	0.1459	0	0	0
P5	VS	52	40	<b>76.9</b>	<0.0001	0	0	0
P6	VS	56	39	<b>69.2</b>	0.0013	0	0	0
P7	VS	57	43	<b>75.4</b>	<0.0001	0	0	0
P8	VS	53	41	<b>77.4</b>	<0.0001	0	1	1
P9	MCS	60	36	60.0	0.0569	0	0	0
P10	MCS	53	42	<b>79.2</b>	<0.0001	0	0	1
P11	MCS	53	42	<b>79.2</b>	<0.0001	0	1	1
P12	MCS	54	41	<b>75.9</b>	<0.0001	0	0	1
P13	MCS	52	45	<b>86.5</b>	<0.0001	1	2	2

nontarget numbers were labelled as +1 and −1, respectively. Furthermore, the classification model was updated after each test block using the data from this test block. For example, we used the data from Block 2 to re-train/update the SVM model for the test in Block 3. We used this model training method because (i) the blocks were conducted on separate days for each patient; and (ii) the patients' statuses did not allow them to perform a training run before each block. For each test trial, the trained SVM was applied to the two feature vectors corresponding to the two buttons, and the predicted result corresponded to the button with the higher score.

For each patient, the online classification accuracy was calculated as the ratio of the number of trials with correct responses to the total number of presented trials. To assess the significance of the accuracy, we used a binomial test based on Jeffreys' Beta distribution in which the significance level in a two-class paradigm is calculated as follows [19, 20]:

$$\lambda = \left\{ a + \frac{2(N - 2m)z\sqrt{0.5}}{2N(N + 3)} \right\} + z\sqrt{\frac{a(1 - a)}{N + 2.5}} \quad (1)$$

where  $N$  is the number of actual trials (in this experiment,  $N = 60$ —the numbers of discard trials);  $m$  is the expected number of successful trials (in this study,  $m = N/2$  for a two-class problem);  $a$  is the expected accuracy (0.5 in this study);  $\lambda$  is the accuracy rate; and  $z$  is the  $z$ -score based on the standard normal distribution. At a significance level of 0.05 for a one-sided test,  $z$  is 1.65. Using this formula, for our two-class BCI, the accuracy rate  $\lambda$  corresponding to the significance level was 60.4% for 60 trials in this study. If given an accuracy rate  $\lambda$ , we can calculate  $z$  using this formula and obtain the  $p$  value based on the standard normal distribution.

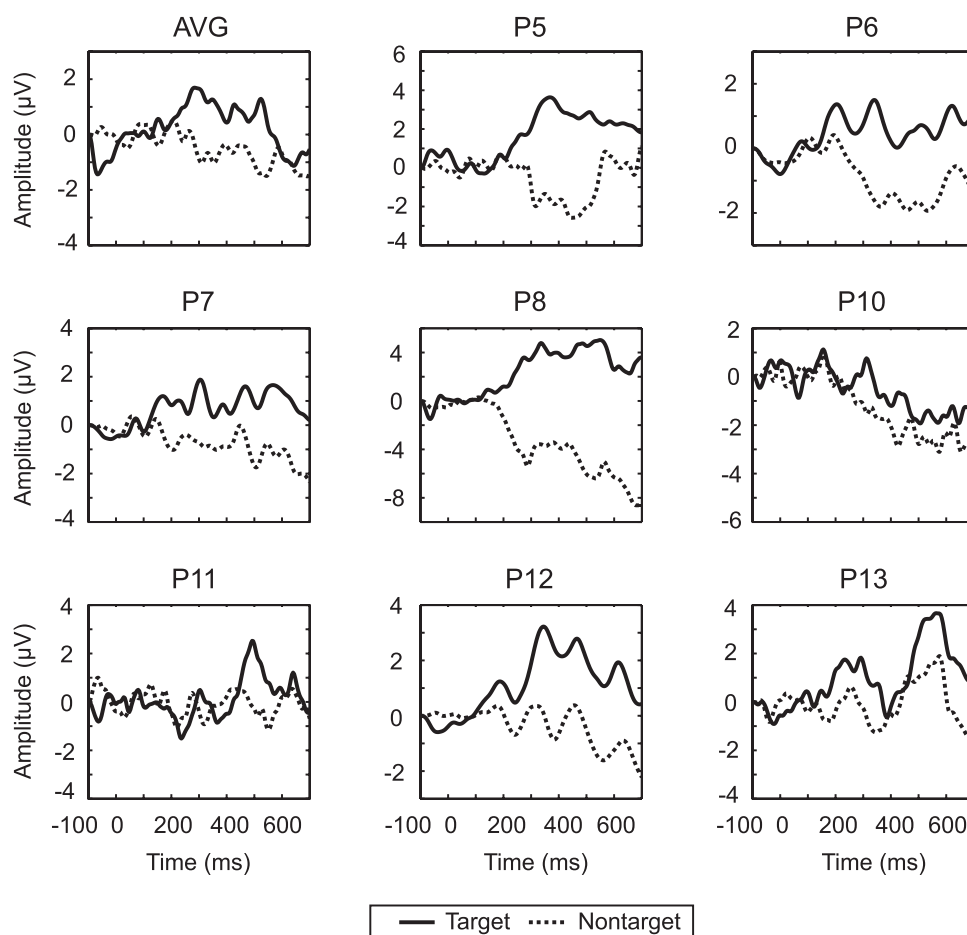
### 3. Results

Table 2 summarizes the online accuracy of the BCI experiment and the JFK CRS-R communication subscale scores for each patient. One MCS patient (P13) achieved a score of 1

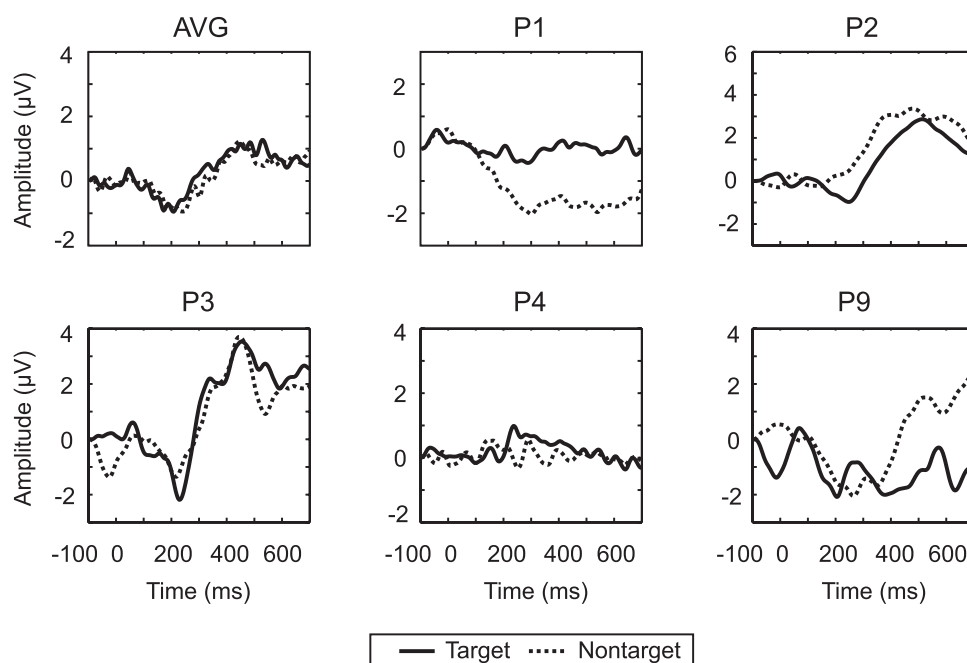
on the JFK CRS-R-based communication assessment and an accuracy of 86.5% in the BCI-based assessment. Five patients obtained unresponsive results (four VS and one MCS; P1, P2, P3, P4 and P9) in both the BCI-based and JFK CRS-R-based assessments. The other seven patients (four VS and three MCS; P5, P6, P7, P8, P10, P11 and P12), who all had scores of 0 on the JFK CRS-R communication subscale before the experiment, achieved accuracies ranging from 65.5% to 86%, which were significantly higher than the chance level ( $p < 0.05$ , the binomial test). In addition, the results of the two JFK CRS-R-based communication assessments after the experiment showed that four patients (P8, P10, P11 and P12) who achieved responsive results in the BCI-based assessment improved on the scores they obtained in the JFK CRS-R-based assessment before the experiment.

Based on whether the online classification accuracy of the BCI experiment was significantly higher than the chance level we classified the patients into two groups: the responsive group, which included patients P5, P6, P7, P8, P10, P11, P12 and P13, and the unresponsive group, which included the remaining patients. The average accuracy of the responsive group with standard deviation was  $77.3\% \pm 4.88$ . For the patients in the responsive group, the group-average ERP waveforms and each patient's waveform from the 'Cz' channel are shown in figure 3. The ERP waveforms of each patient were extracted by time-locked averaging of the EEG signal across all trials in the test run for each stimulus type, whereas the group-average waveforms were obtained by averaging the ERP waveforms of all patients in that group. As shown in figure 3, obvious P300 responses were observed for the target stimuli in the group-average ERP waveforms and in each patient's waveforms.

Figure 4 shows the group-average ERP waveforms and each patient's waveforms in the unresponsive group. In the group-average waveforms, P300 responses seemed to be elicited by both target and nontarget stimuli, but they could not be distinguished. Furthermore, for patients P1 and P4, no P300 responses were observed for the target stimuli. For



**Figure 3.** The group-average ERP waveforms and ERP waveforms for each patient in the responsive group of the BCI-based assessment from the 'Cz' channel. The solid curves correspond to the target stimuli, whereas the dashed curves correspond to the nontarget stimuli.



**Figure 4.** The group-average ERP waveforms and ERP waveforms for each patient in the unresponsive group of the BCI-based assessment from the 'Cz' channel. The solid curves correspond to the target stimuli, whereas the dashed curves correspond to the nontarget stimuli.

**Table 3.** Results of the comparison between the responsive and unresponsive groups in the BCI-based assessment based on clinical parameters.

Clinical parameters	Responsive group mean (SD)	Unresponsive group mean (SD)	<i>p</i> value (ANOVA)
Age (years)	41 (13.617)	31.2 (7.694)	0.174
Time since onset (months)	2.375 (1.217)	3 (3.102)	0.613
JFK CRS-R scores before the experiment	7.75 (1.643)	6.2 (1.643)	0.299
JFK CRS-R scores shortly after the experiment	12.375 (4.596)	6.2 (1.643)	0.016
JFK CRS-R scores two months after the experiment	14.625 (5.041)	6.2 (1.643)	0.004

patients P2, P3 and P9, P300 responses were elicited by both target and nontarget stimuli, but these responses could not be distinguished.

The results of the comparison between responsive and unresponsive groups based on several clinical parameters (age, time since onset, JFK CRS-R total scores) are reported in table 3. There were no significant differences between the responsive and unresponsive groups in age, time since onset and JFK CRS-R scores before the experiment. However, a significant difference was observed between the two groups in JFK CRS-R scores shortly after the experiment ( $p < 0.05$ ) and the JFK CRS-R scores two months after the experiment ( $p < 0.01$ ). This result implies that the patients in the responsive group recovered better than those in the unresponsive group.

#### 4. Discussion

BCI-based methods could evaluate DOC patients' responses to external stimuli based on brain signals instead of behaviors and may provide an assistive tool for clinical evaluation. Previous studies mainly focused on using BCI to detect the awareness of DOC patients [13–15]. These studies determined whether patients were aware by verifying whether the patients followed commands. However, BCI-based methods have not been used to assess the single consciousness function included in behavioral scales, such as the 6 subscales in the JFK CRS-R. In this paper, we proposed an audiovisual BCI-based communication assessment system to supplement the JFK CRS-R in testing DOC patients' communication ability. The combination of the JFK CRS-R- and BCI-based assessments may provide more sensitive and precise diagnosis results than the JFK CRS-R alone.

In our previous study [15], we proposed an audiovisual BCI system for awareness detection in patients with DOC based on semantically congruent audiovisual stimuli, including visual and spoken numbers. Our experimental results for healthy subjects showed that the audiovisual BCI outperformed the corresponding visual-only and auditory-only BCIs, whereas the experimental results for the patients demonstrated satisfactory performance of the system in awareness detection. In this study, we used an audiovisual BCI that was a variant of the BCI system described in a previous study [15], to assess communication in DOC patients. Thirteen patients (eight VS and five MCS) participated in our experiment. Among these patients, one MCS patient (Patient 13) who received a score of 1 in the JFK CRS-R communication subscale before the

experiment and a score of 2 in the JFK CRS-R-based assessment shortly after the experiment achieved an accuracy of 86.5% in the BCI experiment, which demonstrated the effectiveness of our BCI-based communication assessment system. The other twelve patients (eight VS and four MCS) received a score of 0 in the JFK CRS-R communication subscale before the experiment. However, seven of these patients (four VS and three MCS) achieved accuracies that were significantly higher than the chance level in the BCI-based assessment. This means that the seven patients could understand the examiner's situation questions and answer these questions through our BCI system.

In the JFK CRS-R communication subscale, a score of 1 indicates that the patient can answer some questions, and a score of 2 indicates that the patient can answer all questions correctly. Therefore, scores of 1 and 2 in the JFK CRS-R communication subscale are responsive results, and a score of 0 is an unresponsive result.

We now compare the sensitivity of the two assessment methods. First, we assumed that the responsive samples in the JFK CRS-R-based assessment were actual positives. The true positive rate (TPR) is often used to estimate the sensitivity of a test and can be defined as follows: the number of true positives (TP) / (number of true positives (TP) + number of false negatives (FN)). The number of total actual positive samples (TP + FN) is denoted by  $a$  and the number of TP in the BCI-based assessment is denoted by  $TP_b$ . Then, the TPR corresponding to the BCI-based assessment is  $TP_b/a$ , denoted by  $TPR_b$ . The number of TP in each JFK CRS-R-based assessment are denoted by  $TP_{c1}$ ,  $TP_{c2}$ , and  $TP_{c3}$ , and the TPR in each JFK CRS-R-based assessment is denoted by  $TPR_{c1}$ ,  $TPR_{c2}$ , and  $TPR_{c3}$ . Because  $TP_{c3} > TP_{c2} > TP_{c1}$  (see table 2),  $TPR_{c3} > TPR_{c2} > TPR_{c1}$ . Furthermore, all responsive results in the three JFK CRS-R-based assessments were detected by the BCI-based method (see table 2). Therefore, we had  $TP_b \geq TP_{c3} > TP_{c2} > TP_{c1}$ , and  $TPR_b \geq TPR_{c3} > TPR_{c2} > TPR_{c1}$ . However, as the number of responsive patients was small in this study, the statistical test for the above results was intractable. But we could still conclude that the discrepancy between the BCI- and JFK CRS-R-based assessments suggested a potentially higher sensitivity of the BCI-based assessment in consciousness detection.

The specificity of a test can be evaluated by the true negative rate (TNR), which is calculated as follows: the number of true negatives (TN) / (the number of true negatives (TN) + the number of false positives (FP)). The number of FP in the BCI- and JFK CRS-R-based assessments is denoted by  $FP_b$  and



$FP_c$ , and the TNRs in the BCI- and JFK CRS-R-based assessments are denoted by  $TNR_b$  and  $TNR_c$ . In this study, since we assumed that the responsive patients in the JFK CRS-R were true positives,  $FP_c = 0$ , and  $TNR_c = 1$ . Furthermore,  $FP_b \geq 0$ , thus  $TNR_c \geq TNR_b$ . However, because we had no evidence of FP in the BCI-based assessment, it is difficult to compare the specificity of the two methods.

Among eight patients in the responsive group in the BCI experiment, one patient (P13) obtained responsive results in the JFK CRS-R-based assessment before the experiment. The JFK CRS-R-based assessments shortly after and two months after the experiment confirmed that four patients (one VS and three MCS; P8, P10, P11 and P12) could communicate to some extent. Although three other patients (P5, P6 and P7) who achieved significant scores in the BCI-based assessment still had scores of 0 in the JFK CRS-R communication subscale, their total JFK CRS-R scale scores later improved. These results indirectly indicate the effectiveness of our BCI-based assessment. An interesting observation in this study is that all eight patients in the responsive group in the BCI experiment had improved JFK CRS-R scores after the experiment (as shown in tables 1 and 3). Specifically, two MCS patients (P10 and P12) from the responsive group of BCI experiment, who obtained unresponsive results in the JFK CRS-R-based assessments before and shortly after the experiment but responsive results in the JFK CRS-R two months after the experiment, recovered and were released from the hospital three month after the experiment. Furthermore, as shown in figure 3, the ERP waveforms of the patients in the responsive group showed that the target stimuli elicited obvious P300 responses, whereas the nontarget stimuli did not. For patients in the unresponsive group, no P300 response was elicited by the target stimuli (see figure 4). These results and those in table 3 show that our BCI paradigm might indicate the prognosis for DOC patients to some extent but requires further study.

Recent studies have explored the effectiveness of BCI technology as a communication tool for patients with amyotrophic lateral sclerosis (ALS) [21, 22]. For example, Sellers *et al* evaluated the effectiveness of a BCI operated by detecting a P300 elicited by one of four randomly presented stimuli (i.e. Yes, No, Pass and End) in three ALS patients and three controls. The offline analysis indicated that two of the three ALS patients' classification rates were equal to those achieved by the controls [22]. Nijboer *et al* reported that six ALS patients achieved mean online and offline accuracies of 62% and 82%, respectively, using a visual  $6 \times 6$  P300 speller [21]. Several studies have also reported successful attempts to restore communication with brain-computer interfaces for LIS patients [17, 23]. In Lulé *et al*'s study [17], eighteen patients (three VS, thirteen MCS and two LIS) were evaluated with a four-choice (Yes, No, Stop and Go) auditory oddball BCI. One LIS patient obtained a significant online accuracy of 60%, whereas the online accuracies for the other patients were not significantly higher than the chance level. De Massari *et al* proposed a two-class (Yes and No) paradigm based on the slow cortical potentials for communication [23]. The offline analysis results involving three LIS patients indicated that one patient obtained an accuracy (70%) significantly higher than

the chance level, whereas the other two did not. If the subjects achieved an overall accuracy higher than or equal to 80%, they were deemed proficient in the BCI [24]. In this study, an MCS patient obtained an accuracy higher than 80% in the BCI experiment, whereas the others did not. Therefore, only this patient could proficiently use the system for communication.

## 5. Conclusion

Overall, our experimental results demonstrate that the BCI-based method can provide some degree of correction for a doctor's clinical diagnosis and that a more precise evaluation could be obtained for DOC patients by combining the JFK CRS-R- and BCI-based assessments. Furthermore, one MCS patient (P13 in this study) could proficiently use this simple communication BCI system to exchange messages with the outside world. In the future, we will establish a practical communication system for this patient.

## Acknowledgments

This work was supported in part by the National Natural Science Foundation of China under Grant 61633010 and Grant 91420302, in part by the National Key Basic Research Program of China (973 Program) under Grant 2015CB351703, and in part by Guangdong Natural Science Foundation under Grant 2014A030312005.

## References

- [1] Seel R T *et al* 2010 Assessment scales for disorders of consciousness: evidence-based recommendations for clinical practice and research *Arch. Phys. Med. Rehabil.* **91** 1795–813
- [2] Jennett B and Plum F 1972 Persistent vegetative state after brain damage a syndrome in search of a name *Lancet* **1** 734–7
- [3] Giacino J T *et al* 2002 The minimally conscious state: definition and diagnostic criteria *Neurology* **58** 349–53
- [4] Haig A J, Katz R T and Sahgal V 1987 Mortality and complications of the locked-in syndrome *Arch. Phys. Med. Rehabil.* **68** 24–7
- [5] Monti M M, Laureys S and Owen A M 2010 The vegetative state *BMJ* **341** c3765
- [6] Schnakers C, Vanhaudenhuyse A, Giacino J, Ventura M, Boly M, Majerus S, Moonen G and Laureys S 2009 Diagnostic accuracy of the vegetative and minimally conscious state: clinical consensus versus standardized neurobehavioral assessment *BMC Neurol.* **9** 35
- [7] Giacino J T, Kalmar K and Whyte J 2004 The JFK coma recovery scale-revised: measurement characteristics and diagnostic utility *Arch. Phys. Med. Rehabil.* **85** 2020–9
- [8] Giacino J T, Schnakers C, Rodriguez-Moreno D, Kalmar K, Schiff N and Hirsch J 2009 Behavioral assessment in patients with disorders of consciousness: gold standard or fool's gold? *Prog. Brain Res.* **177** 33–48
- [9] Demertzi A, Vanhaudenhuyse A, Bruno M-A, Schnakers C, Boly M, Boveroux P, Maquet P, Moonen G and Laureys S 2008 Is there anybody in there? detecting awareness in disorders of consciousness *Expert Rev. Neurotherapeutics* **8** 1719–30

- [10] Andrews K, Murphy L, Munday R and Littlewood C 1996 Misdiagnosis of the vegetative state: retrospective study in a rehabilitation unit *BMJ* **313** 13–6
- [11] Childs N L, Mercer W N and Childs H W 1993 Accuracy of diagnosis of persistent vegetative state *Neurology* **43** 1465–7
- [12] Wolpaw J R, Birbaumer N, McFarland D J, Pfurtscheller G and Vaughan T M 2002 Braincomputer interfaces for communication and control *Clin. Neurophysiol.* **113** 767–91
- [13] Cruse D, Chennu S, Chatelle C, Bekinschtein T A, Fernández-Espejo D, Pickard J D, Laureys S and Owen A M 2012 Bedside detection of awareness in the vegetative state: a cohort study *Lancet* **378** 2088–94
- [14] Pan J, Xie Q, He Y, Wang F, Di H, Laureys S, Yu R and Li Y 2014 Detecting awareness in patients with disorders of consciousness using a hybrid braincomputer interface *J. Neural Eng.* **11** 056007
- [15] Wang F, He Y, Pan J, Xie Q, Yu R, Zhang R and Li Y 2015 A novel audiovisual brain-computer interface and its application in awareness detection *Sci. Rep.* **5** 9962
- [16] Li Y, Pan J, He Y, Wang F, Laureys S, Xie Q and Yu R 2015 Detecting number processing and mental calculation in patients with disorders of consciousness using a hybrid brain-computer interface system *BMC Neurol.* **15** 259
- [17] Lul D *et al* 2013 Probing command following in patients with disorders of consciousness using a braincomputer interface *Clin. Neurophysiol.* **124** 101–6
- [18] Belitski A, Farquhar J and Desain P 2011 P300 audio-visual speller *J. Neural Eng.* **8** 025022
- [19] Berrar D, Bradbury I and Dubitzky W 2006 Avoiding model selection bias in small-sample genomic datasets *Bioinformatics* **22** 1245–50
- [20] Noirhomme Q, Lesenfants D, Gomez F, Soddu A, Schrouff J, Garraux G, Luxen A, Phillips C and Laureys S 2014 Biased binomial assessment of cross-validated estimation of classification accuracies illustrated in diagnosis predictions *NeuroImage Clin.* **4** 687–94
- [21] Nijboer F *et al* 2008 A p300-based brain-computer interface for people with amyotrophic lateral sclerosis *Clin. Neurophysiol.* **119** 1909–16
- [22] Sellers E W and Donchin E 2006 A p300-based brain-computer interface: initial tests by als patients *Clin. Neurophysiol.* **117** 538–48
- [23] De Massari D, Ruf C A, Furdea A, Matuz T, van der Heiden L, Halder S, Silvoni S and Birbaumer N 2013 Brain communication in the locked-in state *Brain* **136** 1989–2000
- [24] He B, Baxter B, Edelman B J, Cline C C and Wenjing W Y 2015 Noninvasive brain-computer interfaces based on sensorimotor rhythms *Proc. IEEE* **103** 907–25